The Systematicity of the Sign: Modeling Activation of Semantic Attributes from Nonwords

Padraic Monaghan
Centre for Research in Human Development and Learning, Department of Psychology, Lancaster University, UK

Gary Lupyan
Department of Psychology, University of Wisconsin-Madison, USA

Morten H. Christiansen
Department of Psychology, Cornell University, USA

Abstract

There are numerous studies demonstrating that people's judgments about meanings of words can sometimes derive from their sound – a phenomenon often referred to as sound symbolism. A recent comprehensive assessment of English demonstrates that some small amount of systematicity exists between form and meaning. Is this small level of systematicity in language sufficient to drive the observed behavioral effects of sound symbolism? In this study we first tested the extent to which similarities amongst the sounds of words was sufficient to drive sound symbolic effects. We then tested whether a computational model that learned to map between form and meaning of English words better accounted for the observed behavior. We found that phonological similarity alone was sufficient to account for several effects of sound symbolism (without reference to meaning at all), but that the form-meaning mapping model was able to reproduce additional key behavioral effects of sound symbolism.

Keywords: form-meaning mappings, arbitrariness of the sign, brand names, word learning.

Introduction

The sound of a word bears an arbitrary relationship to its meaning, such that the phonological properties of the word generally carry no information about its intended meaning (de Saussure, 1916; Hockett, 1960). Arbitrariness in language provides a host of potential advantages for communication: It permits the speaker and listener to extract themselves from the immediate situation, allowing language to express concepts distant in space and time, as well as hierarchical, abstract relationships (Clark, 1998); arbitrariness promotes learnability, in that acquisition of a new form-meaning relationship is not constrained by previous learning (Gasser, Sethuraman, & Hockema, 2010) and also ensures that the amount of information present in the environment for learning, and communicating the mapping is maximized (Monaghan, Christiansen, & Fitneva, 2011); arbitrariness enables direct mapping from word to concept, which iconic symbols do not (Lupyan & Thompson-Schill, 2012); finally, arbitrariness ensures that duality of patterning – where the relationships between phonemes constitute a word, and then the relationships between words constitute meaning, independent of those phonemes – is sustained.

Yet, there are very many apparent exceptions to the rule of arbitrariness observed in spoken language. Morphology carries information about form-meaning mappings at a grammatical category level – even in languages with such impoverished systems of morphology as English, there are indicators in the spoken word about the general syntactic role of a word, in turn supporting inferences about semantic properties (Seidenberg & Gonnerman, 2000: Monaghan, Chater & Christiansen, 2005). Furthermore, there are many instances of sound-meaning relationships (Hinton, Nichols, & Ohala, 1994), such as phonoaesthemes (certain phonological clusters relating to meaning, e.g., for English, *spr-* relates to fast movement, or *–ump* relates to roundedness). Such phonoaesthemes have been shown to be significantly expressed in corpora of English (Otis & Sagi, 2008), and have processing consequences in word reading (Bergen, 2004). Given these mixed messages available from the literature, it is unclear precisely how arbitrary language is.

In a recent corpus analysis, Monaghan Shillcock, Christiansen & Kirby (2014) provided a first comprehensive assessment of the extent to which form and meaning is systematic or arbitrary in the vocabulary of English. They correlated the similarities between words in terms of their phonological form and determined the extent to which those phonological distances could predict distances in meaning space, where meaning was derived from two alternative representations. The first semantic representation was constructed on the basis of contextual co-occurrence vectors, similar to Latent Semantic Analysis (Landauer, Foltz, & Laham, 1998). The second was based on semantic features derived from WordNet (Miller, 1990). Both types of semantic representation resulted in significant but very small amounts of systematicity between form and meaning, even when morphology and etymology was controlled.

The actual systematicity between form and meaning is very small indeed, and, though intellectually intriguing, the small amounts of variance accounted for in the semantics by the phonology may be practically unimportant. However, studies of semantic attributes related to nonwords have typically adopted a forced-choice design, and this forced-choice may be sufficient to pick up on very subtle distinctions in semantic priming from sound.

One of the most famous examples of sound symbolism is the relationship between speech sounds and shape. Köhler (1929) showed participants two shapes – a rounded and a spikey object, and asked them which of the two nonwords *kiki* and *bouba* related to which object. In this frequently replicated study (see, e.g., Ramachandran & Hubbard, 2001; Maurer, Pathman, & Mondloch, 2006), participants typically map *kiki* to the angular object and *bouba* to the rounded object. In terms of the phonological properties of the nonwords, studies have demonstrated that both the phonological features of the vowel as well as of the consonants contribute to the effect (Monaghan, Mattock, & Walker, 2012; Nielsen & Rendall, 2013).

Such sound influences on meaning appear to generalize to numerous semantic attributes. In a series of studies on brand name choices, Klink has shown that nonwords varying in consonant and vowel properties relate to a whole range of brand decisions (Klink, 2000, 2003; Klink & Wu, 2013). In his most comprehensive study, Klink (2000) tested a small set of nonwords that varied in terms of whether they contained fricative or stop consonants, voiced or unvoiced consonants, or front or back vowels. A small set of nonwords that manipulated each of these phonological features was tested for the extent to which participants judged the nonword to be an appropriate brand name for promoting a variety of semantic attributes in a questionnaire study.

Klink (2000) found that nonwords containing front vowels were judged to be smaller, lighter, milder, thinner, softer, faster, colder, more bitter, more feminine, friendlier, weaker, and prettier than nonwords containing back vowels (e.g., *detal* versus *dutal*). He also found that fricatives were smaller, faster, more feminine and lighter than plosives (e.g., *fazz* versus *kazz*), and that unvoiced consonants were smaller, softer, faster, more feminine, lighter, and sharper than voiced consonants (e.g., *faruck* versus *varuck*).

Insofar as there exists some systematicity between certain sounds and certain semantic dimensions in English, a model that is trained to map between phonological and semantic representations of English vocabulary should be able to account for the meanings people attribute to various nonwords.

A further advantage of the computational model is that it permits testing of various phonological properties of words simultaneously, rather than measuring vowels or consonants features separately. In Klink's (2000) studies, he tested vowel position by contrasting nonwords containing the letters *i* and *e* (front) from those containing *o* and *u* (back). However, these vowels differ not only in position but also in height (high and low, respectively). Furthermore, Klink tested fricatives versus plosives by a comparison conflating unvoiced fricatives and voiced plosives, so several the stimuli differed on two dimensions. In addition, the combinations of certain consonants and vowels may drive participants' judgments rather than the individual properties of vowels or consonants. There are practical constraints in testing a large set of nonwords in behavioral studies, but ensuring that effects are generalizable is possible in a computational model tested on a large set of stimuli, which is not possible with behavioral studies. Thus, we investigated the simultaneous contribution of consonant manner and voicing and vowel position and height for their relationship to different semantic attributes.

However, the model results have to be considered in terms of whether observed behavioral effects are due to regularities in the form-meaning mappings, or whether they are due to other uncontrolled contributions to decisions about meaning of nonwords. One possibility is that the effects may just be due to phonological similarity between nonwords and the meaning of known words. If this is so, then behavioral effects could be captured by assessing the similarity between the phonology of nonwords and the phonology of semantic attributes directly, without any role of semantics in participants' decisions. We first describe how we assessed phonological similarity effects before presenting the model of form-meaning mappings.

Determining phonological similarity effects

Some observed effects of sound symbolism, in behavioral as well as computational studies, may be due, not to the systematicity that exists between form and meaning, but rather to analogies between the sound of the attribute word and the nonword being assessed. In experimental studies of the effect of sound symbolism nonword stimuli are sometimes, but by no means always, controlled for the extent to which they remind participants of existing words. In Klink's (2000) studies, for instance, stimuli that reminded participants of words were not used in the main study. However, the extent to which implicit associations between sounds of nonwords and sounds of attribute words may still be affecting performance. For example, asking participants about the extent to which a nonword elicits the attribute *cold* may be influenced by whether or not the nonword contains a plosive. If it does then it may resonate with the /k/ in the onset of *cold*, whereas if the nonword does not contain a plosive then it may be judged to be distinct, again solely based on comparisons in sound similarity.

Materials

In order to investigate the effects of phonological similarity on judgments about meaning, a corpus of nonwords was generated. We would have liked to have employed the precise materials from previous experimental studies of attribute selection from nonwords. However, these were unsuitable for our analyses, primarily because they tended to use polysyllabic words which are not available in the training set for the phonology to semantics model reported below, but also because the word sets were small – the power from these previous studies derives from the large number of participants used in the studies. Thus, we required a large set of nonwords to be tested, which would be impractical in a behavioral study but possible in a modeling context.

To create a large set of nonwords, all the single fricative/plosive onsets, vowels, and single phoneme codas from the set of 6229 words used to train the form-meaning model (see below) were selected. Then, these were joined together to form a candidate set of nonwords. These were then pruned by detecting any of the CVC sequences that were actual words, which were then omitted. This resulted in 2142 nonwords.

Each nonword was encoded for analysis in terms of its phonological properties. The manner of articulation (plosive or fricative) of the onset consonant, and whether the onset consonant was voiced or unvoiced were recorded. For the vowel, the position (front/back) and height (close/open) were recorded. Nonwords with close-mid, or open-mid vowels were classified as close and open, respectively. The characteristics of the nonwords are shown in Table 1.

Table 1: Number of nonwords with each phonological property.

	Manner		Voicing	
Consonants	Plosive	Fricative	Voiced	Unvoiced
	682	1460	1061	1081

Vowels	Position		Height	
	Front	Back	Close	Open
	1473	669	1541	601

Attributes were taken from the set used by Klink (2000, 2003), with antonym pairs derived for each attribute. Only antonym pairs that were both monosyllabic were included in the analysis. The derived pairs, corresponding to the original attributes used by Klink (2000) were: small-large, light-dark, mild-harsh, thin-thick, soft-hard, fast-slow, cold-warm, weak-strong, and sharp-blunt.

Testing

In order to determine the effect of phonological similarity, we measured the edit distance between the phonology of each nonword and each pair of attribute words, such that small values indicate the nonword was similar to the attribute word's phonology, and larger values indicate greater distinctiveness in the sound (see Monaghan, Christiansen, Farmer, & Fitneva, 2010, for discussion of this and similar measures of phonological similarity). Differences between the phonological similarity for each nonword and each antonym pair of attributes were entered as the dependent variable into the analysis.

To determine the effect of form-meaning mappings on judgments about semantic properties of nonwords, we next constructed a model that learned to map between phonology and semantics for a large set of English words.

Modeling form-meaning mappings

Architecture

The model was an extract of the connectionist triangle model developed by Harm and Seidenberg (2004). We implemented only one pathway from their model: the mapping from phonological input to semantic output. The model is shown in Figure 1. At input the model had 8 slots, and each slot contained 25 units to represent each phoneme. The input phonology layer was fully connected to a set of 1000 hidden units, which were in turn fully connected to an output semantic layer comprising 2446 units. The output semantic layer was connected to and from a set of 50 "cleanup" units which assisted in increasing the fidelity of the output semantic representation. The model differed from that of Harm and Seidenberg's (2004) model only in terms of the number of hidden units – the original model contained 500 hidden units, but we found in pilot simulations that this was inadequate for accurate learning.

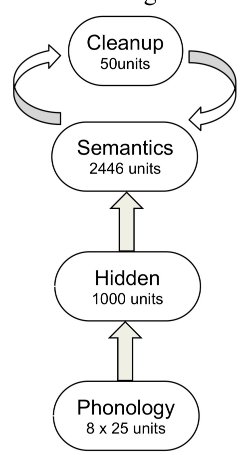

Figure 1: Model of form-meaning mappings.

Training

The model was trained on 6229 monosyllabic words. Phonology for each word was represented by 25 binary phonological features for each phoneme. There were 3 slots for the onset consonants, one slot for the vowel (diphthongs were represented in terms of a single vowel), and 4 slots for the coda consonants. Phonemes were always adjacent to the vowel, so that, if the word contained one onset consonant, that consonant occurred in the third slot for the word, and if the word contained two onset consonants, then those occupied the second and third slot.

For the semantic representations, words were encoded in terms of semantic features, derived from WordNet (Miller, 1990). Each word activated a subset of the 2446 features, and the target representation for each word in terms of semantic features was binary. Use of this semantic representation was appropriate in order to test the extent to which small amounts of systematicity in form-meaning mappings may drive behavioral judgments about semantic attributes of nonwords, because Monaghan et al. (submitted) tested exactly these semantic features in their corpus analysis.

The model was trained by randomly sampling words according to square root compression of frequency taken from the Wall Street Journal corpus (Marcus, Santorini, & Marcinkiewicz, 1993), presenting the phonological representation for the word at input, and requiring the model to learn to produce the semantic representation at the output, using the continuous recurrent backpropagation learning rule (Harm & Seidenberg, 2004). The model had 12 time steps in order to generate the semantics, at which point error was propagated through the network according to the cross-entropy of the difference between the target and actual semantic layer activations for that word. The learning rate was 0.2.

After 5M patterns, the model had reached asymptote in learning to produce the correct semantic output for 94.3% of the words, and we proceeded to test the model. Of the errors made by the model, they were generally low frequency words that were not effectively learned where no relationship between phonological or semantic form could be discerned, together with a small number of phonological errors (e.g., deigned → deign, seize → sea) and a small number of semantic errors (e.g., beers → ales, dross → waste).

Testing

In order to determine that the model was able to learn to map between phonology and semantics for the words in the training set, we determined for each word whether the output semantic representation was closer to the target semantic representation than to the semantics of any other word. Several words in the dataset were homophones (e.g., *beet, beat*), and if the word activated a semantic representation that was consistent with the phonology then this was also accepted as accurate.

The same set of nonwords, as described in Table 1, were presented to the model, and the output activations across the semantic layer were recorded. The semantic output for the nonword was compared against the target vector for the same set of semantic attributes as were tested in the phonological similarity analyses, using cosine distance. Differences between the cosine distance for the nonword and each word in the semantic attribute antonym pair was taken, and the difference scores were entered as the dependent variable into the analysis.

Results

We first report the results of the phonological analogical similarity for the extent to which they can account for the behavioral effects, before testing the form-meaning mapping model against the behavioral results. Finally, we report the effects of the form-meaning mapping model on the nonwords predicting preferences for semantic attributes when the phonological similarity effects are partialled out.

Phonological similarity effects

We ran linear mixed effects models using restricted maximum likelihood for difference scores for each attribute, with nonword as random factor, and vowel position, manner of the onset (fricative/plosive), and voicing of the onset as fixed factors. The dependent variable was the difference between the phonological similarity between each nonword and each semantic attribute antonym pair.

The results are shown in Table 2, which shows all the significant effects ($p < .05$) for each of the semantic attributes. In the results tables, bold indicates that the effect is significant and consistent with Klink (2000), gray indicates a significant effect in Klink (2000) and consistent but not significant effect in the model. * indicates an effect in Klink (2000) contradicted by the model, and normal font indicates a comparison not tested directly in the behavioral study, but significant in the model.

Table 2: Summary of the effects of phonological analogical similarity for comparisons of antonym pairs.

Attribute	Vowel position	Manner	Voicing	Vowel height
Smaller	front	**fricative**	**unvoiced**	
Lighter	front	**fricative**	**unvoiced**	
Milder	front			
Thinner	front			
Softer	front		unvoiced	
Faster	front	fricative	unvoiced	open
Colder	front			
Weaker	**front**			close
Sharper			**unvoiced**	

The results demonstrate that several of the sound symbolic semantic attributes reported in the behavioral literature may be due to similarity in terms of just the sound of nonwords and the attributes being assessed. This was particularly true of the properties of the consonants, where two of the three attributes associated with manner of articulation, and three of the five voicing effects, were replicated in the phonological similarity results. Only one of the eight vowel position effects was found to be significant in the phonological similarity results. However, vowel height was also found to be differently distributed in terms of phonological similarity for several of the semantic attribute antonym pairs.

Phonology to semantics effects in the model

The same linear mixed effects analyses as were applied to the phonological analogical results were tested on the output of the form-meaning mapping model. Table 3 reports all the significant effects ($p < .05$) resulting from the mixed models analyses, with the direction of the effect illustrated. For example, from the first row of the Table, for manner of the consonant onset, fricatives related more closely to the attribute *small* than to the attribute *large*, in comparison with plosives. Similarly unvoiced onset consonants related more closely to *small* than *large*, in comparison with voiced onset consonants. In addition, closed vowels related more closely to *small* than *large*, compared to open vowels.

As with the phonological similarity effects, several of the behavioral effects from Klink (2000) were replicated in the form-meaning mapping model, with particularly strong effects of vowel position, as well as vowel height, though not all effects were in the same direction as the behavioral observations.

Table 3: Summary of the effects of the phonology to semantics computational model for comparisons of antonym pairs.

Attribute	Vowel position	Manner	Voicing	Vowel height
Smaller	front	**fricative**	**unvoiced**	closed
Lighter	front*	fricative	unvoiced	open
Milder	**front**			closed
Thinner	**front**			open
Softer	front*		unvoiced	
Faster	front*	fricative	unvoiced	open
Colder	**front**			
Weaker	front*			open
Sharper			unvoiced	

Table 4: Effects of the phonology to semantics computational model with phonological similarity as a covariate.

Attribute	Vowel position	Manner	Voicing	Vowel height
Smaller	**front**	fricative	**unvoiced**	
Lighter	front*	fricative	unvoiced	open
Milder	**front**			close
Thinner	**front**			
Softer	front*		unvoiced	
Faster	**front**	fricative	unvoiced	close
Colder	**front**			
Weaker	front*			open
Sharper			unvoiced	

Phonology to semantics without phonological similarity effects

The final analyses tested the effects of the form-meaning mappings when the phonological similarity between nonwords and attributes was accounted for. This simulates the assumption in the study that participants are not influenced by the particular phonological similarity between a nonword and the given semantic attribute.

The same linear mixed effects models were run as before with vowel position, vowel height, consonant manner and consonant voicing as factors and the difference between the form-meaning mapping semantic activation related to each semantic antonym pair as the dependent variable, except with the addition of the difference in the phonological analogical similarity between the nonword and each pair of semantic antonyms as a covariate. The results are shown in Table 4.

Removing the effect of phonological similarity from the analyses enhanced the match between the computational model's results and the behavioral studies for the vowels – 5 of the 8 attributes were now aligned with the behavioral studies. However, the effects for consonants were weaker.

Discussion

There is a modicum of systematicity between form and meaning in the vocabulary of English. This paper aimed to determine whether capturing this systematicity in a model trained to map between form and meaning for a large set of English words was able to reproduce behavioral studies of apparent sound symbolic effects, where certain phonological features of nonwords affected judgments about a range of semantic attributes.

We first tested the extent to which similarities within the phonological forms of words alone could explain the behavioral effects. These analyses demonstrated that this was the case for many of the relationships between properties of consonants and semantic attributes. However, effects associated with the phonological features of vowels were not effectively simulated just by computing phonological similarity. Hence, sound symbolic effects are not due only to phonology.

Rather, several of the behavioral effects of the properties of vowels were reproduced only when the form-meaning regularities were taken into account in the simulations. Furthermore, determining the pure effect of the sound-meaning mappings, when the phonological similarity was partialled out of the analyses, resulted in the best fit between the behavioral and the modeling effects.

However, not all of the sound symbolic effects were predictable from our analyses. This could be either because the modeling did not accurately represent the psycholinguistic properties of the phonological or semantic similarity, or because these effects are due to sound symbolic decisions that are not encapsulated within the structure of the language. Using WordNet semantic features to represent semantic similarity is certainly imprecise, but this is more likely to result in an absence of effect, whereas some of the observed effects are reverse to the behavioral studies. Intriguingly, participants have to suppress the regularities within the language (as exhibited by the model) in order to make their decisions.

A further alternative, highlighted by the computational modeling presented here, is that *combinations* of phonological properties of nonwords may be driving decisions in sound symbolism studies. The modeling results show that there are many regularities in vowel height, as well as vowel position, that correspond to semantic distinctions, which have not yet been appraised in behavioral studies, and yet are conflated in the design of these studies. Furthermore, the properties of the nonwords' codas have not yet been systematically assessed, either in behavioral or our computational studies. This is a further likely source of influence on semantic judgments which merits further investigation. We do know, for instance, that vowels in first versus second syllables of a nonword affect performance differently (Klink & Wu, 2013). Thus, it is likely that further subtle patterns of interactions between phonological properties are the likely drivers of effects.

Our computational framework presented here provides a starting point for generating predictions about complex interactions between phonological properties resulting in sound symbolism decisions. We have shown that consideration of phonological similarity alone, as well as pure measures of form-meaning systematicity, are able to predict a range of behavioral studies of sound symbolic judgments. Testing the interactions in the computational model, rather than only, as we have so far done, testing main effects of individual phonological properties, will enable us to uncover the potent complexity of the systematicity of the sign.

Acknowledgements

We are grateful to Mike Harm and Mark Seidenberg for making the code for their 2004 model available online.

References

Bergen, B. (2004). The psychological reality of phonaesthemes. *Language*, 80, 291-311.

Clark, A. (1998). Magic words: How language augments human computation. In P. Carruthers & J. Boucher (Eds), *Language and thought: Interdisciplinary themes*, 162-183. Cambridge, UK: Cambridge University Press.

de Saussure, F. (1916). *Course in general linguistics*. New York: McGraw-Hill.

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

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

Hinton, L., Nichols, J., & Ohala, J. (Eds.). (1994). *Sound symbolism*. Cambridge: Cambridge University Press.

Hockett, C. F. (1960). The origin of speech. *Scientific American,* 203, 89-96.

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

Klink, R. R. (2003). Creating meaningful brands: The relationship between brand name and brand mark. *Marketing Letters, 14*, 143–157.

Klink, R.R. & Wu, L. (2013). The role of position, type, and combination of sound symbolism imbeds in brand names. *Marketing Letters*, 1-12.

Köhler, W. (1929). Gestalt psychology. New York, USA: Liveright.

Landauer, T., Foltz, P., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25,* 259– 284.

Lupyan, G. & Thompson-Schill, S.L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*. *141*, 170-186.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound–shape correspondences in toddlers and adults. *Developmental Science, 9*, 316-322.

Miller, G.A. (1990). WordNet: An on-line lexical database. *International Journal of Lexicography, 3,* 235-312.

Marcus, M., Santorini, B., & Marcinkiewicz, M.A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, *19*, 313-330.

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorization. *Cognition, 96,* 143–182.

Monaghan, P., Christiansen, M.H., Farmer, T.A. & Fitneva, S.A. (2010). Measures of phonological typicality: Robust coherence and psychological validity. *The Mental Lexicon, 5*, 281-299.

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

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

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*, 153-163.

Otis, K. & Sagi, E. (2008). Phonoaesthemes: A corpora-based analysis. In Love, B.C., McRae, K., & Sloutsky, V.M. (Eds.), *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*. Boston, MA.

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

Seidenberg M.S. & Gonnerman, L.M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences, 4,* 353-361.