Brief article

Sound to meaning correspondences facilitate word learning

Lynne C. Nygaard *, Allison E. Cook, Laura L. Namy

Department of Psychology, Emory University, 4532 Kilgo Circle, Atlanta, GA 30322, USA

Article info

Article history:
Received 26 January 2009
Revised 28 March 2009
Accepted 2 April 2009

Keywords:
Spoken language
Arbitrariness
Sound symbolism
Word learning

Abstract

A fundamental assumption regarding spoken language is that the relationship between sound and meaning is essentially arbitrary. The present investigation questioned this arbitrariness assumption by examining the influence of potential non-arbitrary mappings between sound and meaning on word learning in adults. Native English-speaking monolinguals learned meanings for Japanese words in a vocabulary-learning task. Spoken Japanese words were paired with English meanings that: (1) matched the actual meaning of the Japanese word (e.g., “hayai” paired with fast); (2) were antonyms for the actual meaning (e.g., “hayai” paired with slow); or (3) were randomly selected from the set of antonyms (e.g., “hayai” paired with blunt). The results showed that participants learned the actual English equivalents and antonyms for Japanese words more accurately and responded faster than when learning randomly paired meanings. These findings suggest that natural languages contain non-arbitrary links between sound structure and meaning and further, that learners are sensitive to these non-arbitrary relationships within spoken language.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

A fundamental assumption regarding spoken language is that the sound structure of words bears an essentially arbitrary relationship to meaning (de Saussure, 1959; Hockett, 1977). Arbitrariness is considered a hallmark of spoken language that sets it apart from other types of communication systems, both human and non-human, providing the compositional power that gives spoken language its flexibility and productivity (Gasser, 2004; Monaghan & Christiansen, 2006). Indeed, linguistic pairings of sound to meaning are largely arbitrary. Across languages, for example, a variety of word forms that set apart from other types of communication systems, both human and non-human, providing the compositional power that gives spoken language its flexibility and productivity (Gasser, 2004; Monaghan & Christiansen, 2006). Indeed, linguistic pairings of sound to meaning are largely arbitrary. Across languages, for example, a variety of word forms (e.g., dog, chien, ferro, cane, kelb) can refer to the same object, action, or concept (e.g., dog), exhibiting seemingly little relationship between particular sounds and properties of the referent.

Despite this apparent arbitrariness, reliable correspondences between sound structure and linguistic category exist in natural language, and listeners seem to be sensitive to these relationships. One of the most obvious examples is onomatopoeia. Resemblance between a word and the sound that it represents (e.g., boom, meow) is common across languages, although different languages capture the sounds using different sets of phonemes (e.g., woof woof, au au, ham ham, gav gav). Likewise, languages such as Japanese utilize mimetics – words that resemble their meanings and are often used in interactions with children and in poetry (e.g., goro goro meaning “thunder” or “stomach rumbling”; Imai, Kita, Nagumo, & Okada, 2008; Kita, 1997). Another example is the existence of phonesthemes, particular sound sequences that are used across a range of words within a language to reflect semantically related meanings (e.g., the consonant cluster “gl” to imply things that shine: glisten, gleam, glitter, glow; Bergen, 2004; Hutchins, 1998). Although these phenomena are common across natural languages, they are not systematic or pervasive within any specific language and may well be special cases.

* Corresponding author.
E-mail address: lnygaar@emory.edu (L.C. Nygaard).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.04.001
A growing literature suggests a class of subtler, but more pervasive sound-to-meaning correspondences that fundamentally challenge assumptions regarding the exclusive arbitrariness of natural language. One of these sound-to-meaning links consists of correspondences between phonological form and lexical category (Cassidy & Kelly, 2001; Farmer, Christiansen, & Monaghan, 2006; Kelly, 1992; Sereno & Jongman, 1990). For example, listeners displayed a response-time advantage in processing sentences that contained nouns and verbs with category-typical phonological structure over sentences that contained nouns and verbs with atypical structure (Farmer et al., 2006). Shi, Werker, and Morgan (1999) found that even newborns are sensitive to phonological properties that differentiate function and content words.

Sound–meaning pairings may also include non-arbitrary relationships between sound and semantics (Nuckolls, 1999), suggesting that systematic relationships also exist between what words sound like and what they mean (Bergen, 2004; Berlin, 1994; Cassidy, Kelly, & Sharoni, 1999; Kunihira, 1971; Nygaard, Herold, & Namy, 2009). Köhler (1947) and more recently, Ramachandran and Hubbard (2001), Maurer, Pathman, and Mondloch (2006), and Westbury (2005) have found that English-speaking adults and children interpret nonwords such as ‘maluma’ and ‘boubu’ as referring to round, amoeboid shapes and words like ‘takete’ and ‘kiki’ as referring to angular figures. These findings suggest that a resemblance or relationship between the sound structure of these nonwords and perceptual properties of the figures constrains meaning, and raises the possibility that these relationships may facilitate both word learning and lexical processing.

Although correspondences between phonological characteristics and linguistic categories call into question the exclusively arbitrary nature of language, these relationships may reflect within-language conventions that have evolved over time through generalization and borrowing within a language. However, cross-linguistic tasks suggest that sensitivity to these correspondences is not entirely the result of conventional mappings learned by speakers of a particular language (Berlin, 1994; Brown, Black, & Horowitz, 1955; Brown & Nuttall, 1959; Imai et al., 2008; Tsuru & Fries, 1933; Weiss, 1966). For example, Kunihira (1971) reported that when native English speakers were presented with Japanese antonym pairs (e.g., “akurai” paired with “bright”); in the Random condition, Japanese words were paired with the English equivalent of their antonym (e.g., “akurai” paired with “dark”); in the Random condition, Japanese words were randomly paired with English translations of other antonyms from the learning set (e.g., “akurai” paired with “wide”). At test, learners were asked to identify the meaning of the word that they had learned in a speeded choice task. Response speed served as the dependent measure, allowing us to examine whether putative sound symbolic relationships influenced the fundamental processing efficiency of accessing newly acquired vocabulary items.

In general, we expected that learners would acquire the learned equivalents quickly and as accuracy reached ceiling, any sensitivity to sound–meaning relationships would be reflected in response times. Thus, if sound symbolic relationships exist in spoken language and word learners are sensitive to non-arbitrary relationships between sound and meaning, response times to identify the meanings of newly acquired words would be faster to Japanese vocabulary words that were paired with their actual English equivalents than to Japanese vocabulary words that were paired with an unrelated English word. We also considered the possibility that the learning of antonyms for the English equivalents would result in a processing benefit as well. Because antonyms highlight the same conceptual domain, differing along a single dimension (Murphy & Andrew, 1993), response times to antonyms might be expected to benefit as well from sound symbolic relationships. However, if learners do not utilize sound–meaning correspondences to learn novel words, variations in word-meaning mappings should have little effect on word learning performance.

2. Method

2.1. Participants

Participants were 104 native speakers of American English who reported no familiarity with the Japanese language and had no known hearing or speech disorders. Participants either received course credit (n = 68) or were paid $10 for their participation (n = 36).
2.2. Stimulus materials

The stimuli consisted of 21 Japanese antonym pairs used by Kunihira (1971) and recorded by a native female Japanese speaker (see Appendix). The Japanese antonyms were divided into three groups of seven words equated for frequency of occurrence of their English equivalents (Kucera & Francis, 1967). Three within-subject experimental conditions were constructed. 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 antonyms. In the Random condition, Japanese words were randomly paired with the correct meanings of other words in that same learning condition (see Appendix for pairings). Half the subjects learned one of the Japanese antonyms (List 1 in Appendix) from each pair and half learned the other set (List 2 in Appendix). Across participants, Japanese words were rotated through conditions (Match, Opposite, and Random) so that each Japanese word appeared in each condition. Words were presented in random order in each learn–test cycle.

To create distractors for use in the test phase, in addition to the English words that were learned as equivalents, a second unrelated English word was randomly assigned to each Japanese word, with the caveat that word frequency was equated across pairings. The 21 distractors were English translations of the Japanese antonyms that the participants did not encounter otherwise. The target and distractor words appeared as fixed pairings in all three test blocks.

2.3. Procedure

Three learning phases alternated with three test phases. In each learning phase, participants were asked to listen carefully to each Japanese word and to pay attention to the visually presented English word with which it was paired. Participants were told that they would subsequently be asked to identify the English “translation” for each Japanese word. On each trial, participants saw a fixation cross (500 ms) and then heard a Japanese word through the headphones. The word was then repeated while the English “translation” appeared on the computer screen for 2000 ms.

Following each learning phase, participants completed a test phase. Participants heard the Japanese word once and then saw the two possible English “translations”, the target and distracter, side by side on the computer screen. For example, in the Match condition, the Japanese word “ue” was presented auditorily and participants chose between up (target) and walk (distractor). Participants were asked to make their choice as quickly as possible using a standard button box.

3. Results

3.1. Accuracy

Although most participants readily learned the novel words and reached near ceiling performance across learn–test blocks, a small subset of learners had difficulty with the task. Because our assessment of the processing efficiency of new Japanese vocabulary items required that participants had learned those items, we set a learning criterion of 80% correct across training blocks for inclusion in the study. Based on this criterion, 90 participants were included in the final analysis.

In order to evaluate learning accuracy, participant (F1) and item (F2) analyses of variance (ANOVAs) with block (1, 2, and 3) and condition (Match, Opposite, and Random) as factors were conducted on percent correct identification. The analyses revealed a significant main effect of block, F1(2, 178) = 35.19, partial η² = .283, p < .001; F2(2, 82) = 30.38, partial η² = .426, p < .001. As expected, performance significantly improved over the three training cycles (Mblock1 = 89.2, Mblock2 = 93.9, Mblock3 = 95.8). A significant main effect of condition was also found for participants, F1(2, 178) = 3.17, partial η² = .034, p < .05; F2(2, 82) = 1.93, partial η² = .045, p = .152. No significant interaction was found between block and condition.

Planned comparisons were conducted to evaluate the main effect of condition. Accuracy was significantly better in the Match (M = 94.3) than in the Random condition (M = 91.6) by both participants (p < .01) and items (p < .04). No significant differences were found between Match and Opposite (M = 93.0) or between Opposite and Random conditions by either participants or items.

3.2. Response time

Fig. 1 shows mean response time as a function of condition and training cycle. Response times 2 SDs above or below the mean were removed, which accounted for elimination of less than 5% of the data. Participant (F1) and item (F2) ANOVAs with block and condition as factors revealed a significant main effect of block, F1(2, 178) = 109.42, partial η² = .551, p < .001; F2(2, 82) = 156.59, partial η² = .792, p < .001, indicating an expected overall decrease in response times across the three blocks. A marginally significant main effect of condition was found by participants, F1(2, 178) = 2.91, partial η² = .032, p < .06, but not by items, F2 < 1. However, a significant interaction between block and condition was found, F1(4, 356) = 2.53, partial η² = .028, p < .04; F2(4,164) = 2.823, partial η² = .064, p < .03, indicating that the pattern of response times across blocks changed as a function of condition, Match, Opposite, or Random. Simple main effects conducted for each block indicated a significant effect of condition in Block 3, F1(2, 178) = 4.37, partial η² = .047, p < .02; F2(2, 82) = 3.04, partial η² = .069, p = .053, but not in Blocks 1 and 2. Planned comparisons between conditions for Block 3 revealed significantly better performance in the Match than in the Random conditions, p < .02, by both participants and items. There was a significant difference between Opposite and Random conditions by participants (p < .05), but not items. No differences between Match and Opposite were found either for participants or items. Participants responded more quickly overall to Japanese words paired with their correct English equivalents.
drew (1993) have shown that a word’s antonym is their antonym may have benefited from non-arbitrary words that were paired with the English equivalent of associating novel words with meanings. Within the complex, and arguably taxing, cognitive task of language, but they also recruit this sensitivity only do learners encode and represent the sound proper-

ess of mappings between sound and meaning. Not constraints that facilitate on-line acquisition and pro-

arbitrary or sound symbolic properties of speech provide meaning was unrelated. This finding implies that non-

form and meaning matched than when the learned

Learners responded significantly faster when the word relationships emerged in the response time measure. During the first learning block, identification perfor-
mance was more accurate when sounds and meaning matched than when sounds and meaning were randomly paired. Across learning–test blocks, as identification performance neared ceiling, the effect of sound symbolic relationships existed in the response time measure. Learners responded significantly faster when the word form and meaning matched than when the learned meaning was unrelated. This finding implies that non-arbitrary or sound symbolic properties of speech provide constraints that facilitate on-line acquisition and processing of mappings between sound and meaning. 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.

 Interestingly, the results also suggest that Japanese words that were paired with their correct English equivalents during a garden-variety vocabulary-learning task were responded to more quickly and recognized more accurately than incorrect English-to-Japanese mappings. During the first learning block, identification performance was more accurate when sounds and meaning matched than when sounds and meaning were randomly paired. Across learning–test blocks, as identification performance neared ceiling, the effect of sound symbolic relationships emerged in the response time measure. Learners responded significantly faster when the word form and meaning matched than when the learned meaning was unrelated. This finding implies that non-arbitrary or sound symbolic properties of speech provide constraints that facilitate on-line acquisition and processing of mappings between sound and meaning. 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.

Interestingly, the results also suggest that Japanese words that were paired with their correct English equivalents of their antonym may have benefited from non-arbitrary sound to meaning mappings as well. Murphy and Andrew (1993) have shown that a word’s antonym is highly similar conceptually, differing in meaning only along a single value (see also, Keysar & Bly, 1995). Consequently, if sound-to-meaning correspondences facilitate access to a particular semantic domain, then related concepts or meanings, particularly antonyms, should also benefit from that facilitation. Thus, learners in the current task appeared sensitive both to the sound symbolic characteristics of the Japanese words as well as to the relatedness of learned antonym meanings.

These findings are consistent with other demonstrations of sensitivity to mappings between the sound structure of language and lexical class or meaning (Bergen, 2004; Berlin, 1994; Brown et al., 1955; Farmer et al., 2006; Imai et al., 2008; Kelly, Springer, & Keil, 1990; Kunihira, 1971). However, the particular instantiation of sound symbolic facilitation demonstrated here occurred for mappings that cut across two languages, English and Japanese, from distinct lineages and with unique phonologies. That native English speakers were sensitive to non-arbitrary sound to meaning mappings in Japanese suggests that performance on this task cannot be explained exclusively by invoking learned within-language regularities. Although previous work has shown that language users can guess meanings cross-linguistically (Berlin, 1994; Kunihira, 1971), our findings demonstrate a genuine functional role in language learning for these cross-linguistic mappings. In the current task, non-arbitrary correspondences between sound and meaning had an implicit effect on the accuracy and efficiency with which learners identified the learned English equivalents of Japanese words. Sound symbolic relations influenced basic on-line lexical and learning processing in spoken language.

It should be noted that learners may have also been sensitive to reliable cues to meaning in the prosodic contours of the Japanese utterances. Indeed, previous work has shown that listeners can infer word meaning from prosody, suggesting that prosody can convey semantic information both within and outside the domain of emotion (Kunihira, 1971; Nygaard et al., 2009; Shintel, Nusbaum, & Okrent, 2006). Indeed, Kunihira (1971) found superior performance when his Japanese antonyms were pronounced with “expressive” prosody. However, he also found that printed and “monotone” versions of Japanese antonyms were matched to their respective referents at above chance rates, indicating that phonemic content alone was sufficient to disambiguate word meaning. Although it is difficult to rule out a possible contribution of prosody, Kunhira’s findings foreshadow our learning effects and suggest that our learners were sensitive to the sound symbolic properties inherent in the Japanese words, in addition to prosodic information. Future research will be needed to determine the relative contribution of these two aspects of spoken language to learners’ ability to associate words with their referents.

One potential mechanism by which sound symbolic properties may facilitate word-learning cross-linguistically is that these mappings may reflect a sensitivity to general perceptually-based cross-modal links between
sounds and aspects of meaning (Marks, 1978; Martino & Marks, 1999; Melara, 1989). Properties of the sound structure of language may activate or resemble characteristics inherent in the object, event, or property to which the sound sequence refers, and language learners may use this inherent cross-modal mapping to facilitate the retrieval of word meaning. Such cross-modal correspondences may be achieved via some literal or figurative resemblance between the sound and meaning (e.g., vowel height may correlate with relative size), or may reflect an embodied representation involving simulation of the actual meaning.

Although it is unclear which sound properties of the 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 the task of acquiring and accessing novel words in a large lexicon with a rich, complex semantic space, non-arbitrary relationships may not necessarily be specific. Instead, they may operate in a general and flexible manner to facilitate the mapping between word forms and meanings. Thus, reliable mappings between sound and meaning, whether they be probabilistic or inherently cross modal, could work to constrain novel word learning and subsequent word retrieval and recognition by guiding processing to properties and meaning within a particular semantic context.

The current investigation provides one of the first demonstrations that learners can use sound symbolism to derive meaning during spoken language processing. These results challenge the traditional assumption that words bear an exclusively arbitrary relationship to their referents and that lexical processing relies solely on this arbitrary relationship. Rather, the sound structure of spoken language may engage cross-modal perceptual-motor correspondences that permeate the form, structure, and meaning of linguistic communication (Martino & Marks, 2001; Ramachandran & Hubbard, 2001). As such, sound symbolic properties of language may be evolutionarily preserved features of spoken language that have functional consequences not only for word learning and processing in adults, but also for the acquisition of spoken language in young children. Although arbitrariness certainly remains a central design characteristic of linguistic structure, these results indicate that language users also can and do exploit non-arbitrary relationships in the service of word learning and retrieval.

Acknowledgements

We thank three anonymous reviewers for helpful comments and suggestions on a previous draft of this paper. We also thank Jessica Alexander and Stig Rasmussen for stimulus and protocol preparation. An Emory College Scholarly Research and Inquiry at Emory (SIRE) undergraduate research grant to the second author supported a portion of this research.

Appendix

Japanese Word | Match | Opposite | Random
--- | --- | --- | ---
**List 1**
akarui | bright | dark | catch
nageru | throw | catch | wide
nureta | wet | dry | bright
omoi | heavy | light | sit
semai | narrow | wide | up
suwaru | sit | stand | dry
ue | up | down | light
amai | sweet | sour | cool
katai | hard | soft | rise
futoi | fat | thin | hard
neru | lie | rise | sharp
nibii | blunt | sharp | fat
suzushii | cool | warm | bad
ii | | | 
ara | rough | smooth | short
aruku | walk | run | slow
hayai | fast | slow | rough
mijikai | short | long | walk
tetsu | iron | gold | move
tsuioi | strong | weak | iron
ugoku | move | stop | weak
kurai | | | 
toru | catch | throw | light
kawaita | dry | wet | narrow
karui | light | heavy | catch
hirooi | wide | narrow | blunt
tatsu | stand | sit | down
shita | down | up | dark
suppai | sour | sweet | warm
yawarakai | soft | hard | lie
hosoi | thin | fat | blunt
okiru | rise | lie | soft
surudo | sharp | blunt | bad
atatakai | warm | cool | sweet
warui | bad | good | thin
nameraka | | | 
hashiru | run | walk | smooth
osoi | slow | fast | weak
negai | long | short | run
kin | gold | iron | long
yowai | weak | strong | gold
tomaru | stop | move | fast

*Words that were modified from the original list in Kunihira (1971).*

References


