The Bouba-Kiki Phenomenon Tested via Schematic Drawings of Facial Expressions: Further Validation of the Internal Simulation Hypothesis

Sethu Karthikeyan and Bianca Rammairone
Communication Sciences and Disorders Program, Pace University, NY, USA

Vijayachandra Ramachandra
Department of Communication Sciences and Disorders, Marywood University, Scranton, PA, USA

Abstract
Sound-shape associations involving consistent matching of nonsense words such as ‘bouba’ and ‘kiki’ with curved and angular shapes, respectively, have been replicated in several studies. The purpose of the current study was to examine the robustness of previously noted sound-shape associations when shape variations (angular and curvy) are embedded in schematic expressions of emotions (sad and happy). Results revealed consistent matching tendencies based on sound-emotion expression mapping irrespective of the particular shape of the expressions. We suggest that internally simulating the facial expressions/oral gestures may have played a significant role in driving the matching preferences.

Keywords
bouba-kiki, sound-shape mapping, cross-modal activation, mirror neurons

Recent studies that reveal consistency in different kinds of sound-meaning associations, and the increasing scope of multimodal activations in the brain raise questions about the extent of arbitrariness in language (e.g., see review in Schmidtke, Conrad, & Jacobs, 2014). When angular and rounded shapes are presented with nonsense words such as bouba and kiki, 95% of adults match the rounded shape with bouba and the jagged shape with kiki—the bouba-kiki phenomenon (Ramachandran & Hubbard, 2001; for original work, Corresponding author:
Sethu Karthikeyan, Communication Sciences and Disorders Program, Pace University, 163 William Street, NY 10038, USA. Email: skarthikeyan@pace.edu

Creative Commons CC-BY: This article is distributed under the terms of the Creative Commons Attribution 3.0 License (http://www.creativecommons.org/licenses/by/3.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (http://www.us.sagepub.com/aboutus/openaccess.htm).
see Köhler, 1929), an effect that has been replicated cross-linguistically (e.g., Bremner et al., 2012). As well, toddlers and infants display similar matching tendencies to those of adults (Maurer, Pathman, & Mondloch, 2006; Ozturk, Krehm, & Vouloumanos, 2012) suggesting that the matching biases may be innate and not learned via experience. Whereas a few studies have found consonant-driven matching patterns (e.g., Nielsen & Rendall, 2013), Spector and Maurer (2013) found that even when the consonant environment was kept constant, toddlers demonstrated consistent vowel-shape matches of /i/ (as in beet) and /o/ (as in boat) with angular and curvy images, respectively.

Consider also that the articulatory gestures for /i/ and /o/ are similar to the lip movements in a smile and a frown (vocal tract shortening and lengthening) leading to comparable acoustic characteristics—raising and lowering of filtered frequency components referred to as formants (e.g., Raphael, Borden, & Harris, 2011). Listeners are able to accurately identify speech samples spoken with a smile and those spoken with a frown (Tartter & Braun, 1994). Additionally, the vowels /i/ and /o/ have been associated with pleasantness and gloominess, respectively (e.g., Newman, 1933). These findings point to the possibility that nonsense words with vowel sounds /i/ and /o/, as those mentioned earlier, may be non-arbitrarily linked to smile and frown expressions. The purpose of the current study was to examine the robustness of previously noted sound-shape associations when shape variations were embedded in schematic expressions of emotions (see Figure 1).

Ramachandran and Hubbard (2001) proposed a mirror neuron-based cross-modal activation hypothesis to explain the bouba-kiki effect; the internal simulation of the appropriate articulatory gesture of the auditory stimulus is mapped onto specific phonemic inflections, which then are non-arbitrarily linked with specific shapes or images. In the current study, while the facial expression of a smile or a frown may be internally simulated in addition to the motor patterns of the aurally presented words, the exact shape of the expressions may not be amenable to simulation because angularity variations such as the ones examined in the study are not naturally occurring or socially relevant (e.g., Oberman & Ramachandran, 2007). The resulting matching tendencies, therefore, must reflect consistencies between the corresponding acoustic characteristics of the emotional expression and of the auditory stimulus. For example, a sad angular/curvy face matched with words containing the /o/ vowel sound. If, on the other hand, simulations of facial expressions of emotions do not occur, the variations of the facial expressions in the angularity curviness dimension must influence the matching preferences in a similar manner to the established matching tendencies between sounds and random angular-curvy shapes in earlier studies. For example, a sad/happy angular face matched with words containing the /i/ vowel sound.

Each of the four computerized schematic drawings of faces was presented individually three or four times with randomly chosen word pairs (see Table 1). The aural presentation of the word pairs was counterbalanced across the four faces for the order of /i/ and /o/ vowel sounds and for the consonant environment. The institutional review board approved the study; 50 participants between 18 to 24 years (45 females and 5 males) were asked to match the happy/sad face with one of the words in the presented word pair by clicking on the appropriate text options: WORD 1 or WORD 2.

As Figure 2 shows, the happy angular face and the happy curve face was each matched more frequently with words containing the /i/ vowel sound than with words containing the /o/ vowel sound. The preference for the /i/ vowel sound did not differ between these two faces (paired t test, $t_{(49)} = 1.59, p > .05$). The sad curve face and the sad angular face was each matched more frequently with words containing the /o/ vowel sound. The preference for the /o/ vowel sound did not differ between these two faces (paired t test, $t_{(49)} = 0.35, p > .05$).
Table 1. Word Pairs Created Using Cepstral David (Swifttalker) Separated by 500 Milliseconds.

| Consonant environment 1 | Consonant environment 2 |
|-------------------------|-------------------------|
| bibi-bobo               | fifi-fofo               |
| bobo-bibi               | fofo-fifi               |
| didi-dodo               | kiki-koko               |
| dodo-didi               | koko-kiki               |
| lili-lolo               | titi-toto               |
| lolo-lili               | toto-titi               |
| mimi-momo               | zizi-zozo               |
| momo-mimi               | zozo-zizi               |

Figure 1. Schematic drawings presented via SuperLab 5; (a) happy angular, (b) sad angular, (c) happy curve, and (d) sad curve.
If the acoustic characteristics associated with the exact shape of the facial expressions were being mapped onto the acoustic characteristics of the words presented aurally, the shape differences within each emotional category (curved vs. angular versions of happy and sad faces) would have resulted in a notable difference in the choices of vowel sounds; however, a difference did not emerge. We cannot rule out the potential influence of emotional contagion (e.g., Lundqvist & Dimberg, 1995), in that the happy “feeling” on seeing a happy face may have been mapped onto the “pleasant” sounding /i/ and the sad feeling on seeing a sad face was matched with the gloomy sounding /o/ (e.g., Newman, 1933).

In sum, the current study in conjunction with the previous ones on sound-symbolism demonstrate that cross-modal matching may allow for non-arbitrary associations of the vocal–verbal signal with aspects of inanimate and animate entities (including the self and others). When two or more aspects co-occur, these may compete with one another or one may supersede the others in guiding the associations; in the current study, the facial expressions took precedence over the angularity or curviness of the expressions. Also, there exists the possibility of covert imitation of the oral gestures alone (retraction of lips and rounding or protrusion of lips), devoid of any emotional meaning, to aid this kind of auditory visual matching task (see Studdert-Kennedy, 2000, 2002, for a discussion on the evolution of vocal imitation as a step toward promoting arbitrary linkages between signals and messages or referents). Considering the larger scheme of things, it is reasonable to deduce that a communication system that is based heavily on non-arbitrariness or iconicity could lead to ambiguity (e.g., Pinker & Bloom, 1990), and limited expressive scope (e.g., Studdert-Kennedy, 2000, 2002), a few possible factors that may have favored arbitrariness in the evolution of the human capacity for language.

Acknowledgment

The authors wish to thank Hisham Abboud (SuperLab) for help with setting up the experiment.
Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The project was supported by Pace University College of Health Professions New Faculty Development funds given to the first author.

References

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2012). “Bouba” and “Kiki” in Namibia? A remote culture makes similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition, 122*, 80–85.

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

Lundqvist, L. O., & Dimberg, U. (1995). Facial expressions are contagious. *International Journal of Psychophysiology, 9*, 203–211.

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

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

Nielsen, A. K., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology, 67*, 153–163.

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin, 133*, 310–327.

Ozturk, O., Krehm, M., & Vouloumanos, A. (2012). Sound symbolism in infancy: Evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology, 114*, 173–186.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707–784.

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

Raphael, L. J., Borden, G. J., & Harris, K. S. (2011). *Speech Science Primer*. Philadelphia, PA: Lippincott Williams & Wilkins.

Schmidtke, D. S., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. *Frontiers in Psychology, 5*, 1–6.

Spector, F., & Maurer, D. (2013). Early sound symbolism for vowel sounds. *i-Perception, 4*, 239–241.

Studdert-Kennedy, M. (2000). Evolutionary implications of the particulate principle: Imitation and the dissociation of phonetic form from semantic function. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language* (pp. 161–176). Cambridge, England: Cambridge University Press.

Studdert-Kennedy, M. (2002). Mirror neurons, vocal imitation and the evolution of particulate speech. In M. Stamenov, & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 207–227). Amsterdam, The Netherlands: John Benjamins.

Tartter, V. C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *The Journal of the Acoustical Society of America, 96*, 2101–2107.
Author Biographies

Sethu Karthikeyan is an assistant professor in the Communication Sciences and Disorders program, Pace University. Sethu is interested in examining speech-based social evaluations and language development from an evolutionary perspective.

Bianca Rammairone graduated summa cum laude from Pace University’s Pforzheimer Honors College with a BA in Communication Sciences and Disorders and a minor in Psychology, as well as NY State Certification for Teaching Students with Speech and Language Disabilities. She is currently pursuing an MS in Communicative Sciences and Disorders at NYU Steinhardt’s graduate school. Bianca will be continuing her applied research studies abroad this summer in the Department of Logopedics and Phoniatrics at the Medical School of Lund University in Sweden.

Vijayachandra Ramachandra is an associate professor in the Department of Communication Sciences and Disorders, Marywood University. Vijay conducts research in the area of language and cognition, and is interested in exploring the relationship between working memory and novel word learning in children and adults, theory of mind and language in people with brain damage, nonverbal emotions, and linguistic and cognitive aspects of synesthesia.