The Role of Sound Symbolism in Language Learning

Padraic Monaghan, Karen Mattock & Peter Walker

Centre for Research in Human Development and Learning

Department of Psychology

Lancaster University, UK

Word count: 7524

Corresponding author:

Padraic Monaghan

Department of Psychology

Lancaster University

Lancaster

LA1 4YF

UK

E-mail: p.monaghan@lancaster.ac.uk

Tel: +44 1524 593813

Fax: +44 1524 593744

Abstract

Certain correspondences between the sound and meaning of words can be observed in subsets of the vocabulary. These sound-symbolic relationships have been suggested to result in easier language acquisition, but previous studies have not explicitly tested effects of sound symbolism on word learning, but only on learning category distinctions. In two word learning experiments, we varied the extent to which phonological properties related to a rounded-angular shape distinction, and distinguished learning of categories from learning of individual words. We found that sound-symbolism resulted in an advantage for learning categories of sound-shape mappings, but did not assist in learning individual word meanings. These results are consistent with the limited presence of sound symbolism in natural language. The results also provide a reinterpretation of the role of sound symbolism in language learning and language origins, and a greater specification of the conditions under which sound symbolism proves advantageous for learning.

The Role of Sound Symbolism in Language Learning

The relationship between the sound of a word and its meaning has long been considered to be arbitrary (de Saussure, 1916; Gasser, 2004; Hockett, 1960). Thus, even if all words in the language were known to the listener, from a new word's sound it would not be possible to determine its meaning. Such a dissociation between sound and meaning might provide numerous advantages for communicative systems by promoting the development of abstract terms and by generating the ability to talk about events distant in time and space (Clark, 1998). It is also useful for maximising the information in the environment determining the intended referent (Monaghan, Christiansen, & Fitneva, 2011). However, there are exceptions to this arbitrariness. Languages with expressed morphology reflect aspects of meaning within a word's phonology (Bybee, 1985), and arbitrariness may give way at a more localised level in the vocabulary, such as the language-general occurrence of phonaesthemes, e.g., words beginning with *sn*- tend to relate to the nose (e.g., *sneeze*), and *str*- commences words often referring to collisions (e.g., *strike*) (Bergen, 2004; Bloomfield, 1933; Wallis, 1699).

Such instances of sound-symbolism, where systematicity between sounds and meanings of words can be observed (we use sound-symbolism in the broad sense of systematic preferences for certain sound-meaning mappings, rather than a narrower sense of an iconic representation between sound and meaning), have been proposed to be critically important for language evolution and language acquisition. Ramachandran and Hubbard (2001) suggested that sound-symbolism tells us something profound about the origins of language, and is an important factor in language evolution. Given that certain sounds in words elicit a high degree of agreement in terms of the extent to which they reflect semantic distinctions, this suggests that such sound-meaning relationships could implicitly influence the labels developed for referents, and these labels may be those utilised in the very first words of human language (Paget, 1930). However, this view raises a conundrum. If sound-symbolism is so vital for language learning and a driving force in language origins and language evolution, then why do we observe so very few correspondences between sounds of words and meaning distinctions in natural languages (Brown, 1958; Newman, 1933; Newmeyer, 1993; Otis & Sagi, 2009)?

In this paper we address this tension between sound-symbolism as an aide to language learning and the observed arbitrariness of the sign. We first review previous studies demonstrating sound-symbolism in language processing and determine precisely what these studies have shown about language learning. In particular, we distinguish studies demonstrating learning labels for categories from learning labels for particular referents. We then report the results of a novel experimental paradigm that enables the simultaneous testing of the extent to which participants learn names for categories of shape and names for particular shapes. We conclude with a corpus analysis of words for the angular/rounded distinction in English to assess the extent to which sound-symbolism supports a category distinction in natural language.

Sound-Symbolism and Learning Words or Forming Categories

In order to assess the role of sound-symbolism for language learning it is necessary to assess which aspects of learning previous studies have addressed. In particular, it is important to distinguish between sound symbolism influencing category learning from acquisition of individual word-meaning mappings. This is important because there is evidence that at broad category levels there is substantial sound-symbolism, such as grammatical categories being reflected in certain patterns of phonological and prosodic cues (Monaghan, Christiansen, & Chater, 2007). However, in terms of learning individual word referents, there is evidence that systematicity in form-meaning mappings may in fact impede learning (Monaghan et al., 2011). Also, studies that draw attention to distinctions in meaning or sound may actually train participants in the distinctions to be measured in the study, thereby affecting the results. It is also of importance to discern whether studies are testing preferences for certain mappings, with the implicit inference that these preferences would facilitate acquisition of the sound-referent mapping (e.g., Maurer, Pathman, & Mondloch, 2006), or whether they explicitly test the acquisition of mappings.

Sapir (1929) and Köhler (1929) were the first to experimentally test preferences for certain speech sounds relating to semantic distinctions. Sapir (1929) gave participants a forced choice of two words varying by a single vowel or consonant, and asked them to say which more appropriately labeled a large and which a small table. Close vowels were more likely to be related to the small table and open vowels more likely to be related to the large table, and Newman (1933) replicated these studies. Köhler (1929) presented participants with two shapes, one with rounded and the other with angular sides, and two words, one with a majority of high front vowels and unvoiced plosives (*takete*), and the other containing contrasting phonemes (*baluma*). Participants almost unanimously paired *takete* with the angular shape and *baluma* with the rounded shape. These forced-choice test results have been replicated in numerous studies, with slight variations to the words used (*takete*/*maluma*: Köhler, 1947; *takete*/*uloomo*: Davis, 1961; *takete*/*baluma*: Holland & Wertheimer, 1964; *kiki*/*bouba*: Ramachandran & Hubbard, 2001; *tuhkeetee*/*maaboomaa* (amongst others): Maurer et al., 2006; *taakootaa*/*muhbeemee* (amongst others): Nielsen & Rendall, 2011). In all these studies, the sound and the shape distinction was explicitly presented, and therefore attention may have been drawn to the correspondence during the experimental study. The forced-choice test, by measuring decisions between categories, also means that it is not known whether participants were mapping a word to the category or were mapping onto particular referents.

Other tasks have also revealed systematic preferences for certain sounds relating to the angular/rounded distinction. Wertheimer (1958) asked participants to judge sets of real words to the extent that they fitted their meaning according to several scales of binary semantic distinctions. He found that "fitting" words were generally judged more extreme on several scales of binary semantic distinctions (including angular/rounded) than were "non-fitting" words. In another paradigm, Westbury (2005) tested the implicit influence of angular/rounded shapes on processing of words with varying phonological characteristics. In a lexical decision task, words with stop consonants or continuants were presented in an angular or a rounded frame. The continuant nonwords in a curved frame, and the stop nonwords in an angular frame, were responded to most quickly.

Imai, Kita, Nagumo, and Okada (2008) presented participants with actions varying along dimensions of fast-slow and heavy-light, which are two principal dimensions of mimetic forms in Japanese (Kita, 1997) and words varying in the extent to which these two dimensions are symbolised in speech (e.g., *batabato* was a fast and heavy action, *tokutoku* was normal speed and light, and *nosunosu* was slow and heavy). In a forced-choice task, participants were given a word and two actions and both Japanese and English speaking participants performed better than chance (see also Kantartzis, Kita, and Imai, 2009).

Once again, in each of these studies the forced-choice alternatives provide the category distinction to the participant and, furthermore, do not distinguish between participants forming a link between categories of sound and meaning from participants forming links between particular words and referents. For individual referent learning to be tested requires the category distinction to be unavailable during testing. Furthermore, in all the studies thus far reviewed, there has been no explicit test of learning, but rather only implicit preferences for correspondences.

In a study that did test learning, Kovic, Plunkett, and Westermann (2010) presented the angular/rounded distinction to participants in a paradigm that "captures the processes involved in natural language interpretation" (p.19). Participants were asked to classify rounded/angular shapes into one of two categories labelled by the words *mot* or *rif*. After learning with feedback in either a congruent (*mot* = rounded) or incongruent (*mot* = angular) condition, participants were tested on both congruent and incongruent pairings. Participants in the congruent condition were faster and more accurate at responding than participants in the incongruent condition. In a second experiment, participants demonstrated the same congruency effect with the category words *dom* and *shik*. These studies provided the category contrast to which sound distinctions are to be mapped and again were not able distinguish between learning categories and learning individual word labels.

Nygaard, Cook, and Namy (2009) also attempted an explicit test of learning, by presenting English-speaking participants with Japanese-English antonym words, where the pairings were either congruent, incongruent, or random between the languages. During training, participants heard the Japanese word and read the English word on a computer screen. During testing, the Japanese word was heard and two alternatives (the paired word and a random word) were presented on a computer screen. Participants in the congruent condition were not more accurate than the incongruent condition, but were quicker to respond. These effects suggest that word learning may be more effective (in that recognition is quicker) due to sound-symbolism, though the test could be interpreted as forming category distinctions rather than learning individual word pairings.

In summary, the experimental evidence for the role of sound-symbolism in word learning fails to distinguish learning categories of word-meaning mappings from learning particular word-referent mappings, due to the forced-choice paradigm. In a series of behavioural and computational studies, Monaghan et al. (2011) trained participants to pair words with pictures of actions and objects where there was either a systematic or an arbitrary relationship between phonological features of words and semantic features of the pictures. They found that a systematic relationship assisted in learning the categories of the pictures, which is consistent with the studies reviewed here in terms of forming a category distinction through a forced-choice test. However, for learning individual words for pictures (tested by presenting all pictures at once and providing a single word), an arbitrary relationship was optimal, particularly under more naturalistic word-learning conditions where additional contextual information was present for the learner to assist in meaning identification.

An additional problem in forced-choice tests of the effect of sound-symbolism is that learning may be dynamic – the observed preferences for certain sound-meaning mappings may be a consequence of exposure to a distinction along a certain semantic property (such as angular/rounded) and exposure to a distinction in phonological features. Thus, in the case of angular versus rounded shapes, the presentation of an angular and a rounded shape simultaneously may draw the participant's attention to this distinction without it generally being sufficiently important or salient in the environment to require the contrast to be noted, and consequently, sound-symbolic preferences may be an artefact of the experimental design rather than reflections of tendencies in naturalistic learning situations. Though preferences between certain semantic and phonological features may be consistently demonstrated, as so strikingly indicated since Sapir's (1929) and Köhler's (1929) pioneering work, this does not demonstrate that such patterns are either used or usable in natural language where such distinctions in phonology and semantics are seldom clear cut. Indeed, as Monaghan et al. (2011) point out, words of the same semantic category tend to co-occur in similar environmental situations, meaning that broad semantic distinctions are less likely to be apparent in learning situations.

In the following two experiments, we distinguish whether sound-symbolism can be shown to assist in learning individual word-referent mappings, from whether it assists only in learning categories of referent. Monaghan et al. (2011) found a correspondence between sound similarity and grammatical category distinctions but not individual referents, but it is possible that a distinction within a particular grammatical category (for example, names for shapes) may mean that sound symbolism is more generally advantageous, as the sound symbolism and language learning studies have been taken to suggest. However, another alternative is that the advantage for systematicity between sound and meaning only applies at the category level, applying to both subcategories of object as well as grammatical categories, and not at the individual word level.

In this paper, we focus on the angular versus rounded distinction and relate this to phonological distinctions involving consonants (Experiment 1) and vowels (Experiment 2). We distinguish these in separate studies to ensure that our effect is robust, but also to determine the relative role of vowels and consonants in the potential relationship between sound and angular/rounded shape. In previous studies, excepting the careful distinguishing of consonants and vowels in Nielsen and Randall (2011), both consonants and vowel types are co-varying in the words for the two shape categories (such as for *maluma/takite*). In our experimental design, we distinguish learning trials where the participant can utilise semantic category information to learn the word, and trials where this semantic category information is not available. Our design also ensures that participants do not learn from correspondences between word sounds and meanings during the course of the experiment.

Experiment 1: Sound-symbolism in Consonants and Word Learning

In this experiment, participants learned to pair a set of words with a set of objects. This was done using an adaptation of what has been termed the cross-situational learning paradigm (Monaghan & Mattock, 2009; Yu & Smith, 2007). In this paradigm, participants view a set of objects and hear one or more words. From a single trial, the participant does not know which word refers to which object, but over multiple trials, participants have the opportunity to learn the correspondences between particular words and objects. Studies of this type have been used to simulate the process of word learning in early language learning (Horst, Scott, & Pollard, 2010; Smith & Yu, 2008; Yu & Smith, 2007) in that children do not have privileged information about which of multiple words refers to which of several objects and events but are able to acquire the mapping from several learning instances. The advantage of learning from this paradigm is that explicit information about correct pairings does not have to be given, as they can be acquired from the cross-situational statistics between objects and words.

In our study, participants viewed two objects and heard a word, and had to make a judgement about which object was referred to by the word. After several trials, participants will have experienced one of the two objects always occurring when a given word was heard. The words varied in terms of whether they contained plosives or continuants, and the objects varied in terms of whether they had an angular or rounded characteristic. We predicted that congruent pairings (according to established sound-symbolic correspondences) would result in better learning than incongruent pairings. We also predicted that if sound-symbolism assists word learning then this benefit would be general across trials whether the two presented shapes are drawn from the same category or from different categories. If sound-symbolism only supports the learning of category distinctions, however, then the congruent advantage should only be observed for trials where the two shapes are of contrasting type.

Method

Participants. Twenty four staff and students (13 female, 11 male) at Lancaster University were each paid £3 for participating. All participants reported speaking English as their first language and had no hearing or visual impairments. Mean age of participants was 20.6 years (range 18-26).

Materials. We used 16 nonwords, each of which was paired with one of 16 nonsense shapes. Eight of the nonsense shapes had angular edges, and the other eight had rounded edges. All the shapes were created using a virtual pattern comprising 6 concentric rings superimposed on 24 spokes. Each angular shape was designed around a random subset of eight spokes. Each of these spokes was then linked to a particular ring, with their point of intersection marking the alternate points of convexity and concavity of the shape. Working in a clockwise direction, spokes were linked to a ring selected alternately from the outer and inner three rings. Straight lines connected the eight points of intersection defined by the spoke-ring pairings. For each angular shape created in this way, an equivalent rounded shape was created using the Reshape function in the AppleWorks Draw package. This version was rotated and/or flipped so that it would not be identified as a rounded version of the pointed shape from which it had been derived. Example stimuli are shown in Figure 1. When all 16 shapes had been created, the overall size of each round shape was adjusted so that it matched the equivalent pointed shape in the subjective judgment of 10 viewers, none of whom took part in the main experiment. The shapes appeared on a white background, and a shadow was added to enhance their perception as objects.

For the speech stimuli, we generated 16 nonsense words of the form consonant-vowel-consonant. Eight of the nonwords had plosives (/k/, /g/, /t/, /d/, /p/, /b/) in onset and coda position, one of which was voiced the other unvoiced. The remaining 8 nonwords had nasals (/m/, /n/, /ŋ/), liquids or approximants (/l/, /ɹ/, /w/) in onset and coda (nasals, liquids and approximants are all continuant consonants). Across the set of nonwords, each consonant occurred either two or three times, and no more than two times in onset or coda position. The lists of nonwords are reported in Appendix 1. The vowels were balanced across the two types of nonword, with /æ/, /ɛ/, /ɪ/, and /ɒ/ each used in two of the plosive nonwords and two of the continuant nonwords. The nonwords were designed to be congruent with either an angular or a rounded shape, given previous studies of forced-choice preferences for pairing certain speech stimuli with angular or rounded shapes (Davis, 1961; Holland & Wertheimer, 1964; Köhler, 1929, 1947; Kovic et al., 2009; Ramachandran & Hubbard, 2001).

The nonwords were recorded by a female native speaker of British English who was instructed to read the nonwords in a monotone to minimise prosodic differences between the speech stimuli. We compared the plosive and continuant words on a large number of acoustic properties measured using Praat (Boersma & Weenink, 2009), described in Appendix 2. In particular, we tested that the groups of words were not distinct in terms of intensity or fundamental frequency (pitch), or of the first and second formant (which are the frequency components of speech that distinguish vowels) and ensured that distinctions were in terms of the qualitative properties of the consonants. As shown in Table 1 we found that the significant differences between the plosive and the continuant nonwords were in terms of the higher pitch at the onset for continuants, which was possibly due to the overall earlier onset of voicing in the continuants, and greater consonant intensity for continuants, which again was due to the earlier onset of voicing in the continuants. These properties reflect distinctions between plosives and continuants in terms of the "amplitude envelope" whereby plosives result in a more sudden increase in amplitude than do continuants. Distinctions in shape of the amplitude envelope has been taken to be symbolic of angular versus rounded distinctions (Rhodes, 1994).

Half the plosive nonwords were paired with angular shapes, and the other half were paired with rounded shapes. We refer to these as plosive-angular and plosive-rounded, respectively. The continuant nonwords were paired with the remaining shapes, so four were continuant-angular and four continuant-rounded. Based on previous studies of sound-symbolism we hypothesised that the plosive-angular and continuant-rounded pairings would be congruent and the other pairings incongruent. To avoid biases from any existing preferential associations between particular sounds and shapes, the word-shape pairings were randomised in 16 different ways, and no more than two participants were exposed to any one set of pairings.

Procedure. Participants learned the mapping between nonwords and shapes using a cross-situational learning paradigm. This enabled us to expose participants to correspondences between particular nonwords and shapes without explicitly labelling particular shapes and without drawing attention to distinctions in sounds or shapes. Participants were seated at a computer and instructed that they were to learn an alien language by determining to which of two shapes a word referred. For each trial, two shapes appeared on the screen. Then, after 500ms, participants heard a nonword, which always accompanied one of the shapes, and had to press either the “1” or “2” on a keyboard to indicate whether they felt the word referred to either the left or the right shape. No feedback was given to participants. There were four blocks of 64 trials. Within each block, each nonword was heard 4 times. Each picture was presented 8 times, four times as the target and four times as a foil. Within each block, each type of word-shape pairing (plosive-angular/plosive-rounded/continuant-angular/continuant-rounded) appeared twice with two foil word-shape pairings from each of the four types, thus angular and rounded shapes occurred equally often accompanied by other angular or rounded shapes. The target shape occurred an equal number of times on the left and the right of the screen within each block. The probability of the target object occurring with a given word was 1.0, the probability of another object (labelled by one of the other words) occurring with a given word was .14.

Results and Discussion

Accuracy of participants’ responses was recorded for each block of learning, with a correct response recorded when the target object corresponding to the heard word was selected (i.e., the object that always co-occurred with the word). We hypothesised that if participants’ learning was supported by sound-symbolism then they would be more accurate with the congruent (plosive-angular and continuant-rounded) pairings as these are consistent with the proposed sound-symbolic relationship derived from previous studies, than with the incongruent (plosive-rounded and continuant-angular) pairings. If sound-symbolism facilitated learning of individual word-referent pairings as well as learning word-category relationships, then learning should be more effective for the congruent pairings when the target and foil shapes are from the same category (so both rounded or both angular) as well as from different categories. However, if sound-symbolism affects only learning of word-category relationships, then congruent pairings should only demonstrate an advantage over incongruent pairings when the target and foil shapes are from different shape categories.

We performed an ANOVA with block (1-4), congruence of the word-shape pairing (congruent versus incongruent), and whether the target and foil shapes were from the same or different shape categories, as within-participants factors with accuracy as the dependent variable.

There was a significant main effect of block, $F(3, 69) = 22.87$, $p < .001$, $\eta_p^2 = .50$, with accuracy increasing with time, as shown by the linear contrast, $F(1, 23) = 39.99$, $p < .001$, $\eta_p^2 = .64$. The main effect of congruence was significant, $F(1, 23) = 4.32$, $p = .049$, $\eta_p^2 = .16$, indicating that over all trials sound-symbolism affected learning word-object pairings. The main effect of same-or-different categories of shape was also significant, $F(1, 23) = 5.89$, $p = .023$, $\eta_p^2 = .20$, with better performance when the shapes were drawn from different categories.

There was a significant interaction between block and congruence, $F(3, 69) = 5.25$, $p = .003$, $\eta_p^2 = .19$, with post-hoc tests indicating that congruence had a greater effect for blocks 2 and 3 ($p = .018$, $p = .008$, respectively) than for blocks 1 and 4 ($p = .258$, and $p = .779$, respectively, see Figure 2). This was because learning increased for congruent trials from block 1 to block 2, $p < .001$, but thereafter stabilised (block 2 and block 3, $p = .392$; block 3 and block 4, $p = .960$), whereas for incongruent trials, learning improved between each block (block 1 and block 2, $p = .041$; block 2 and block 3, $p = .008$; block 3 and block 4, $p < .001$).

The interaction between block and same-or-different shapes was significant, $F(3, 69) = 3.77$, $p = .014$, $\eta_p^2 = .14$, with same-or-different shapes having the largest difference for the second block of learning (differences between same-or-different shape category trials: block 1: $p = .572$; block 2: $p < .001$; block 3: $p = .148$, block 4: $p = .065$, see Figure 3). Critically, the interaction between congruence and same-or-different shape category was significant, $F(1, 23) = 15.14$, $p < .001$, $\eta_p^2 = .40$, with congruence effective when the shapes were different, $p = .003$, but not when they were the same category, $p = .337$ (see Figure 4). The three-way interaction (block x congruence x same-or-different shape category) was marginally significant, $F(3, 69) = 2.70$, $p = .052$, $\eta_p^2 = .11$, which suggested that the greater effect of congruence in the medial blocks of learning was emphasised when the two shapes were drawn from different categories.

The significant main effect of same-or-different shapes could indicate that the greater discriminability of shapes in the different shape condition resulted in better performance. However, main effects must be interpreted in terms of higher-order interactions, and so the interaction between same-or-different shape category and congruence must be considered. If visual discriminability was driving the results, then we would expect that in the different shapes conditions, both congruent and incongruent mappings would demonstrate an advantage over the same shape category conditions. However, only the congruent condition showed this advantage, meaning that visual discriminability is unlikely to be the source of the observed results.

We also tested whether participants learned the plosive or continuant words better, and found no significant difference, $F < 1$, indicating that one nonword type was not more salient or more easily acquired than the other. When comparing learning of the rounded versus angular shapes there was also no significant advantage for learning either type of shape, $F < 1$.

The results of Experiment 1 demonstrate that participants are able to learn to pair nonwords with abstract objects using cross-situational statistics. Furthermore, we demonstrated an effect of sound-symbolism on word learning, but only when the two shapes in the learning situation were of different types, as shown in the interaction between congruence and same-or-different shape categories (Figure 4). This suggests that sound-symbolism is useful for learning not individual word-object pairings, but rather for learning correspondences between categories of sound and categories of object. The fact that learning of congruent sound-meaning relationships was only advantageous over learning incongruent relationships when the two objects presented were from different categories suggests that participants use sound-symbolism to guide category learning rather than as assistance in learning individual words.

Experiment 1 tested effects of sound symbolism in terms of relationships between shape and consonant quality. However, in many previous studies of sound-symbolism both consonants and vowels have been varied between stimuli. Experiment 2 therefore tested whether properties of the vowel can elicit a greater influence on individual word learning than was observed with consonants. Experiment 2 tested the same hypotheses as Experiment 1 but focused on the distinction between front and back vowels in the nonwords relating to the angular versus rounded shape distinction.

Experiment 2: Sound-symbolism in Vowels and Word Learning

Method

Participants. Twenty four students (17 female, 7 male) at Lancaster University participated for course credit or payment of £3. All participants reported speaking English as their first language, and had no reported hearing or visual impairments. The mean age of participants was 19.9 years (SD = 3.9, range 18-36).

Materials. We employed the same shapes as used in Experiment 1. We generated 16 new nonwords, each of which was paired with one of the 16 nonsense shapes. As in Experiment 1, the nonwords were of the form consonant-vowel-consonant. Eight of the nonwords had closed front vowels (/i/, /ɛ/, /ɪ/, /eɪ/), and the other eight contained open back vowels (/ɑ:/, /ɔ:/, /ɒ/, /ʌ/). Each vowel occurred two times in the set of nonwords. The consonants /ɡ/, /d/, /b/, /m/, /n/, /l/, and /ɹ/ occurred once in onset and once in coda position in each set of nonwords. The consonant /w/ occurred once in onset and /ŋ/ once in coda position in each set of nonwords. The nonwords are shown in Appendix 1. The nonwords were recorded by the same female native speaker of British English who recorded the nonwords for Experiment 1. She was again instructed to read the nonwords in a monotone.

The groups of nonwords were compared in terms of their acoustic properties (Appendix 2) as with the nonwords in Experiment 1. The results are shown in Table 2. We found that the difference between the sound stimuli was in terms of vowel quality rather than general features of the speech such as amplitude or pitch. This was reflected in the higher pitch of F2 in the high/front vowels, which has previously been identified as a potential sound-symbolic property of speech with respect to indicating size (Fischer-Jørgensen, 1978; Ohala, 1994), and reflects differences in front or back position of the tongue.

Half the closed/front nonwords were paired with angular shapes, and the other half were paired with rounded shapes. We refer to these as front-angular and front-rounded, respectively. The open/back nonwords were paired with the other shapes, so four were back-angular and four back-rounded. The front-angular and back-rounded pairings were considered congruent, and the other pairings were considered incongruent, according to sound-symbolism relations described in the literature (Köhler, 1947; Holland & Wertheimer, 1964; Ramachandran & Hubbard, 2001). To avoid biases from any existing associations between particular sounds and shapes, the word-shape pairings were randomised in 16 different ways, and no more than two participants were exposed to any one set of pairings.

Procedure. The procedure was identical to Experiment 1.

Results and discussion

We performed an ANOVA with block (1-4), congruence of the word-shape pairing (congruent/incongruent), and whether the target and foil shapes were from the same or different shape categories as within-participants factors, with accuracy as the dependent variable.

The results were very similar to those of Experiment 1. There was a significant main effect of block, $F(3, 69) = 13.56$, $p < .001$, $\eta_p^2 = .37$, with accuracy increasing with time, as reflected in the linear contrast, $F(1, 23) = 18.94$, $p < .001$, $\eta_p^2 = .45$. The main effect of congruence was significant, $F(1, 23) = 4.59$, $p = .043$, $\eta_p^2 = .17$. The main effect of same-or-different shape category was also significant, $F(1, 23) = 5.91$, $p = .023$, $\eta_p^2 = .20$, with an advantage for learning from situations with different types of shape. There was a significant interaction between block and congruence, $F(3, 69) = 2.90$, $p = .041$, $\eta_p^2 = .11$, which was due to a significant difference between congruent and incongruent mappings in block 2, $p = .005$, but not in the other blocks, all $p > .2$, see Figure 5. The stabilising of performance earlier for the congruent than incongruent trials seen in Experiment 1 was not so clearly exhibited here. For congruent trials, there was an improvement from block 1 to block 2, $p < .001$, and then no significant effect of learning (block 2 and block 3, $p = .661$; block 3 and block 4, $p = .076$). For incongruent trials, there was no significant improvement between any two sequential blocks (block 1 and block 2, $p = .473$; block 2 and block 3, $p = .081$; block 3 and block 4, $p = .158$).

The interaction between block and same-or-different shape category was also significant, $F(3, 69) = 6.51$, $p < .001$, $\eta_p^2 = .22$, with an advantage for different shape trials in blocks 2 and 3 ($p < .001$, and $p = .04$, respectively), but not for blocks 1 and 4, both $p > .18$ (see Figure 6).

The critical interaction between congruence and same-or-different shape category was also significant, $F(1, 23) = 7.54$, $p = .012$, $\eta_p^2 = .25$, with learning for congruent mappings when presented with different shapes significantly better than learning from incongruent mappings, or from congruent mappings with same shape categories (see Figure 7), indicating that, as in Experiment 1, sound-symbolism in vowels assisted in learning categories of word-object relationships rather than individual word-referent pairings.

The three-way interaction (block x congruence x same-or-different shape category) was not significant, $F < 1$. We also tested whether participants learned the closed/front or open/back vowel nonwords better, but this was not significant, $F < 1$, and when comparing learning of the rounded versus pointed shapes there was also no significant effect, $F(1, 23) = 3.02$, $p = .096$, $\eta_p^2 = .12$.

The results showed that the effect of sound-symbolism in word learning was also influential when the sound-category distinction was reflected by distinctions in the properties of the vowel, however this effect was due to effects on learning at the category level, and no advantage for learning of individual word-object pairings was observed in the experiment. Finding sound-symbolism at the category level for vowel-shape correspondences contrasts somewhat with the findings from Nielsen and Rendall (2011) who found that consonants rather than vowels were tracked for sound-shape correspondences. But this is perhaps because consonants and vowels conflicted in Nielsen and Rendall's (2011) study, and because the vowel distinction was in terms of rounded-unrounded, whereas the vowels in our study were distinguished in terms of vowel position and height.

Given that sound-category relationships seem to support learning of the category distinction, the question arises as to the extent to which these relationships are observed in natural language. The next study addresses this issue.

Corpus analysis: Correspondences between speech sounds and the angular-rounded distinction in English

If sound-symbolism is useful for language learning, it should be possible to observe relationships in natural language between speech sounds and a whole range of category level distinctions (Hinton, Nicols, & Ohala, 1994). There have been some valiant attempts to assess the prevalence of sound-symbolism in natural language. First amongst these was Newman's (1933) study that found little evidence for clusters of phonemes relating to thesaurus entries. Otis and Sagi (2008) tested the extent to which a range of phonoaesthemes in English reflected semantic categories greater than chance. They found that, of 22 prefixing phonaesthemes proposed in the literature (e.g., Bergen, 2004), 12 proved to be statistically significantly clustered together in terms of context of usage. Abelin (1999) undertook a thesaurus study in Swedish to determine whether words entered in the same category within a thesaurus exhibited clusters of phonaesthemes. The results were not systematically assessed, but there was some evidence for certain pre-specified phonaesthemes corresponding to certain categories of meaning. Gasser, Hockema, and Sethuraman (2010) investigated thesaurus entries in Tamil and in Japanese for mimetics of certain patterns of motion, and found support for systematicity at the phoneme level within categories of movement. These studies suggest that for certain specific contrasts, some small instances of sound-symbolism can be observed. For other sources of sound symbolism, such as mimetics or expressives, there appears to be evidence for their prominence in some languages (Gasser et al., 2010; Imai et al., 2008).

Method

Corpus preparation. We conducted an analysis of sound-symbolism for the angular versus rounded distinction as defined in Roget's Thesaurus (1911; using the Project Gutenberg online edition, 1991), selecting the category "Form" under the class "Space". In this category, we selected all the groups of terms that related to distinctions in terms of roundedness or angularity of shape. For roundedness, the groups were *curvature*, *circularity*, *rotundity*, *convexity*, *concavity*, *bluntness*, and *smoothness*. For angularity, the groups were *angularity*, *straightness*, *sharpness*, and *roughness*. We omitted direct morphological derivations (e.g., *round/rounded*) such that the root form (*round*) only remained, but indirect derivations (e.g., *round/rotund*) were maintained in the list. We also omitted terms from foreign languages, those classified as obsolete, phrases of more than one word, alternative spellings, and if a word occurred in both the angular and the rounded lists then we omitted one, as appropriate.

Phonological forms of the words were derived from the CELEX English word lemma database (Baayen, Pipenbrock, & Gulikers, 1995), and words which did not occur in the database were omitted. There were 198 angular words and 311 rounded words.

Corpus Analysis. Angular and rounded terms were compared in terms of the proportion of consonants that they contained with distinct manner and place features. The manner features were plosives, nasals, laterals/approximants, fricatives, and affricates, and the place features were bilabial, labiodental, dental, alveolar, postalveolar, palatal, velar, and pharyngeal. We also measured the proportion of consonants that were voiced. For the vowels, we measured the mean height and position of vowels in the word, with close vowels scoring 0, close-mid vowels scoring 1, open-mid vowels scoring 2, and open vowels scoring 3. Front vowels scored 0, central vowels 1, and back vowels 2. We also measured the number of vowels and consonants in each word. A similar approach has been applied to discovering systematicities between sound and grammatical categories in a range of languages (Monaghan et al., 2007).

Results and Discussion

The values for proportion of rounded words and angular words that contained each phoneme property were compared using Mann-Whitney U-tests, in order to determine whether certain phonological properties were more prevalent in one set of words compared to the other. The results are shown in Table 3. With no correction for multiple comparisons, the only significant differences found were for proportion of velars, with more in the angular than rounded words, $p = .014$, and voicing, with angular words more likely to be unvoiced than rounded words, $p = .006$.

There is therefore some limited evidence that sound-symbolism may be present in natural language, in terms of place and voicing of consonants, though the effects are very weak (neither is significant if corrected for multiple comparisons), even though there seems to be behavioural evidence for the relationship.

General Discussion

Studies of sound-symbolism have attracted considerable recent interest in terms of their potential role in language acquisition (Spector & Maurer, 2009) and language evolution (Hubbard & Ramachandran, 2001). The observed effects on language processing are certainly striking. In our study, as in previous investigations of certain sound-shape correspondences, we found strong preferences for pairing nonwords containing plosives and front vowels with angular shapes, and continuants and back vowels with rounded shapes. This proved to be a robust effect, even when the distinction between shapes was not highlighted during the course of the experimental study itself. In our study, angular and rounded shapes occurred the same number of times with shapes with similar (within-category) and distinct (between-category) features during the learning trials.

However, the effect of sound-symbolism on language *learning,* in terms of pairing words and shapes, proved to be far more constrained than has been characterised in previous studies of sound-symbolism. Our study confirmed that forced-choice experiments contrasting rounded and angular shapes alongside certain speech sounds results in a preference for certain sound-category pairings to be derived. In both our experiments, participants could more accurately learn to pair nonwords with shapes under the congruent condition than the incongruent condition but only when the shape category distinction was evident in the learning trial – i.e., when one rounded and one angular shape appeared on the screen. When the shapes appearing on the screen were of the same category, then the effect of a congruent or incongruent pairing had no influence on participants' performance, indicating that congruency did not assist in learning the referent for specific, individual words.

Our experiments suggest that previous demonstrations of sound-symbolism effects in language learning facilitate acquisition of word-category mappings rather than individual word-referent pairings. This is entirely consistent with previous studies of the systematicity of natural language in terms of coherence among phonological and prosodic properties of words that belong to the same grammatical category (Cutler, 1993; Farmer, Monaghan, & Christiansen, 2006; Kelly, 1992). For English, there are now identified over 20 speech sound features that relate to grammatical category distinctions (Monaghan & Christiansen, 2008) which together can provide a high degree of accuracy of classification (Monaghan et al., 2005). Similar properties can be found in other, unrelated languages (Monaghan et al., 2007).

However, such category-level correspondences between sound and meaning are not generally the offered interpretation of the apparent sound-symbolic effects on language acquisition for words. This is perhaps because sound-symbolism was previously taken to be a contributor to learning individual words and their meaning rather than making recognition judgments about category distinctions (Imai et al., 2008; Kantartzis et al., 2009; Maurer et al., 2006; Nygaard et al., 2009). Our experiments show these claims to be as yet unsupported – we suggest instead that previous demonstrations of apparent word learning from sound symbolism are likely to be demonstrations of responding to category distinctions.

However, there is some evidence that sound-symbolism may assist in guessing definitions of words. Parault and Schwanenflugel (2006) and Parault and Parkinson (2008) tested participants' ability to guess definitions for obsolete English words that did or did not begin with a phonoaestheme (derived from Bloomfield, 1933, and Ciccotosto, 1991). In these studies, there was no forced choice between categories, and responses were collected according to guessing of definitions of words, and then later by selecting a correct definition. They found more accurate guessing of definitions for the obsolete words which contained identified phonoaesthemes, where an accurate definition was determined if the given definition contained a synonym for the obsolete word. In the latter task, the phonoaestheme relationship may have been highlighted by the task, but in the definition guessing task this was not possible (see also Parault, 2006). Yet, in 45% of cases for the sound-symbolic obsolete words in Study 1 and 51% of Study 2 of Parault and Schwanenflugel (2006), the definition included a word with the same phonaesthemes, which suggests that the results are affected by analogy among word sounds. However, guessing of definitions was not quite the same as learning the words, and it is as yet unknown whether learning of definitions is facilitated by an advantage for guessing definitions.

Another reason why sound-symbolic effects at the word level have attracted interest is because they have been taken to provide insight into how natural correspondences between sound and meaning have influenced choices about words' sounds in the origins of language (Brown, 1958; Hubbard & Ramachandran, 2001; Plato, 1961). However, any observed cross-modal correspondences may just be learned extremely rapidly (e.g., Shankar, Levitan, Prescott, & Spence, 2009) rather than be expressions of innate preferences (e.g., Walker et al., 2010). Thus, evidence from experimental studies may merely be demonstrating effects of prior learning rather than exhibiting forces contributing to language evolution. We have established in our experimental studies that certain mappings between sounds and categories of meaning are easier to process and learn, consistent with learning as a selective pressure for language evolution (Christiansen & Chater, 2008; Nowak and Krakauer, 1999), though if these sound-category correspondences promote learning it remains to be answered why so little of the sound-category relationship is observed in the English language, as demonstrated by our corpus analysis.

So, why is there not more systematicity in the English language, given that it seems to be advantageous for learning? Why does language, as we observed in the introduction, remain largely arbitrary? Establishing that sound-symbolism is at the category level means that we can predict when sound-symbolism is likely to be observed, and specify which aspects of language learning it is likely to benefit. Based on previous studies of word learning, category-level distinctions will only be supported in the language when it is not critical to identify the precise referent of a word, or when the referent is usually disambiguated by context (Monaghan et al., 2011). Gasser et al. (2010) make a case for this applying in terms of expressives. The precise definition of terms such as *large, humungous, enormous* is seldom important to distinguish and they can thus possess sound-symbolic properties to provide a category contrast with expressives of small size (tiny, little, miniscule). Such sound-symbolism has been claimed to be a language universal (Ohala, 1994).

Similarly, for the role of mimetics in language reported by Kita (1997) and experimentally investigated by Imai et al. (2008), it may be the case that the context provides a great deal of information about the intended referent – most mimetics are for actions and it is plausible that few potential actions occur simultaneously in a child’s language learning environment. In contrast, multiple objects are commonplace in a child’s environment (Yu & Ballard, 2007). When there is a single action, sound-symbolism may be advantageous in providing category-level information, as precise meaning is not required to be conveyed. Alternatively, it may be that the precise meaning of the referent for the mimetic is not required, regardless of whether the context disambiguates the intended referent. The principles explored in our experiments thus provide some potential guidance in predicting when sound-symbolism will be observed in natural languages, and when this may be utilised by the child in language learning.

In summary, we have shown that sound-symbolism has a role for language learning, but only in terms of learning relationships between categories of speech sounds and categories of meaning, and it does not have an influence on learning individual word meanings. Such sound-symbolism may have an effect on language learning but can only be restricted to aspects of learning where category distinctions are important but precise identification of meaning is not. Previous studies on language learning have shown that systematicity in sound-meaning correspondences is indeed an impediment for learning individual words (Monaghan et al., 2011). For this reason, sound-symbolism is likely to be non-pervasive in natural language, and, when it does occur, is likely to be restricted to situations where it is sufficient to convey a general meaning rather than a specific identification. Whether there is an early benefit in language acquisition from exposure to these special cases of sound-symbolism, or whether the origins of language satisfied these conditions, remain open and unanswerable questions, respectively.

References

Abelin, A. (1999). Studies in sound symbolism. *Gothenburg Monographs in Linguistics, 17*. Gothenburg: Gothenburg University.

Baayen, R.H.,Pipenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

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

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

Boersma, P. & Weenink, D. (2009). *Praat: doing phonetics by computer,* Version 5.1.05 [Computer program]. Retrieved June 27, 2009, from http://www.praat.org

Brown, R. (1958). *Words and things*. New York: Free Press.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

Christiansen, M. H. & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences 31*, 489–558.

Ciccotosto, N. (1991). Sound symbolism in natural language. Doctoral dissertation, University of Florida. *Dissertation Abstracts International, 53*, 541.

Clark, A. (1998). Magic words: How language augments human computation. In P. Carruthers & J. Boucher (Eds), *Language and thought: Interdisciplinary themes*, 162-183. Cambridge, UK: Cambridge University Press.

Cutler, A. (1993). Phonological cues to open- and closed-class words in the processing of spoken sentences. *Journal of Psycholinguistic Research, 22*, 109-131.

Davis, R. (1961). The fitness of names to drawings: A cross-cultural study in Tanganyika. British Journal of Psychology, 52, 259–268.

de Saussure, F. (1916). *Course in general linguistics*. New York: McGraw-Hill.

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.

Fischer-Jørgensen, E. (1978). On the universal character of phonetic symbolism with special reference to vowels. *Studia Linguistica, 23*, 80-90.

Gasser, M. (2004). The origins of arbitrariness of language. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, 434-439. Mahwah, NJ: Lawrence Erlbaum.

Gasser, M., Sethuraman, N., & Hockema. S. (2010). Iconicity in expressives: an empirical investigation. In J. Newman & S. Rice (Eds.), *Empirical and experimental methods in cognitive/functional research,* pp.163-180. Stanford, CA: CSLI Publications.

Hinton, L., Nichols, J., & Ohala, J. (Eds.). (1994). *Sound symbolism*. Cambridge: Cambridge University Press.

Holland, M. K., & Wertheimer, M. (1964). Some physiognomic aspects of naming, or, maluma and takete revisited. Perceptual and Motor Skills, 19, 111–117.

Horst, J. S., Scott, E. J. & Pollard, J. A. (2010). The role of competition in word learning via referent selection*. Developmental Science*, *13*, 706-713.

Hockett, C. F. (1960). The origin of speech. *Scientific American,* 203, 89-96.

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

Kantartzis, K., Kita, S., & Imai, M. (2009). Japanse sound-symbolism facilitates word learning in English speaking children. *Proceedings of the 31st Annual Conference of the Cognitive Science Society*.

Kelly, M.H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*, 349-364.

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

Köhler, W. (1929). Gestalt psychology. New York, USA: Liveright.

Köhler, W. (1947). Gestalt psychology (2nd ed.). New York, USA: Liveright.

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

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

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

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorization. *Cognition, 96*, 143–182.

Monaghan, P., & Christiansen, M. H. (2008). Integration of multiple probabilistic cues in syntax acquisition. In H. Behrens (Ed.), *Corpora in language acquisition research: History, methods, perspectives*, 139-164. Amsterdam: John Benjamins.

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The Phonological Distributional Coherence Hypothesis: Cross-linguisitic evidence in language acquisition. *Cognitive Psychology, 55*, 259-305.

Monaghan, P., Christiansen, M.H., & Fitneva, S.A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, *140*, 325-347.

Monaghan, P. & Mattock, K. (2009). Cross-situational language learning: The effects of grammatical categories as constraints on referential labelling. *Proceedings of the 31st Cognitive Science Society Conference*. Mahwah, NJ: Lawrence Erlbaum.

Newman, S. (1933). Further experiments in phonetic symbolism. *American Journal of Psychology, 45*, 53–75.

Newmeyer, F. J. (1993). Iconicity and generative grammar. *Language, 68*, 756–796.

Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic *takete-maluma* phenomenon. *Canadian Journal of Experimental Psychology, 65*, 115-124.

Nowak, M.A. & Krakauer, D.C. (1999). The evolution of language. *Proceedings of the National Academy of Sciences, 96*, 8028-8033.

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

Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In Hinton, L., Nichols, J., & Ohala, J. (Eds.), *Sound symbolism* (pp.325-347). Cambridge: Cambridge University Press.

Otis, K. & Sagi, E. (2008). Phonoaesthemes: A corpora-based analysis. In Love, B.C., McRae, K., & Sloutsky, V.M. (Eds.), *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*. Boston, MA.

Paget, R. (1930). *Human speech*. New York: Harcourt.

Parault, S.J. & Parkinson, M. (2008). Sound symbolic word learning in the middle grades. *Contemporary Educational Psychology, 33*, 647-671.

Parault, S.J. & Schwanenflugel, P.J. (2006). Sound-symbolism: A piece in the puzzle of word learning. *Journal of Psycholinguistic Research, 35*, 329-351.

Plato (1961). Cratylus, in Hamilton, E. & Cairns, H. (Eds.), *The collected dialogues*. Princeton, NJ: Princeton University Press.

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

Rhodes, R. (1994). Aural images. In Hinton, L., Nichols, J., & Ohala, J. (Eds.), *Sound symbolism* (pp.276-292). Cambridge: Cambridge University Press.

Roget. P.M. (1911/1991). *Thesaurus*. Project Gutenberg Edition, http://www.gutenberg.org.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology, 12*, 239–255.

Shankar, M. U., Levitan, C. A., Prescott, J., & Spence, C. (2009). The influence of color and label information on flavor perception. *Chemosensory Perception, 2*, 53-58.

Smith, L. & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*, 1558-1568.

Spector, F. & Maurer, D. (2009). Synesthesia: A new approach to understanding the development of perception. *Developmental Psychology, 45*, 175-189.

Tamariz, M. (2008). Exploring systematicity between phonological and context-cooccurrence representations of the mental lexicon. *The Mental Lexicon, 3*, 259-278.

Walker, P, Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21*, 21-25.

Wertheimer, M. (1958). The relation between the sound of a word and its meaning. *American Journal of Psychology, 71*, 412–415.

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

Yu, C. & Ballard, D.H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing, 70*, 2149-2165.

Yu, C. & Smith, L. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*, 414-420.

Table 1. Acoustic properties of speech stimuli in Experiment 1. Values in bold are statistically significant.

Property	Continuant nonwords (SD)	Plosive nonwords (SD)	*t*(14) =	*p*
F0 at onset (Hz)	218.4 (3.9)	227.6 (4.9)	**-4.15**	.001
F0 midpoint (Hz)	215.5 (5.2)	219.4 (5.5)	-1.45	.169
F0 offset (Hz)	198.5 (37.6)	217.2 (19.0)	-1.26	.229
F0 Mean (Hz)	207.2 (23.6)	214.1 (18.2)	-.652	.525
F0 Mean onset, midpoint, offset (Hz)	209.9 (15.6)	219.6 (9.2)	-1.52	.151
F0 Minimum (Hz)	171.2 (53.3)	191.0 (45.1)	-.80	.437
F0 Maximum (Hz)	231.2 (11.8)	241.6 (12.8)	-1.69	.113
F0 Range (Hz)	33.5 (47.1)	67.1 (66.0)	-1.17	.260
F1 at onset (Hz)	556.0 (92.1)	599.9 (37.2)	-1.25	.232
F1 at midpoint (Hz)	585.0 (190.1)	721.8 (298.5)	-1.09	.293
F1 at offset (Hz)	505.9 (161.6)	588.5 (103.1)	-1.22	.243
F1 mean (Hz)	549.0 (140.0)	636.7 (117.3)	-.89	.391
F1 mean onset, midpoint, offset (Hz)	551.0 (149.9)	601.6 (60.6)	-1.36	.196
F2 at onset (Hz)	1638.5 (445.7)	1921.6 (430.4)	-1.29	.217
F2 midpoint (Hz)	1695.9 (488.1)	1888.4 (423.5)	-.84	.414
F2 offset (Hz)	1491.8 (458.3)	1772.5 (372.3)	-1.35	.200
F2 mean (Hz)	1646.8 (424.4)	1843.1 (392.0)	-.96	.353
F2 mean onset, midpoint, offset (Hz)	1608.7 (441.5)	1860.0 (383.7)	-1.22	.244
Duration (ms)	441.5 (62.3)	415.4 (71.8)	.78	.450
Transition duration (ms)	52.5 (54.5)	78.2 (44.4)	-1.04	.317
Intensity consonants (dB)	70.9 (3.1)	65.1 (6.4)	**2.29**	**.038**
Rise time (ms)	107.9 (29.3)	97.1 (56.4)	.48	.640
dB low	68.0 (3.8)	62.1 (8.7)	1.77	.098
dB high	72.7 (2.0)	71.8 (2.7)	.74	.472
dB difference	4.7 (3.2)	9.7 (6.4)	-2.00	.066
Intensity vowel at onset (dB)	71.7 (1.7)	71.9 (2.4)	-.23	.822
Intensity vowel midpoint (dB)	70.0 (2.3)	71.5 (1.9)	-1.47	.163
Intensity vowel offset (dB)	68.7 (2.3)	67.7 (2.6)	.84	.414
Intensity vowel mean (dB)	70.2 (1.7)	69.2 (2.4)	.91	.378
Intensity vowel mean onset, midpoint, offset (dB)	70.1 (2.0)	70.3 (1.9)	-.21	.836

Table 2. Acoustic properties of speech stimuli in Experiment 2. Values in bold are statistically significant.

Property	Closed/front nonwords (SD)	Open/back nonwords (SD)	*t*(14) =	*p*
F0 at onset (Hz)	222.6 (12.0)	221.8 (5.3)	.19	.853
F0 midpoint (Hz)	201.9 (56.9)	248.8 (90.7)	-1.24	.236
F0 offset (Hz)	223.2 (23.6)	253.9 (92.0)	-0.91	.377
F0 Mean (Hz)	222.2 (18.9)	221.2 (7.1)	.14	.890
F0 Mean onset, midpoint, offset (Hz)	217.5 (23.8)	236.4 (46.6)	-1.02	.325
F0 Minimum (Hz)	216.5 (20.9)	213.0 (7.3)	.45	.661
F0 Maximum (Hz)	230.6 (21.0)	231.5 (9.6)	-0.11	.916
F0 Range (Hz)	14.1 (9.9)	18.5 (11.9)	-0.80	.439
F1 at onset (Hz)	527.6 (96.8)	536.8 (88.6)	-0.20	.847
F1 at midpoint (Hz)	581.8 (128.4)	506.1 (123.0)	1.20	.249
F1 at offset (Hz)	513.9 (75.5)	517.0 (155.6)	-0.05	.960
F1 mean (Hz)	541.1 (78.2)	520.0 (116.1)	.43	.676
F1 mean onset, midpoint, offset (Hz)	546.2 (87.7)	518.5 (119.6)	.53	.605
F2 at onset (Hz)	1314.8 (231.1)	1989.5 (440.5)	**-3.84**	**.002**
F2 midpoint (Hz)	1279.4 (219.7)	1802.5 (667.5)	-2.11	.054
F2 offset (Hz)	1322.1 (228.8)	1801.9 (385.0)	**-3.03**	**.009**
F2 mean (Hz)	1311.0 (188.3)	1850.8 (310.4)	**-4.20**	**.001**
F2 mean onset, midpoint, offset (Hz)	1307.1 (184.4)	1859.5 (290.4)	**-4.54**	**<.001**
Duration (ms)	429.8 (54.8)	451.2 (51.8)	-0.81	.433
Transition duration (ms)	68.0 (49.3)	82.6 (51.9)	-0.58	.572
Intensity consonants (dB)	71.8 (2.7)	70.9 (2.4)	.70	.498
Rise time (ms)	92.0 (57.2)	112.2 (50.3)	-0.75	.465
dB low	69.1 (6.0)	65.9 (7.0)	.98	.344
dB high	74.7 (1.6)	75.5 (1.3)	-1.21	.245
dB difference	5.6 (5.5)	9.7 (6.7)	-1.32	.207
Intensity vowel at onset (dB)	74.2 (1.2)	73.8 (1.7)	.69	.499
Intensity vowel midpoint (dB)	72.2 (1.4)	71.8 (2.0)	.58	.568
Intensity vowel offset (dB)	70.2 (1.6)	69.4 (3.4)	.66	.518
Intensity vowel mean (dB)	72.0 (1.4)	72.0 (1.7)	.00	1.000
Intensity vowel mean onset, midpoint, offset (dB)	72.2 (1.1)	71.7 (1.7)	.66	.518

Table 3. Mean (SD in parentheses) proportion of consonants with each manner and place feature, voicing, vowel position and height, and length in consonants and in vowels, for words classified as angular and rounded. Significant distinctions are shown in bold font.

Phonological Property	Angular words	Rounded Words	*Z*	*p*
Consonant Manner Features				
Plosive	.42 (.26)	.41 (.29)	1.00	.318
Nasal	.15 (.21)	.14 (.21)	.38	.702
Approximant	.22 (.23)	.23 (.23)	.58	.560
Fricative	.19 (.23)	.20 (.24)	.46	.643
Affricate	.03 (.10)	.02 (.10)	.89	.373
Consonant Place Features				
Bilabial	.14 (.21)	.18 (.26)	1.21	.228
Labiodental	.04 (.12)	.06 (.16)	1.81	.070+
Dental	.01 (.08)	.01 (.05)	.82	.413
Alveolar	.55 (.27)	.55 (.30)	.44	.663
Postalveolar	.02 (.08)	.01 (.08)	.58	.563
Palatal	.01 (.06)	.01 (.04)	.04	.966
Velar	**.18 (.23)**	**.14 (.21)**	**2.46**	**.014**
Pharyngeal	.02 (.09)	.02 (.08)	1.11	.266
Voiced	**.52 (.29)**	**.59 (.31)**	**2.74**	**.006**
Vowel Properties				
Height	1.73 (.92)	1.62 (.86)	1.38	.167
Position	.93 (.73)	.92 (.72)	.01	.989
Number of vowels	1.88 (1.10)	1.75 (.92)	1.01	.310
Number of consonants	3.16 (1.21)	3.00 (1.16)	1.29	.199

Figure Captions

Figure 1. Examples of angular and rounded shapes used in Experiments 1 and 2.

Figure 2. Interaction between learning block and sound-shape congruence for Experiment 1. Error bars show Standard Error of the Mean (SEM).

Figure 3. Interaction between learning block and same-or-different shape category for Experiment 1. Error bars show SEM.

Figure 4. Interaction between sound-shape congruence and same-or-different shape category for Experiment 1. Error bars show SEM.

Figure 5. Interaction between learning block and sound-shape congruence for Experiment 2. Error bars show SEM.

Figure 6. Interaction between learning block and same-or-different shape category for Experiment 2. Error bars show SEM.

Figure 7. Interaction between sound-shape congruence and same-or-different shape category for Experiment 2. Error bars show SEM.

Figure 1.

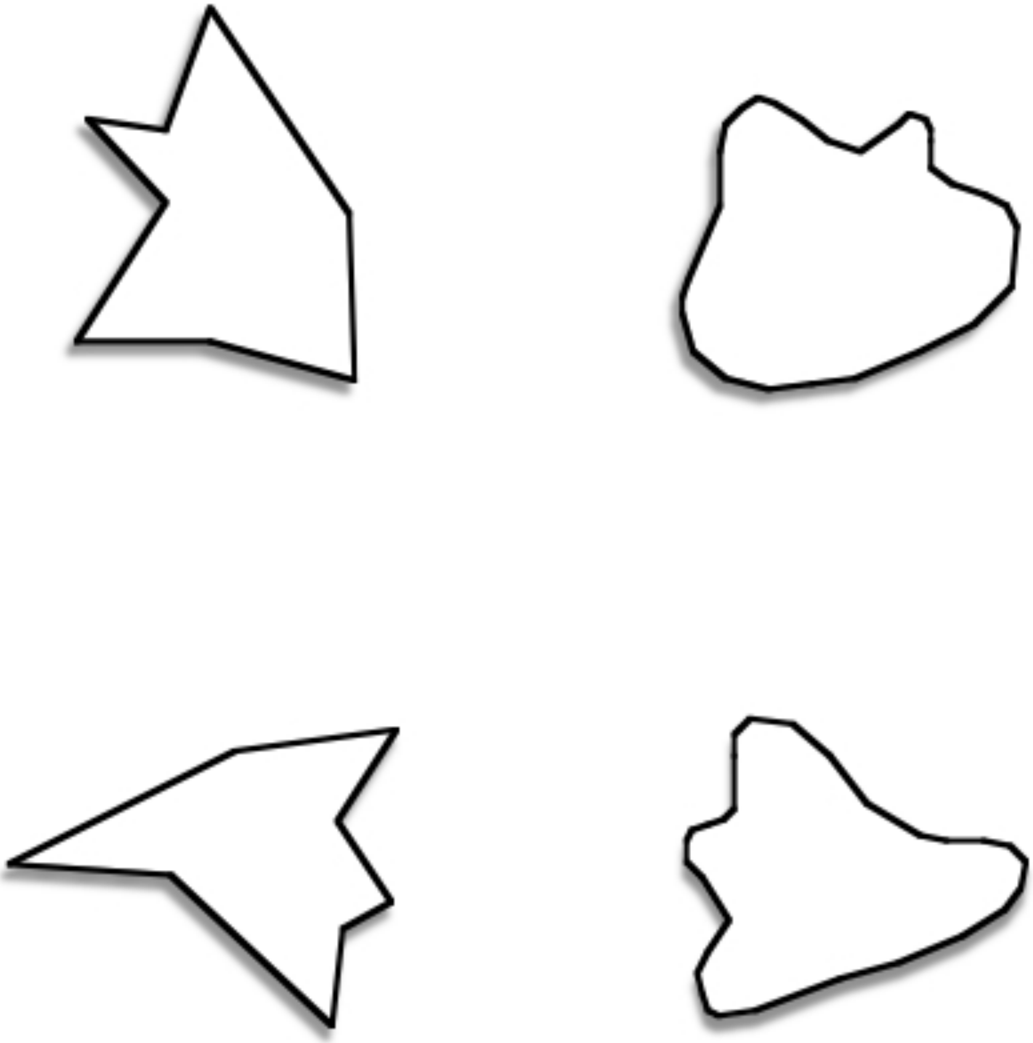
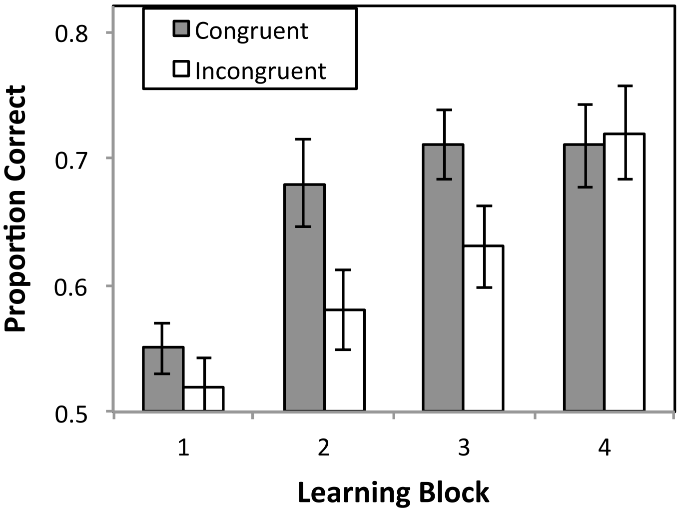
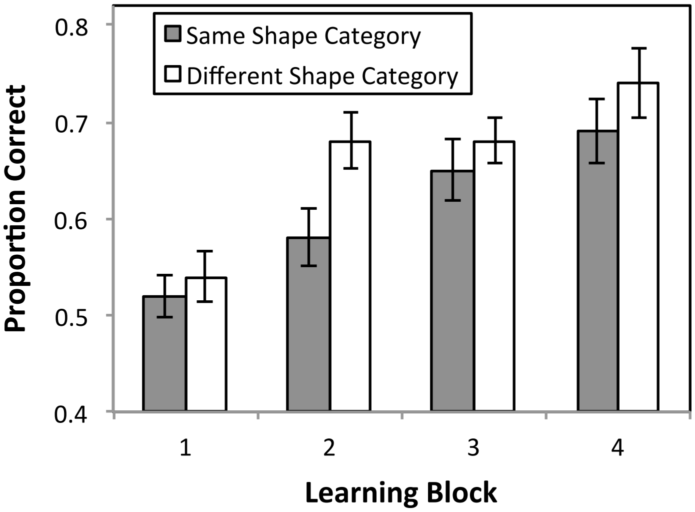


Figure 2.



Congruent
Incongruent
0.8
0.7
0.6
0.5
Proportion Correct
1
2
3
4
Learning Block

Figure 3.



Same Shape Category
Different Shape Category
Proportion Correct
0.8
0.7
0.6
0.5
0.4
1
2
3
4
Learning Block

Figure 4.

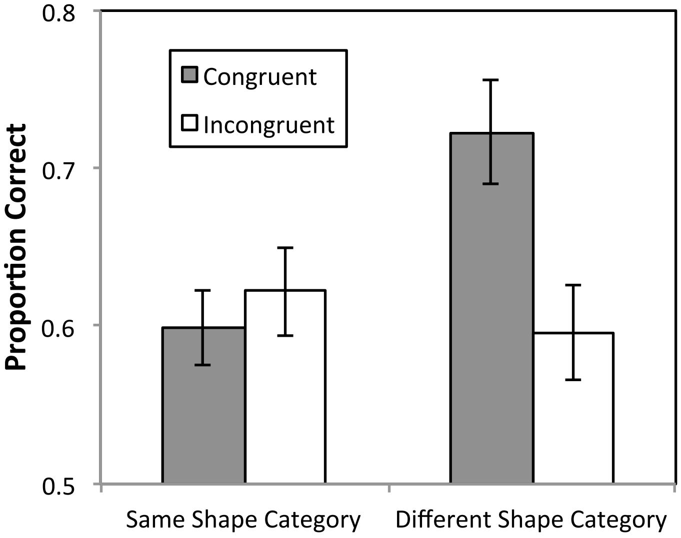

0.8
0.7
0.6
0.5
Proportion Correct
Congruent
Incongruent
Same Shape Category
Different Shape Category

Figure 5.

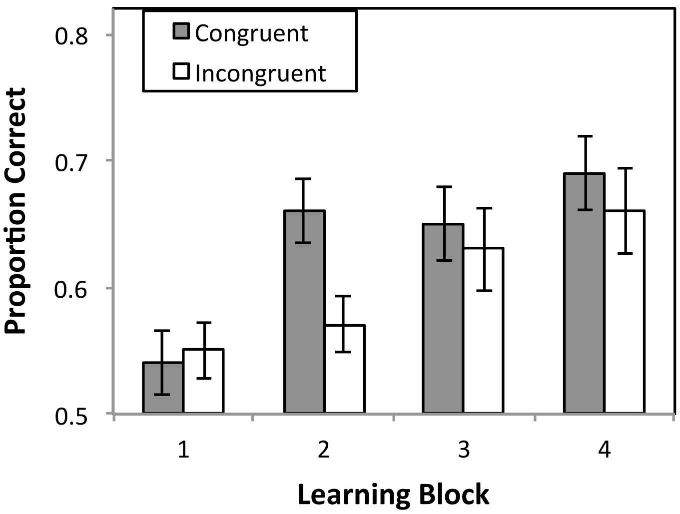
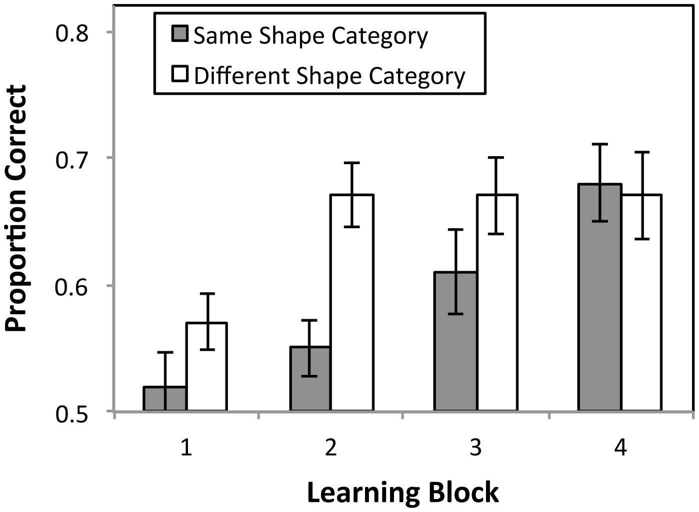
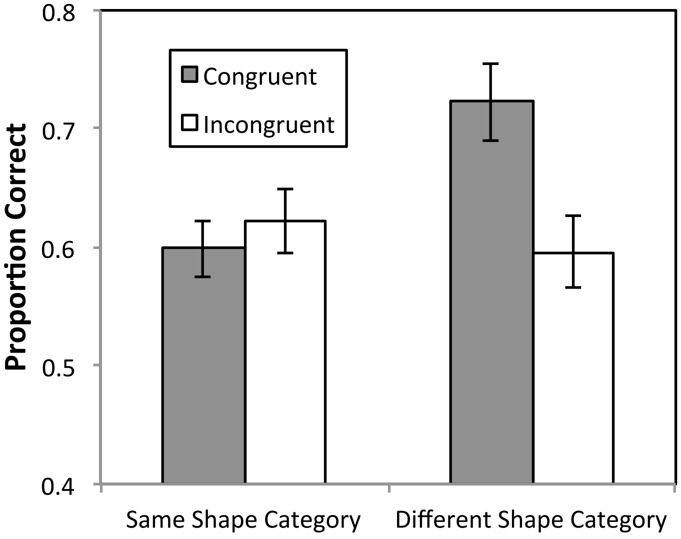

Congruent
Incongruent
Proportion Correct
0.8
0.7
0.6
0.5
1
2
3
4
Learning Block

Figure 6.



Same Shape Category
Different Shape Category
Proportion Correct
0.8
0.7
0.6
0.5
1
2
3
4
Learning Block

Figure 7.


0.8
0.7
0.6
0.5
0.4
Proportion Correct
Congruent
Incongruent
Same Shape Category
Different Shape Category

Appendix 1: Words used in Experimental Studies

Experiment 1:

Plosive words: /kɪb/, /ɡæt/, /tɛɡ/, /dɒp/, pɛd/, /bɪk/, /tɒb/, /kæɡ/.

Continuant words: /mɒŋ/, /nɪm/, /læn/, /ɹɛŋ/, /wɒl/, /wɛm/, /ɹɪn/, /næl/.

Experiment 2:

Front/closed vowels: /ɡin/, /dɪŋ/, /bɛm/, /meɪp/, /nib/, /leɪɡ/, /ɹɛl/, /wɪd/

Back/open vowels: /ɡɑ:n/, /dɒm/, /bɔ:ŋ/, /mʌb/, /nɔ:d/, /lɑ:ɡ/, /ɹʌd/, /wɒb/

Appendix 2: Acoustic properties assessed for experiment materials

F0 at onset (Hz): Fundamental frequency (voice pitch) measured at onset of periodic portion in word (periodic portion is region of the speech signal for which pitch can be reliably detected)

F0 midpoint (Hz): Fundamental frequency measured at midpoint of periodic portion

F0 offset (Hz): Fundamental frequency measured at offset of periodic portion

F0 Mean (Hz): Mean Fundamental frequency of whole periodic portion

F0 Mean onset, midpoint, offset (Hz): Mean fundamental frequency of onset, midpoint and offset values

F0 Minimum (Hz): Minimum fundamental frequency of whole periodic portion

F0 Maximum (Hz): Maximum fundamental frequency of whole periodic portion

F0 Range (Hz): Fundamental frequency range as indexed by F0 maximum minus F0 minimum

F1 at onset (Hz): First formant (frequency component of speech relating to tongue height) at onset of vowel

F1 at midpoint (Hz): First formant at midpoint of vowel

F1 at offset (Hz): First formant at offset of vowel

F1 mean (Hz): Mean first formant of whole vowel

F1 mean onset, midpoint, offset (Hz): Mean first formant of onset, midpoint and offset values

F2 at onset (Hz): Second formant (frequency component of speech relating to tongue position in terms of front/back) at onset of vowel

F2 midpoint (Hz): Second formant at midpoint of vowel

F2 offset (Hz): Second formant at offset of vowel

F2 mean (Hz): Mean second formant of vowel

F2 mean onset, midpoint, offset (Hz): Mean second formant of onset, midpoint and offset values

Duration (ms): Duration of whole word (ms)

Transition duration (ms): Duration of transition between initial consonant and vowel which reflects articulatory movements, for example, change from bilabial plosive to open vowel, and rate of spectral change

Intensity consonants (dB): Amplitude (loudness/intensity) of consonants

Rise time (ms): Measure of the time taken (duration) for the initial consonant to reach its peak amplitude (loudness/intensity; dB)

dB low: Minimum amplitude (loudness/intensity; dB) of a word

dB high: Maximum amplitude (loudness/intensity; dB) of a word

dB difference: Amplitude range as indexed by Maximum minus Minimum amplitude

Intensity vowel at onset (dB): Measure of vowel intensity (dB) at vowel onset

Intensity vowel midpoint (dB): Measure of vowel intensity (dB) at vowel midpoint

Intensity vowel offset (dB): Measure of vowel intensity (dB) at vowel offset

Intensity vowel mean (dB): Mean vowel intensity across word

Intensity vowel mean onset, midpoint, offset (dB): Mean vowel intensity of onset, midpoint and offset values