Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Signature:

_______________________________ _______________

Kaitlyn R. Bankieris 4/20/10

Sound Symbolism and the Perception of Shape and Implied Motion

by

Kaitlyn R. Bankieris

Adviser

Lynne Nygaard, PhD

Program in Neuroscience and Behavioral Biology

____
Lynne Nygaard, Ph.D.
Adviser

____
Laura Namy, Ph.D.
Committee Member

____
James Rilling, Ph.D.
Committee Member

____
4/20/10

Sound Symbolism and the Perception of Shape and Implied Motion

By

Kaitlyn R. Bankieris

Adviser

Lynne Nygaard, Ph.D.

An abstract of
A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Program in Neuroscience and Behavioral Biology

2010

Abstract

Sound Symbolism and the Perception of Shape and Implied Motion

By Kaitlyn R. Bankieris

A growing body of sound symbolism research is beginning to challenge the classic linguistic assumption that the pairing of word to referent is arbitrary. The present study investigated this sound symbolism phenomenon during on-line processing and the extent to which sound symbolism relies on inter-sensory cross-activations. A cross-modal priming task was used, with foreign words meaning *pointy*, *round*, *still*, or *moving* as auditory primes and pictures varying in shape (pointy or round) or implied motion (still or moving) as targets. Participants classified pictorial targets (pointy or round, still or moving) during trials presented across three conditions: 1) Match, in which the prime was sound symbolic for the perceptual property of the pictorial target; 2) Mismatch, in which the prime was sound symbolic for the perceptual property opposite of the pictorial target; and 3) Control, in which the prime was not sound symbolic. The results showed that symbolic (Match and Mismatch) trials facilitated processing in the motion domain, but interfered with processing in the shape domain. The current results provide preliminary evidence that non-arbitrary sound to meaning mappings in natural language may affect on-line processing of pictures and may result from low-level cross-modal connections.

Keywords: sound symbolism, arbitrariness, cross-modal, priming

Sound Symbolism and the Perception of Shape and Implied Motion

By

Kaitlyn R. Bankieris

Adviser

Lynne Nygaard, Ph.D.

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Program in Neuroscience and Behavioral Biology

2010

Acknowledgements

This research was funded by a Scholarly Inquiry and Research at Emory (SIRE) grant. I thank Lauren Clepper for assistance with running participants and analyzing data as well as the members of the Speech and Perception Lab and Language Group for thoughtful insights into experimental design and results.

Table of Contents

Sound Symbolism and the Perception of Shape and Implied Motion

Plato's philosophical discussion of the correctness of names in *Cratylus* is one of the earliest works addressing a question of significant importance to linguists today: what is the relationship between a word and its referent? Cratylus declares a correctness of names, explaining that an entity and its name are naturally connected (Plato, 1988). Expressing the opposite position, Hermogenes says the only connection between an object and its name is necessarily formed by communal agreement. Without convention, Hermogenes believes, there is no connection between an object and its label. Cratylus and Hermogenes present this issue to Socrates who synthesizes the opposing theories. Socrates concludes that at the deepest level, names are naturally connected to their referents, but on the surface level convention dictates names.

In contrast to Socrates' conclusion that both arbitrary and non-arbitrary sound to meaning correspondences exist in natural language, linguists have classically regarded the *arbitrary* relationship between sound and referent as a fundamental feature of human language (de Saussure, 1959; Hockett, 1977). Indeed, arbitrariness has been argued to be a necessary design feature of human language. The decoupling of the sound structure of words from characteristics of their related referents is assumed to confer a referential power to language, in principle allowing for an infinite number of sound to meaning pairings (de Saussure, 1959; Hockett, 1977). Gasser (2004) demonstrates this predicted benefit of arbitrariness in language with a computational model of language learning. Presented with a set of lexical items to learn, Gasser's model makes fewer mistakes when these items are arbitrarily paired to their meanings compared to when they are iconically paired. These findings suggest that arbitrary mappings of phonological composition to meaning are advantageous in learning a large lexicon. Support for de Saussure's principle is also evident in the variety of word forms that refer to the same object, action, or event across languages. For example, the same object is referred to as *book, knijge, libro,* and *kitap* by speakers of different languages. If the sound structure of words corresponded to or resembled characteristics of referents, this wide variance across languages should not occur. Rather, similar inventories of sounds should reliably correspond to similar types of meaning across languages.

Although abundant support exists for arbitrariness in spoken language, several counter examples suggest that the arbitrariness assumption may not entirely characterize sound to referent mappings in natural language. The most obvious example of reliable correspondences between sound and meaning (*sound symbolism*) is onomatopoeia (e.g. *oink, bam*). These non-arbitrary mappings use the sound structure of words to mimic the sounds of the referent. Onomatopoeia is common across languages, but involves different language-specific constraints. For example, different languages refer to the sound a dog makes with various phonemic combinations such as *woof, gav, ham,* and *au.* Furthermore, onomatopoeia represents a small proportion of all words, suggesting this type of sound symbolism is a restricted phenomenon found in a narrow set of situations.

Mimetics, a sound-symbolic word class in Japanese, is similar to onomatopoeia but is more prevalent and extends beyond mimicry. Mimetics include onomatopoeiac words mimicking the sound they represent, but move beyond this sound-to-sound correspondence encompassing a larger proportion of words referring to tactile, visual, and emotional experiences in which sound is not inherently incorporated (Kita, 2001). The mimetics *goro, koro, guru,* and *kuru* meaning "a heavy object rolling, a light object rolling, a heavy object rotating around an axis” and “a light object rotating around an object,” respectively, illustrate the characteristics of the sound symbolism in this class of words. For example, Kita (2001) has found that combinations of “g/k” and “r” appear to represent rotation, with an initial voiced consonant denoting a large mass and an initial voiceless consonant signaling a small mass. Since the above examples do have sounds associated with the referent's action (e.g., the sound of a heavy object rolling), one may argue that mimetics are an extended example of onomatopoeia. However, mimetics without associated sounds also exist. Mimetic words referencing states or actions with no associated sound (e.g., *nurunuru* “tactile sensation caused by a slimy object”) support the sound-to-meaning correspondence present in this class of words and highlight potential cross-modal or inter-sensory correspondences (Kita, 1997). Additional support for potential inter-sensory correspondences comes from native speakers’ reports that hearing or reading mimetic words creates a simulation of the sensory experience caused by the actual referent (Kita, 1997).

As a class of words with sound symbolic properties, mimetics is not unique to Japanese. Similar word categories exist in other languages across the world. Sound symbolic classes are found in many East Asian languages, South East Asian languages (expressives), sub-Saharan African languages (ideophones), and Northern Aboriginal Australian languages (Voeltz & Kilian-Hatz, 2001). McGregor (2001) reports sound symbolism similar to that of Japanese mimetics found in some Northern Aboriginal Australian languages’ uninflected verbs. In these languages, this class of uninflected verbs is distinguished by having fewer syllables than typical of other words, ending in consonants and consonant clusters, and exhibiting a tight link between final phoneme and semantics. For instance, uninflected verbs ending in *-ng* refer to processing involving blunt or bent objects, or hollow or resounding sounds. In contrast, words with a final lateral consonant refer to actions involving liquids such as *duburl* “to swim” and *dulul* “to pour” (McGregor, 2001). Like Japanese mimetics, these sound symbolic word classes found across several languages express temporal structure and affective states as well as convey information from various perceptual modalities (Imai, Kita, Nagumo, & Okada, 2008).

Sound symbolism outside of onomatopoeia is also found in English and Indo-European languages, although it is not readily identifiable by a distinct grammatical class of words. Phonaesthemes are combinations of sounds frequently occurring in words with a shared semantic meaning (Bergen, 2004). For example, the consonant cluster *sn-* commonly occurs in words related to “nose, mouth” such as *snore, snack, snout, snarl,* and *sniff.* The lack of formal identification of these reliable sound-to-meaning pairings does not diminish their psychological reality. Phonaesthemic priming, which cannot be attributed to shared morphology or semantics alone or in combination, demonstrates listeners’ sensitivity to these sound symbolic correspondences (Bergen, 2004). These sound symbolic occurrences, however, are not highly prevalent throughout multiple grammatical classes in any given language and appear to be the result of within language conventional mappings between sound and meaning.

A growing body of research suggests that subtler forms of sound-meaning pairings occur frequently within specific languages outside of particular specialized word classes. These examples challenge the marginal nature of sound symbolism and the exclusive arbitrariness of language. One such non-arbitrary connection is a link between grammatical class and phonological characteristics. For example, through a corpus analysis, Farmer, Christiansen and Monaghan (2006) found that English nouns and verbs have category-typical phonological properties, with characteristics of nouns and verbs forming relatively distinct clusters. That is, nouns are phonologically closer to other nouns than they are to verbs whereas verbs are phonologically closer to other verbs than they are to nouns. Furthermore, Farmer, et al. found that listeners were sensitive to these properties during on-line processing tasks. Participants displayed a response-time advantage in naming category-typical nouns and verbs. In addition, Farmer, et al. found that a word's phonological typicality influenced on-line sentential processing, with incongruence between phonological form and word class inhibiting word and sentence processing (but see Staub, Grant, Clifton, & Rayber, 2009).

Supporting the phonological distinction between grammatical classes, nouns and verbs have been found to contain category typical vowels (Sereno & Jongman, 1990) and number of syllables (Cassidy & Kelly, 1991). Cassidy and Kelly (1991) reported that both adults and children used the syllable distinction to assign meaning to nonwords, suggesting a psychological reality to these statistical patterns. These phonological differences across grammatical class also influence word-learning in children (Cassidy & Kelly, 2001) and even young infants appear to discriminate between nouns and verbs on the basis of non-arbitrary phonological structure (Shi, Werker, & Morgan, 1999). These findings suggest that reliable correspondences between sound structure and word class exist statistically in natural languages and that these relationships influence the processing of spoken language.

Consistent correspondences between the sound structure of words and *semantics* have also been demonstrated. Köhler (1929) found that English-speaking adults reliably match nonsense words such as *baluma* to rounded shapes and nonsense words such as *takete* to angular shapes. Arbitrary mappings between phonological composition and meaning do not predict these findings. Replicating and extending Köhler's research, Maurer, Pathman, and Mondloch (2006) report that adults and children as young as 2.5 years old demonstrate the same phenomenon. In a forced choice task, participants labeled round objects with nonsense words such as *bouba* and angular objects with nonsense words such as *kiki*. Since these studies did not systematically alter phonological composition of the nonwords used, it is unclear whether the differing consonants, vowels, or a combination thereof created the dichotomous labeling of round and angular items. Regardless of the phonological characteristics contributing to this effect, these results demonstrate consistent mappings between word structure and meaning. Further supporting these consistent mappings from word form to semantics, Westbury (2005) presented English-speaking adults with printed nonwords containing stop consonants or liquids inside of rounded or angular shapes. In a lexical decision task, participants were slower to reject nonwords containing stop consonants presented inside of angular shapes and nonwords with liquids inside of rounded shapes. These results suggest that the relationship between phonological or orthographic word forms and surrounding context influenced on-line speech processing. That Westbury's experiment did not require participants to make decisions about the correspondence between shape and word form suggests an automatic sensitivity to these correspondences.

Extending beyond nonwords, sound symbolism is also found cross-linguistically. Kunihira (1971) presented native English speaking adults with Japanese antonym pairs in three conditions: 1) expressive voice, 2) monotone without expressive voice, and 3) printed in Romanized characters. Participants were instructed to guess the meanings of the Japanese antonym pairs with the corresponding English antonym pairs presented. Kunihira found that participants selected the correct meanings significantly above chance in all three conditions. The ability of English speakers to assign the correct meaning to Japanese words using only the phonological properties (monotone without expressive voice) suggests that the sounds of the words themselves allowed English speakers to determine their meanings. Other studies have found that native English speakers also correctly identified meanings of foreign dimensional adjectives in Chinese, Czech, Hindi, Japanese, and Tahitian suggesting a universal presence of sound symbolism (Brown, Black, & Horowitz, 1955; Klank, Huang, & Johnson, 1971).

Berlin (1994) demonstrated that native English speakers' sensitivity to sound symbolism extends beyond antonyms to words pertaining to bird and fish names in Huambisa, a language spoken in Peru. Native English-speaking participants were presented with pairs of bird and fish names both spoken and written. When asked to choose which of two words corresponded to the bird name, native English speakers reliably chose the correct bird name at rates significantly above chance. Berlin analyzed the acoustic characteristics of the words and found that high frequency segments appeared to characterize bird names and low frequency segments characterized fish names. Thus, native English speakers were able to detect the phonological patterns distinguishing birds from fish present in Huambisa. These cross-linguistic studies demonstrate the prevalence of sound symbolism in languages across the world. The ability of language users to recognize corresponding sound to meaning mappings in languages other than their own suggests that these mappings may be consistent across languages.

Beyond recognizing sound to meaning correspondences in foreign languages, language users also appear to use sound symbolism in word learning. Nygaard, Cook, and Namy (2009) demonstrated listeners' cross-linguistic sensitivity to and functional use of sound symbolism in a novel word-learning task. Native English speakers learned English meanings of the Japanese antonym pairs from Kunihira (1971) in one of three conditions: 1) English meanings matched Japanese meanings; 2) English meanings were the antonym of Japanese meanings; or 3) English meanings were unrelated to Japanese meanings. Participants then identified the learned meanings of the Japanese words in a speeded choice task. The results indicated that word learning was facilitated when foreign words were paired with their actual English equivalent or the English antonym of their English equivalent over pairings with a non-related English word. That word learning was facilitated for foreign words not necessarily selected to be sound symbolic supports the existence of cross-linguistic sound to meaning correspondences and, further, suggests that listeners use these correspondences during word learning processes.

The benefit of sound symbolism for word learning has also been found in children. Imai, et al. (2008) presented 3-year-old Japanese children with a verb learning task involving novel sound symbolic and non-sound symbolic verbs. Children were tested to determine if they would generalize the meaning of a verb to a different actor on the basis of the sameness of action. The 3-year-olds did not successfully generalize the meanings of non-sound symbolic verbs across actors but did correctly generalize the meanings of sound symbolic verbs on the basis of sameness of action. Thus, sound symbolism seems to play a facilitatory role in early verb learning in children in addition to later vocabulary learning in adults.

Sound symbolism has also been reported to facilitate categorization (Kovic, Plunkett, & Westermann, 2010). Kovic, et al. demonstrated that without explicitly drawing participants’ attention to sound-meaning correspondences, sounds symbolism affects categorization behavior. In this experiment, participants learned to label two groups of nonsense objects as *mots* and *riffs* either in a congruent condition in which *mots* had rounded characteristics or an incongruent condition in which *mots* had angular characteristics. During test trials, participants were presented with an object and a label and decided if each pair was a match or a mismatch. Behavioral results indicated that participants were faster to identify a match and slower to identify a mismatch for congruent label-object pairings. Thus, participants appeared to be sensitive to the associations between characteristics of the object and the sound symbolic labels. These results may approximate the effect of sound symbolism in every day language use as participants’ attention was not drawn to the sound symbolic connections and they reported being unaware of any such correspondences.

Event related potential (ERP) data from this study provides neurophysiological evidence that language users' sensitivity to sound symbolic relationships between labels and objects may have been due to sensory-perceptual cross-modal mappings. ERP signals in the parietal-occipital regions differed from congruent to incongruent conditions as early as 140-180ms following presentation of the visual object. The observed difference was that congruent conditions displayed an early negative component present only weakly in incongruent conditions. Molholm, et al. (2002) report similar results suggesting that this negative ERP component may indicate audio-visual integration. This study reports differences in congruent and incongruent audio-visual presentations beginning 145ms post stimuli presentation. Akin to Kovic, et al., the observed difference was a stronger negativity in the parietal-occipital area for congruent conditions compared to incongruent conditions. According to these findings, both response time and neural patterns support the psychological reality and facilitatory role of sound symbolism during categorization. Furthermore, these results suggest that sound symbolism may arise from cross-modal integration.

Ramachandran and Hubbard (2001) have suggested that links between phonological properties of labels and perceptual properties of referents do arise from inter-sensory neural cross-activations. Specifically, Ramachandran and Hubbard offer three hypotheses for the contribution of neuronal cross-activations to non-arbitrary sound to meaning mappings. First, they conjecture that sound and meaning may be connected via the nature of articulatory patterns during vocalization and visual properties of the referent. For example, words referring to smallness of an object may encode this visual property in the sensory experience of motor movements used to produce them. The vowel /i/ is commonly found in words meaning “small” in English (e.g. *petite*, *teeny*) and is produced by narrowing the lips and vocal tract. Another possibility proposed by Ramachandran and Hubbard is that connections between two motor maps create consistent correspondences between meaning and word structure. Using the example above, the lip movement associated with the production of /i/, and therefore smallness, may mimic the finger movement used to gesture smallness. Finally, reliable correspondences in language could result from non-arbitrary mappings between phonological representations and visual object properties. This proposal is of particular relevance as such inter-sensory connections between auditory and visual cortices would explain the consistent correspondences between shape and word form that have been reported (Köhler, 1929; Kovic, et al., 2010; Maurer, et al., 2006). This cross-modal activation theory of sound symbolism is supported by Ramachandran and Hubbard's report of an anomic aphasic that did not consistently map words like *bouba* and *kiki* to round and spiky shapes, respectively. This patient suffered damage to the left angular gyrus, an area known to be involved with cross-modal associations, highlighting the possible role of cross-modal activations in sound symbolism.

Listeners' sensitivity to cross-modal interactions supports this inter-sensory explanation for sound symbolism and suggests that sound symbolism may reflect a more general sensitivity to cross-modal correspondences between auditory and other sensory modalities. Investigating cross-modal correspondences, Marks, Ben-Artzi, and Lakatos (2003) presented bimodal stimuli varying in pitch and brightness, or loudness and brightness and asked participants to discriminate the stimuli based on pitch, or on loudness. Unrelated variations in brightness and loudness were found to affect participants' pitch discrimination suggesting that cross-modal interactions between pitch and brightness and loudness and brightness exist, and affect on-line processing. Additional research by Melara and Marks (1990) has demonstrated links between semantics and varied pitch and loudness. In a speeded classification task, participants were visually presented with the words *high* and *low* paired with sounds varying in pitch or loudness. Classification reaction times to congruent pairings, *high* presented with a loud or high frequency sound or *low* presented with a soft or low frequency sound, were faster than reaction times to incongruent pairings suggesting that auditory-visual cross-modal mappings influenced perceptual processing. These cross-modal mappings between sound and visual properties have been demonstrated in preverbal infants as young as 3- to 4-months old (Walker, et al., 2010). In a preferential looking paradigm, preverbal infants demonstrated sensitivity to correspondences between auditory pitch and height as well as auditory pitch and sharpness, suggesting that these cross-modal mappings may be unlearned. These experiments demonstrate that consistent cross-modal auditory-visual mappings do exist and that they automatically affect on-line processing. It is possible that the presence of correspondences such as these in language could contribute to the sound symbolism phenomenon.

Reliable mappings between other types of auditory properties of language and visual properties have been demonstrated more directly. Shintel, Nusbaum, and Okrent (2006), suggest that these mappings are an intricate part of language use, automatically affecting both production and perception. Shintel, et al. conducted an experiment in which one set of participants described the motion of a dot moving at varying speeds by saying, “It is moving right/left.” Notably, the speed of the dot was not linguistically encoded in the utterance. A second set of participants judged the speed of the dots using the first set of participants’ sentences, which described the direction of the dot’s movement only. Listeners were able to accurately judge the speed of the dots, suggesting that speakers naturally encode meaningful properties extracted from visual input into the acoustic characteristics of their speech and that listeners are sensitive to this encoded cross-modal information. In a similar experiment, Shintel and Nusbaum (2007) presented participants with sentences varying in speaking rate. The content of the sentences in this task included references to particular objects, but did not include any information about motion or speed. After presentation of the sentence, listeners then viewed a picture with or without implied motion (e.g., a galloping versus a standing horse) and were asked to identify if the object (e.g., the horse) had been mentioned in the previous sentence. Faster recognition occurred when the motion implied by the speaking rate matched that implied by the picture. These results demonstrate listeners' sensitivity to the semantic information encoded in speaking rate and the processing advantage resulting from congruency. These experiments are consistent with findings demonstrating listeners' ability to detect cross-modal connections both conveyed in and outside language processing, demonstrating a general processing benefit for congruent connections across modalities.

These cross-modal connections evident between non-linguistic and visual properties may serve as one potential mechanism for sound symbolic or non-arbitrary connections between the sound structure of words and semantics. To date, it is unclear whether inter-sensory connections underlie sound-to-meaning correspondences in language and if these connections are activated automatically during on-line language processing. The current study aimed to address listeners' automatic sensitivity to cross-modal correspondences between sounds and visual features of an object during on-line processing. A word to picture priming paradigm was used to address these questions.

The set of priming words consisted of sound symbolic foreign words that had been judged to correspond to the meanings *still, moving, round,* or *pointy*. A set of control items consisted of non-sound symbolic foreign words drawn from the same languages. Target images were pictures of objects or animals that displayed shape characteristics of round or pointy or implied motion characteristics of still or moving. Participants first heard the foreign word prime and then were presented with the picture. They were instructed to make a shape (round or pointy) or implied motion (still or moving) discrimination for each target image presented. Importantly, they were not directed to attend to the prime or pay attention to any relationship between the sound and visual properties of the image. This lack of explicit attention allowed for an examination of the automaticity of sound symbolism. The study evaluated whether sound symbolic properties of the foreign primes would directly influence processing time during the picture classification task. Thus, the investigation provides insight into the extent to which sound symbolism in language relies on cross-modal connections between sensory-perceptual domains.

We predicted that sound-symbolic foreign primes that are judged as consistent with the meanings *still, moving, round,* and *pointy* would influence response times during the picture decision task, suggesting that sound symbolic properties are mapped cross-modally to specific aspects of the visual scene and used during on-line processing. Specifically, we expected that response times in conditions with congruent relationships (Match) between primes and targets would be faster than those in conditions with incongruent (Mismatch) or no relationship (Control) between primes and targets. If reliable sound-to-meaning mappings prime the picture decision task, it would suggest that cross-modal connections exist and are activated during on-line processing. According to this prediction, when a prime judged to be sound symbolic for the meaning *round* is presented, auditory regions processing the sound symbolic phonetic components will be activated. The activation of these phonetic processing areas will cause activation of non-arbitrarily connected visual areas processing round properties. Thus, the auditory input will prime the visual discrimination of the picture target and facilitate picture processing and response times. We predicted two possible outcomes for the Mismatch conditions. One, the Mismatch condition could produce the slowest response times. These slow response times would reflect inhibited processing as a result of the prime and target having opposite meanings. Alternatively, if sound symbolism has a more general facilitative effect, response times to mismatch trials could fall in between response times of control and match conditions. This intermediate response time would indicate that although the meanings of prime and target are incongruent for specific meaning, they are semantically related in the shape or motion domain.

Method

Participants

73 native English speakers between the ages of 18 and 35 with no reported speech or hearing disorders and no familiarity with any of the ten languages used participated in this study. Participants were compensated with $15 or received course credit for their participation.

Stimuli

In order to examine the presence of sound symbolism in naturally occurring languages, words from ten different foreign languages were used as auditory primes. Using a variety of languages addresses the general presence of cross-modal sound symbolism, rather than sound to meaning mappings restricted by the conventions of a particular language. The priming stimuli consisted of a set of foreign words sound symbolic for the meanings *still*, *moving*, *round*, and *pointy* (n = 40) and a set of foreign words that are non-sound symbolic with the same meanings (n = 20) (Appendix A). The primes were selected from a larger database in which native speakers of ten foreign languages (Albanian, Dutch, Gujarati, Indonesian, Korean, Mandarin, Romanian, Tamil, Turkish, and Yoruba) nominated synonyms for the meanings of a set of nine dimensional adjective pairs. These words were then digitally recorded by native speakers of each language and edited into separate files for presentation to listeners. Words from this larger set that were used in the present study as sound symbolic primes were judged by separate groups of 15 native English speakers to mean *still* or *moving* or to mean *round* or *pointy* with at least 80% consistency. Those items that were used as non-sound symbolic primes were judged at approximately chance (50%) on the still/moving or round/pointy classification. In this rating task and in the current experiment, primes were always presented auditorally over headphones. Presenting the primes auditorally rather than visually, addressed the connection between phonological structure and visual form rather than between orthographic and visual form.

Target stimuli were 15 pairs of static pictures depicting either implied or no motion and 15 pairs of pictures depicting objects that are round or pointy (Appendices B & C). The 15 pairs of pictures representing motion and shape were chosen from 20 pairs of motion pictures and 20 pairs of shape pictures. A separate group of 40 participants judged each picture in a still/moving or round/pointy discrimination task. For the implied motion classification, the 15 still and moving pictures with the highest combined accuracy were selected to form the set of motion targets. Likewise, the 15 round and pointy pictures with the highest combined accuracy were selected as the shape targets. All stimuli were classified with at least 90% accuracy. Table 1 shows mean percent correct classification performance and standard error for each picture target type.

Three conditions varying the relationship between prime and target, Match, Mismatch, and Control, were manipulated within subjects. The Match condition paired a sound symbolic word with a related picture while the Mismatch condition paired a sound symbolic word with an image implying the opposite meaning. For example, a Match trial presented a foreign prime judged to be sound symbolic for the meaning *round* (e.g., *bombat*) followed by an image of a round object. A Mismatch trial paired a foreign prime judged to be sound symbolic for the meaning *pointy* (e.g., *kesici*) with a picture of a round object. The Control condition consisted of non-sound symbolic primes paired with the picture targets. Shape and motion trials were blocked and the presentation order of blocks as well as the assignment of target picture to condition were counterbalanced across subjects. Trials within each shape or motion block were presented randomly to participants.

Procedure

Testing occurred in sound attenuated rooms using E-Prime Version 2.0 software. Before beginning, participants were informed that they would be hearing words and seeing images during the experiment. They were told that their task was to discriminate the pictures based on shape or implied motion, deciding whether each image presented was still or moving, or round or pointy. Importantly, participants were not directed to attend to the prime or pay attention to any relationship between the sound and visual properties of the image. Before each block, instructions on the computer screen notified participants which domain (motion or shape) they would be making their decisions in. During each trial, participants saw a fixation cross (750ms), heard the priming word over headphones (average duration 681ms), and then immediately saw a target image (interstimulus interval 0ms). Placing their index fingers on two buttons labeled still and moving or pointy and round, participants made a shape or implied motion judgment on the target image with a button box. Location (left/right) of endpoint labels was counterbalanced across subjects and the image disappeared when a response was made. 500ms elapsed between each trial. Response times were measured from the onset of the picture presentation and only the response times for correct decisions were analyzed.

Results

Because the main dependent variable of interest was response time (RT), a criterion of 80% accuracy across domains was set to insure participants understood and were focused on the task. This resulted in the exclusion of three participants, leaving 70 participants in the final analysis. All analyses were conducted by participant.

Response accuracy

Table 2 reports mean proportion correct and standard error across domain and condition. To evaluate any changes in accuracy across domain (motion, shape) and condition (match, mismatch, control), a 2 x 3 repeated measures analysis of variance (ANOVA) was conducted on mean proportion correct. A significant main effect of domain was found, $F(1, 69) = 11.93$, *partial* $\eta^2 = .15$, $p < .001$. Participants' judgments were more accurate for shape trials (M = .97) than for motion trials (M = .95). Thus, although performance was highly accurate on both the shape and motion discrimination tasks, the shape discrimination task was easier than the implied motion discrimination task. In addition, the analysis did not reveal a main effect of condition or an interaction between domain and condition, suggesting that discrimination performance did not differ as a function of these variables.

Response time

Incorrect trials and response times of correct trials 2.5 standard deviations (SD) above or below mean RT were removed by participant, which eliminated less than 3% of the data. Mean response time and standard error across domain and condition are reported in Table 3. To evaluate whether performance varied as a function of meaning domain and word to picture matching condition, a 2 x 3 repeated measures ANOVA was conducted on mean reaction times with the within subjects variables of domain (motion, shape) and condition (match, mismatch, control). The analysis revealed a significant main effect for domain, $F(1, 69) = 44.89$, *partial* $\eta^2 = .39$, $p < .001$. Participants were faster overall to respond on shape trials (M = 821) than on motion trials. (M = 926). Neither a main effect of condition nor an interaction between domain and condition was found. Performance did not appear to differ significantly depending on the sound symbolic relationship between word and meaning for either domain.

Although the initial analysis did not reveal an effect of matched or mismatched sound symbolism, previous research has found that a sound symbolic relationship between a foreign word and both its correct English equivalent and its antonym results in facilitated performance (Nygaard, Cook, & Namy, 2009). In order to determine if sound symbolism in general resulted in picture priming, a combined sound symbolism variable was created in which all sound symbolic trials (match + mismatch) were averaged. A 2 x 2 repeated measures ANOVA with the within subjects factors of domain (motion, shape) and sound symbolism (control, symbolic) was conducted to analyze whether sound symbolism in general influenced performance. Again, a significant main effect of domain was found, $F(1, 69) = 53.5$, $partial\ \eta^2 = .44$, $p < .001$. Overall, participants responded faster on shape trials (M = 817) than on motion trials (M = 930). Although no significant main effect of sound symbolism was found, a significant interaction between domain and sound symbolism condition did appear, $F(1, 69) = 4.47$, $partial\ \eta^2 = .06$, $p < .05$. As seen in Figure 1, for the motion domain, sound symbolic trials ($M_{MotionS} = 921$) displayed a RT advantage over control trials ($M_{MotionC} =938$). The opposite pattern was observed for the shape domain. Control trials ($M_{ShapeC} = 804$) displayed an RT advantage over sound symbolic trials ($M_{ShapeS} = 829$). T-tests were conducted to evaluate the effect of sound symbolism within each domain. For the motion domain, the difference between symbolic and control trials did not reach significance, $t(69) = 1.30$, $p = .20$. For the shape domain, the difference between control and symbolic trials was marginally significant, $t(69) = -1.76$, $p = .09$. These analyses reveal that sound symbolism in general seems to influence picture categorization performance, with response times varying as a function of whether priming words were sound symbolic or not sound symbolic. This influence, however, appeared to operate in a unique manner for each domain. For motion, the sound symbolic relationship appeared to facilitate performance while for shape, the sound symbolic relationship appeared to inhibit or interfere with performance on the picture judgment task.

In order to determine if sound symbolic primes differentially affected the processing of individual domain endpoints, separate ANOVAs were conducted for the motion and shape domains with domain endpoint and sound symbolism condition as within subjects factors. Table 4 reports mean response time and standard error across endpoint and sound symbolism condition. For the motion domain, a 2 x 2 repeated measures ANOVA was conducted on response time with domain endpoint (moving, still) and sound symbolism condition (control, symbolic) as factors. A significant main effect of endpoint was found, $F(1, 69) = 7.98$, *partial* $\eta^2 = .10$, $p < .01$, with participants responding faster to *moving* trials (M = 900) than to *still* trials (M = 963). No effect of sound symbolism or interaction of endpoint and sound symbolism was found. For shape, we conducted another 2 (pointy, round) x 2 (control, symbolic) ANOVA. Again, a significant main effect of endpoint was found, $F(1, 69) = 4.77$, *partial* $\eta^2 = .07$, $p < .05$. RTs for *round* trials (M = 797) were faster than those for *pointy* trials (M = 841). No effect of sound symbolism or interaction of endpoint and sound symbolism was found. It appears that domain endpoints are not significantly affected by sound symbolism condition.

Discussion

The aims of this study were twofold: 1) to investigate the role and automaticity of sound symbolism during on-line language processing and 2) to determine the extent to which sound symbolism relies upon cross-modal inter-sensory connections. The study employed a forced choice task with participants labeling images as still or moving, or round or pointy. Pictures were preceded by foreign words judged to be sound symbolic or non-sound symbolic for the meanings *still, moving, round,* or *pointy*. The results showed an influence of prime types on response times to the picture targets, although the effects of sound symbolic primes differed as a function of meaning dimension. These findings are consistent both with previous cross-modal research demonstrating that auditory input influences speeded visual classification tasks (Marks, et al., 2003; Molholm, Ritter, Javitt, & Foxe, 2004) and suggested non-arbitrary links between sound structure and perceptual properties in language (Köhler, 1929; Kovic, et al., 2010, Maurer, et al., 2006).

The current results provide some preliminary evidence that non-arbitrary sound to meaning mappings in natural language affect on-line processing of pictures and potentially result from low-level cross-modal connections. A significant interaction between domain and sound symbolic condition was found in the cross-modal priming task. Foreign word primes that were sound symbolic for meanings in the motion domain displayed a non-significant trend toward facilitating the processing of pictorial targets varying in visual properties of implied motion. The opposite pattern was marginally statistically significant for the shape domain. Relative to non-sound symbolic control primes, foreign word primes that were sound symbolic for the meanings *round* and *pointy* led to lengthened response times to round or pointy visual targets. The findings suggest a possible inter-sensory or cross-modal perceptual basis to sound symbolism. Properties of the sound structure of the prime words appeared to influence the processing of pictures with visual features that varied along particular motion and shape dimensions.

Although the sound symbolic nature of the foreign word primes influenced response times to the picture targets, the direction and size of the effect differed across domains. One possible reason for this difference may have been that overall performance differed significantly for the shape and motion judgments. Participants were significantly faster to make shape judgments than to make motion judgments in this task, perhaps reflecting the ease of perceiving veridical presentation of shape compared to the inferential perception of motion from static pictures. This significant difference in response time could have affected whether the primes enhanced or interfered with the processing of pictures across shape and motion conditions. Thus, these response time differences across domain suggest a possible role of the timecourse of processing on the direction of the effect of sound symbolism.

Supporting this possible explanation, previous research examining the mental simulation of perceptual and motor experiences during language comprehension (Kaschak, et al., 2005; Stanfield & Zwaan, 2001; Zwaan, Madden, Yaxley, & Aveyard, 2004) suggests that the faciliatory or inhibitory effects of simulation appear to depend on the time between linguistic and visual stimulus presentation. For example, Zwaan, et al. (2004) presented participants with sentences describing motion toward or away from them, followed 750ms later by two pictures of an object varying slightly in size. This minimal size difference created the illusion of motion either toward or away from the participant. When asked to judge whether the two pictured objects were the same, participants responded more quickly when the motion mentioned in the sentence matched the simulated motion of the visually presented objects. These results suggest not only that listeners engage in automatic mental simulation during language comprehension, but also that the preceding congruent information presented in a different modality facilitated picture processing.

In a similar experiment by Kaschak, et al. (2005), participants heard sentences describing motion toward or away from them and *simultaneously* saw a black and white image creating the perception of motion toward or away from them. In contrast to the Zwaan, et al. (2004) results, inhibitory effects were seen when the motion described in the sentence and that created by the visual stimulus was congruent. Converging evidence (Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002) suggests that simultaneous presentation of congruent perceptual and linguistic information results in inhibition due to increased demands placed on the same set of cognitive resources. That is, when a sentence and visual percept of motion toward oneself are presented simultaneously, processing is inhibited because two separate demands are being placed on the same set of processing mechanisms. When time between linguistic and perceptual presentation is increased, however, facilitatory priming effects are observed.

This simulation literature suggests that presentation of the visual targets immediately following the presentation of the auditory primes in the current study may have encouraged competition rather than facilitation between the two types of stimuli. Indeed, sound symbolic primes appeared to interfere with processing of the pictures depicting differences in shape. Participants responded more slowly on sound symbolic than on control trials. The nominal reversal of the inhibitory effect in the motion trials may have been a consequence of the significantly slower response times during motion trials. Certainly, if time is required both to access the cross-modal information assumed to be elicited by the sound symbolic trials and to encourage a facilitatory priming effect, the delayed responding for motion might have resulted in the relative facilitatory versus inhibitory nature of the influence. Taking the response time difference into consideration, the results from the literature examining simulation accounts of language offer a possible explanation for the contrasting effects of sound symbolism in shape and motion trials.

Although sound symbolism did appear to influence processing, caution must be used when interpreting the current results, as only the interaction between domain and sound symbolism condition was significant. The hypothesized response time differences across match, mismatch, and control conditions were not found. This lack of congruency effect could reflect that the sound symbolic foreign words primed each domain, shape or motion, *as a whole*, regardless of the particular property (e.g., round or pointy) within domain. Significant differences between sound symbolic and non-sound symbolic conditions were not found within domains either, further suggesting that the sound symbolic words primed domain rather than endpoint. Findings from Nygaard, et al. (2009) support this possibility, as a general facilitatory effect of sound symbolism, both in congruent and incongruent conditions, was found in a word learning study.

The lack of a robust effect across all findings may have occurred for several reasons. One possible explanation is that the task design did not encourage participants to use inter-sensory cues during responding. Participants did not need to pay attention to the auditory prime to successfully complete the task. Indeed, participants could ignore the prime entirely and perform with high accuracy on the visual classification task. This design was used purposefully to assess the automaticity of sound symbolism and the possibility that the phenomenon arises from low-level cross-modal connections. The results suggest that some level of attention to auditory stimuli is necessary to elicit sound symbolic effects.

Makovac and Gerbino (2010) offer an alternative explanation drawing on task difficulty rather than focused attention. Makovac and Gerbino report that multimodal sensory experience does enhance the processing of a unimodal stimulus. However, when performance for responses to a unisensory percept is already optimal, this effect is greatly reduced or absent. Since accuracy in the current study was nearly at ceiling, it is possible that the effects of a sound symbolic auditory prime were attenuated. Indeed, there is a trend in the data suggesting that lower accuracy in the motion trials (*M* = .95) led to a non-significant facilitatory role of sound symbolism whereas the high accuracy of shape trials (*M* = .97) may have prevented sound symbolism from influencing processing of the inter-sensory perceptual features of the picture targets. Perhaps a more taxing unimodal task would reveal both inhibition and facilitation based on congruency of stimuli presented in a second modality.

It is also possible that the nature of the targets, rather than the task itself, contributed to the lack of a robust priming effect in this study. The specific images used to convey shape and implied motion could have varying degrees of success in depicting these domains and endpoints. Recall that picture targets were chosen based on participants' accuracy at simply identifying the pictures as round, pointy, still, or moving in a relatively easy two-alternative forced choice task. Perhaps a more sensitive measure of the extent to which each picture represents the meaning dimension might encourage inter-sensory or cross-modal integration. Assessing the images’ representation of endpoint and domain through a Likert rating scale may address this issue although it is possible that static, two-dimensional stimuli in general do not represent shape and motion domains equally. Altering the targets to increase and equalize difficulty may better reveal the effects of sound symbolism during on-line processing.

Another consideration for the lack of robust results is the timing between auditory presentation of the prime and visual presentation of the target. As the pictorial target was displayed immediately after the offset of the auditory prime, it is possible that participants did not have time to access cross-modal representations prior to making their decision. The reversal of the effect of sound symbolism from slow motion trials to fast shape trials suggests this explanation may have some merit. However, it should be noted that this explanation would support less direct cross-modal connections than those hypothesized. If cross-modal connections are truly automatic, a time delay should not be necessary to elicit priming effects. The possible necessity of a time delay may support another mechanism, such as simulation, underlying sound symbolism.

Lastly, variation in simulation and target is another candidate for the lack of robust effects. Although the meanings of the primes and targets may be shared, the precise simulation elicited by a prime may differ from the specific perceptual presentation of a target. For instance, a prime sound symbolic for the meaning *round* may create a mental simulation for a globular, amoeboid shape. If the pictorial target displays a spherical object, there is not specific congruency between the mental simulation induced by the prime and the perceived target. It is possible that the specificity of sound to meaning mappings is quite narrow, causing intended congruent pairs to be incongruent. Supporting this possible explanation, Zwaan, et al. (2002) found that listeners not only simulate objects mentioned in sentences, but also represent orientation and shape of objects in mental simulations. For example, when hearing the sentence *The ranger saw the eagle in the sky,* participants responded faster to a picture of an eagle with outstretched wings than to an eagle in a perched position. These results suggest that a mental simulation with different perceptual details than the target can affect priming.

Finally, it is possible that sound symbolism does not arise from low-level connections between sensory cortices. Further studies altering attention to prime, timing between prime and target, and the specificity of task and pictorial stimuli should be conducted before rejecting the possibility of low-level cross-modal activation. Using brain-imaging techniques to monitor neural activation in low-level sensory areas during processing of sound symbolic language would directly address the cross-modal hypothesis as well. Synaesthetes offer another way in which to approach the investigation of inter-sensory connections and sound symbolism. Synaesthetic connections between sensory modalities are thought to arise either from a lack of pruning between neurons (Maurer & Maurer, 1988; Ramachandran & Hubbard, 2001) or a lack of inhibition in neural connections (Grossenbacher & Lovelace, 2001). This unique neural pattern may cause the effects of sound symbolism to be exaggerated in such individuals. An amplification of sound symbolic effects would support cross-modal interactions as the basis for non-arbitrary sound to meaning mappings and also allow for a greater understanding of the phenomenon.

Overall, this study suggests that sound symbolism affects on-line processing and that this effect may arise from cross-modal activation of low-level sensory areas. The findings encourage future research in several directions to discover the intricacies of sound symbolism and its effects on language and perceptual processing.

References

Bergen, B. K. (2004). The psychological reality of phonaesthemes. *Language, 80*(2), 290-311.

Berlin, B. (1994). Evidence for pervasive synaesthetic sound symbolism in ethnozoological nomenclature. In L. Hinton, J. Nichols & J. Ohala (Eds.), *Sound Symbolism* (pp. 77-93). New York: Cambridge University Press.

Brown, R. W., Black, A. H., & Horowitz, A. E. (1955). Phonetic symbolism in natural languages. *Journal of Abnormal Psychology, 50*(3), 388-393.

Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language, 30*(3), 348-369.

Cassidy, K. W., & Kelly, M. H. (2001). Children's use of phonology to infer grammatical class in vocabulary learning. *Psychonomic Buletin & Review, 8,* 519-523.

de Saussure, F. (1959). *Course in General Linguistics*. New York: Philopsophical Library.

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.

Gasser, M. (2004). *The Origins of Arbitrariness in Language.* Paper presented at the Proceedings of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends in Cognitive Science, 5*(1), 36-41.

Hockett, C. F. (1977). *The view from language: Selected essays 1948-1974*. Athens: The University of Georgia Press.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition, 109*(1), 54-65.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., et al. (2005). Perception of motion affects language processing. *Cognition, 94*(3), B79-89.

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics, 35*, 379-415.

Kita, S. (2001). Semantic schism and interpretive integration in Japanese sentences with a mimetic: a reply to Tsujimura. *Linguistics*, *39*(2), 419-436.

Klank, L., Huang, Y.-H., & Johnson, R. (1971). Determinants of success in matching word pairs in tests of phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior, 10*, 140-148.

Köhler, W. (1929). *Gestalt Psychology*. New York: Liveright Publishing Corporation.

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, *114*(1), 19-28.

Kunihira, S. (1971). Effects of Expressive Voice on Phonetic Symbolism. *Journal of Verbal Learning and Verbal Behavior, 10*(4), 427-429.

Makovac, E., & Gerbino, W. (2010). Sound-shape congruency affects the multisensory response

enhancement. *Visual Cognition*, 18, 133-137.

Marks, L. E., Ben-Artzi, E., & Lakatos, S. (2003). Cross-modal interactions in auditory and visual discrimination. *International Journal of Psychophysioogyl, 50*(1-2), 125-145.

Maurer, D., & Maurer, C. (1988). *The world of the newborn*. New York: Basic Books.

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.

McGregor, W. (2001). Ideophones as the source of verbs in northern Australian languages. In F. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones*. Amsterdam: John Benjamins.

Melara, R. D., & Marks, L. E. (1990). Dimensional interactions in language processing: investigating directions and levels of crosstalk. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 539-554.

Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cerebral Cortex, 14*(4), 452-465.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Research: Cognitive Brain Research, 14*(1), 115-128.

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

Plato. (1998). *Cratylus* (C. D. C. Reeve, Trans.). Indianapolis: Hackett Publishing Company, Inc.

Ramachandran, V., & Hubbard, E. M. (2001). Synaesthesia - A window into perception, thought and language. *Journal of Consciousness Studies, 8*(12), 3-34.

Sereno, J. A., & Jongman, A. (1990). Phonological and Form Class Relations in the Lexicon. *Journal of Psycholinguistic Research, 19*(6), 387-404.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition, 72*(2), B11-21.

Shintel, H., & Nusbaum, H. C. (2007). The sound of motion in spoken language: visual information conveyed by acoustic properties of speech. *Cognition, 105*(3), 681-690.

Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language, 55*, 167-177.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153-156.

Staub, A., Grant, M., Clifton, C., & Rayner, K. (2009). Phonological Typicality Does Not Influence Fixation Durations in Normal Reading. *Journal of Experimental Psychology, 35*(3), 806-814.

Voeltz, F., & Kilian-Hatz, C. (2001). *Ideophones*. Amsterdam: John Benjamins.

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., et al. (2010). Preverbal Infants' Sensitivity to Synaesthetic Cross-Modality Correspondences. *Psychological Science, 21*(1), 21-25.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language*, *93*(1), 10-19.

Zwaan, R,A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: dynamic representations in language comprehension. *Cognitive Science*, *28*, 611-619.

Zwaan, R.A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168-171.

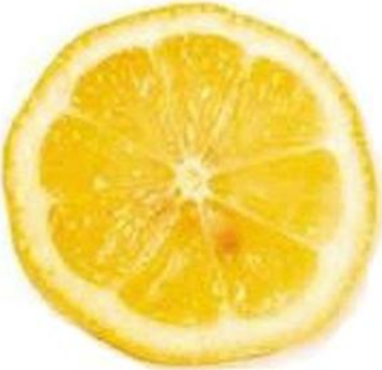
Appendix A: Primes

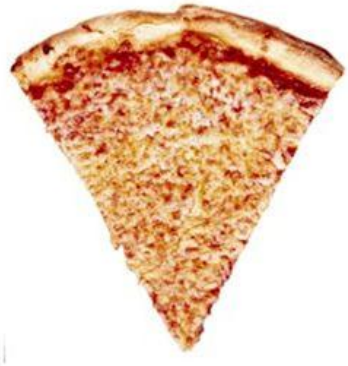
Word	Language	Symbolism	Domain	SoundsLike	ActualMeaning
palevizshem	Albanian	Symbolic	Motion	moving	still
bergesas-gesas	Indonesian	Symbolic	Motion	moving	moving
berjalan	Indonesian	Symbolic	Motion	moving	moving
pu dong sa se i da	Korean	Symbolic	Motion	moving	still
um ji gi he ga nun	Korean	Symbolic	Motion	moving	moving
inmarmurit	Romanian	Symbolic	Motion	moving	still
achaivillatha	Tamil	Symbolic	Motion	moving	still
nilaiyana	Tamil	Symbolic	Motion	moving	still
hareketli	Turkish	Symbolic	Motion	moving	moving
calisan	Turkish	Symbolic	Motion	moving	moving
stheer	Gujarati	Symbolic	Motion	still	still
chaltu	Gujarati	Symbolic	Motion	still	moving
gerak	Indonesian	Symbolic	Motion	still	moving
teu da	Korean	Symbolic	Motion	still	moving
yi dong	Mandarin	Symbolic	Motion	still	moving
katham	Tamil	Symbolic	Motion	still	moving
asai	Tamil	Symbolic	Motion	still	moving
notr	Turkish	Symbolic	Motion	still	still
sare	Yoruba	Symbolic	Motion	still	moving
fo	Yoruba	Symbolic	Motion	still	moving
ne ecuri	Albanian	Control	Motion	neither	moving
tharelu	Gujarati	Control	Motion	neither	still
jalan	Indonesian	Control	Motion	neither	moving
dung dung te eu	Korean	Control	Motion	neither	moving
gu ding	Mandarin	Control	Motion	neither	still
a duce	Romanian	Control	Motion	neither	moving
kayin	Tamil	Control	Motion	neither	moving
tari	Tamil	Control	Motion	neither	still
kosma	Turkish	Control	Motion	neither	moving
duroje	Yoruba	Control	Motion	neither	still
dhembezuar	Albanian	Symbolic	Shape	pointy	pointy
mprehte	Albanian	Symbolic	Shape	pointy	pointy
ajubaju	Gujarati	Symbolic	Shape	pointy	round
berbentuk lingkaran	Indonesian	Symbolic	Shape	pointy	round
bergerigi	Indonesian	Symbolic	Shape	pointy	pointy
geu jo ka da	Korean	Symbolic	Shape	pointy	pointy
tun gu tru ma da	Korean	Symbolic	Shape	pointy	round
on yong hi da	Korean	Symbolic	Shape	pointy	round
ding zi ban	Mandarin	Symbolic	Shape	pointy	pointy
kesici	Turkish	Symbolic	Shape	pointy	pointy
maje	Albanian	Symbolic	Shape	round	pointy
bute	Albanian	Symbolic	Shape	round	round
goad	Gujarati	Symbolic	Shape	round	round
bulat	Indonesian	Symbolic	Shape	round	round

lun	Mandarin	Symbolic	Shape	round	round
yuan	Mandarin	Symbolic	Shape	round	round
bombat	Romanian	Symbolic	Shape	round	round
urunta	Tamil	Symbolic	Shape	round	round
mu	Yoruba	Symbolic	Shape	round	pointy
gbun	Yoruba	Symbolic	Shape	round	pointy
ashper	Albanian	Control	Shape	neither	pointy
rrethor	Albanian	Control	Shape	neither	round
ani	Gujarati	Control	Shape	neither	pointy
runcing	Indonesian	Control	Shape	neither	pointy
bu na da	Korean	Control	Shape	neither	pointy
hun yuan	Mandarin	Control	Shape	neither	round
inelar	Romanian	Control	Shape	neither	round
kaciram	Tamil	Control	Shape	neither	round
sarmal	Turkish	Control	Shape	neither	round
sonso	Yoruba	Control	Shape	neither	pointy

Appendix B: Shape targets

Pointy	**Round**
	
	
	

	
	
	
 RAIL ROAD CROSSING	 R R

Appendix C: Implied motion targets

Moving	Still
	
	
	
	

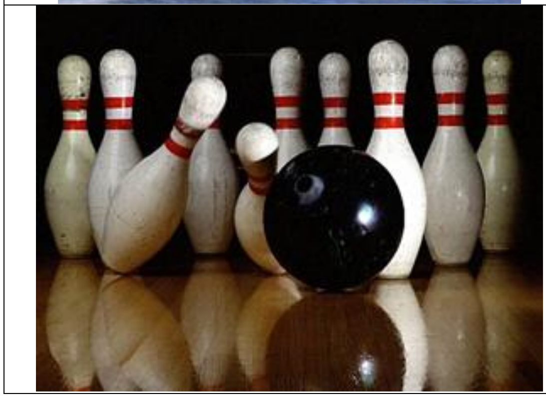
	DAREDEVIL GAMEDISC 175-GRAM ULTIMATE DISC
	
	
	

Table 1

Average proportion correct identification and standard error for pictorial targets

	Round	**Pointy**	**Still**	**Moving**
M	.96	.96	.97	.95
SE	.01	.01	.01	.01

Table 2

Mean proportion correct and standard error as a function of domain and congruency condition

		Symbolic		**Control**
		Match	**Mismatch**	
Motion	M	.93	.95	.96
	SE	.01	.01	.01
Shape	M	.97	.97	.98
	SE	.01	.01	.01

Table 3

Mean response time (in ms) and standard error as a function of domain and congruency condition

		Symbolic		**Control**
		Match	**Mismatch**	
Motion	M	931	910	938
	SE	33	28	30
Shape	M	830	828	804
	SE	34	35	27

Table 4

Mean response time (in ms) and standard error as a function of domain endpoint and sound symbolism condition

		Symbolic	**Control**
Moving	M	883	917
	SE	26	31
Still	M	965	961
	SE	38	35
Pointy	M	852	830
	SE	37	30
Round	M	808	786
	SE	33	32

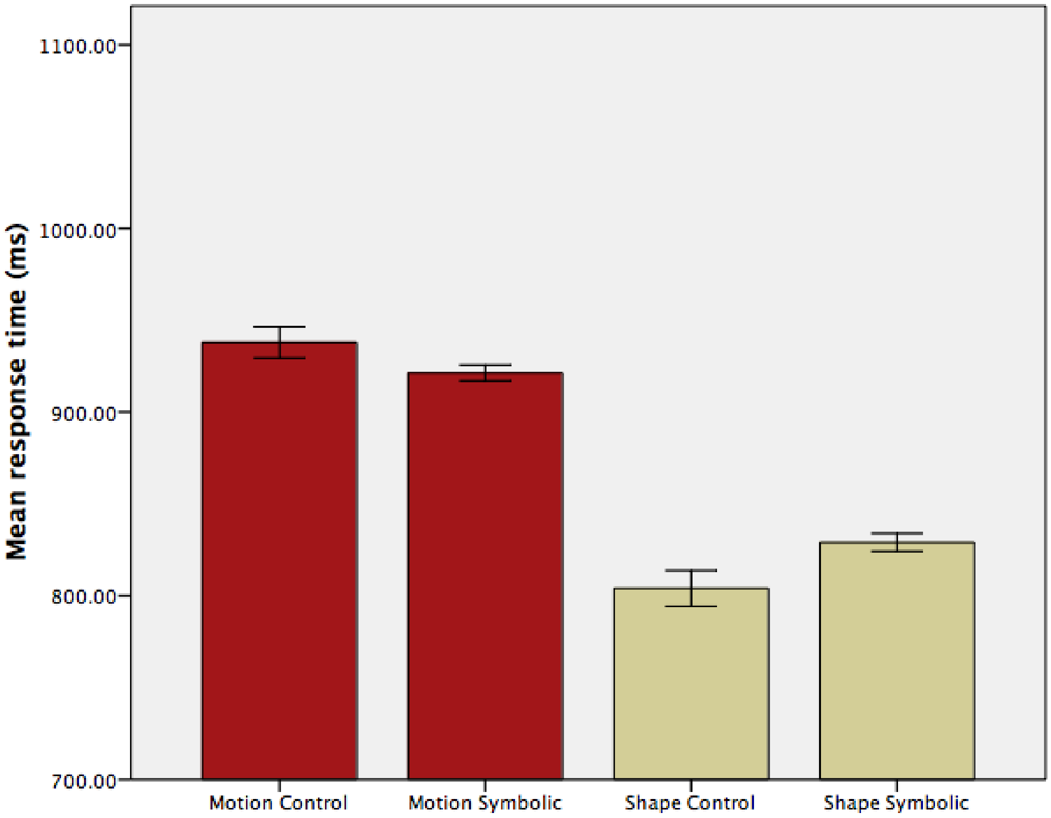


Figure 1. Mean response time (in ms) as a function of domain and sound symbolism condition. Error bars represent ± 1 standard error.