Title

Eye Movements Reveal Sensitivity to Sound Symbolism
Early and Late in Word Learning

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors

Revill, Kate Pirog
Namy, Laura L
Nygaard, Lynne C

Publication Date

2015

Eye Movements Reveal Sensitivity to Sound Symbolism Early and Late in Word Learning

Kate Pirog Revill (krevill@emory.edu)
Facility for Education and Research in Neuroscience, 36 Eagle Row
Atlanta, GA 30322 USA

Laura L. Namy (lnamy@emory.edu)
Department of Psychology, 36 Eagle Row
Atlanta, GA 30322 USA

Lynne C. Nygaard (lnygaar@emory.edu)
Department of Psychology, 36 Eagle Row
Atlanta, GA 30322 USA

Abstract

Although the relationship between sound and meaning in language is arbitrary, reliable correspondences between sound and meaning have been found in natural language. These sound symbolic relationships affect word learning, but less is known about how sound symbolism affects online processing during learning or for well-learned stimuli. We use the visual world paradigm and an artificial lexicon featuring carefully controlled sound symbolic correspondences to examine the effects of sound symbolism on the online processing of novel and well-learned stimuli. Initially, participants chose novel shapes matching the sound symbolic properties of the word above chance, reliably fixating consistent shapes around word offset. As learning approached ceiling, accuracy and reaction time differences between matching and mismatching stimuli disappeared but a disadvantage in the online processing of mismatching stimuli persisted in the form of lagging target fixations. This suggests that sound symbolism affects the online processing of spoken stimuli even for well-learned words.

Keywords: sound symbolism; eyetracking; visual world paradigm; artificial lexicon

Introduction

Despite the apparent arbitrariness of the relationship between words and their meanings, both historical and recent evidence suggests that non-arbitrary correspondences between linguistic structure and categories of meaning exist in natural language, and that language users are sensitive to these correspondences (Köhler, 1947; Maurer et al., 2006; Nygaard et al, 2009; Ramachandran & Hubbard, 2001; Revill et al., 2014; Sapir, 1929). For example, Maurer, Pathman, and Mondlock (2006) found that both adults and children (2.5-year-olds) readily associated nonwords such as 'maluma' and 'bouba' with round, amoeboid shapes and words such as 'kiki' and 'takete' with sharp, spiky shapes (see also Köhler, 1947; Ramachandran & Hubbard, 2001). Similarly, Sapir's (1929) classic study demonstrated that adults reliably judged the nonword 'mal' to refer to large objects and the nonword 'mil' to refer to small objects. These sound-to-shape biases have been demonstrated across many languages and cultures (Bremner et al., 2013) and across development (Maurer et al., 2006). These types of reliable correspondences between sound and meaning have been dubbed *sound symbolism*.

Correspondences between phonological form and grammatical or semantic class have been shown to facilitate spoken sentence and word processing (Farmer, Christiansen & Monaghan, 2006; Reilly et al, 2012). These correspondences have also been found to benefit learning. For example, Nygaard et al. (2009) taught native English speakers the Japanese translations of English antonyms. Learners responded more quickly and accurately when the Japanese words were paired with their true English equivalents during training than when they had been paired with a mismatched meaning. However, to date, little work has examined the consequences of sound-to-meaning correspondences for *online* processing during word learning or for the subsequent lexical access of well-learned words. To address this question, we use the visual world paradigm in which fixation duration and latency on the visual referent of a word and its competitors can be used as implicit measures of real-time lexical processing (Creel, Tanenhaus, & Aslin, 2006; Revill, Tanenhaus, & Aslin, 2008).

If sound symbolism facilitates online lexical and semantic processing, listeners should more rapidly fixate potential referents when the objects possess visual characteristics consistent with the sound symbolic auditory features of the words. This study investigates the extent to which visual orienting to objects is influenced by the sound symbolic characteristics of novel labels, both at initial presentation and as learning approaches ceiling. More specifically, we investigated the effects of sound symbolic mappings when sound properties match (e.g., round labels paired with rounded objects) or mismatch (e.g., round labels paired with pointy objects) listeners' off-line judgments. We used an artificial lexicon paradigm in which language users acquired a novel lexicon by learning label-object pairings over the course of a brief training session (e.g., Revill et al., 2008). An artificial lexicon allows us to precisely manipulate the correspondence between the auditory or linguistic properties of object labels and the visual properties of object referents in order to evaluate the role of sound to meaning correspondences in processing.

Materials & Methods

Participants

Twenty four members of the Emory University community took part in the study (8M/16F, age 21.7±3.5). Data from two participants were excluded due to failure to comply with task instructions, and eyetracker malfunction resulted in the loss of eyetracking data from one additional participant, leaving N=22 for accuracy and reaction time analyses and N=21 for eyetracking analyses. All participants were native English speakers (n=17) or early bilingual (n=4) English speakers for whom English is the dominant language (the pattern of results reported here did not differ reliably when these 4 participants' data were removed). All participants had normal hearing and normal or corrected-to-normal vision and no history of language or learning disabilities.

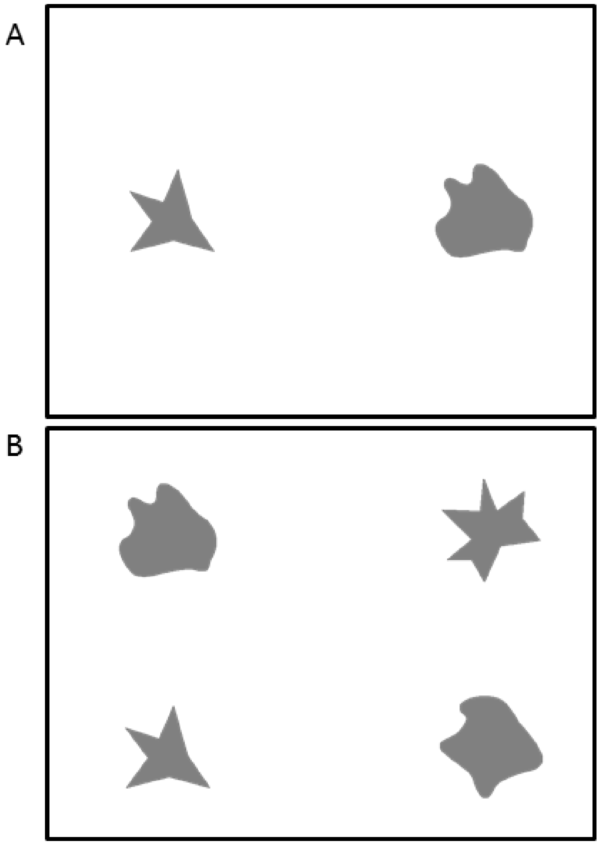

Figure 1. Screen layouts for the (A) 2AFC pretest and (B) 4AFC test blocks.

Materials

During training, participants learned to pair 24 novel CVCV word forms with 24 unfamiliar shapes. Both the verbal and visual stimuli were drawn from a larger set of stimuli previously normed by a separate group of participants (List, 2014; McCormick et al., 2015). All pseudowords were recorded by a female speaker of American English and were edited into separate files and amplitude normalized for presentation. From this set, we selected eight novel words that had been previously rated as highly ‘rounded’ (on average, 9.8% of 34 norming participants selected ‘pointy’ in a 2AFC task), eight that were highly ‘pointy’ (92.4% selected ‘pointy’), and eight showing no evidence of sound symbolism (50.0% selected ‘pointy’). Although ‘bouba’ and ‘kiki’ were not among the stimuli, we refer to the words rated to sound highly rounded as ‘boubas’ and the words rated as pointy-sounding as ‘kikis’ since these words are canonically associated with the sound-to-shape matching paradigm. Words that lacked a strong shape selection bias are termed ‘nonsymbolic’. Phonemic transcriptions of the full list of pseudoword stimuli appear in Table 1. Average word duration was 558ms and did not differ among bouba, kiki, and nonsymbolic categories ($F(2, 21) = 2.14$, $p > .1$).

Shape stimuli consisted of 24 line drawings of abstract rounded and angular shapes with 4-6 protuberances drawn from a larger set of abstract shapes previously rated for roundedness/pointiness by a separate group of 34 participants. Twelve of the shapes had previously been rated as highly rounded (mean rating 2.1 on a Likert scale where 1 = very rounded and 7 = very pointy) and twelve as highly pointy (mean rating 5.7). Four of each of the shape stimuli were paired with words that had a matching sound symbolic bias (rounded shapes → bouba words, pointy shapes → kiki words), four with mismatching words (rounded shapes → kiki words, pointy shapes → bouba words), and four of each with nonsymbolic words, for a total of eight *match*, eight *mismatch*, and eight *nonsymbolic* stimuli. To control for possible learnability biases, six different stimulus lists were created with different word-shape pairings, with individual words and shapes rotating through conditions.

Table 1: Word Stimuli

Round-Biased “boubas”	Pointy-Biased “kikis”	Nonsymbolic
bubo	kɛte	bɛde
gubu	piki	sefi
lʊlu	tɛpi	dʒuzo
mʊnu	fItʃe	tʃufo
bugu	kiti	tʃɛse
lomu	pIke	leni
mumo	teki	gɛgi
nʊlo	tite	sotʃu

Procedure

Participants were seated comfortably in front of the display screen with their chins in a chinrest at a viewing distance of 60cm. Stimulus presentation was controlled by E-Prime 2.0. Visual shape stimuli subtended 5.5 degrees of visual angle. Spoken stimuli were presented over Sennheiser HD280 Pro headphones at a comfortable listening volume. Eye movements were monitored using a table-mounted Eyelink 1000 eyetracker (SR Research). A nine-point calibration was performed before beginning the experiment and drift correction was performed before the start of each eyetracked test block. See Figure 1 for examples of the stimulus displays during the pretest and test blocks.

Pretest Following eyetracker calibration, participants completed 24 trials in a 2AFC pretest. Participants clicked on a central fixation cross to begin each trial. Two shape stimuli (one round, one pointy) appeared on the screen. After 250ms, participants heard the name of one of the displayed shapes and used the mouse to click on the shape that they thought had been named, which ended the trial. No feedback was provided during the pretest block. Participants were instructed that guessing was fine and that they should just listen to the word and decide which shape they thought it named.

Participants heard each word once during the pretest. The “correct” shape, i.e., the shape that would be paired with the word during training, was one of the two shape options available; the other item was pseudorandomly drawn from the opposite shape category so that one round and one pointy shape was present on each trial and each shape appeared onscreen twice during the pretest block (once as the target and once as a distracter stimulus.)

Training Following the pretest, participants completed sixteen interleaved blocks of training and testing. Each of the eight training blocks consisted of 48 2AFC trials. On each trial, two shapes were displayed on the screen. After 250ms, participants heard the name of one of the displayed shapes and used the mouse to click on the shape that they thought had been named. Regardless of whether they selected the correct or incorrect choice, the incorrect shape disappeared and the correct shape remained on screen for 1000ms while its name was repeated.

Participants heard each word twice during each training block. Each shape was presented four times during each training block, twice as the target stimulus and twice as a distracter stimulus. Unlike the pretest block, there was no requirement that both a rounded and a pointy shape appear on each trial so participants had to decide between two rounded or two pointy shapes half the time to make the visual discrimination more challenging.

Testing A test block occurred immediately after each of the eight training blocks. Each of the eight testing blocks consisted of 24 4AFC trials. Four shapes appeared on each trial: the target shape, one distracter from the same shape category as the target, and two distracters from the opposite shape category, so that two rounded and two pointy shapes were onscreen during every trial. After 250ms, participants heard the name of one of the displayed shapes and used the mouse to click on the shape that they thought had been named. No trial-by-trial feedback was given during test blocks, though participants were given a score (e.g. 18/24 correct) at the end of each test block. Each word was presented once per test block, and each shape appeared four times per test block; once as the target, once as a same-shape distracter, and twice as an opposite shape distracter.

Results

Full analysis of the data from the training and initial testing blocks is beyond the scope of this report; here we focus on data from the pretest and final two test blocks. Linear and logistic mixed effects models were used to analyze reaction time and choice/accuracy data respectively (Jaeger, 2008) using R (v3.1.1) and lme4 (v1.1-7). Maximal random effects (random effects of subject on the intercept and slope) were included in all models. Fixations and saccades were automatically detected by the Eyelink software and combined into gazes starting from the beginning of a saccade to the end of the subsequent fixation. Only signal-driven fixations (i.e., gazes beginning 200ms after the onset of the spoken word to account for eye movement planning) are shown.

Pretest

Participants showed clear sensitivity to the sound symbolic properties of the pseudowords during the initial pretest. Shape choice was strongly associated with the sound symbolic properties of the word, with participants choosing round shapes after hearing a ‘bouba’ word 69% of the time and choosing a pointy shape 73% of the time after a ‘kiki’ word. These tendencies mean that prior to any training, participants chose the ‘target’ shape that would be learned during training on 73% of match trials, 31% of mismatch trials, and 52% of nonsymbolic trials. Including a fixed effect of word category (match/mismatch/nonsymbolic) in the model significantly improved the model fit (χ^2 (2) = 19.2, $p < .001$) over the baseline model which contained only a fixed effect of intercept and had random effects of subject on the intercept and category slope term. Reaction times during the pretest were not significantly affected by word category whether sorted by eventual match status (RT_{match} = 1350ms, $RT_{mismatch}$ = 1384ms, $RT_{nonsymbolic}$ = 1419ms, χ^2 (2) = 0.71, $p > .1$) or sound structure (RT_{bouba} = 1357ms, RT_{kiki} = 1367ms, $RT_{nonsymbolic}$ = 1419ms, χ^2 (2) = 0.35, $p > .1$). Thus the best fitting reaction time model contained only an intercept term in the fixed effects along with random effects of subject on both the intercept and slope (category or match) term.

Participants’ eye movements during pretest were also affected by the sound symbolic properties of the word. The difference in fixation proportions between the target and the distracter shape was calculated for match, mismatch, and

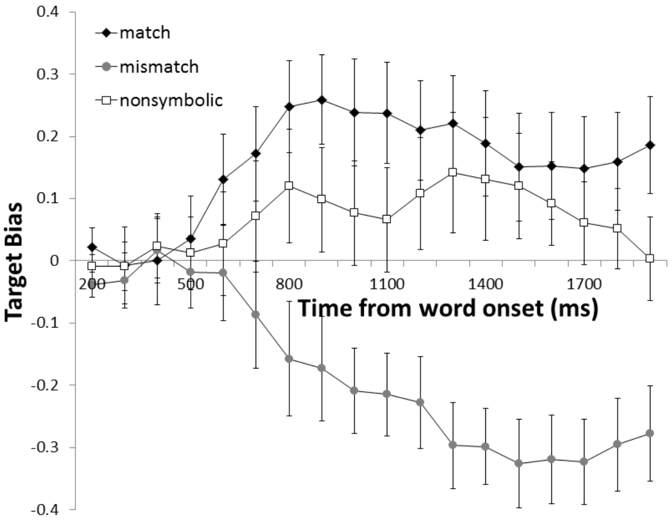

Figure 2. Pretest fixation proportion difference curves (target – distracter). Mean word offset is 558ms; mean RT ~1400ms. For display purposes, data has been binned into 100ms windows. Error bars indicate SEM.

nonsymbolic words without regard to the final click decision; all trials were included, whether they ended in the participant correctly guessing the target stimulus or clicking on the distracter shape, since the participant had no basis for knowing which were the correct pairings during the pretest block. Mean fixation proportions were calculated across the time window extending from 200ms after word onset (the first signal-driven fixations) to 1400ms (the average RT across conditions). Single sample t-tests were used to determine whether the difference in fixation proportions between the target and distracter significantly differed from zero; i.e., if participants showed a significant bias in looking to either target or distracter shapes. As seen in Figure 2, participants showed a strong bias to fixate shapes consistent with the sound symbolic properties of the spoken word. In the match condition, participants fixated the sound symbolically consistent target shape more than the distracter shape ($M_{match\_difference} = 0.15$, $t(22) = 5.28$, $p < .01$). They also preferred the shape consistent with the sound symbolic properties of the word in the mismatch condition, fixating the sound symbolically consistent distracter shape more than the target ($M_{mismatch\_difference} = -0.14$, $t(22) = -3.19$, $p < .01$). Importantly, the difference in fixations between target and distracter items was not significant for the nonsymbolic stimuli ($M_{nonsymbolic\_difference} = 0.07$, $t(22) = 1.51$, $p > .1$).

Final Test

By the final two test blocks, participants were approaching ceiling performance on the 4AFC task with high accuracy in all conditions ($M_{match} = 91\%$, $M_{mismatch} = 87\%$, $M_{nonsymbolic} = 93\%$). However, inclusion of word category in the accuracy model marginally improved model fit (χ^2 (2) = 4.9, $p = .09$) over a baseline model which contained only a fixed effect of intercept and had random effects of subject on the intercept and category slope term. Contrast analysis suggests that this effect is carried by a slight accuracy advantage for the nonsymbolic items over the mismatch items ($b = -0.063$, $SE = 0.023$, $p_{norm\_approx} = .024$), with match items intermediate and not significantly different from either.

Analysis of reaction times also suggested an advantage for nonsymbolic items (1699ms) relative to match (1971ms) and mismatch items (2018ms), with inclusion of word category as a regressor significantly improving model fit ($\chi^2(2) = 6.2$, $p < .05$) over a baseline model which contained only a fixed effect of intercept and had random effects of subject on the intercept and category slope term and contrast analysis showing faster reaction times for nonsymbolic items relative to both match ($b = 271.4$, $SE = 128.4$, $p_{norm\_approx} = .034$) and mismatch items ($b = 318.4$, $SE = 159.0$, $p_{norm\_approx} = .045$). Importantly, reaction times for matching and mismatching items did not differ ($b = 47.0$, $SE = 178.5$, $p_{norm\_approx} = .8$).

Although participants' behavioral responses to match and mismatch items no longer differed by the end of training, a disadvantage for mismatching word/shape pairings is still apparent in the eye movement data. Figure 3 shows target fixation proportions beginning 200ms after the onset of the word for matching, mismatching, and nonsymbolic stimuli. Only data from trials where the participant ultimately selected the correct shape are shown. We fit a four-parameter logistic function to each subject's average fixation proportion curve for the match, mismatch, and nonsymbolic conditions following methods described by McMurray & colleagues (McMurray et al., 2010; Farris-Trimble et al., 2014). The four parameters include lower and upper asymptotes (representing baseline and peak fixations), the crossover point (the timepoint where the function's rate of change is maximal), and the slope at that timepoint. The resulting parameter estimates for each combination of subject and condition were analyzed in separate ANOVAs.

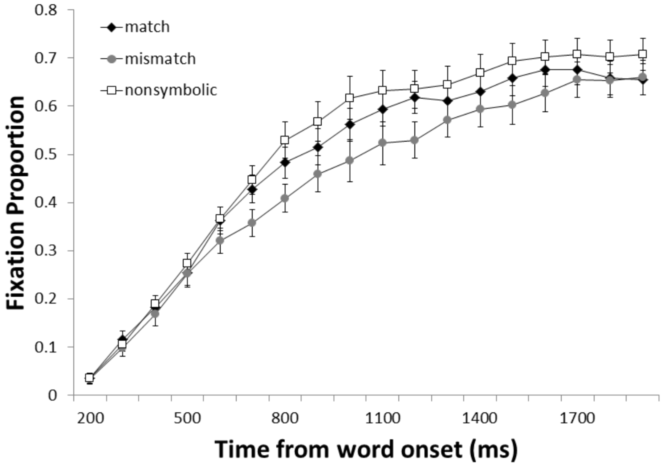

Figure 3. Fixations to target items in final 4AFC test block for correct trials only. For display purposes, data has been binned into 100ms windows. Error bars indicate SEM.

We found a significant main effect of condition on the function's slope parameter (*F*(2, 40) = 4.83, *p* < .05). Pairwise comparisons show a lower slope (slower rate of increase in fixations to the target) for the mismatch condition relative to both the match (*t*(20) = 3.1, *p* < .01) and nonsymbolic (*t*(20) = 2.92, *p* < .01) conditions, which do not differ from each other (*t*(20) = 0.01, *p* > .1). The upper asymptote and the crossover point parameters were not significantly affected by condition. The lower asymptote was artificially constrained to be zero by the choice to include only signal driven fixations and was therefore not analyzed.

Discussion

As expected, participants encountering novel words for the first time showed a consistent bias in pairing unfamiliar sound symbolic stimuli with novel shapes, matching bouba-like words to shapes with curved contours and kiki-like words to shapes with sharp edges at above-chance rates. This effect was apparent in the pretest block both in participants' choices and in their eye movements. A clear bias to fixate shapes consistent with the sound symbolic properties of the words began to emerge approximately 700ms after word onset. Given the mean duration of the word stimuli and the roughly 200ms it takes to plan and launch a saccade, this suggests that these effects emerge rapidly, near word offset and several hundred milliseconds before participants give an overt behavioral response. Words without sound symbolic properties were not associated with particular types of shapes, with participants' overt responses at chance and no significant bias evident in their eye movements.

These findings are consistent with past studies in our lab and others' and provide a mechanism by which sound symbolic properties of a word affect word learning. Prepotent biases to associate particular sounds with particular meanings increased the likelihood of a learner making the correct word-to-meaning mappings in these cases, and words with more sound-to-meaning systematicity appear to have an earlier age of acquisition (Monaghan et al., 2014). However, there has been little evidence that sound symbolism continues to impact lexical processing of words that are well-learned (Kunihira, 1971; Nygaard et al., 2009). One possibility is that behavioral measures like accuracy and reaction time are not sensitive enough to detect subtler effects that might occur during online processing of the well-learned stimuli. Indeed, by the end of training in the current study, participants achieved around 90% accuracy across all conditions with little evidence that whether the sound symbolic properties of the word matched or mismatched the physical properties of the referent affected either accuracy or reaction time, despite the fact that at pretest, learners had exhibited a strong bias to choose shapes with matching properties. However, participants' eye movements exhibited a persistent processing disadvantage for mismatching stimuli, with a significantly slower latency to fixate targets that mismatched the sound symbolic properties of the word relative to both matching and nonsymbolic stimuli. This suggests the sound symbolic properties of the word were still affecting online processing of well-learned stimuli. That this effect manifested as a disadvantage for mismatching word-shape pairings rather than an advantage for matching pairings at the end of training suggests that concordance between sound and meaning may facilitate early but not later stages of learning whereas interference from discordant sound to meaning mappings persists. Indeed, examination of accuracy and reaction time data from earlier testing blocks (Figure 4) suggests that participants are initially slower and less accurate in pairing shapes that mismatch the sound symbolic properties of the word, though further exploration of sound symbolic effects over the entire timecourse of learning is beyond the scope of this report. Future experiments will be needed to determine whether these effects persist indefinitely with overlearned stimuli or whether the eye movement effects are learning-specific and are only present because accuracy, while high, may not yet have reached asymptote for all participants.

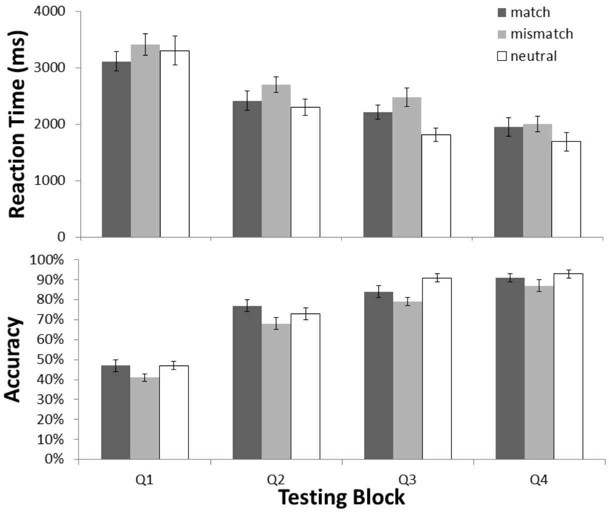

Figure 4. Reaction time and accuracy data for all 4AFC test blocks. Only results of Q4 block are reported in detail here. Error bars indicate SEM.

One unexpected result that emerges from the final testing blocks is the advantage for nonsymbolic stimuli over matching and mismatching stimuli in both accuracy and reaction times late in learning. This was unexpected given previous research showing a learning advantage for sound symbolic stimuli. However, a closer examination of the word materials in Table 1 suggests that the nonsymbolic stimuli in this experiment may be more phonologically distinct from each other than items within the 'bouba' or 'kiki' stimulus groups, as the eight nonsymbolic stimuli contained combinations of 10 consonants and 5 vowels while each of the groups of sound symbolic stimuli drew from a set of only 5 consonants and 3 or 4 vowels. Words with sparser phonological neighborhoods are recognized faster and more accurately than words from denser neighborhoods (Luce & Pisoni, 1998), which may explain the nonsymbolic advantage seen here. Future work will need to control for the phonological makeup of the nonsymbolic stimuli as well as the symbolic stimuli to ensure approximately equal neighborhood densities. Nevertheless, initial exposure to the nonsymbolic stimuli during the pretest confirms that there were no sound symbolic biases facilitating learning from the words designated as nonsymbolic, and examination of Figure 4 suggests that this effect emerges late in learning, with no advantage for nonsymbolic words in the first half of training. Further, direct comparison of the match and mismatch stimuli late in learning reveal an effect of congruence of sound and meaning on visual fixation independent of performance on the nonsymbolic items.

Non-arbitrary correspondences between the sound of a novel word and the shape of a potential referent appear to promote an initial pairing between the word and referent that may speed the learning process. Here we demonstrate that this initial bias can also be seen in participants' eye movements, a rapid and implicit measure of online processing. Furthermore, eye movements show evidence for a continued cost when there is a mismatch between sound and shape late in learning, even when the effect is no longer evident in accuracy or reaction time measures. This effect appears to emerge during or immediately following the presentation of the spoken word and is resolved by the time an overt behavioral response is made, emphasizing the importance of the availability of online, continuous measures of processing. This technique may therefore prove useful for examining the subtler effects of sound symbolism in natural or well-learned language stimuli and in situations where an explicit judgment from the participant may be difficult to obtain due to task, strategy, or participant age.

Acknowledgments

We thank Kelly McCormick, Christina Tzeng, and Sara List for sharing their sound and shape stimuli and norming data, and to the Dilks lab for eyetracker access. Laura Namy's effort on this project was supported by the National Science Foundation. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

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,* 165-172.

Creel, S.C., Tanenhaus, M.K., & Aslin, R.N. (2006). Consequences of lexical stress on learning an artificial lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 15-32.

Farmer, T.A., Christiansen, M.H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences, 103,* 12203-12208.

Farris-Trimble, A., McMurray, B., Cigrand, N., & Tomblin, J. B. (2014). The process of spoken word recognition in the face of signal degradation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 308-327.

Jaeger, T.F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434–446.

Köhler, W. (1947). *Gestalt psychology (Second Edition).* New York: Liveright.

Kunihira, S. (1971). Effects of the expressive voice on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior, 10,* 427-429.

List, S. M. (2014). The sound of shape: Functional and neural correlates of sound to shape mapping in natural language. (Unpublished honors thesis). Emory University, Atlanta.

Luce, P. A. & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19,* 1-36.

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

McCormick, K., Kim, J. Y., List, S., & Nygaard, L. C., (2015). Sound to meaning mappings in the Bouba-Kiki effect. *Proceedings of the 37th Annual Conference of the Cognitive Science Society,* Pasadena, CA.

McMurray, B., Samelson, V. M., Lee, S. H., Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology, 60,* 1-39.

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

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

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

Reilly, J., Westbury, C., Kean, J., & Peelle, J.E. (2012). Arbitrary symbolism in natural language revisited: When word forms carry meaning. *PLoS ONE, 7,* e42286.

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

Revill, K. P., Tanenhaus, M. K., & Aslin, R. N. (2008). Context and spoken word recognition in a novel lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1207-1223.

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