Proceedings of the 4th International Conference on Literature, Language, and Culture Development
DOI: 10.54254/2753-7064/2025.23309

Evidence to the Ubiquity of Sound Symbolism - Reviewing Recent Experimental Studies on Crossmodal Sound-meaning Association

Yueshan You

School of Foreign Studies, Nankai University, Tianjin, China
2210335@mail.nankai.edu.cn

Abstract: Sound symbolism, the study of the direct relationship between sound and meaning, challenges the traditional notion of the arbitrariness of linguistic symbols. This paper reviews recent experimental research, particularly studies utilizing nonce words, to explore the mechanisms and universality of sound-meaning associations and discuss the thematic and methodological innovations and restraints. Our synthetic review highlights the recently-discussed aspects including the robustness and ubiquity, the inner mechanism and potential undermining factors of the well-known bouba/kiki effect, suggesting that sound symbolism is deeply rooted in human cognition, although different methodological realizations may lead to contradicting findings. In light of the importance of sound symbolism in language evolution, acquisition, and use, this research underscores the crucial role of relevant bouba/kiki experiments as probe, which can offer a cross-linguistic perspective applicable to a wide range of languages and speakers.

Keywords: Sound symbolism, iconicity, sound-meaning correspondence, segmental feature, orthography

1. Introduction

Last century the relation between form and meaning of linguistic symbols was widely assumed to be arbitrary, a principle asserted that phonemes and sound features of a word are essentially meaningless apart from their phonemic representations and morphemes they are part of [1]. For example, there is no logical reason why the word "dog" in English should refer to the four-legged hairy animal it denotes while that signified object correspond to ‘Hund’ in Deutsch, hence the association being purely conventional.

However, growing research on sound symbolism challenges this notion, suggesting a direct correspondence between sound and meaning. The growing literature has depicted a diverse perspective on a synthesis of domains, including linguistics, psychology, anthropology, etc.

Sound symbolism has been used to cover myriad concepts, including onomatopoeia, ideophones, phonaesthesia, phonosemantics, phonosyntactic restrictions, and most importantly, the iconicity in speech [2-6]. Iconicity refers to a direct form-meaning association acrooss modalities, which is manifested by the sound-meaning association when comes to speech [2]. Recent studies highlight the role of iconicity in language use, evolution, and acquisition, extending beyond mimetic words to basic vocabulary [7,8]. This review aims to synthesize recent studies on the sound-meaning

DOI: 10.54254/2753-7064/2025.23309

association, focusing on experiments using nonce words to observe the iconic phenomenon. The primary question is not what sound symbolism should cover, but what has been mentioned and attested in relevant experiments under the broader notion of sound symbolism. The study primarily reviews experimental quantitative research, particularly those using nonce words, to understand the non-arbitrary form-meaning associations. These experiments observe the iconic phenomenon between sound and meaning, providing empirical evidence across different linguistic domains such as size and spatial deixis. This study offers a broader, cross-linguistic perspective applicable to a wider range of languages and speakers. It underscores the pivotal role of iconicity in language, suggesting it operates as a cognitive scaffold enabling intuitive form-meaning connections.

2. The bouba/kiki effect and its robustness across languages

Iconicity in speech has been tested through nonce words as a common practice [2]. Investigations into the mental reflections associated with specific segmental features of language date back to the earliest epochs of linguistic history. In 1929, Sapir conducted a classic experiment in which participants were presented with a pair of nonse words including mil and mal [9]. The result showed that English speakers tend to associate larger sized table with mil and similar pseudowords containing high front vowels, while matching smaller sized objects with mal and similar words containing low back vowels. The findings prompted diverse discussions regarding the effects attributed to the substitution of the nucleus vowel. Phonetically, the high front vowel /i/ in ‘mil’ possesses a higher fundamental frequency and elevated F2 frequency, potentially evoking the perception of small-sized objects. This impression is grounded in the premise that smaller animals and objects tend to produce higher-frequency sounds compared to larger ones, aligning with the articulatory characteristics of the /i/ vowel [10]. Another interpretation interlinked the image of a fully-opened mouth when pronouncing the low back vowel/a/ with the image of larger objects [11]. Here the correspondence between sound and images, as A. Ćwiek suggested, is not direct, but is mediated by the imagined visual impression of making such sound, hence indicated that the iconicity in spoken languages doesn’t simply come from resemblance between sounds and auditory impressions but also resemblances between speech sounds and other sensory impressions that are mediated through crossmodal correspondences [2].

Iconic phenomenon in speech has been largely observed by utilizing pseudowords (a kind of non-lexical vocable) as stimuli in experimental studies. The bouba/kiki effect is a well-discussed phenomenon that highlights the crossmodal associations between speech sounds and visual shapes, illustrating a form of iconicity in language [2]. This effect, widely studied in the realm of psycholinguistics, demonstrates that people tend to associate the pseudoword "bouba" with round shapes and "kiki" with spiky shapes [12]. This association is believed to stem from the inherent qualities of the sounds themselves, where the smooth, rounded sounds of "bouba" contrast with the sharp, angular sounds of "kiki." First identified in the early 20th century, the bouba/kiki effect suggests that there are natural constraints on how sounds can be mapped to meanings, potentially offering insights into the origins and evolution of language. Despite its intriguing implications, it was not until recent years that several untested variables, which may contribute to the bouba/kiki effect beyond sound-meaning correspondence, had been investigated [6,13,14].

In 2022, A. Ćwiek raised a possible disturbing factor, which is the spelling system most of the bouba/kiki experiments used. Whilst it is plausible to deduce that the bouba/kiki distinction encompasses variations in vowel formants, vowel-intrinsic fundamental frequency, consonant-induced fundamental frequency perturbation, duration, consonant voicing, voice onset time, vowel rounding, and place of articulation—all of which potentially influence the bouba/kiki effect to varying degrees—the orthographic shape of these words might have perplexed participants.

DOI: 10.54254/2753-7064/2025.23309

This is particularly relevant given that the majority of experiments were conducted with Western participants.

To address this concern, Ćwiek's study took a wider range with a diverse sample of speakers from 25 different languages and 9 language families that use 10 different scripts. To measure the effect across languages, this study fitted a Bayesian logistic regression model. To better demonstrate the cross-linguistic consistency of the bouba/kiki effect, the required matching task was presented in sequential order, and responses were counted as congruent only when participants produced corresponding answers for both 'bouba' and 'kiki' trials. The sequential order of the bouba/kiki matching tasks and the scripts used by participants, which exhibited a binary effect based on whether a language uses the Roman alphabet, were therefore labeled as fixed effects.

The result reported an overall 72% congruent response, which indicated that the cross-cultural effect of sound-meaning iconicity is robust, while languages that predominantly used Roman orthography showed a stronger effect than others. Nonetheless, the biasing effect of orthography was reported to be weak in general. Participants using non-Romanized scripts in their L1 still produced an above-chance probability of congruent matching, a finding that supports the bouba/kiki effect's independence from orthography.

3. The inner mechanism of bouba/kiki effect

As mentioned above, whether such crossmodal correspondence comes from the direct linkage between sound and the presented visual images, or it is essentially mediated by the association between sounds and their imagined mouth shapes is the key question in the relevant studies. More recently, Passi and Arun examined the effectiveness of the bouba/kiki effect in unpronounceable sounds and illuminated the whirling debate. In their serial experiments, mouth-shape hypothesis was tested after a confirmatory analysis examining the validity of the original bouba/kiki effect in the sampled population[15]. By playing each word backward in time, reversed words with identical phonetic properties such as the frequency content were created. Real object sounds were also incorporated to create a mixture with normal and reversed sounds. The real object sounds comprised low- and high-frequency sounds made by real-world objects when they were struck (e.g., pillow vs. metal objects). Since reversed sounds and real object sounds are hard to pronounce, the sensory visual images are therefore assumed to be nonexistent. If the results had demonstrated a significant difference from those of the previous normal group, the mouth-shape hypothesis would thereby have been validated.

The results indicated that the bouba/kiki effect remained robust even when applied to unpronounceable sounds, suggesting that pronounceability and associated mouth shape imagery are not essential prerequisites for this phenomenon. This finding contradicts the mouth-shape hypothesis. Additionally, the authors observed that spectral properties of sounds, such as mean frequencies, are indicative of the effect's strength, aligning with the hypothesized direct correspondence between sound and shape. Although the effect observed in unpronounceable words could be initially learned through the resemblance between mouth shape and sounds, the author considered it unlikely but with no further explanation. Other potential explanations for the bouba/kiki effect include articulatory properties such as tongue movements and the shape of the oral cavity, as well as shared neural properties, such as larger neural responses to angular shapes and high-frequency sounds. Passi and Arun have argued that further experiments are needed to characterize the underlying neural representations or to test these phenomena across different species. This study did not establish an exact source for this association, as the authors proposed a plausible alternative explanation: the mental reflection evoked by specific-shaped objects that resemble our acoustic perception may be material-dependent. For instance, a wooden object with spike-like features still produced low-frequency sounds, whereas a metallic object with sharp features produced higher-frequency

DOI: 10.54254/2753-7064/2025.23309

sounds, corresponding to high front vowels like ‘kiki’. This speculation was supported by contemporary research findings, which suggested an association between visual spatial frequencies and auditory temporal frequencies, possibly mediated through the common modality of touch [16, 17].

4. Bouba/kiki effect: the cross-modal nature beyond sound-shape association and the impact of orthography

By far we have gathered that the bouba/kiki effect is seemingly robust across cultures and languages, and the mechanism is largely attributed to the sound properties as the mean frequency but not the related articulatory properties as the mouth shape. Just as discussed above, the high front vowel /i/ was believed to share a phonetic resemblance to spiky objects in terms of the frequency, and the low back vowel /a/ vice versa. But what about the consonants? The effects of consonants and vowels on sound perception have been investigated both jointly and separately. An open question remains: if vowels are kept constant, will changes in consonants produce similar effects?

Furthermore, given that the crossmodal correspondence between sound and visual images, such as shapes, has been extensively examined, it is worth exploring whether other modalities also correlate with our perception of sound. To investigate these questions, Lo conducted a quasi-bouba/kiki experiment in 2017 with 45 undergraduate participants at a Hong Kong university. The stimuli consisted of Cantonese character strings, adhering to a fixed pattern of vowel distribution and lexical tones. Within each pair, both labels shared identical vowels and tones, with the only difference being the plosive or fricative consonants. An example was given by the author: ‘翻慳山’ was pronounced as /faan1/ /haan1/ /saan1/, while the characters in the other label contained initial plosive consonants, ‘班攀灘’was pronounced as /baan1/ /paan1/ /taan1/. The two sections of the experiment investigated both the visual sound-meaning iconicity and the textile sound-meaning correspondence. As the author hypothesized, since the turbulence of the airflow caused by the articulation of fricatives had been associated with the spiky drawing, the drawing may simultaneously connote a roughness feeling like touching a hard rock.

To our surprise, the sound-shape correspondence wasn’t significantly observed when initial fricative/plosive consonants were tested, while the second experiment with the same auditory stimuli led to a strong preference to match the initial fricative nonce words to rough texture material and vice versa, among participants. In other words, the sound-symbolic crossmodal correspondence on tactile stimuli was supported.

Several considerations should be addressed to explain these results. The weak sound-shape correspondence, which contradicts prevailing observations, may be attributed to the distinctive experimental design. Notably, the experiment was rooted in a Cantonese linguistic context, as all participants were recruited from a single university in Hong Kong. Therefore, it is reasonable to assume, while acknowledging the potential for sampling bias, that participants possessed at least a basic level of familiarity with Cantonese, both phonologically and semiotically, even though their specific linguistic backgrounds were not detailed. The stimuli consisted of Cantonese morphemes. Although these were combined into nonce phrases representing meaningless content, participants were likely to recognize the heard phonemes and thereby associate the acoustic features of those phonemes with the imagined morphemes they represent. This possibility was further supported and perplexed by the special tonal system of Cantonese. Although all the stimuli were consistent in high-level tones(marked in number ‘1’), the recognizable tonal pitch and length of articulation could not only inform the participants with indexed Cantonese symbols, which may not be identical to the morphemes put in the stimuli but varied individually, but also add new ingredient to the current

DOI: 10.54254/2753-7064/2025.23309

analysis as the contradicted finding may result from the interrupting tonal information and related high articulatory frequency [18].

Secondly, as the author discussed, the control of vowels may have an unprecedented impact on the bouba/kiki effect. This is because the segmental features of vowels and consonants, which may contribute to this effect, have previously been examined both separately and jointly without controlling for unwanted variables [13]. In contrast, the stimuli prepared for Lo's experiment imposed a strict constraint on vowels. The author argues that the weaker effect observed with consonants is consistent with previous findings, which suggest that imposed segmental effects are more significant in vowels than in consonants.

The third consideration reopens the debate on the influence of orthography. The author cites a reported lower phonological awareness among Cantonese speakers, suggesting a potential reliance on a more pronounced association between sound and visual concepts [19]. It is crucial to acknowledge that the linguistic backgrounds of the selected participants were not disclosed; however, given their status as undergraduates in Hong Kong, it is reasonable to assume their proficiency in the Cantonese writing and speaking system. Consequently, their mental representations of heard sounds—phonemes that represent morphemes in Cantonese, characterized by tones—may inevitably align with Cantonese written symbols. In the initial experiment discussed, the robustness of such crossmodal correspondence was observed independently of orthographic influence. Nevertheless, the findings may not be definitive, as Cantonese speakers were not included in the study. Furthermore, the question of whether distinct phoneme-represented sounds heard by speakers of a particular language are influenced by written scripts warrants further investigation.

Although Lo's experiment did not establish a direct sound-shape correspondence, it revealed that the manner of articulation of consonants evokes different tactile perceptions. This suggests that the iconicity in speech may extend beyond auditory and visual domains to resemble meanings from other modalities. Specifically, fricative-only three-morpheme combinations were found to strongly correspond with rough-textured material, whereas those with plosives were more likely to be associated with smooth-textured material.

5. Discussion

Relying on observation from sound-meaning mapping tasks, we found that iconicity in speech has been examined in a variety of methodologies, stimuli, and sampled populations. While the bouba/kiki effect has become a cornerstone of research on sound-meaning iconicity, recent experimental advances reveal both its robustness and its limitations, offering critical insights into the mechanisms underlying such crossmodal associations.

The bouba/kiki effect's persistence across diverse linguistic and orthographic systems supports the existence of universal, biologically grounded sound-shape mappings [2]. This aligns with theories positing that acoustic properties—such as spectral frequency and sharpness—directly evoke perceptual analogies (e.g., high-frequency /i/ evoking spikiness) rather than relying solely on articulatory simulations like mouth shape [15]. Such results strengthen the argument for *natural iconicity*, where sound-meaning links arise from innate perceptual biases rather than cultural convention. However, the weaker bouba/kiki effect in non-Roman script users suggests that orthography and linguistic familiarity may modulate, though not fully determine, these associations [2]. The potential impact of orthography on phonological awareness, as Lo accounted for in his experiment's failure to observe a stable sound-shape correspondence, may stem from the use of different stimuli settings. Specifically, the use of Cantonese single-morpheme combinations as auditory material for undergraduates, who were likely to be pre-familiar with Cantonese, could have influenced the results. As Cantonese has no direct linkage from grapheme to morpheme, the author explained that Cantonese speakers may need more direct signs in order to relate to other modalities.

DOI: 10.54254/2753-7064/2025.23309

Contrary to the bouba/kiki paradigm, Lo's study highlights the nuanced role of consonants in crossmodal iconicity [13]. While consonant-driven sound-shape mappings proved weak, fricatives (e.g., /s/, /f/) elicited strong tactile associations with roughness, likely due to their turbulent acoustic profiles. This divergence underscores the multidimensionality of sound symbolism: vowels may dominate in visual-shape mappings due to their spectral salience, whereas consonants, with their articulatory complexity, may better encode tactile or textural qualities. Such findings align with Jakobson's emphasis on phonetic symbolism, where articulatory or acoustic features (e.g., voicing, turbulence) map to sensory experiences [20].

The reviewed literature has also pointed out possible unreconciled methodologies and further influential factors to be examined to state a more precise sound-meaning correspondence. The failure to replicate sound-shape iconicity in Lo's Cantonese-based experiment raises critical questions about methodological and linguistic variables [13]. First, the use of tonal morphemes in Cantonese may have introduced competing perceptual cues (e.g., pitch contours), diluting the effect of segmental features. Second, participants' familiarity with Cantonese phonology and orthography—where characters often encode semantic radicals—might have primed them to interpret pseudowords through lexical rather than crossmodal associations. This contrasts with Ćwiek et al.'s cross-linguistic design, which minimized lexical interference by testing speakers unfamiliar with the stimuli's script [2].

Furthermore, Passi and Arun's demonstration of the bouba/kiki effect in unpronounceable sounds challenges the articulatory hypothesis, favoring instead a direct acoustic-perceptual linkage [15]. However, their reliance on spectral frequency as an explanatory variable does not fully account for the role of learned associations (e.g., cultural exposure to high-pitched sounds from sharp objects). This tension between innate and learned mechanisms remains unresolved, echoing broader debates in sound symbolism research.

6. Conclusion

Through a review of relevant literature, we have elucidated recent groundbreaking findings pertaining to the mechanism underlying the bouba/kiki effect, a crossmodal correspondence between sound and meaning that substantiates the presence of sound symbolism. The robustness and universality of iconicity in speech across languages have been examined; however, orthography may influence the direct sound-shape resemblance, as evidenced by Ćwiek's study, which encompassed a diverse range of participants with varying linguistic backgrounds, and by Lo's findings. Furthermore, Passi and Arun's research has principally investigated the prime-driven motivation behind the observed sound-shape correspondence, wherein high front vowels are associated with spiky shapes and low back vowels with rounded shapes. The articulatory and acoustic features of sounds, such as those in the bouba/kiki effect, have been examined as the intrinsic mechanism, suggesting a direct linkage between sound and shape rather than an association mediated by the imagined mouth shape of such pronunciations. Nevertheless, potential methodological limitations and variability in sampling have left the discussion incomplete, necessitating further investigation in this area. This is particularly important given that iconicity in speech, or sound symbolism, is believed to play an essential role in shaping the languages we use today.

References

[1] *Saussure FD (1916) Cours de Linguistique Générale, eds Bally C, Sechehaye A, Riedlienger A (Payot, Paris).*

[2] *Ćwiek, A., Fuchs, S., Draxler, C., Asu, E.L., Dediu, D., Hiovain, K., Kawahara, S., Koutalidis, S., Krifka, M., Lippus, P., Lupyan, G., Oh, G.E., Paul, J., Petrone, C., Ridouane, R., Reiter, S., Schümchen, N., Szalontai, Á., Ünal-Logačev, Ö., Zeller, J., Perlman, M., & Winter, B. (2021). The bouba/kiki effect is robust across cultures and writing systems. Philosophical Transactions of the Royal Society B: Biological Sciences, 377.*

DOI: 10.54254/2753-7064/2025.23309

Blake, B. J. (2017). Sound Symbolism in English: Weighing the Evidence. AUSTRALIAN JOURNAL OF LIN GUISTICS, 37(3), 286–313. https://doi.org/10.1080/07268602.2017.1298394

[3] *Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. Language, 92(2), e117–e133. ht tps://doi.org/10.1353/lan.2016.0034*

[4] *Feist, J. (2013). "Sound symbolism" in English. Journal of Pragmatics, 45(1), 104–118. https://doi.org/10.10 16/j.pragma.2012.10.008*

[5] *Johansson, N. E., Anikin, A., Carling, G., & Holmer, A. (2020). The typology of sound symbolism: Defining macro-concepts via their semantic and phonetic features. Linguistic Typology, 24(2), 253–310. https://doi.org /10.1515/lingty-2020-2034*

[6] *Joo, I. (2020). Phonosemantic biases found in Leipzig-Jakarta lists of 66 languages. Linguistic Typology, 24(1), 1–12. https://doi.org/10.1515/lingty-2019-0030*

[7] *Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. Cognition, 109(1), 54–65. https://doi.org/10.1016/j.cognition.2008.07.015*

[8] *Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound–shape correspondences in toddlers and adults. Developmental Science, 9(3), 316–322. https://doi.org/10.1111/j.1467-7687.2006.00495.x*

[9] *Sapir, E. (1929). A study in phonetic symbolism. Journal of Experimental Psychology, 12(3), 225–239. https://doi.org/10.1037/h0070931*

[10] *Ohala, J. J. (1983). Cross-Language Use of Pitch: An Ethological View. Phonetica, 40(1), 1–18. https://doi. org/10.1159/000261678*

[11] *Tarte, R. D., & Barritt, L. S. (1971). Phonetic symbolism in adult native speakers of English: Three studies. Language and Speech, 14(2), 158–168. https://doi.org/10.1177/002383097101400206*

[12] *Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. Proceedings of the National Academy of Sciences, 113(39), 10818–10823. https://doi.org/10.1073/pnas.1605782113*

[13] *Lo, L.-Y., Luk, H. M., & Thompson, N. (2017). An Investigation of Sound-Symbolism in the Context of Tactile Feeling. The Journal of General Psychology, 144(2), 157–167. https://doi.org/10.1080/00221309.2016.1276047*

[14] *Winter, B., & Perlman, M. (2021). Size sound symbolism in the English lexicon. Glossa: A Journal of General Linguistics, 6(1), Article 1. https://doi.org/10.5334/gjgl.1646*

[15] *Passi, A., & Arun, S. P. (2024). The Bouba-Kiki effect is predicted by sound properties but not speech properties. Attention, Perception & Psychophysics, 86(3), 976–990. https://doi.org/10.3758/s13414-022-02619-8*

[16] *Fryer, L., Freeman, J., & Pring, L. (2014). Touching words is not enough: How visual experience influence s haptic–auditory associations in the "Bouba–Kiki" effect. Cognition, 132(2), 164–173. https://doi.org/10.101 6/j.cognition.2014.03.015*

[17] *Hamilton-Fletcher, G., Pisanski, K., Reby, D., Stefańczyk, M., Ward, J., & Sorokowska, A. (2018). The role of visual experience in the emergence of cross-modal correspondences. Cognition, 175, 114–121. https://doi. org/10.1016/j.cognition.2018.02.023*

[18] *De Carolis, L., Marsico, E., & Coupé, C. (2017). Evolutionary roots of sound symbolism. Association tasks of animal properties with phonetic features. Language & Communication, 54, 21–35. https://doi.org/10.1016/ j.langcom.2016.10.003*

[19] *Yeung, P. S., Ho, S. H. C., Chik, P. M. P., Lo, L. Y., Luan, H., Chan, W. O. D., & Chung, K. K. H. (2011). Reading and spelling Chinese among beginning readers: What skills make a difference? Scientific Studies of Reading, 15(4), 285–313. https://doi.org/10.1080/10888438.2010.482149*

[20] *Jakobson, R., & Waugh, L. R. (2011). The Sound Shape of Language. De Gruyter Mouton. https://doi.org/10. 1515/9783110889451*