Changing Signs: Testing How Sound-Symbolism Supports Early Word Learning

James Brand[1] j.brand@lancaster.ac.uk
Padraic Monaghan[1,2] p.monaghan@lancaster.ac.uk
Peter Walker[1,3] p.walker@lancaster.ac.uk
[1]Department of Psychology, Lancaster University, UK
[2]Department of Psychology of Language, Max Planck Institute for Psycholinguistics, NL
[3] Department of Psychology, Sunway University, MY

Abstract

Learning a language involves learning how to map specific forms onto their associated meanings. Such mappings can utilise arbitrariness and non-arbitrariness, yet, our understanding of how these two systems operate at different stages of vocabulary development is still not fully understood. The Sound-Symbolism Bootstrapping Hypothesis (SSBH) proposes that sound-symbolism is essential for word learning to commence, but empirical evidence of exactly how sound-symbolism influences language learning is still sparse. It may be the case that sound-symbolism supports acquisition of categories of meaning, or that it enables acquisition of individualized word meanings. In two Experiments where participants learned form-meaning mappings from either sound-symbolic or arbitrary languages, we demonstrate the changing roles of sound-symbolism and arbitrariness for different vocabulary sizes, showing that sound-symbolism provides an advantage for learning of broad categories, which may then transfer to support learning individual words, whereas an arbitrary language impedes acquisition of categories of sound to meaning.

Keywords: Sound-symbolism, Language learning, Vocabulary development, Word learning

Introduction

Words are a fundamental unit of human language, enabling the representation of specific meanings through the production of a discrete signal. But what determines how a particular word form should come to represent a specific meaning?

Such a question has been the topic of debate for many years, with Plato (390-370BCE/1971) questioning whether names themselves can be meaningful, or are just simply used to signify meaning. Some researchers have posited that any relationship between form and meaning is by necessity arbitrary (de Saussure, 1916; Hockett, 1960). Yet, more contemporary schools of thought have challenged such a view, putting forward substantial evidence that demonstrates non-arbitrariness is a fundamental part of human language (Dingemanse, et al. 2015; Monaghan et al., 2014; Perniss, Thompson & Vigliocco, 2010). Often referred to as sound-symbolism, this non-arbitrariness has been proposed to be crucial for language learners, as they can exploit the association between form and meaning, allowing the learner to ground their communicative system and bootstrap their way into language (Imai & Kita, 2014).

Consider a scenario where somebody is trying to communicate the meaning of ‘dog’ to an infant. The phonological sounds comprising the English word form *dog* offer very little information about the intended referent – it is an arbitrary mapping. In contrast, using a sound-symbolic form, such as *woof*, where the sound of the dog itself is iconically imitated, carries within the form itself referential information about what the intended meaning of the utterance is. Thus, the use of sound-symbolic forms, could function to aid the learning of specific form-meaning mappings, but also to allow the infant to begin to understand that word forms actually refer to things in the world around them (Spector & Maurer, 2009).

Recent evidence from corpus analyses has provided further support for the claim that sound-symbolism plays an important role in early language acquisition (Laing, 2014; Monaghan et al, 2014; Perry, Perlman & Lupyan, 2015). Such studies have demonstrated that non-arbitrariness is found predominately in the words acquired earliest in vocabulary development, suggesting that these types of words have a privileged status within the vocabulary, and potentially indicate that sound-symbolism is essential for the process of language acquisition.

Experimental studies have also contributed to the evidence that sound-symbolism benefits learning. Such studies have demonstrated how learning of adjectives (Nygaard, Cook & Namy, 2009), verbs (Kanartzis, Imai & Kita, 2011; Imai et al., 2008) and nouns (Maurer, Pathan & Mondloch, 2006; Nielsen & Rendall, 2012), are all boosted when the form-meaning mapping is sound-symbolically congruent. However, the experimental designs implemented in these studies tend to only test learning of categorical distinctions, i.e. where the meanings are semantically distinct from one another, as was the case in Köhler’s (1947) classic bouba/kiki experiment, or occur in antonymic pairs. So, any claims relating to the benefits of sound-symbolism for vocabulary acquisition rather than categorisation, i.e., where word meanings are *not* drawn from the same category, are not necessarily supported.

Addressing this issue, Monaghan, Mattock, and Walker (2012) designed an experiment that assessed whether sound-symbolic mappings could aid not only learning of categorical distinctions, but also individuated meanings. Participants had to learn 16 form-meaning mappings that were either sound-symbolically congruent or incongruent for two different trial types – where the meanings were distinct from one another (categorical learning) and where the meanings were closely related (individual word learning), see Figure 1. Their results

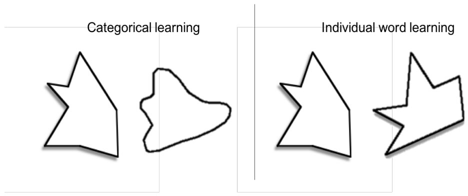

Figure 1. Example of categorical learning trial, where the two meaning are drawn from different categories (angular/round) and individual word learning, where the two meanings are from the same category (round/round).

showed a benefit for sound-symbolism for categorical learning trials, consistent with previous reports, but no benefit when participants had to identify meanings drawn from within the same category, i.e. individual word learning. This indicated that sound-symbolism may only be beneficial for more broad learning purposes.

But does the same apply for early language acquisition? The Sound-Symbolism Bootstrapping Hypothesis (SSBH) (Imai & Kita, 2014), claims that sound-symbolism should benefit word learning at the earliest stages of vocabulary development. In a computational model of learning by Gasser (2004), the size of the vocabulary was shown to determine whether a sound-symbolic or an arbitrary system was optimal for learning. Consistent with the SSBH, sound-symbolism was shown to be beneficial when the vocabulary size was small, but arbitrariness was preferred for a larger vocabulary size.

Indeed, the importance of arbitrariness in language has been highlighted in a series of corpus, computational and behavioural studies by Monaghan, Christiansen, and Fitneva (2011). They reported that when form-meaning mappings are arbitrary, the learner will benefit from the ability to individuate meanings from one another, by reducing ambiguity in expression. Take for example the words *cog* and *dog*, although phonologically similar they are semantically quite distinct, and so unlikely to occur in similar contexts, whereas *dog* and *wolf* are semantically related, but phonologically distinct. Thus, in an arbitrary system with a large vocabulary size, individual meanings can be learnt and communicated much more efficiently than a fully sound-symbolic system by maximising the distinctiveness of context and phonological form.

Importantly however, Monaghan et al.'s (2011) arbitrary advantage was only observed when there was a contextual cue present. Only when there was a systematic cue that occurred with the word form, such as a marker word that identified the syntactic category of the word meaning (e.g. noun/verb), would learning outperform a sound-symbolic language. Such contextual cues have been found to occur in natural language in the form of statistical regularities, where words from the same grammatical category share certain phonological properties, which has been shown to aid word learning and processing (Farmer, Christiansen & Monaghan, 2006; Monaghan, Christiansen & Chater, 2007).

Further support for the changing role of sound-symbolism came from Brand, Monaghan and Walker (2017), who modified the experimental design of Monaghan et al. (2012) to investigate whether manipulating the vocabulary size in the experiment would reveal a learning advantage for sound--symbolic mappings for individual word learning, and indeed, that is what they found. Confirming Gasser's (2004) computational modelling predictions, congruent sound-symbolism facilitated learning of individual words over incongruent mappings, but only when the vocabulary size was small. When it was large, only categorical learning but not individual word learning was found to benefit from congruent sound-symbolism. This indicated a division of labor for the role of sound-symbolism in the structure of the vocabulary – operating early on in acquisition to help bootstrap the acquisition of individual vocabulary items, but then as the vocabulary grows, functioning to aid the learning of broader categories in the language.

However, in these artificial language learning studies, mappings that are either congruent or incongruent with sound-symbolism are typically the only mappings tested. Thus, tests of the SSBH are required to compare a fully sound-symbolic language to a fully arbitrary language. Here, we present two experimental studies that build on Brand et al.'s (2017) paper, by assessing how learning of individual meanings and categories of words differs in languages that adopt i) an entirely sound-symbolic system and ii) an entirely arbitrary system, within three different vocabulary sizes – small, medium and large. This will allow for a direct comparison to be made between the two systems and offer an insight into how vocabulary size influences the way the languages are learnt.

Moreover, this approach will allow us to re-examine the extent to which previously reported claims about sound-symbolism and categorisation hold up when compared to a fully arbitrary condition. This will address concerns that the salient distinction present in the meanings (i.e. angular and rounded shapes), is in fact driving the categorisation effects found for sound-symbolism. Likewise, we can also investigate the effects of arbitrariness on word learning, by testing whether the increased distinctiveness in the sound space, which should make identification of the referent more efficient, provides a learning advantage over sound-symbolism when the vocabulary size is large.

We first hypothesise that sound-symbolism will promote the learning of categories within the language - more so than arbitrariness - across all vocabulary sizes, confirming the results from congruent and incongruent sound-symbolism mappings. Second, these studies enable a more specific version of the SSBH to be specified – whether bootstrapping into language from sound-symbolism occurs from initially supporting categorisation then vocabulary acquisition, or whether vocabulary acquisition precedes category learning, or whether individual vocabulary items are benefited by the systematicity of the language. Finally, we predict that as the vocabulary size grows, arbitrariness will provide a more suitable system for learning individual words, consistent with Gasser's (2004) results.

Experiment 1: Sound Symbolic Language

Method

Participants

Seventy-two participants took part in the experiment (48 female), with 24 assigned to each vocabulary size condition. Participants were undergraduate students from Lancaster University, with a mean age of 19.4 years (SD = 1.57, range 18-26). It was not required that participants spoke English as their first language (English first language speakers: n = 50), but all participants spoke English competently.

Materials

The same visual and auditory stimuli were used in the present experiment as those used in Brand et al. (2017) and Monaghan et al. (2012). This comprised an inventory of 16 visually presented shapes (8 angular and 8 rounded) and 16 auditorily presented non-words (all monosyllabic and had a CVC structure, with 8 non-words generated using plosive consonants and 8 using continuant consonants.

Each of the non-words was reliably mapped to one of the shapes. Using these mappings, we generated 3 different artificial languages, each of which differed in the number of mappings it used. This was either a small vocabulary (8 mappings), medium (12 mappings) or large (16 mappings), all with an equal number of angular and rounded shapes, drawn from the inventory of 16 shapes and non-words. Critically, in the present experiment all form-meaning mappings were presented to reflect only a congruent sound symbolic relationship (unlike Brand et al. (2017) where a mixture of congruent and incongruent mappings was used).

This meant that all sounds with continuant consonants were mapped exclusively to rounded shapes, likewise all sounds with plosive consonants were mapped exclusively to angular shapes. Which shape the sound was mapped to during the experiment was randomly selected from the set of rounded/angular shapes.

Procedure

A cross-situational learning paradigm was used during the experiment, where two of the visual stimuli would appear on the screen per trial, accompanied by one of the auditory stimuli, which reliably mapped onto one of the presented shapes throughout the experiment. After hearing the word, participants had to choose which image they believed the word was referring to, by pressing '1' for the image on the left, or '2' for the image on the right.

Each mapping was presented 4 times throughout the experiment, over the course of 4 blocks. Trials were designed to assess two distinct learning scenarios: learning of broad categories (by presenting two images that differed in their shape, such as one rounded and one angular), or learning of individuated meanings (by presenting two images from within the same shape category, such as two angular shapes), see Figure 1 for an example of each trial type.

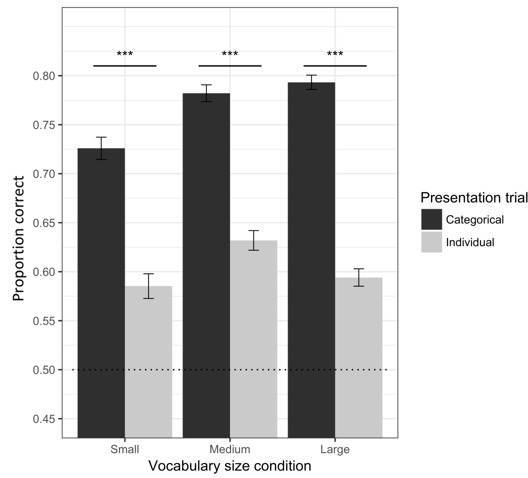

Figure 2. Proportion of correct responses for categorical and individual word learning trials for Experiment 1. Error bars show SEM. *** p < .001. Dotted line shows 50% chance level

Performance was measured online throughout the experiment, with participants responding accurately when they chose the image the sound was being reliably mapped to. No feedback was given at any point during the experiment.

Results and discussion

A series of generalized linear mixed-effects models (glmer) were performed on the data, predicting the dependent variable of response accuracy (correct or incorrect). The models were built up incrementally, adding in fixed effects and performing likelihood ratio tests after the addition of each new fixed effect term (following Barr, Levy, Scheepers & Tily, 2013). Random effects of participant and sound stimuli were included in all reported analyses.

The inclusion of the block effect significantly improved model fit ($\chi^2(1) = 136.36$, $p < .001$), indicating that accuracy increased significantly over the course of the experiment. Further analyses using one-sample t-tests showed that for all experimental conditions, performance was significantly above the 50% chance level at the last block of the experiment. The inclusion of vocabulary size did not significantly improve model fit ($\chi^2(2) = 2.570$, $p = .278$), indicating that there were no significant differences in accuracy for the different vocabulary size conditions. There was a significant improvement to model fit when presentation type (category or individual word learning) was added, ($\chi^2(1) = 500.930$, $p < .001$), indicating that accuracy was significantly higher for categorical than individual word learning trials.

Finally, the inclusion of the interaction term vocabulary size x presentation type significantly improved model fit ($\chi^2(4) = 17.529$, $p = .002$). In line with our first hypothesis, there was a significant change in accuracy for different vocabulary sizes as the presentation type (categorical or individual) varied. Additional analyses on the separate vocabulary size conditions, revealed that this difference between categorical and individual word learning was present in all three vocabulary sizes but increased in the larger vocabularies (small: $\chi^2(1) = 73.394$, p <.001; medium: $\chi^2(1) = 137.73$, p <.001; and large: $\chi^2(1) = 310.94$, p <.001), with no interaction with the effect of block (all *p*'s > .05), see Figure 2 for results.

The results presented in this experiment demonstrate a clear advantage for learning to distinguish between distinct categories, when the language comprised only sound-symbolic mappings. This is consistent with previous reports that suggest sound-symbolism, facilitates the learning of broad categorical boundaries (Farmer et al., 2006; Monaghan et al., 2011, 2012).

Experiment 2: Arbitrary Language

Method

Participants

Seventy-two participants took part in the experiment (45 female), with 24 assigned to each vocabulary size condition. Participants were undergraduate students from Lancaster University, with a mean age of 19.2 years (SD = 2.13, range 18-27). It was not required that participants spoke English as their first language (English first language speakers: n = 45), but all participants spoke English competently.

Materials

The same set of 16 angular/rounded shapes that were used in Experiment 1 were used as the visual stimuli for the present experiment. For the auditory stimuli however, a new set of 16 non-words were created. This was done in order to remove any possible relationship between sound and meaning, be it congruent or incongruent. To achieve this, we first generated a new inventory of auditory non-words. Thirty monosyllabic CVC non-words were recorded by the same native English speaker who recorded the auditory stimuli for Experiment 1. The non-words in this inventory were created using consonants from a set including plosives (/g/, /d/, /p/), continuants (/m/, /n/, /l/) and fricatives (/f/, /v/, /s/), with contrasting consonants being used for each word (i.e. a plosive would be used in onset position, but only a continuant or fricative would be used in coda position). Additionally, one of five vowels (/ʌ/, /ɛ/, /i/, /ɒ/, /a/) was used in the non-word, with a total of 30 non-words generated, all of which were intended to have no dominant phonological property associated with shape (in contrast to the stimuli used in Experiment 1). To ensure that there were no differences between the acoustic properties of the recorded sounds, the properties of intensity, fundamental frequency (pitch), first, second and third formants were normalized using Praat (Boersma & Weenink, 2015), this was consistent with the properties of the auditory stimuli used in Experiment 1.

Data were then collected from a short norming questionnaire, where 22 additional Lancaster University undergraduate participants were presented with the auditory non-words over a pair of headphones, along with a 7-point Likert scale anchored at each end by rounded and angular shapes. Participants heard all 30 of the newly generated non-words, in addition to the 16 non-words presented during

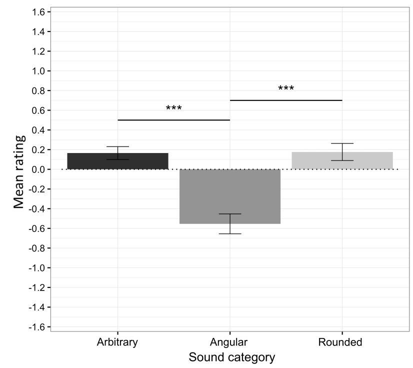

Figure 3. Mean ratings for presentation of auditory stimuli designed to incorporate either rounded or angular phonetic characteristics (Experiment 1) and when there is no intended relationship (Experiment 2). Positive values represent preference for rounded shapes, negative values represent preference for angular shapes. Error bars show SEM. *** p<.001.

Experiment 1, allowing us to make a comparison between the two sets of stimuli.

Participants were asked to rate each of the sounds based on how strong they felt it corresponded to either the rounded or angular shapes. This was done by selecting '0' for no correspondence, '1' for a weak correspondence, '2' for a slightly strong correspondence or '3' for a strong correspondence. Based on the mean ratings for the newly generated non-words, the 16 rated closest to 0 were selected.

To assess whether non-words were rated differently from each other, the 16 new non-words were compared to both the plosive (angular) and continuant (rounded) non-words used in Experiment 1. We ran mixed-effects models with sound category (angular/rounded/no relationship) as a fixed effect, and questionnaire response as the dependent variable. The addition of sound category to the model significantly improved the fit ($\chi^2(2) = 23.634$, $p < .001$), with a significant difference between angular and rounded sounds (estimate = .73, t = 4.882, p < .001) and angular and no relationship sounds (estimate = .72, t = 5.550, p < .001), however there was no significant difference between the rounded and no relationship sounds (estimate = .01, t = .088, p <.931), see Figure 3 for results. This final set of 16 non-words were then mapped randomly to one of the 16 rounded or angular shapes.

Procedure

The procedure was identical to that of Experiment 1.

Results and discussion

Following the same analysis as used in Experiment 1, a sequence of glmer models were fitted to our data to predict response accuracy. The inclusion of block significantly improved model fit ($\chi^2(1) = 58.336$, p <.001), indicating that over the course of the experiment, participants were improving the accuracy of their responses. Further analyses using one-sample t-tests showed that for all experimental conditions, performance was significantly above the 50% chance level at the last block of the experiment. The inclusion of vocabulary size did not significantly improve model fit ($\chi^2(2) = 3.917$, p <.141), indicating that there were no

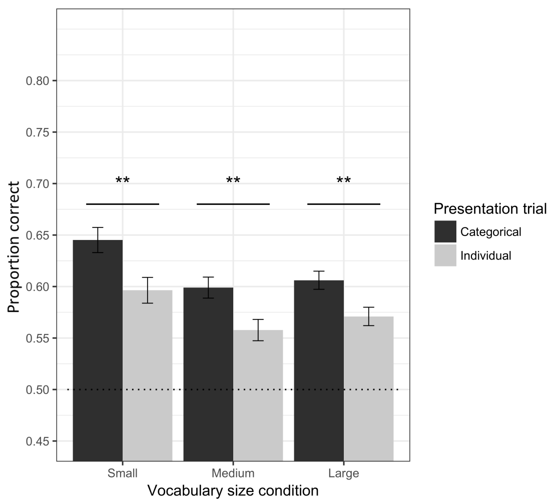

Figure 4. Proportion of correct responses for categorical and individual word learning trials for Experiment 2. Error bars show SEM. ** p < .01. Dotted line shows 50% chance level

significant differences in overall accuracy across vocabulary size conditions. The inclusion of presentation type (categorical/individual) significantly improved model fit ($\chi^2(1) = 23.520$, $p < .001$), indicating that accuracy on categorical trials was higher than individual learning trials. This was the case for each of the three vocabulary size conditions, with small, medium and large conditions showing significantly higher accuracy for categorical trials (small: $\chi^2(1) = 7.740$, $p < .005$; medium: $\chi^2(1) = 8.572$, $p < .003$; and large: $\chi^2(1) = 8.167$, $p < .004$), see Figure 4 for results). However, we did not find a significant improvement to model fit when the interaction term vocabulary size x presentation type was added ($\chi^2(4) = 4.520$, $p = .340$), indicating that as vocabulary size varied, the differences in accuracy for categorical and individual learning trials were stable.

Across Experiment Comparison

The presence of the two-way interaction between vocabulary size and presentation type in the first experiment but not the second experiment suggests that the SSBH may apply to learning by promoting both categorisation and individual word learning in the initial small vocabulary, but being beneficial primarily for categorisation as the vocabulary size increases. In contrast, the arbitrary language shows no relative change in categorisation and individuation as the vocabulary grows. In order to examine the results of the two experiments together, we conducted further analyses on a combined version of both datasets, using glmer. The addition of experiment as a fixed effect revealed a significant improvement to model fit ($\chi^2(2) = 30.66$, $p < .001$), with accuracy in the fully congruent condition being significantly greater than the arbitrary condition. Furthermore, the addition of the three-way interaction between vocabulary size x presentation type x experimental condition also significantly improved model fit ($\chi^2(8) = 22.16$, $p = .005$). In follow up analyses on the categorical trials, the inclusion of experimental condition significantly improved model fit for all vocabulary sizes (small: $\chi^2(1) = 4.13$, $p = .042$; medium: $\chi^2(1) = 26.85$, $p < .001$; large: $\chi^2(1) = 24.35$, $p < .001$), indicating that accuracy was higher for the fully sound-symbolic condition, when compared to the arbitrary condition. See Figure 5 for results.

In the analyses on the individual word learning trials, the inclusion of experimental condition did not significantly improve model fit for the small or large vocabulary sizes (both *p*'s > .05), indicating that there was no difference in accuracy between the fully congruent and arbitrary conditions. However, there was a significant improvement to model fit for the medium vocabulary size ($\chi^2(1) = 6.75$, $p = .009$), indicating that accuracy was higher for the fully congruent condition, when compared to the arbitrary condition. See Figure 5 for results.

General Discussion

We conducted two novel experiments that examined the role of sound-symbolic and arbitrary languages for different sizes of vocabulary, to mimic different stages of language learning. We implemented a design that allows us to investigate whether these two types of form-meaning mappings may be beneficial for learning individual vocabulary items, or whether the source of the sound-symbolic advantage for language acquisition resides in supporting sound-category mappings to be forged.

Furthermore, we have applied an experimental methodology that examines a purely arbitrary set of form-meaning mappings, addressing issues concerning previously reported effects of sound-symbolic and arbitrary mappings being driven by congruent versus incongruent sound-symbolic comparisons. However, even when using this entirely arbitrary set of form-meaning mappings, the key effects reported in the previous literature are still found here.

Our results demonstrate a clear advantage for learning to distinguish between categories when there is a sound-symbolic relationship present in the form-meaning mapping, an effect that is significantly reduced in the arbitrary

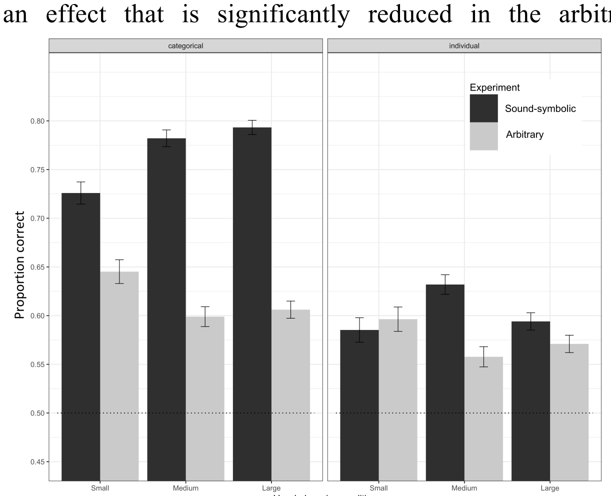

Figure 5. Proportion of correct responses for Experiment 1 (fully congruent) and Experiment 2 (no relationship). Error bars show SEM. Dotted line shows 50% chance level.

language. This further indicates that the observed effects of sound-symbolism on category learning, presented here and in other previous studies, was not simply an artefact of the presence of a salient distinction between visual stimuli (in our experiments, the distinction was two contrasting shape categories, angular and rounded), because we observed a dramatic difference in learning when the sound-symbolism was present in the language compared to when it was not.

Furthermore, even though there were greater potential distinctions available in the sound space of the arbitrary language, which could make identification of the intended mapping easier, we saw no advantage in learning when this distinctiveness in sound was present. For smaller vocabulary sizes, the distinctiveness would be even greater for the arbitrary language, and yet there was no change in the learning advantage with vocabulary size for this language. Thus, the learning advantage is inherent in the mappings between modalities, rather than within the structure of the visual or the sound space.

Critically, we observed no effect of arbitrariness for individual word learning, contrary to the results reported by Gasser (2004). One possible explanation for this contrast could come from Monaghan et al's (2011) results, which indicate that if there is to be an arbitrary advantage, then some systematicity is necessary for the effect to be observed. Future research could aim to assess the conditions under which sound-symbolism and arbitrariness co-exist in form-meaning mappings to provide optimal division of labor as the vocabulary develops and grows.

The results presented here are consistent with the SSBH: sound-symbolism does improve learning. But our results invite greater specificity for this hypothesis. We suggest that sound-symbolism bootstraps language learning by initially promoting categories of sound-meaning to be formed, which can then precipitate vocabulary acquisition. Based on our data, and on the results of other sound-symbolic language learning studies, we cannot say for certain whether sound-symbolism is essential for language to be acquired, but we have shown that it benefits the grounding of broad distinctions in word forms to be linked to general referential categories, thus extending our understanding of the mechanisms involved in the acquisition of our first words, both in first and second language acquisition.

Acknowledgments

This work was supported by the International Centre for Language and Communicative Development (LuCiD) at Lancaster University, funded by the Economic and Social Research Council (UK) [ES/L008955/1].

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3), 255-278.

Boersma, P. & Weenink, D. (2009). *Praat: doing phonetics by computer*, Version 6.0.07 [Computer program]. Retrieved December 12, 2014, from http://www.praat.org

Brand, J., Monaghan, P., & Walker, P. (2017). The Changing Role of Sound-Symbolism for Small Versus Large Vocabularies. *Cognitive science*.

De Saussure, F. (1916). *Course in general linguistics*. New York: Columbia University Press.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, 19(10), 603–615.

Farmer, T.A., Christiansen, M.H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *PNAS*, 103, 12203-12208.

Gasser, M. (2004). The origins of arbitrariness in language. In *Proceedings of the Annual Conference of the Cognitive Science Society* (p. 434 - 439). Lawrence Erlbaum.

Hockett, C. (1960). *The origin of speech*. Scientific American. 203, 88-96.

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

Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. *First Language*, 34(5), 387-405.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound–shape correspondences in toddlers and adults. *Developmental Science*, 9(3), 316–322.

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive psychology*, 55(4), 259-305.

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(3), 325–347.

Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(5), 1152-1164.

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130299.

Nielsen, A., & Rendall, D. (2012). The source and magnitude of sound-symbolic biases in processing artificial word material and their implications for language learning and transmission. *Language and Cognition*, *4*, 115–125.

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009a). Sound to meaning correspondences facilitate word learning. *Cognition*, 112(1), 181–186.

Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, 1, 227.

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(9), e0137147.

Plato (1971). *Cratylus*. In Hamilton, E. & Cairns, H. (Eds.), The collected dialogues of Plato, including the letters. Vol. 71. Princeton University Press.

Spector, F., & Maurer, D. (2009). Synaesthesia: a new approach to understanding the development of perception. *Developmental Psychology*, 45(1), 175.