Sound Symbolism in Word Learning

Lynne C. Nygaard (lnygaard@emory.edu)
Department of Psychology, Emory University
Atlanta, Georgia 30322 USA

Allison E. Cook (allisonecook@gmail.com)
Department of Psychology, Emory University
Atlanta, Georgia 30322 USA

Laura L. Namy (lnamy@emory.edu)
Department of Psychology, Emory University
Atlanta, Georgia 30322 USA

Abstract

Although numerous examples of sound symbolism exist in spoken language, little is known about whether listeners recognize and utilize these non-arbitrary relationships between sound and meaning to guide their interpretation of novel words. The present experiments examined the extent to which listeners are able to constrain meaning solely on the basis of the segmental sound structure of language. Native English-speaking monolinguals learned meanings for Japanese words in a vocabulary-learning task. Spoken Japanese words were paired with English meanings that either matched the actual meaning of the Japanese word (e.g., “ue” → up), were antonyms (e.g., “ue” → down), or were randomly paired (e.g., “ue” → slow). The results revealed a processing benefit in learning matching word-meaning pairs over mismatching pairs, as evidenced by greater accuracy and faster reaction times. These findings suggest that adult word learners are sensitive to non-arbitrary relationships between the sound structure of spoken language and word meaning.

Keywords: Speech; word learning; sound symbolism; arbitrariness.

Introduction

The current project examines the relationship between meaning and the sound structure of language to assess its potential role in spoken language learning. Traditionally, it has been assumed that the sounds of language, roughly the consonants and vowels of a language, bear an arbitrary relationship to their referents (de Saussure, 1959; Hockett, 1977). For example, across languages, dogs are referred to by a variety of names (e.g., dog, chien, perro, cane, kelb) and the sound structure of those names is thought to have little relationship to properties of the referent (e.g., dog). Linguistic communication clearly relies on a basic set of arbitrary sound-meaning pairings, however mounting evidence suggests that a significant proportion of sound-meaning pairings also include various types of non-arbitrary relationships (Nuckolls, 1999). In addition to examples such as onomatopoeia, multiple examples have been found both within and across languages that suggest that systematic relationships exist between what words sound like and what they mean (Bergen, 2004; Berlin, 1994; Cassidy, Kelly & Sharoni, 1999; Nygaard, Herold, & Namy, in press; Kunihira, 1971). For example, Berlin (1994) presented English-speaking listeners with a list of bird and fish names from Huambisa, a language spoken in north central Peru, and asked them to classify each word as a bird or fish. Despite being unfamiliar with the language, English speakers were able to classify the Huambisa terms into the correct categories significantly above chance. This finding suggests that the sound structure of the Huambisa words is systematically related to their referents. English-speaking listeners were able to pick up on this systematicity in order to correctly classify terms. If these types of relationships between sound and meaning prove to be ubiquitous in spoken language, and if listeners are sensitive to these relationships, then basic theoretical assumptions of the nature and structure of linguistic reference will need to be revised.

The current experiments investigated both the extent to which the sounds of language are systematically related to the objects, actions, and concepts to which they refer and the ways the relationship between sound and meaning might influence spoken language processing. Of interest is why non-arbitrary links between sound and meaning, sound symbolism, might exist and how these links might be useful in the understanding of spoken language. We argue that one reason for the preservation of non-arbitrary links is to facilitate word learning. That is, language learners, either children first acquiring a set of sound-meaning pairings or second language learners acquiring a non-familiar mapping between sound and meaning, could exploit non-arbitrary associations between sound structure and meaning. This mechanism could facilitate both first and second language learning. The present experiments constitute a first step at addressing this question.

Little previous research has examined a potential role for sound symbolism in word learning in particular, or in language processing more generally. However, classic work by Kohler (1949) and more recently, by
Ramachandran and Hubbard (2001), Maurer, Pathman, & Mondloch (2006), and Westbury (2005) suggest that English-speaking adults and children will interpret nonwords such as ‘maluma’ and ‘bouba’ as referring to round, amoeboïd shapes and words like ‘takete’ and ‘kiki’ as referring to angular figures. This work suggests that the sounds in these non-words bear some type of resemblance or relationship to the properties of particular figures and that this resemblance can be exploited by English speakers to create a consistent mapping across individuals. In a more challenging word matching task, Kunihira (1971) studied the effects of expressive voice on the recognition of sound-meaning correspondences. Kunihira presented native English speakers with 23 Japanese language antonym pairs and asked subjects to match the Japanese words with the correct English translation. The Japanese words were either presented in printed form, spoken with a monotone voice, or spoken with expressive tone of voice or prosody. Kunihira found that in all conditions, English speakers unfamiliar with Japanese were able to reliably match the Japanese word with its corresponding English equivalent. Participants performed best in the expressive voice condition suggesting that in addition to using the sound structure of each word to derive its meaning, they also used the speakers’ tone of voice.

The current experiments were designed to examine the contribution of sound symbolism to a novel word-learning context. We investigated the extent to which sound symbolism would influence monolingual native English speakers’ ability to acquire new vocabulary items in Japanese. Specifically, we investigated whether sound symbolic relationships aid in the acquisition of new vocabulary items. Using the same Japanese antonyms that Kunihira (1971) had shown to have sound symbolic properties, we examined whether learners were able to use sound-meaning relationships to link unfamiliar words to English meanings. Vocabulary learning was tested in three conditions using a repeated learn-test paradigm, in which learning and test cycles alternated over three separate blocks. Each learning cycle included three conditions: In the Match condition, participants learned Japanese words that were paired with their actual English equivalents (e.g., “ue” paired with up); in the Opposite condition, Japanese words were paired with the English equivalent of their antonym (e.g., “ue” paired with down); in the Random condition, Japanese words were paired with an unrelated English word (e.g., “ue” paired with wide). Retrieval accuracy and speed were the dependent measures and were recorded after each of three learning blocks.

If sound symbolic relationships exist in spoken language and word learners are sensitive to those non-arbitrary relationships between sound and meaning, then we hypothesized that when Japanese vocabulary words were paired with their actual English equivalents, word learning should occur more quickly and accurately. Additionally, if the sound structure of a word evokes a specific semantic domain as represented by the antonym pair dimension (e.g., fast/slow, up/down), Japanese words paired with an opposite meaning should be learned more quickly and with greater accuracy than Japanese words paired with an unrelated English word. However, if learners are not sensitive to sound-meaning correspondences in natural language, then variations in word-meaning mappings should have little effect on word learning performance.

Experiment 1

Experiment 1 evaluated participants’ ability to learn novel words with varying degrees of sound-meaning correspondence. Participants completed three sets of the learn-test cycles in which paired-associate learning was alternated with tests of word retrieval. Specifically, during the learning phase, participants heard each Japanese word presented twice and at the same time, saw the English equivalent on the screen. During each test phase, participants were prompted with the spoken Japanese word and asked to type the corresponding English word. Recall accuracy was used as an index of whether participants were able to exploit phonetically-based sound symbolic relationships when storing and referencing information about unfamiliar words.

Method

Participants

Participants were 60 native speakers of American English with no familiarity with the Japanese language and no report of hearing or speech disorders. All participants received credit towards course requirements for their participation.

Stimulus Materials

The stimuli consisted of 21 Japanese antonym pairs used by Kunihira (1971). A native female Japanese speaker recorded each of the 42 words in list format. The stimuli were recorded in a sound attenuated room on a SONY Digital Audio Tape-corder TCD-D7 and were then re-digitized onto computer, edited into individual files for presentation, and amplitude normalized.

The 21 Japanese antonym pairs were divided into three groups equated for the frequency of occurrence of their English equivalents (Kucera & Francis, 1967). Three within-subject experimental conditions were constructed using the three word groups. In the Match condition, Japanese words were paired with their English equivalents. In the Opposite condition, Japanese words were paired with the English equivalent of their antonym pair. In the Random condition, Japanese words were paired with a
random, unrelated English word. Antonym pairs and word pairing conditions were counterbalanced across participants such that the words in each group appeared equally often in each word pairing condition, Match, Opposite, and Random. Thus, each participant heard only one of the words from each antonym pair. Words were presented in random order to each participant in each learn-test cycle.

Procedure
Participants were informed that they would be performing a vocabulary learning task with foreign words. Participants were asked to listen carefully to each Japanese word, presented auditorily over headphones, and to pay attention to the visually presented English word, with which it was paired. Participants were also informed that the second part of the task involved recalling the English “translation” after hearing only the spoken Japanese word. Participants engaged in three blocks of learning phases alternated with three blocks of test phases. Stimulus presentation was controlled on-line via E-Prime software (Psychology Software Tools, Pittsburgh, PA) on a Dell Optiplex computer. During each learning phase, participants saw a fixation cross (500 ms) and then heard a Japanese word through the headphones (Beyerdynamic DT100) while the monitor displayed a blank screen. The word was then repeated while the English “translation” appeared on the computer screen for 2000 ms. The inter-trial interval was 3500 ms. This sequence repeated for each of the 21 Japanese words.

Immediately following completion of each learning phase, participants began the test phase. During each trial in all three test phases, participants saw a fixation cross (500 ms) and then heard a Japanese word through the headphones. Participants then typed the word’s English equivalent in response to an arrow prompt presented on the screen. The next trial was automatically initiated 500 ms after each response. Based on pilot testing, in order to reduce working memory load and facilitate word retrieval performance, a word list containing all the given English translations was presented on the screen during the test phases. Before beginning the experiment, participants were informed that a list of all the English translations they saw during the learning phase would appear on the screen during the test phases, and they were reminded about the list again before each individual test phase. The possible answers were listed in a random order on the left hand side of the screen and appeared when the participants were prompted with the spoken Japanese words and were asked to recall the paired English word. The list did not appear on the screen at any other point during the experiment. After the test phase was complete, the next learning phase began.

Results and Discussion
Figure 1 shows the percent correct word choices for each condition (Match, Opposite, Random) across the three training cycles. Percent correct scores were converted using an arcsine transform to correct for negative skew in the distribution of scores. A 3x3 repeated measures analysis of variance (ANOVA) on the transformed scores with condition and training cycle as within-subject variables revealed a significant main effect of training cycle, $F(2, 118) = 239.65, p < .001$. Participants’ overall word retrieval accuracy improved over the course of the three training cycles. In addition, the effect of condition neared significance, $F(2, 188) = 2.91, p = .06$, suggesting that learning varied as a function of condition.

![Figure 1: Percent correct word choices in Experiment 1 as a function of training cycle and word learning condition.](image-url)

Planned comparisons revealed that participants performed significantly better on the Match words than the Random words, $t(59) = 2.54, p = .01$, in the third block. No significant difference was found between Match and Opposite ($p = .32$) or between Opposite and Random conditions ($p = .20$). Although no significant differences were found among the three conditions in the first block, the difference between the Match and Random conditions in the second training block approached significance as well, $t(59) = 1.85, p = .07$.

These findings suggest that learners were sensitive to connections between sound and meaning and that these connections may influence vocabulary learning. Learners appeared to store and retrieve sound-meaning relationships during word learning and the functional benefit of these correspondences emerged with increased exposure.
Experiment 2

Although the results of Experiment 1 suggest that sound symbolic relationships may be used by word learners when acquiring a new vocabulary, the task in Experiment 1 may have encouraged learners to look for any clues to the correct vocabulary item. That is, learners may have behaved strategically, explicitly attending to any relationship between sound pattern and reference and choosing items with evident sound-meaning correspondences. Although even strategic responding would suggest that learners are sensitive to sound symbolic relationships, the question remains regarding the extent to which learners are actually incorporating sound-meaning relationships into their on-line word retrieval processes.

To examine the effects of sound symbolism on the word learning process and on novel word retrieval, we examined a situation in which word selection and retrieval performance was highly accurate and consequently, time to respond became the dependent measure to determine if participants were able to identify correct translations faster when a sound-meaning relationship was present than when no relationship was present. Thus, instead of transcribing the English words at test, participants were presented with two possible English word choices and asked to choose which word they believed was the correct match. By presenting two options, the test reduces the short-term memory demands of the test and provided an alternative metric for evaluating whether non-arbitrary relationships influence the immediate processing of language.

Method

Participants

Participants were 60 native speakers of American English with no familiarity with the Japanese language and no hearing or speech disorders. Participants either received course credit for their participation (n=24) or were paid for their participation (n=36).

Stimulus Materials

The same stimuli and experimental conditions used in Experiment 1 were also used in Experiment 2. The same division of word groups was maintained for Experiment 2. However, in addition to the English words that were learned as equivalents in each of the experimental conditions, a second unrelated English word was randomly assigned to each Japanese word to be used as distractors in each test phase of the experiment. The 21 distractors were English translations of the Japanese antonyms from antonym pairs that the participants did not hear and were randomly paired with target words. Word frequency was equated across pairings. The target and distractor words appeared as fixed pairings in all three test blocks.

Procedure

All procedures were the same as in Experiment 1, except that a two alternative forced-choice paradigm was used in the test phases. After completing each learning phase, participants were informed that the second part of the task involved choosing the correct English equivalent for each Japanese word from two options using a standard button box. The participants heard the Japanese word once and were presented with the two possible English “translations” on the computer screen, the correct target word and the distracter word. For example, in the Match condition, the Japanese word “ue” was presented auditorily and participants were asked to choose between up (the correct target word) and walk (the distractor). Participants were asked to make their choice as quickly as possible. The two English words were presented side by side on the computer screen, and the sides on which the English words appeared were counterbalanced for each test phase and across participants. As in Experiment 1, the learning phases and test phases alternated and repeated a total of three times.

Results and Discussion

Accuracy. Percent correct responses for each condition across the three blocks are presented in Figure 2. An arcsine transform was again used to correct for negative skew. A 3x3 repeated measures ANOVA with block and condition as within-subject variables revealed a significant main effect of block, F(2, 118) = 27.04, p < .001. Participants’ performance significantly improved over the three training cycles. However, although the pattern is consistent with better performance in the Match Condition, overall performance did not vary significantly as a function of condition and there was no interaction between block and condition. The lack of a significant difference among the experimental conditions likely reflects the expected ceiling effect that emerged in the second and third blocks. In order to evaluate this possibility, we conducted planned comparisons among conditions in the first block in which ceiling had not yet been reached. Similar to Experiment 1, participants performed significantly better in the Match than in the Random condition t(59) = 2.23, p = .02. No significant differences were found between Match and Opposite or between Opposite and Random conditions.

Response times. Response times for correct responses were analyzed to determine if latency of response varied as a function of experimental condition. Figure 3 shows mean response times as a function of condition and training cycle. A 3x3 repeated measures ANOVA with block and condition as the within-subject variables revealed a significant main effect of block, F(2, 118) = 39.57, p < .001, indicating an overall decrease in response times across the three blocks. In addition, a significant main effect of condition was found, F(2, 118) = 4.19, p < .02, indicating that the nature of the
pairing, Match, Opposite, or Random, significantly influenced how quickly learners chose word meanings.

![Figure 2: Percent correct word choices for Experiment 2 as a function of training cycle and learning condition.](image)

Follow-up simple effects analyses on each block revealed a significant effect of condition in Block 3, $F(2, 118) = 3.56, p < .05$. As learners’ performance approached ceiling in terms of accuracy in Block 3, effects of sound symbolic relationships on speed of responding emerged. There were no significant differences in performance as a function of condition in Blocks 1 and 2. Only after accuracy approached ceiling performance, did participants demonstrate differences in processing speed as a function of condition. Pairings with non-arbitrary sound-to-meaning mappings appeared to be processed more quickly than arbitrary pairings.

### General Discussion

The current investigation examined the potential influence of sound symbolism on the ability to learn new lexical items. Our results indicate that sound symbolic properties of spoken language influenced the encoding and retrieval of the meaning of unfamiliar words. In Experiment 1, participants’ ability to choose the correct translation from a list of possible candidates varied significantly as a function of experimental condition (Match, Opposite, Random). Learners chose correct vocabulary equivalents more often when the word form and meaning matched than when the learned meaning was unrelated to the actual meaning. Similarly, in Experiment 2, learners responded significantly faster when word-meaning pairs matched than when the pairings were mismatched. In this paradigm, once word meanings became readily accessible, online word processing and retrieval exhibited a facilitation effect for correct mappings implying processing constraints from non-arbitrary relationships between sound and meaning. It should be noted as well that accuracy in the Opposite condition did not differ significantly from either Match or Random conditions in either Experiment 1 or Experiment 2, suggesting that having a word pairing within the same semantic domain, albeit an antonym (e.g., walk/run), may result in some benefit in this vocabulary learning task.

These results fundamentally challenge the traditional assumption that words bear an exclusively arbitrary relationship to their referents and that linguistic representation and processing relies solely on this arbitrary relationship. Sensitivity to the relationship between the sound structure of language and meaning implies that a strict dichotomy between linguistic structure and processing and non-linguistic form may be difficult to maintain. Learners appear to incorporate both sound properties of language and linguistic form into the process of associating unfamiliar words with meanings.

Planned comparisons revealed that participants performed significantly better overall in the Match than in either the Opposite, $t(59) = 1.82, p < .05$, or Random conditions, $t(59) = 1.97, p < .05$. No significant difference was found between Opposite and Random conditions, $p = .38$. Participants responded more quickly to Japanese words paired with their English equivalents than to Japanese words paired either with a random, unrelated English word or with the English equivalent of the word’s antonym.

These findings are consistent with other demonstrations of consistent mappings between the sound structure of language and meaning (Bergen, 2004; Berlin, 1994; Bybee, 1985; Farmer, Christiansen, & Monaghan, 2006; Kelly, Springer, & Keil, 1990). It is likely that reliable, statistical associations between sound sequences and meanings within
a natural language family can explain some instances of sound symbolism. However, this particular instantiation of sound symbolic facilitation occurred for mappings that cut across two languages, English and Japanese, from distinct lineages and with unique phonologies, decreasing the likelihood of experience-based statistical mechanisms as an account of the phenomenon. That native English speakers were sensitive to non-arbitrary sound to meaning mappings in Japanese suggests that these regularities are not strictly conventions exclusive to Japanese, but rather reflect general perceptually based cross-modal relationships. Although it is unclear at this point which sound properties of these Japanese words are related to aspects of word meaning, it is clear that these properties were available to native speakers of English during word learning.

Given that language is not exclusively comprised of arbitrary relationships, these findings suggest that non-arbitrary relationships and sound symbolic structure have psychological functionality during language processing. Not only do learners encode and represent the sound properties of language, but they also recruit this sensitivity within the complex, and arguably taxing, cognitive task of associating novel words with meanings. This influence of sound-meaning associations emerged both in a word selection task that emphasized controlled, strategic processing as well as in a speeded forced choice task that emphasized processing efficiency. The effects of sound symbolism on both types of tasks and processing highlight a potential role for these constraints in word learning.

The current investigation provides one of the first demonstrations that learners can use sound symbolic relationships to derive meaning during spoken language processing. Although arbitrariness certainly remains a central design characteristic of linguistic structure, these results indicate that language users can and do exploit non-arbitrary relationships in the service of word learning and retrieval.

Acknowledgements
The authors would like to thank Jessica D. Alexander and Stig Rasmussen for their help preparing stimuli and developing protocols. A portion of this research was supported by an Emory College Scholarly Research and Inquiry at Emory (SIRE) undergraduate research grant to the second author.

References


