Title
Sound to Meaning Mappings in the Bouba-Kiki Effect

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

Authors
McCormick, Kelly
Kim, Jee Young
List, Sara
et al.

Publication Date
2015

Peer reviewed

Sound to Meaning Mappings in the Bouba-Kiki Effect

Kelly McCormick (kelly.mccormick@emory.edu)
Department of Psychology, 36 Eagle Row
Emory University, Atlanta, GA 30322 USA

Jee Young Kim (j.kim@yale.edu)
Department of Psychology, 36 Eagle Row
Emory University, Atlanta, GA 30322 USA

Sara List (slist@emory.edu)
Neuroscience Program
Graduate Division of Biological and Biomedical Sciences, 1462 Clifton Road
Emory University, Atlanta, GA 30322 USA

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

Abstract

Sound to meaning correspondences in spoken language are assumed to be largely arbitrary. However, research has identified a number of exceptions to the arbitrariness assumption. In particular, non-arbitrary mappings between sound and shape, the bouba/kiki effect, have been documented across diverse languages and both children and adults are sensitive to this type of sound symbolic mapping. The cognitive basis for the associations between nonword labels and particular shapes remains poorly understood making it difficult to predict how findings generalize beyond the limited stimuli tested. To identify systematic bases for sound-to-shape mappings, we collected ratings of roundedness and pointedness for a large database of pseudowords. We find that attributes of both consonants and vowels are systematically related to judged shape meanings of pseudowords, and offer hypotheses as to the cognitive mechanisms underlying the observed patterns.

Keywords: Sound symbolism; Language; Bouba-Kiki Effect; Multisensory Representation

A central question in the cognition of language is how sounds in language encode and convey meaning. Sound-to-meaning mappings in language are believed to be largely arbitrary, with sounds in words bearing no inherent relationship to the objects, actions, and events in the world that they represent (Gasser, 2004; Monaghan, Christiansen, & Fitneva, 2011; Saussure, 1916). In such a system words come to represent meanings by learned associations, and in principle, any meaning could be represented by any combination of the finite inventory of sounds in a given language. Arbitrary reference is considered powerful because symbolic forms are unconstrained in relation to meaning, supporting referential flexibility (Gasser, 2004; Monaghan & Christiansen, 2005).

But is the sound structure of language completely independent of meaning or are certain speech sounds systematically related to the meanings they represent? Research has found that certain classes of sounds are more likely than others to appear in words with particular meanings and that language users are sensitive to these non-arbitrary sound to meaning mappings. Cross-linguistic research has identified systematic sound-to-meaning mappings in terms describing a range of perceptual experiences such as brightness/lightness (Hirata & Kita, 2011; Kunihira, 1971), texture (Dingemanse, 2011; Kita, 1997; Magnus, 2001), and size (Nuckolls, 1999; Ultan, 1978). For example, across many languages vowels with higher frequency components (such as /i/) tend to be associated with small/diminutive concepts, whereas vowels with lower frequency components (such as /a/) tend to be used to represent large or augmentative concepts (Nuckolls, 1999; Ultan, 1978). Further, listeners appear to be sensitive to these regularities in spoken language and reliably apply these mappings when inferring meanings of unfamiliar words (Mondloch & Maurer, 2004; Sapir, 1929; Spector & Maurer, 2008; Thompson & Estes, 2011).

Perhaps one of the most studied examples of non-arbitrariness, coined the 'bouba-kiki effect', involves the reliable matching of nonsensical names to abstract shapes. Although the specific labels and shapes employed in the task have varied from study to study, the basic finding is that in a forced-choice task people consistently match labels such as *kiki* and *takete* to angular/pointed shapes and match labels such as *bouba* and *maluma* to rounded/amoeboid shapes (Köhler, 1929; Ramachandran & Hubbard, 2001). Robust mappings between certain nonsense words and object shapes have been demonstrated across a variety of languages and cultures, with reported matching rates ranging from approximately 80-95% of respondents (Bremner et al., 2013; Davis, 1961; Köhler, 1929). These are unfamiliar nonsense words and novel shapes, with which the individual does not have prior experience, yet stable associations emerge across individuals.

What makes some sounds more fit than others for representing certain meanings? Although there is widespread documentation of the bouba-kiki effect across languages, cultures, and development, the phenomenon is not well understood (see Nielsen and Rendall, 2011). In the present study, we seek to identify acoustic and phonetic patterns in this particular sound to meaning mapping and work to explain how certain attributes of the linguistic signal might serve to bring online mental representations of meaning.

Experiment

The aim of this study is to identify acoustic and phonemic correlates to meaning, and posit cognitive mechanisms underlying these sensory-conceptual mappings. We examined the specific relationship between the sound structure of language and meanings related to visuo-haptic properties of objects. To assess language users' mappings between pseudowords and rounded and pointed meaning, we conducted three behavioral rating experiments including one forced-choice task and two likert rating tasks, collecting ratings on 570 nonsense words on dimensions of roundedness or pointedness. Numerous studies examining the bouba-kiki effect have found associations between specific speech sounds and visuo-haptic object properties. However, previous research on sound to shape mappings has generally employed a very limited number of pseudowords (but see Fort, Martin, & Peperkamp, 2014), most famously *takete* and *kiki* (which are mapped to pointed shapes) and *maluma*, and *bouba* (mapped to bloblike shapes). Although these studies report high rates of reliability in matching tasks, they offer little in the way of an explanation of the phenomenon. To evaluate the contribution of various properties of spoken language, we build on previous research by systematically sampling the phonetic and acoustic space of reliable sound to shape mappings. To this end, we constructed a database of English-like pseudowords from several classes of speech sounds. The resulting permutations of phonemic combinations allowed us to examine how particular sound properties are mapped to meaning.

We examined speech sounds that varied in *manner of articulation*, including stops, sonorants, and fricatives, and *voicing*. Each consonant class included voiced and voiceless sounds with the exception of sonorants, which only exist in voiced form. The sonorants in our set consist of both nasals and liquids, the obstruents consist of affricates, fricatives and stop consonants. We combined these classes of consonants with different vowel types, rounded and unrounded, in order to study how phonemes combine and interact and how accompanying acoustic and articulatory properties mapped to particular meanings.

Because of the number of possible permutations of phoneme combinations in English, we constrained our set to include classes of sounds that appear to be reliably mapped in bouba-kiki tasks or for which there was a clear possible basis for mapping. For example, stimuli containing unvoiced stop consonants are preferentially mapped to pointed shapes (Bremner et al., 2013; Fort et al., 2014; Köhler, 1929; Nielsen & Rendall, 2011; Ramachandran & Hubbard, 2003; Westbury, 2005). In contrast, sonorants emerge as a class of sounds that are preferentially mapped to bloblike or amoeboid shapes (Fort et al., 2014; Köhler, 1929; Ramachandran & Hubbard, 2003; Westbury, 2005). As an intermediate category between stops (which are completely obstructed when produced) and sonorants (which are completely unobstructed), we included fricative and affricate phonemes. Parise and Spence (Parise & Spence, 2009, 2012) found that tones composed of square waves (and which have a noisy quality) were associated with a more pointed visual object as compared to tones composed of sinusoidal waves (which have a smoother tonal quality), which were associated with the more bloblike shape. With respect to roundedness or pointedness of various vowels, research comparing speakers of English and Czech showed that both groups reliably matched /i/ (as in *neat*) to a triangular shape and /u/ (as in *mood*) to an elliptical shape (Tarte & Barritt, 1971; Tarte, 1974). In an implicit association experiment, Parise and Spence (2012) found that perceivers associate relatively high pitch tones with more acutely pointed shapes and relatively lower pitch tones with more obtusely pointed shapes.

In addition to systematically sampling the acoustic and phonemic space, we also varied the way in which we assessed sound to meaning mappings. Previous research on the bouba-kiki phenomenon has relied heavily on two-alternative forced choice tasks. Although these studies provide robust evidence that certain sounds are preferentially matched to particular meanings, they do not necessarily capture richer, more nuanced information about these mappings. By using likert-type rating tasks in addition to typical forced-choice tasks, we included a measure that was able to capture graded phenomena.

Method

Participants

Participants were 65 members of the Emory University community. All participants were native speakers of English, with normal hearing and no reported history of speech/language impairments. 34 participants participated in the two-alternative forced-choice task, and 31 participants participated in the likert-type rating task.

Materials

Stimuli consisted of audio recordings of 570 pseudowords. All were composed of two syllables with consonant-vowel-consonant-vowel (CVCV) structure and comprised of a subset of sounds from the phonemic inventory of American English. All nonwords in the set contained of one of three classes of consonants: sonorants (/n/, /m/, /l/), affricates/fricatives (/v/, /z/, /d͡ʒ/, /f/, /s/, /t͡ʃ/), or stops (/d/,

/g/, /b/, /t/, /k/, /p/). Half of each type of obstruent were voiced (/v/, /z/, /d͡ʒ/, /d/, /g/, /b/) and half were unvoiced (/f/, /s/, /t͡ʃ/, /t/, /k/, /p/). Within a given nonword, consonants were either both unvoiced (as in *kupo*) or both voiced (as in *gubo*). Nonwords in the set contained either front/rounded (/u/, /ʊ/, /o/) or back/unrounded (/i/, /e/, /ɪ/, /ɛ/) vowels. Stimuli in the set did not contain reduplicated (repeated) syllables, so nonwords such as *kiki* and *lolo* did not appear. Because three vowels in our set (/ɪ/, /ɛ/, /ʊ/) do not appear in word-final positions in English, these vowels appeared only in the first vowel (V1) position in our stimuli. All stimulus items were recorded by a female native English speaker and edited into separate sound files for presentation.

Procedure

Two-alternative forced choice task Participants in the forced-choice task were asked to decide whether each nonword stimulus sounded more rounded or more pointed. Participants sat at a desktop computer in a sound-attenuated room. The experiment consisted of 570 trials, presented in random order. Each trial began with a blank white screen, after 1000 milliseconds the pseudoword played, then the screen changed to show the two response options (Pointed, Rounded) presented on the left and right sides of the screen. Participants responded with a button press corresponding to the response options displayed on the screen. The configuration of response keys was counterbalanced across participants. Auditory stimuli were played through Beyerdynamic DT 100 headphones at approximately 75dB SPL. Only responses made after the word had played in its entirety were registered by the system. This ensured that participants heard the entire pseudoword before judging meaning.

Likert tasks Because *roundedness* and *pointedness* may not be mutually exclusive concepts, we collected separate likert-type ratings for the roundedness and pointedness of each of the 570 words in our stimulus set.

Pointedness rating task. Fifteen participants completed this rating task. The experiment design was exactly as described for the two-alternative forced-choice paradigm, with the exception of the response phase of each trial. After hearing each pseudoword, the participant rated how pointed each nonword sounded on a 7-point scale ('1-not pointed' to '7-very pointed').

Roundedness rating task. A separate group of 16 participants completed this rating task. The experiment design was exactly as described for above, with the exception that in the response phase, the participant responded how rounded each nonword sounded on a 7-point scale ('1- not rounded' to '7- very rounded').

Results

Three independent judges identified items in the stimulus set that were actual words in English. Any item identified by at least two judges as being a real word was excluded from subsequent analyses. We conducted a series of analyses to compare ratings across the classes of sounds, and to evaluate the extent to which various properties of our pseudowords influenced judgments of roundedness and pointedness in each task.

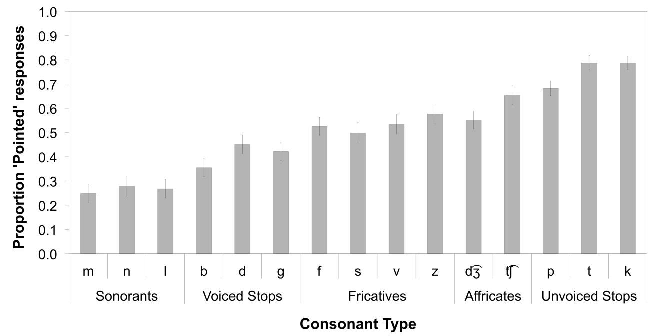

Figure 1: Proportion pointed responses on two alternative forced-choice task by consonant class.

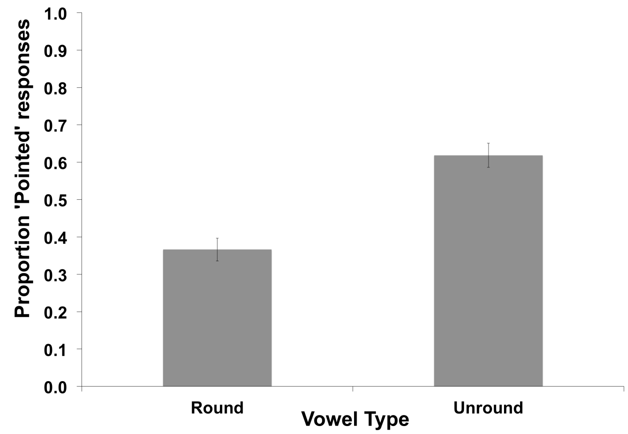

Figure 2: Proportion pointed responses on two alternative forced-choice task by vowel type.

Forced-choice task Figure 1 shows forced-choice performance for each consonant manner and voicing class.

Consonants. We first compared the proportion of pointedness ratings for stimuli consisting of two broad classes of phonemes, sonorants and obstruents. A paired-sample t-test showed that subjects rated nonwords containing obstruents as sounding pointed (M=.58) significantly more than nonwords containing sonorants (M=.27), $t(33)= 7.47$, $p<.001$.

In order to examine the effect of consonant voicing and manner of articulation on judgments of meaning, we conducted a two-way repeated measures ANOVA with manner (affricates/fricatives, stops) and voicing (voiced, unvoiced) as factors. Because sonorants only occur in voiced form, our analysis excluded these items. Although there was no main effect of manner of articulation on round/pointed judgments, there was a significant main effect of consonant voicing, $F(1,33)=46.76$, $p<.001$, partial $\eta^2=.586$. There was also a significant interaction between manner of articulation and voicing $F(1, 33)= 35.30$, $p<.001$, partial $\eta^2=.517$. Means comparisons showed that affricates/fricatives produced comparable ratings in the voiced and unvoiced groups (both classes, M=.56; *p*>.1), whereas unvoiced stops produced significantly more 'pointed' judgments (M=.76) than voiced stops (M=.42; *p*<.001).

Vowels. Figure 2 shows forced-choice performance for rounded and unrounded vowels. Participants judged words composed of rounded vowels as sounding pointed at significantly lower rates (M=.38) than words composed of unround vowels (M=.62), *t*(33)=-5.47, *p* <.001.

Likert-rating tasks

Consonants. Table 1 reports mean roundedness and pointedness ratings for each consonant class. We conducted one-way within-subjects ANOVAs with manner of articulation as a factor for both rounded and pointed likert-rating data. For both rounded and pointed ratings, there was a significant effect of manner of articulation, $F_{rounded}(1.472, 22.08)=10.30$ (Greenhouse-Geisser corrected), $\eta^2=.407$, *p*=002; $F_{pointed}(2,28)=7.82$, $\eta^2=.358$, *p*=.002. Means comparisons of roundness ratings demonstrated that nonwords containing stops were rated as less rounded than nonwords containing sonorants (*p*=.007), and nonwords containing fricative/affricates were rated as less rounded than nonwords containing sonorants (*p*=.001). Affricates and stops did not differ significantly (*p*=.86). Means comparisons of pointedness ratings demonstrated that ratings for nonwords containing sonorants differed significantly from nonwords containing stops (*p*=.001) as well as for nonwords containing affricates/fricatives (*p*=.04). Affricates and stops did not differ significantly (*p*=.18).

Table 1: Likert Ratings (1-7 scale) by consonant class.

	Roundedness ratings			
	Voiced		Unvoiced	
Consonant features	*M*	*SD*	*M*	*SD*
Sonorants	4.22	0.96	~	~
Affricates/Fricatives	3.36	0.77	3.38	0.71
Stops	3.68	0.77	3.11	0.71
	Pointedness ratings			
	Voiced		Unvoiced	
Consonant features	*M*	*SD*	*M*	*SD*
Sonorants	3.34	0.90	~	~
Affricates/Fricatives	4.13	0.72	3.82	0.57
Stops	3.98	0.84	4.73	0.85

In light of the results of our forced-choice task, we expected that consonant voicing might interact with manner of articulation to affect graded judgments of meaning. To test this, we conducted two-way repeated measures ANOVAs for both rounded and pointed datasets, with manner of articulation and voicing as factors. For both the rounded and pointed datasets, the effect of consonant voicing on likert ratings approached significance (*p*=.058 pointed, *p*=.064 rounded). The overall trend was for voiced consonants to produce higher ratings on the rounded scale as compared to unvoiced consonants. Neither rounded nor pointed datasets showed a main effect for manner of articulation. In both analyses, there was a significant interaction between manner and voicing. For the interaction in the pointed ratings F(1,14)=24.82, *p*<.001, partial $\eta^2=.639$, unvoiced stops were judged as more pointed (M=4.73) than voiced stops (M=3.98; *p*=.001), whereas unvoiced fricatives/affricates were judged as less pointed (M=3.82) than voiced fricatives/affricates (M=4.13; *p* <.009). The interaction of manner and voicing in the roundedness ratings, F(1,15)=10.24, p=.006, partial $\eta^2=.406$, appeared to be driven by differences in the stop consonants. The affricates/fricatives were rated similarly on the different levels of voicing (voiced M=3.36, unvoiced M= 3.38; *p*=.83), voiced stops were rated as sounding more rounded than unvoiced stops (voiced M=3.68, unvoiced M=3.11; *p*=.01).

Table 2: Likert ratings (1-7 scale) by vowel type.

	Roundedness ratings	
Vowel features	*M*	*SD*
Round	4.29	0.85
Unround	2.97	0.94
	Pointedness ratings	
Vowel features	*M*	*SD*
Round	3.73	0.56
Unround	4.20	0.62

Vowels. Table 2 reports mean roundedness and pointedness ratings for each vowel type. To determine the relationship between vowel rounding and likert ratings of roundedness and pointedness, we conducted pairwise t-tests on both datasets. Subjects judging roundedness rated words with rounded vowels as sounding more rounded than words with unrounded vowels (M=4.29 and 2.97 respectively) *t*(15)=3.77, *p*=.002. Subjects judging pointedness rated words with unrounded vowels as sounding more pointed (M=4.20) than words with rounded vowels (M=3.73) *t*(14)=-2.52, p=.02.

Discussion

The current set of findings demonstrate that certain classes of speech sounds were readily matched to shape-based meanings and response patterns remained consistent across different task types. Likert ratings of roundedness and pointedness mirrored both one another (words rated higher on the rounded scale were rated lower on the pointed scale and vice versa) and the responses given in the two-alternative forced choice task. This consistency suggests that roundedness and pointedness were treated as contrastive categories in these tasks such that a nonword that was judged as more rounded was also judged to be less pointed. Consistent with previous work, language users readily mapped sound to this domain of meaning, and did so across a range of tasks.

The findings also provide information about what particular correspondences underlie mappings of particular sound to shape-based concepts. Nonword stimuli containing the relatively strident, noisy or discontinuous sounds of obstruents such as /b/, /t/, or /z/ were judged as sounding more pointed than words containing sonorant sounds such as /m/ and /l/. Consonant voicing modulated judgments of pointedness or roundedness of the nonword stimuli as well. In general, nonwords containing voiced consonants were judged as more round than nonwords containing unvoiced consonants. Consonant voicing appeared to influence the strident or discontinuous perceptual nature of the obstruents in our set, perhaps because voiceless phonemes have more abrupt transitions than their voiced counterparts (e.g. /p/ compared to /b/). In line with previous research, we also found that back, rounded vowels were strongly associated with rounded meanings, whereas front, unrounded vowels are more often associated with pointed meanings.

In addition to the individual contribution of phoneme type, these findings also suggest that the *combined* sound attributes of a particular stimulus served to determine the extent to which a mapping would be made. For example, the patterning of responses illustrated in Figure 1 suggests that judgments of shape were graded with respect to the sound structure of each nonword. As listeners encountered stimuli with different collections of pointed versus rounded sounds, judgments of shape varied in a graded fashion. This finding suggests that listeners did not rely exclusively on specific sounds to make their judgments, but rather took into account the entire structure of each nonword stimulus (see Thompson & Estes, 2011).

Taken together, these findings are generally consistent with research on cross-modal associations found outside the realm of language. For example, in an implicit association task, Parise and Spence (2012) found that tones composed of square waves (and which have a noisy quality) were associated with more pointed visual objects as compared to tones composed of sinusoidal waves (which have a smoother tonal quality), which were associated with more rounded shapes. This similarity between findings across linguistic and non-linguistic contexts suggests that sound symbolic mappings in natural language may arise from more general tendencies to associate experiences across sensory-motor and perceptual domains (Namy & Nygaard, 2008).

The question remains, however, regarding how and why these mappings are made. One way in which labels such as *bouba* and *kiki,* and the particular speech sounds examined in our study, may be mapped to visual forms is on a basis of analogous intermodal structural properties, some being smoother and more continuous, and others being more disrupted, discontinuous, strident, or jarring. Such cross-modal mapping may be based in comparison or alignment of shared structural properties such as relative frequency (e.g. spatial, temporal), and alignable differences in the respective domains (Gentner & Markman, 1997; Marks, 1989).

It could also be the case that systematic sound-to-meaning mappings reflect statistical regularities of our experience in the physical world. For example, if pointed objects tend to produce less tonal sounds with more abrupt transitions, and amoeboid forms tend to produce more tonal and continuous sounds, these co-occurrences could be invoked by language. Similarly, Ramachandran and Hubbard (2001) hypothesized that the sound-to-shape mappings in the bouba-kiki effect result from synesthetic connections between perceptual, somatosensory, and motor areas thereby linking representations of speech sounds with the orofacial expressions made when articulating these sounds. On this account, the rounded shape of the lips when articulating an /o/ or /u/, for example, comes to be associated with the sound created.

It may be that individuals are predisposed to make certain inter-sensory or sensory-motor connections and these predispositions may underlie or give rise to the kinds of sound-to-meaning correspondences observed in our study. Maurer, Pathman, and Mondloch (2006) found that toddlers (mean age 2.8 years) exhibit the bouba-kiki effect suggesting that even young children are able to represent and form expectations about multimodal perceptual-motor couplings. This finding is consistent with the view that these are not necessarily learned mappings, but may reflect a more general multisensory integration process.

Regardless of how these mappings arise, if semantic representations are grounded in our sensory systems, sounds in language could serve to re-activate or simulate sensory information encoded by an individual as they experience an object or event (Barsalou, Santos, Simmons, & Wilson, 2008; Bergen & Feldman, 2008; Kita, 1997).

Further research is needed in order to establish whether any of the mappings we document in this study are idiosyncratic to speakers of English, and thus the extent to which such mappings might be a result of language-specific conventions for mapping sounds to meanings. However, the consistency between findings in the present study and research on crossmodal perceptual research outside the linguistic domain suggests the mappings observed in linguistic systems may be based on more general correspondences. However, the perceptual-cognitive basis of these cross-sensory correspondences is itself the subject of much debate.

Conclusion

Although natural language may be largely arbitrary, it is evident that certain sounds are non-arbitrarily associated with particular meanings. Although extensive research has documented the bouba-kiki effect across development and across cultures, the cognitive basis of these mappings is not well understood. By using a large, yet highly controlled set of pseudowords, the present study systematically examined how various factors related to sound structure of words contribute to individuals' sound to meaning mappings. By establishing how specific attributes of the sound structure of language are systematically mapped to meaning, language research will benefit from improved understanding of how sounds evoke meaningful representations. Understanding the cognitive basis of such mappings will ultimately provide insight into language, symbol use, conceptual representation, and cross-sensory cognitive processes.

References

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and Simulation in Conceptual Processing. In *Symbols, embodiment, and meaning* (pp. 245–283).

Bergen, B. K., & Feldman, J. (2008). It's the Body, Stupid: Concept Learning According to Cognitive Science, 1–18.

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). "Bouba" and "Kiki" in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition*, *126*(2), 165–72.

Davis, R. (1961). The fitness of names to drawings. A cross-cultural study in Tanganyika. *British Journal of Psychology*, *52*(3), 259–268.

Dingemanse, M. (2011). Mark Dingemanse, *6*(1), 77–85.

Fort, M., Martin, A., & Peperkamp, S. (2014). Consonants are More Important than Vowels in the Bouba-kiki Effect. *Language and Speech*, 122–140. doi:10.1177/0023830914534951

Gasser, M. (2004). The Origins of Arbitrariness in Language. *Science*.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45–56. doi:10.1037//0003-066X.52.1.45

Hirata, S., & Kita, S. (2011). Cross-modal correspondence between brightness and Chinese speech sound with aspiration. *I-Perception*, *2*(8), 788. doi:10.1068/ic788

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics*, *35*(2), 379–415.

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

Kunihira, S. (1971). Effects of the expressive voice on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior*, *10*(4), 427–429.

Magnus, M. (2001). *What's in a Word? Studies in Phonosemantics*.

Marks, L. E. (1989). On cross-modal similarity: the perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology. Human Perception and Performance*, *15*(3), 586–602.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas : sound – shape correspondences in toddlers and adults. *Developmental Science*, *3*, 316–322.

Monaghan, P., & Christiansen, M. H. (2005). Why Form-Meaning Mappings are not Entirely Arbitrary in Language, (1848), 1838–1843.

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*(3), 325–47.

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective & Behavioral Neuroscience*, *4*(2), 133–6. doi:10.3758/cabn.4.2.133

Namy, L. L., & Nygaard, L. C. (2008). Perceptual-motor constraints on sound to meaning correspondence in language. *Behavioral and Brain Sciences*, *31*, 528–529.

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.

Nuckolls, J. B. (1999). *The Case for Sound Symbolism*. *Reactions*.

Parise, C., & Spence, C. (2009). "When birds of a feather flock together": Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PloS One*, *4*(5), e5664. doi:10.1371/journal.pone.0005664

Parise, C. V, & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. *Experimental Brain Research*, *220*(3-4), 319–33. doi:10.1007/s00221-012-3140-6

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Ramachandran, V. S., & Hubbard, E. M. (2003). The Phenomenonlogy of Synaesthesia. *Journal of Consciousness Studies*, *10*(8), 49–57.

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

Saussure, F. De. (1916). Course in General Linguistics.

Tarte, R. D. (1974). Phonetic symbolism in adult native speakers of Czech. *Language and Speech*, *17*(1), 87–94.

Tarte, R. D., & Barritt, L. S. (1971). Phonetic symbolism in adult native speakers of English: Three studies. *Language and Speech*, *14*(2), 158–168.

Thompson, P. D., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *Experimental Psychology*, (October), 37–41. doi:10.1080/17470218.2011.605898

Ultan, R. (1978). Size sound-symbolism. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of Human Language, Volume 2: Phonology.* (Vol. 2: Phonolo, pp. 527– 568). Stanford, CA: Stanford University Press.

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