JASA ARTICLE

Does sound symbolism need sound?: The role of articulatory movement in detecting iconicity between sound and meaning[a)]

Mutsumi Imai,[1,b)] Sotaro Kita,[2] Kimi Akita,[3] Noburo Saji,[4] Masato Ohba,[5] and Miki Namatame[6]
[1]*Faculty of Environment and Information Studies, Keio University, Fujisawa, Kanagawa 252-0882, Japan*
[2]*Department of Psychology, University of Warwick, Warwick, Coventry CV4 7AL, United Kingdom*
[3]*Department of English Linguistics, Nagoya University, Nagoya, Aichi 464-8601, Japan*
[4]*Faculty of Human Sciences, Waseda University, Tokorozawa, Saitama 359-1192, Japan*
[5]*Graduate School of Media and Governance, Keio University, Fujisawa, Kanagawa 252-0882, Japan*
[6]*Department of Apparel and Space Design, Kyoto Women's University, Kyoto, Kyoto 605-8501, Japan*

ABSTRACT:
Ever since de Saussure [*Course in General Lingustics* (Columbia University Press, 1916)], theorists of language have assumed that the relation between form and meaning of words is arbitrary. However, recently, a body of empirical research has established that language is embodied and contains iconicity. Sound symbolism, an intrinsic link language users perceive between word sound and properties of referents, is a representative example of iconicity in language and has offered profound insights into theories of language pertaining to language processing, language acquisition, and evolution. However, on what basis people detect iconicity between sound and meaning has not yet been made clear. One way to address this question is to ask whether one needs to be able to hear sound to detect sound symbolism. Here, it is shown that (1) deaf-and-Hard-of-Hearing (DHH) participants, even those with profound hearing loss, could judge the sound symbolic match between shapes and words at the same level of accuracy as hearing participants do; and (2) restriction of articulatory movements negatively affects DHH individuals' judgments. The results provided support for the articulatory theory of sound symbolism and lead to a possibility that linguistic symbols may have emerged through iconic mappings across different sensory modality—in particular, oral gesture and sensory experience of the world in the case of speech.
 https://doi.org/10.1121/10.0034832

(Received 27 February 2024; revised 4 December 2024; accepted 6 December 2024; published online 10 January 2025)
[Editor: Susanne Fuchs] Pages: 137–148

I. INTRODUCTION

Where does the sound-meaning relationship of a word come from? In a long tradition of linguistics, it has been assumed that language-specific conventions create arbitrary association between linguistic forms and meanings (Hockett, 1960; de Saussure, 1916). However, words that sound like what they mean have also been recognized. For example, people associate novel words with high front vowels, such as /i/ (e.g., "mil"), with a small object, and words with low back vowels (e.g., "mal") with a large object (Sapir, 1929). People also associate words such as "maluma" and "bouba" with rounded objects and words such as "takete" and "kiki" with spiky objects (Köhler, 1929; Imai *et al.*, 2015). This shape sound symbolism is recognized by adult speakers of a broad range of languages (Cuskley and Kirby, 2013; Spence, 2011). In traditional linguistics, these intrinsic form-meaning relationships are considered to be a "peripheral" feature of language (Newmeyer, 1992).

However, recent psychological and neuroscientific research has established that iconicity, which is perceived form-meaning resemblance in language that underlies many cases of sound symbolism, has a broad impact on language (Perniss and Vigliocco, 2014; Imai and Kita, 2014). Sound symbolic words make language processing more efficient (Perniss *et al.*, 2010). Infants who have not begun active word learning could recognize shape sound symbolism (Asano *et al.*, 2015; Ozturk *et al.*, 2013), and sound symbolism facilitates infants' and toddlers' word learning (Imai *et al.*, 2008; Imai *et al.*, 2015; Kantartzis *et al.*, 2011). Given these findings, some researchers argue that iconicity is a design feature of language (Perniss and Vigliocco, 2014; Dingemanse *et al.*, 2015; Dingemanse *et al.*, 2020).

Thus far, the importance and pervasiveness of sound symbolism in languages have been empirically demonstrated in a number of studies (Voeltz and Kilian-Hatz, 2001; Bremner *et al.*, 2013). However, the essential questions pertaining to the psychological and biological mechanism of sound symbolism sensitivity—e.g., how the sense of

[a)]This paper is part of a special issue on Iconicity and Sound Symbolism.
[b)]Email: imai@sfc.keio.ac.jp

iconicity between word sound and meaning arises—have not yet been uncovered.

Widely recognized sound symbolism may arise from iconicity between sound and properties of the referents of words (Ramachandran and Hubbard, 2001; Sidhu and Pexman, 2018). There are two major non-mutually exclusive hypotheses regarding the mechanisms underlying such iconicity behind sound symbolism. First, sound symbolism could arise because people sense similarity between properties of the referent of a word and acoustic features of speech sounds (the acoustic account; Hinton *et al.*, 1994; Ohala, 1984). An alternative major account is the articulatory account, which maintains that properties of the referent and articulatory movements for speech production are associated with each other (Sapir, 1929; Ramachandran and Hubbard, 2001; Margiotoudi and Pulvermüller, 2020).[1]

Although a number of researchers have discussed the possibility of the two accounts, there is no definitive evidence for the acoustic or articulatory account. In fact, it may not be feasible to isolate the effect of only one of the accounts. For example, the roundness symbolism of sonorant consonants, such as /w, j, ɹ, m/, can be caused by the relatively smooth airstream felt in the vocal tracts (articulation) and gradual amplitude change (acoustics) involved in these sounds. Furthermore, articulatory movements are often reflected in acoustic changes (e.g., a slow articulatory movement generates a slow-changing acoustic signal). In this sense, what is needed in the field may not be to determine which of the two hypothesis is correct.

In this research, instead of attempting to determine which of the two hypotheses is correct, we aim to investigate whether there is evidence for the idea that sound symbolism can arise from iconicity between articulatory movement and a visual property of objects. As the articulatory hypothesis and acoustic hypothesis are not mutually exclusive, providing such proof should not be taken as evidence against the acoustic hypothesis. However, the articulatory hypothesis is important in that it can broaden our discussion on sound symbolism beyond mere iconic mapping between sound and other sensory properties. Humans, from infancy, sense iconicity between different sensory modalities such as size and brightness (Walker and Walker, 2016) or tactile and color (Ludwig and Simner, 2013). If we can establish that articulatory movements are a critical factor, if not a sole factor, in sound symbolism, this allows us to account for sound symbolism as arising from a bias (preference) in mapping between different modalities, which may not solely depend on the ability to hear speech sound.

An effective way to test this possibility is to examine whether people with severe hearing loss could detect sound symbolism and to whom acoustic information for sound symbolism is unavailable or, at best, severely limited. If the articulatory movement plays a critical role in sound symbolism detection, then deaf individuals, even those with profound hearing loss, should be able to recognize sound symbolism in novel words in a way that is similar to how hearing individuals do, relying on iconicity between the movement of articulatory organs and referents. In the study by Eberhardt (1940), deaf children who were orally educated to speak English could correctly judge the meaning of antonym pairs. They were visually presented with phonetic descriptions of antonym pairs that they have not acquired, yet, and the meanings of the anonyms. The antonyms may be English word pairs, such as "long" and "short," or foreign language word pairs, such as "breit (wide)" and "eng (narrow)" in German. They pronounced the word pairs and then judged which word had which meaning. Their judgment was correct above chance. In addition, Eberhart investigated if orally educated deaf children could sense the vowel-size sound symbolism of Sapir (1929). Although deaf children placed some of the vowels in different places in the hierarchy of size connotation, overall, they associated high front vowels to small objects and back vowels to large objects, similar to that of hearing participants in Sapir (1929) and Newman (1933). We can interpret Eberhart's results as supporting evidence for the idea that articulatory movements associated with words can be the basis of sound symbolism.

However, Eberhart's conclusion is limited in two respects. First, children in that study received explicit instruction to pronounce the word, which would lead to specific attention to articulatory movement. Second, Eberhart did not experimentally manipulate the articulatory movement and, thus, the causal role of articulatory movement was not directly established.

The current research investigated the role of articulatory movement in sensing sound symbolism also by asking whether deaf-and-hard-of-hearing (DHH) individuals can sense the shape sound symbolism of Köhler (1929). Like Eberhart, we expected that individuals with severe hearing loss could detect sound symbolism through the sense of iconicity between articulatory movement and the visual property of the object (i.e., shape in this research). Two important methodological changes were made from Eberhart's study to overcome the two limitations. First, our task did not involve explicit instruction to pronounce the stimulus words. Second, in experiment 2, articulatory movement was disturbed such that we could see the role of articulatory movement in sound symbolism detection more directly than that in Eberhart's study.

Experiment 1 examined whether adult DHH participants make similar sound symbolic judgments to hearing participants for various novel words. Experiment 2 examined whether DHH participants' sound symbolic judgment deviates more from that of hearing participants when articulatory movements were restricted to further examine the causal role of articulatory movement in the detection of sound symbolism.

II. EXPERIMENT 1

We compared the two groups of participants—those who are DHH and those with typical hearing capacity—on how they judged the sound symbolic match between novel

https://doi.org/10.1121/10.0034832

words and novel shapes by using the well-established bouba-kiki paradigm (Köhler, 1929; Ramachandran and Hubbard, 2001).

A. Method

1. Participants

It was difficult to determine the sufficient sample size for the DHH selection in advance, but it has been known that shape sound symbolism leads to a highly strong effect in forced choice tasks in the literature. For example, (hearing) adult participants choose the sound symbolically matching item highly consistently in a two alternative forced choice task: about 82% in Maurer *et al.* (2006), 95% in Ramachandran and Hubbard (2001), and 82% in Bremner *et al.* (2013). Thus, we expected at least a moderately strong correlation of $r = 0.5$ between the choices made by the hearing group and those made by the DHH group if individuals in the latter group do detect sound-shape correspondence by articulatory movements. With the alpha level of 0.05, power = 0.8, and $r = 0.5$, the required sample size is $N = 29$. Therefore, we aimed to recruit 30 participants.

From a university in Japan for students with total or severe hearing loss, 34 pupils (15 men and 19 women) volunteered to participate in response to an intraschool advertisement with posters and social media such as Facebook or X (previously Twitter). To be eligible for admission to this institution, which the participants were enrolled in, their hearing must be 60 dB or more for both ears or find it extremely difficult to hear voices at normal loudness even with the use of hearing aids or installed implants. Those who meet this requirement are considered to fit into the class of severest hearing loss. The participants' hearing was severely limited from birth as a result of sensorineural problems, which were caused by damage to the special cells or nerve fibers in the inner ear. To these individuals, hearing aids or implant installation could help hearing to some degree by making the sound louder, but even with such aids, their hearing of speech sound is still severely limited: other people's voices typically sound mumbled or slurred, and it is particularly hard to tell high-pitched sounds (such as "s" or "th") from one another. Thus, our participants' level of hearing is not close to that of typical hearing people, even with implants installed, in volume and clarity of sound. In Japan, implant installation is subsidized only for one ear. Many DHH people in Japan, thus, use a hearing aid on the other ear instead of getting implants installed on both ears. In fact, no participants in this research had implants installed on both ears. Fifteen participants indicated that they used a device and found it useful. However, 17 participants (out of 34) indicated that they did not find their devices useful and could not (or hesitated to) provide the information about the hearing loss range after correction by hearing devices, and 2 participants did not use a device. Hence, we labeled the first group as device effective, the second group as device noneffective, and the third group as "total non-device use" (see supplementary material B). We use these labels because on the post-experiment questionnaire, the DHH individuals who were aware of the improvement in the quality of hearing by the devices tended to use their hearing devices more heavily than those who were not aware.

All participants except one use sign language with their DHH peers or family members (see supplementary material B) but had experienced articulation training as the training is mandatory in schools for DHH students in Japan. Information about hearing and other relevant properties for each individual participant in experiment 1 are given in supplementary material B. These participants were able to read Japanese scripts and knew how one would pronounce the stimulus words. As the two types of phonograms in Japanese, called *hiragana* and *katakana*, roughly represent moras and are taught with a syllabary of onset consonants and vowels in schools, these participants were likely to have some phonological awareness.

An additional 36 individuals (19 men and 17 women) with typical hearing who were enrolled in an introductory psychology course in a university in the Greater Tokyo Area participated to provide the baseline. For both groups, all participants' data were included in the analyses.

2. Stimuli

As the primary purpose of the study was to test whether DHH individuals with the severest hearing loss were able to detect sound symbolism, the stimuli were prepared to ensure that hearing participants would show differential match/mismatch judgments for the round and spiky shapes. As the study did not aim to evaluate the strength of sound symbolism for particular sound-shape combinations or the source of sound symbolic effect, we did not exhaustively and systematically control consonant and vowel combinations in the stimuli. Expanding the stimulus set of D'Onofrio (2014), we created a total of 38 novel words by combining 16 consonants and 5 vowels (Table I). Previous findings (Köhler, 1929) indicate that among these segments, sonorant (/m, n, j, ɾ/) and bilabial consonants (/b, p/) as well as the rounded vowel /o/ are associated with round shapes, whereas non-labial obstruents (/d, t, g, k, z, s, ʑ, ɕ, t, ɕ, dʑ/) and front vowels (/i, e/) are good candidates for spiky shape sound symbolism (D'Onofrio, 2014; Maurer *et al.*, 2006; Nielsen and Rendall, 2011). Possible articulatory features of these segments that may be mapped to shapes include the lip rounding or protrusion of the rounded vowels and bilabials: the phonemes associated with round objects have lip rounding or protrusion of the rounded vowels and bilabials, but the phonemes for spiky objects do not. These consonants and vowels were combined to form words with the /CVCV/ (Consonant-Vowel/Consonant-Vowel) structure, which is a common word form in Japanese. To make our stimuli maximally contrastive, most round-sounding consonants were combined with round-sounding vowels, and most spiky-sounding consonants were combined with spiky-sounding vowels. We also added six well-known stimuli from previous studies examining the shape sound symbolism:

https://doi.org/10.1121/10.0034832

TABLE I. List of stimulus words for experiments 1 and 2, the Roman script, and International Phonetic Alphabet (IPA). The words were presented to the participants in the *katakana* script.

Katakana	Roman	IPA	Katakana	Roman	IPA	Katakana	Roman	IPA
ガグ	gagu	/gagɯ/	タケテ	takete	/takete/	プパ	pupa	/pɯpa/
ゴガ	goga	/goga/	チチェ	chiche	/tɕitɕe/	パプ	papu	/papɯ/
グガ	guga	/gɯga/	チテ	chite	/tɕite/	ポパ	popa	/popa/
キキ	kiki	/kiki/	ブバ	buba	/bɯba/	ノナ	nona	/nona/
キピ	kipi	/kipi/	バブ	babu	/babɯ/	ヌナ	nuna	/nɯna/
キケ	kike	/kike/	ボバ	boba	/boba/	マルマ	maruma	/maɾɯma/
ケキ	keki	/keki/	ブーバ	buuba	/bɯːba/	ムマ	muma	/mɯma/
ジゼ	jize	/ʑize/	ティテ	tite	/tite/	モマ	moma	/moma/
ゼジ	zeji	/zeʑi/	テティ	teti	/teti/	ユヤ	yuya	/jɯja/
ゾズ	zozu	/zozɯ/	テチ	techy	/tetɕi/	ヨヤ	yoya	/joja/
ドゥダ	duda	/dɯda/	シセ	shies	/ɕise/	ルラ	rura	/ɾɯɾa/
ドダ	doda	/doda/	セシ	seshi	/seɕi/	ルロ	ruro	/ɾɯɾo/
			ソス	sosu	/sosɯ/	ロラ	rora	/ɾoɾa/

"maruma" (a Japanese equivalent of "maluma"), "takete" (Köhler, 1929), "buuba" (a Japanese equivalent of "bouba"), "kiki" (Ramachandran and Hubbard, 2001), as well as "moma" and "kipi" (Asano *et al.*, 2015).

Importantly, the round-sounding consonants tend to fall into a category of sounds that have been considered to be relatively easy for individuals with severe hearing loss to distinguish, whereas spiky-sounding consonants tend to be difficult to distinguish (Hillis *et al.*, 2023). Thus, if the level of hearing critically affects sound symbolism sensitivity in DHH people, we might hypothesize that the judgments of sound symbolism for round shapes for DHH people would be more accurate than their judgments for spiky shapes.

3. Procedure

Participants received a two-page booklet. On each page, either a spiky-shaped figure or round-shaped figure was depicted at the top (with the order counterbalanced; see Fig. 1), below which 38 novel words (e.g., "moma" and "kipi") were listed orthographically in *katakana*. Among the two phonographic systems in Japanese, i.e., *hiragana* and *katakana*, we chose to present the stimulus words in *katakana* because nonsense words are more commonly written in *katakana* scripts (see Table I). The order of the figures and words was different across booklets.

The participants were asked to judge whether each word would match the target figure on a three-point scale: (1) good match, (2) neutral, and (3) mismatch. Each participant judged match for round and spiky figures.

B. Results

The results were analyzed in two ways. First, we conducted a by-item correlation analysis to provide basic descriptive statistics, which helps to intuitively grasp how similar responses from DHH participants were to those of hearing participants. However, correlation analysis, which averages responses for each stimulus, does not consider individual differences in the degree of hearing loss. To compensate for this weakness, we employed a generalized linear mixed model that included the hearing difficulty and stimulus shapes of participants as independent variables in addition to the correlation analyses. This allowed us to examine whether the DHH participants' responses remained similar to the hearing participants' responses even when the variability from individual differences of participants and that from stimuli were simultaneously controlled for.

We conducted correlation and model analyses because the two investigations show different sensitivities to data variability; correlation analysis sensitively captures the dispersion in the relationship between two data series, whereas model analysis is good at capturing the magnitude of the slope in the relationship between variables in the linear model fitted to the data. These differences allowed us to uncover the characteristics of the data in a multifaceted way.

1. Correlation analysis

For each word-shape combination, we calculated the proportion of participants in each group who gave "good match" judgment, performed separately for round shape and spiky shapes (Fig. 1). The distribution of match scores across items (word-shape combinations) was highly similar between the DHH and hearing groups ($N = 34$) for the round figures ($r = 0.762$) and the spiky figures ($r = 0.832$). Remember that consonants that were assumed to match spiky figures were considered to be more difficult to distinguish by DHH participants than those assumed to match round shapes. However, the correlations between the two groups were both high, and we did not see that the judgments of spiky sound symbolism by the DHH group were less accurate than those of round sound symbolism. We also analyzed the proportion of participants in each group who gave "mismatch" judgments, and the results were the exact mirror image of the analysis of good match (see supplementary material Fig. S1).

We further examined whether the DHH participants' sound symbolism judgments varied as a function of their degree of hearing difficulty. To index the degree of hearing loss, the hearing loss range measured by audiogram is usually employed. However, as we noted in Sec. II A 1, this index

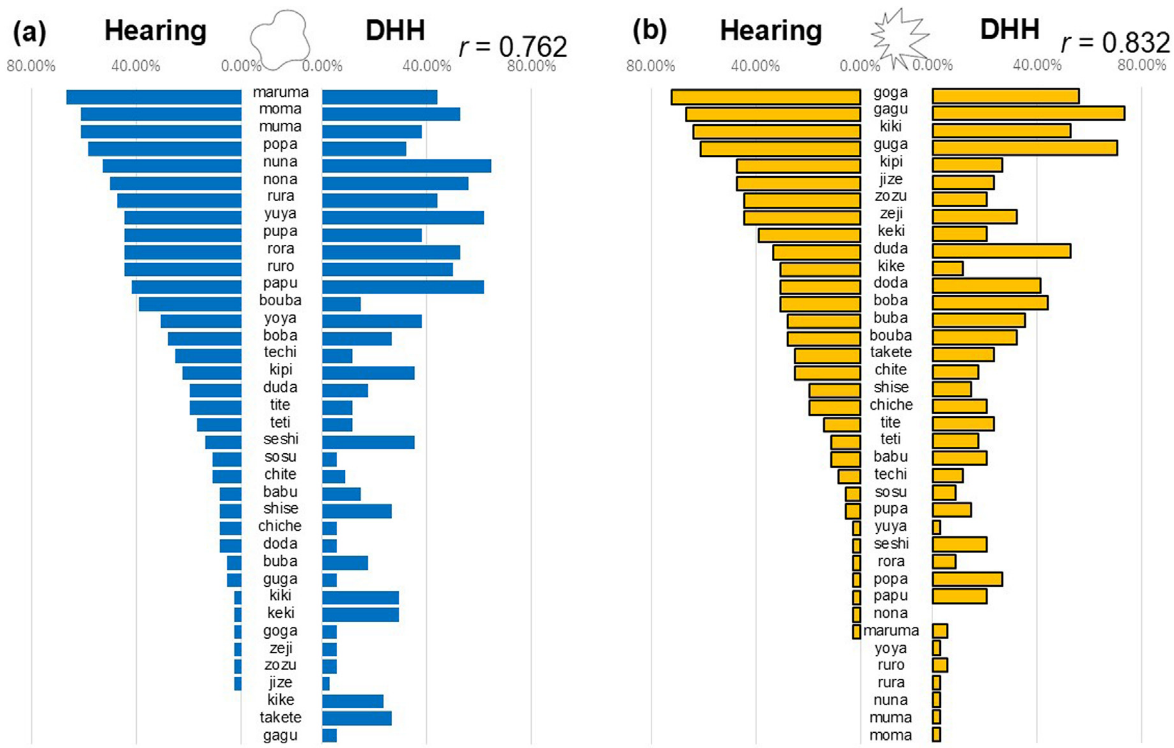

FIG. 1. (Color online) The percentage of participants in the hearing group and those in the DHH group who judged the novel words to be good sound symbolic match with the shapes in experiment 1. The words are arranged in descending order of the percentages for the hearing group.

does not accurately reflect the quality (clarity) of hearing, and some of the participants were not able to provide this value with hearing devices. As described earlier, we classified the DHH participants into three categories: hearing device effective, hearing effect noneffective, and hearing device total nonuse. However, in the following analyses, we labeled the device effective participants as "hearing device users," while combining the latter two categories, and labeled them as "hearing device nonusers," because on the post-experiment questionnaire, the DHH individuals coded as "hearing device effective" who were aware of the improvement of the quality of hearing by the devices tended to use their hearing devices more heavily than those who were not, and the number of participants falling into the third category—"hearing device total nonuse"—were too small for the statistical analyses. The correlation between the device users and hearing participants ($r = 0.72$) and that between the non-device users and hearing participants ($r = 0.79$) were high, and the two correlation values did not differ statistically (from Meng *et al.*, 1992, $z = -1.5874$, $p = 0.1124$).

Taken together, the correlation analyses showed that regardless of the level of hearing loss indexed by hearing device use, the participants in the DHH group judged sound symbolic match in a very similar way to the participants in the hearing group.

2. Model analysis

The correlation analyses examined the effects of hearing device use [i.e., whether participants thought that their hearing was improved by the hearing device(s) they were using; see Sec. II B 1] and shape of stimuli on DHH performance separately, but it is possible that the interaction between the two affects the performance. Here, we examined whether the judgments of hearing groups can predict if DHH participants accepted/rejected each sound-shape pair, even after controlling for hearing device use and stimulus shapes in a single model. Thus, the dependent variable of the model was a binary variable, i.e., whether or not a DHH participant gave a good match judgment to each shape-word pair. The critical independent variables were the proportions of hearing participants who gave a good match judgment for a given shape-word pair, hearing device use (device user vs nonusers), shape (round vs spiky), and the interaction among the three factors. In other words, in this model analysis, we wished to see if (and to what degree) the sound symbolism judgment of the hearing people for each word-shape pair would predict the choice pattern of DHH individuals, and whether or not DHH participants use hearing devices regularly, and the visual property of the referent object (shape) would additionally affect DHH participants' choice pattern. The three factors were included in the model as fixed effects, whereas participant identification (ID) and stimulus ID (the specific sound-shape combination presented) were treated as random intercepts such that the model could consider variabilities from individual differences and characteristics of the stimuli.

We fit a mixed effects logistic regression model with the glmer function in R (R Core Team, 2024). A series of models were fit with all possible pairs of the factors. The

https://doi.org/10.1121/10.0034832

best model (the best combination of the independent variables) was determined by the BIC (Bayesian information criterion; Bhat and Kumar, 2010), using the MuMIn package (Barton, 2020). The best model only included the proportion of hearing participants who gave a good match for a given shape-word pair [$b = 4.36$, standard error (SE) $= 0.273$, $z = 16.0$]. Other factors were not included in the final model, indicating that the performances of DHH participants can predict those of hearing participants regardless of hearing device use and stimulus shapes. Thus, the model analysis echoed the key finding from the correlational analyses: DHH participants and hearing participants gave similar judgments about match between shapes and words even after controlling the hearing quality of DHH participants and stimulus shapes.

C. Discussion

The results of the correlation and model analyses in experiment 1 suggest that individuals with the severest hearing loss can detect inherent correspondence between sound and shape, which is consistent with the articulatory hypothesis for sound symbolism. However, the articulatory movement was not experimentally manipulated in this experiment. To examine the causal role of articulatory movement more directly, we conducted experiment 2. If participants' judgments of sound symbolism deviate when they cannot move their articulators freely, it will provide stronger evidence for the articulatory hypothesis.

III. EXPERIMENT 2

Experiment 2 examined whether disturbance of articulatory movements changes the judgments of sound symbolism, and whether the effect, if any, is different across hearing and DHH groups. Oral movement was restricted in two ways (see Fig. 2). In the on-tongue condition, participants put a spoon in the month and placed it on the tongue, hence, movement within the oral cavity and movement of the lips were restricted. In the between-lips condition, a spoon was placed between the lips, thereby the lip movement was restricted but articulatory movement within the oral cavity was freer. Thus, although both conditions restricted articulatory movements, the on-tongue condition restricted articulatory movement more than the between-lips condition.

We expected that hearing participants and DHH participants should differ in the way that the manipulation affected sound symbolic judgment. Hearing people may additionally recruit acoustic information to compensate for the lack of the articulatory cues. If this is the case, hearing participants should be less affected by the disturbance of articulatory movement than DHH participants. We predicted three outcomes. First, we expected that the DHH participants in experiment 2 should deviate from the baseline judgment by hearing participants in experiment 1 more strongly than DHH participants in experiment 1 (who had no articulatory restrictions). Second, we predicted that the DHH participants in experiment 2, overall, should deviate from the baseline judgment by hearing participants in experiment 1 more strongly than hearing participants in experiment 2. Third, we expected that DHH participants' judgment should deviate from the baseline judgment by hearing participants in experiment 1 more noticeably in the on-tongue condition than in the between-lips condition.

A. Method

1. Participants

Thirty-two DHH students participated, where all had congenital hearing loss and were enrolled in the same university for deaf people as those in experiment 1. Additionally, 61 university students with typical hearing were recruited from the same universities as those in experiment 1 and participated in this study. None of the participants took part in experiment 1. They were randomly divided into the *on-tongue* condition or the *between-lips* condition and judged shape sound symbolism in the same way as in experiment 1. The profiles of the DHH participants in experiment 2 were compatible with those in experiment 1 (range of the hearing loss, 30–130 dB; $M = 98.09$ dB). As in experiment 1, the participants were asked about the situation of hearing devices. It was indicated by 21 participants that they used a device and found it effective, whereas 9 participants indicated that hearing devices were not effective. Two participants did not use hearing devices at all. As in experiment 1, we labeled the first group as device effective, the second group as device noneffective, and the third group as total non-device use. (see supplementary material C). One participant reported that his hearing loss degree was 30 dB when measured by an audiogram at one time but that he had extreme difficulty hearing speech in conversation. All participants except one use sign language in some degree when communicating with their DHH peers. DHH and hearing participants were randomly assigned to either the on-tongue condition or the between-lips condition. The average hearing levels of the DHH participants in the two conditions were not significantly different from each other (on-tongue, range 70–120 dB; $M = 97.88$ dB; between-lips, range 60–130 dB; $M = 98.53$ dB).

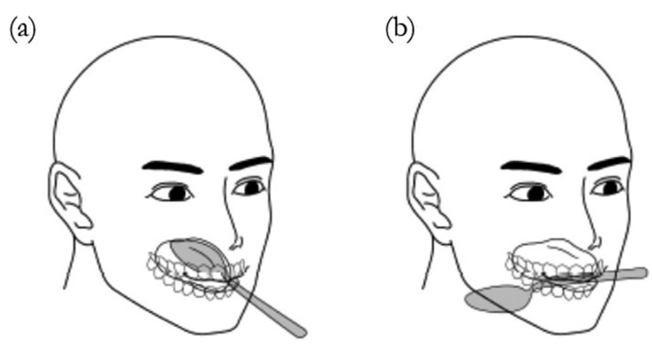

FIG. 2. Illustrations of the (a) "on-tongue condition" and (b) "between-lips condition" in experiment 2.

https://doi.org/10.1121/10.0034832

For both groups, all participant data were included in the analyses.

2. Stimuli and procedure

The stimuli were identical to those used in experiment 1 as was the procedure except that participants were asked to hold a spoon with their mouth (Fig. 2). In the on-tongue condition, they put the spoon in the oral cavity with the concave part of the spoon facing down, placed the concave part on the tongue, and closed the lips to stabilize the handle of the spoon. In the between-lips condition, they did not put the spoon in the oral cavity. They held the handle of the spoon between the lips in such a way that the spoon was kept sideways.

The two conditions (on-tongue vs between-lips) were similar in that the lips were closed. This is because the handle of a spoon was thin, and participants needed to close the lips to stabilize the spoon. In particular, in the on-tongue condition, the lips were closed because the spoon would be unstable and the handle may swing sideways unless the handle was fixed firmly at two points, which were the lips and teeth. Thus, in both conditions, the tongue movement was restricted as the oral cavity was narrow. Importantly, the tongue movement was more restricted in the on-tongue condition than in the between-lips condition because the spoon was fixed at the lips and teeth in the on-tongue condition to fixate the tongue position. Moreover, the spoon was touching the tongue in the on-tongue condition and, thus, the tactile feedback for the tongue position was compromised. In the between-lips condition, the tongue could move, albeit within the narrow oral cavity as a result of closed lips, which may have been sufficient for simulating articulation.

B. Results

As in experiment 1, we first conducted a by-item correlation analysis to grasp how similar DHH and hearing participants in the on-tongue condition and between-lips condition were to the hearing people in experiment 1, which served as the baseline. Then, we conducted model analyses to investigate whether the interaction between the hearing ability and experimental condition predicted the participants' judgment on the sound symbolic relations in two steps. In the first step, as in experiment 1, a model was created exclusively for the DHH group, using the same set of independent variables as employed in experiment 1, to which the experimental condition (on-tongue vs between-lips) was added to the model. In the second step, we examined whether the experimental condition affected performance differently for the DHH and hearing individuals. (See Sec. III B 2 for further details.)

1. Correlation analysis

We performed the same by-item correlation analysis as in experiment 1, where the data from the round figure and the data from the spiky figure are combined. In all analyses, we took the judgments by the hearing participants in experiment 1 (who had no articulatory restrictions) as the baseline and calculated correlations against the baseline to assess how closely the judgments in different conditions and groups aligned with the baseline judgments.

Before we examined how two types of disturbance of articulatory movements affected DHH and hearing participants, we first tested whether the degree of hearing loss influenced the performance of the DHH participants in experiment 2, according to the criterion used in experiment 1 (hearing device users vs nonusers). The analysis showed no sign for this possibility, as reported in Sec. III B. Because the performance of DHH participants did not differ across the two groups in either experimental condition in experiment 2, we used the combined data of the DHH participants for further analyses.

In all four cells arising from the combination of two factors in experiment 2 (see Fig. 3), the correlations with the baseline (the y axis in Fig. 3) were significantly different from chance (DHH on-tongue, $r = 0.639$, $p < 0.001$; DHH between-lips, $r = 0.788$, $p < 0.001$; hearing on-tongue, $r = 0.913$, $p < 0.001$; hearing between-lips $r = 0.882$, $p < 0.001$).

First, we compared the performance of DHH participants in experiments 1 and 2. The correlation with the baseline (the performance of the hearing participants in experiment 1) was stronger for DHH participants in experiment 1 ($r = 0.862$) than for DHH participants in the on-tongue condition ($r = 0.639$; Meng's $z = 4.34$, $p < 0.0001$) and for DHH participants in the between-lips condition ($r = 0.788$; Meng's $z = 1.99$, $p = 0.047$). That is, the DHH participants in experiment 1 gave judgments closer to the baseline than the DHH participants in the two conditions in experiment 2.

Second, we examined the performances of the DHH and hearing participants in experiment 2. In the on-tongue condition, the correlation with the baseline was higher in the hearing group ($r = 0.913$) than in the DHH group ($r = 0.639$; Meng's $z = 5.71$, $p < 0.001$). In the between-lips condition, the correlation with the baseline was also higher in the hearing group ($r = 0.882$) than in the DHH group ($r = 0.788$; Meng's $z = 2.66$, $p = 0.008$). That is, overall, the hearing participants in experiment 2 gave judgments closer to the baseline than the DHH participants in experiment 2, which is consistent with our expectation.

Third, we compared the performances in the between-lips and on-tongue conditions. For the DHH group, the correlation with the baseline was significantly higher in the between-lips condition ($r = 0.788$) than in the on-tongue condition ($r = 0.639$; Meng's $z = 2.67$, $p = 0.007$), which is also consistent with our prediction. For the hearing group, the correlation coefficients in the on-tongue condition ($r = 0.913$) did not significantly differ from those in the between-lips condition ($r = 0.882$; Meng's $z = 1.40$, $p = 0.162$). Thus, again, consistent with our prediction, while the hearing participants were not affected by the disturbance of articulatory movement regardless of whether the spoon was placed on the tongue or between the lips, the

DHH participants were affected, and the effect of articulatory disturbance was stronger in the on-tongue condition than in the between-lips condition.

2. Model analysis

The model analysis examined how closely participants' judgments in experiment 2 can be predicted by the baseline judgment by the hearing participants in experiment 1 for each shape-word pair and the two experimental manipulations in experiment 2, i.e., spoon positions.

In the first analysis, we examined only the DHH group to see whether the interaction between hearing device use and spoon position influenced the participants' judgments on sound symbolic relations. The dependent variable was a binary code indicating whether (or not) a participant gave a good match judgment to each shape-word pair. The critical independent variables were the proportion of good match judgment for each word-shape combination by the hearing participants in experiment 1 (i.e., the baseline), hearing device use (hearing device users vs nonusers), shape (round vs spiky), spoon position (on-tongue vs between-lips), and all interactions among the variables. The first three factors were the same as those included in experiment 1, whereas spoon position was newly introduced to examine the effect of experimental condition. We also incorporated by-participant random slope for spoon position and random intercept for stimuli into the model. As in experiment 1, we fit a mixed effects logistic regression model with the glmer function in R. The series of models with all possible combinations of the independent variables were ranked by BIC. The best model included the main effect of baseline judgment ($b = 4.18$, $SE = 0.307$, $z = 13.6$), indicating that the DHH participants' judgments of sound symbolism were consistent with those of hearing participants in the no-spoon condition in experiment 1. Because the effects of hearing device use were not found, device users and nonusers were combined into a single DHH group for subsequent analyses.

In the second analysis, we examined how DHH and hearing groups differed in how they were affected by the disturbance of articulatory movements. The dependent variable of the second model was the same as that in the first model (i.e., good match judgment by participants). The

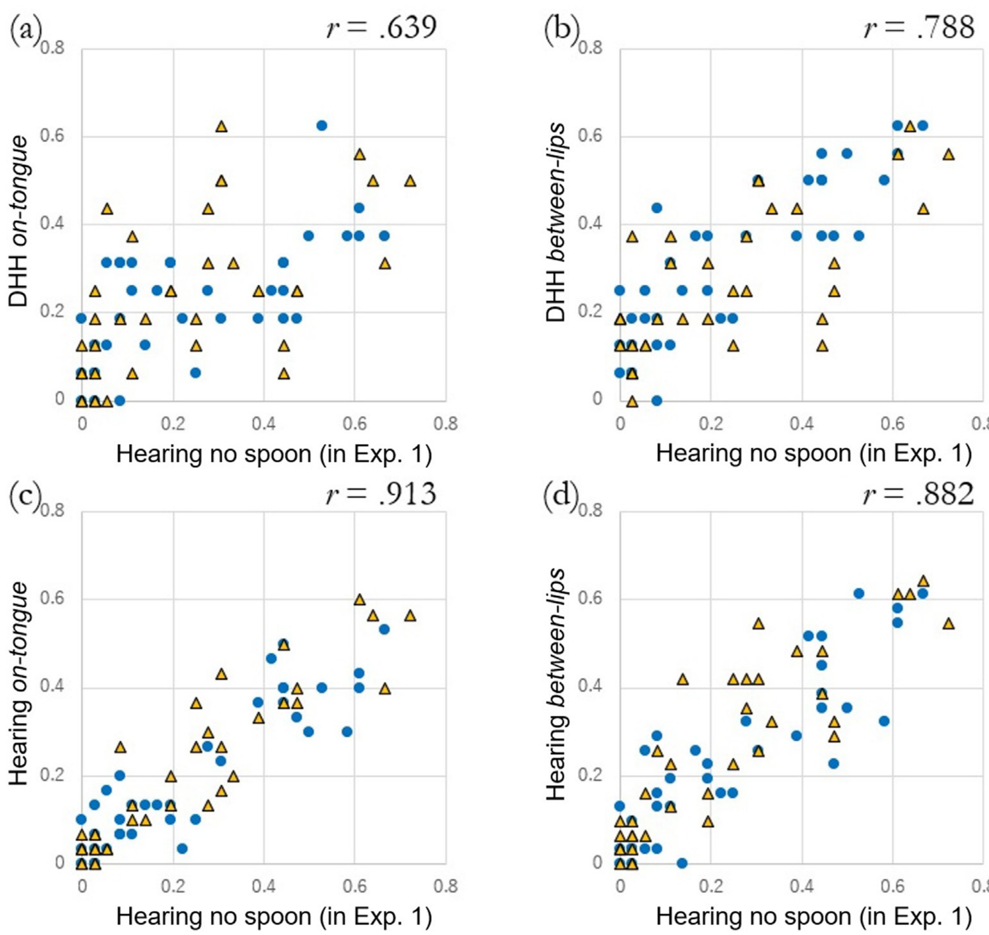

FIG. 3. (Color online) The scatterplots show how well the judgments in experiment 2 corresponded to the judgments by the hearing participants in experiment 1, who did not hold a spoon with their mouth (i.e., baseline judgment). Each data point in the scatterplots is an item (word-figure combination): blue circles represent words combined with the round shape, and yellow triangles represent words combined with the spiky shape. The data from the rounded and spikey shapes were combined in the by-item correlational analyses. The horizontal axis represents the percentage of hearing participants from experiment 1 who judged items to be good sound symbolic matches. The vertical axis represents the percentage of participants in each condition in experiment 2 who judged items to be good sound symbolic matches: (a) DHH on-tongue, (b) DHH between-lips, (c) hearing on-tongue, and (d) hearing between-lips.

independent variables were baseline, group (DHH vs hearing group), spoon position (on-tongue vs between-lips), and all interactions among the variables. As in the first model, by-participant random slope for spoon position and random intercept for stimuli were also incorporated into the model. The effect of shape was excluded from the second model because this factor did not contribute to the model in experiment 1 or the first model in experiment 2.

The results showed a significant main effect of the baseline ($b = 6.17$, SE = 0.256, $z = 24.1$) and the interaction between group and baseline judgment ($b = -1.78$, SE = 0.338, $z = -5.28$) in Table II. The negative coefficients of the interaction indicated that the hearing group in experiment 2 responded more similarly to the hearing group in experiment 1 than to the DHH group in experiment 2.

C. Discussion

The correlation analyses indicated that when articulatory movement was restricted by a spoon, DHH participants' judgments deviated more from the baseline judgments (by hearing participants in experiment 1 with no restrictions on articulation) than the DHH participants in experiment 1. They were also more strongly affected by the experimental manipulations than hearing participants in experiment 2. This deviation in the DHH participants was stronger when the spoon was on the tongue, which restricted the movement of the lips and tongue, in comparison to when the spoon was between the lips, which mainly restricted the movement of the lips and jaw, suggesting that the sensitivity to sound symbolism decreased more as the degree of the restriction of articulatory movements got larger. The hearing participants were not affected by the position of the spoon.[2] The linear mixed effects modeling confirmed the difference between the DHH participants in experiment 1 and DHH participants in experiment 2. It also confirmed the difference between the DHH participants in experiment 2 and the hearing participants in experiment 2, indicating that DHH people were affected more strongly by the disturbance of articulatory movements, which, in turn, suggests that they relied more on the actual articulatory movements than hearing people in sensing iconicity between sound and meaning. However, the model analysis did not confirm the effect of spoon position for the DHH participants, which varied from the correlation analysis.

One possible reason for the discrepancy concerning the effect of the spoon position between the correlational and model analyses is that the two analyses are sensitive to different types of deviation from the baseline judgments, which was noted earlier. The correlation analysis is sensitive to changes in variability in judgments (spread in the data plotted Fig. 3), whereas the model analysis is sensitive to how changes in the judgments in experiment 2 relate to the baseline judgments (slope in the data plotted in Fig. 3). The correlation analyses indicated that the variability in the sound symbolism match judgment was higher in DHH participants than in hearing participants [see Fig. 3(a) vs Fig. 3(c), Fig. 3(b) vs Fig. 3(d), and correlation values]. The variability was also higher in the on-tongue condition than in the between-lips condition for DHH participants (see Fig. 3(c) vs Fig. 3(d) and correlation values). The same difference was found for the DHH participants in experiment 1 (no articulatory restriction) and DHH participants in experiment 2 (restriction by a spoon). Thus, DHH participants' sound symbolism judgment became less stable and deviated more from the baseline judgments by hearing participants as the degree of articulatory restrictions became greater.

IV. GENERAL DISCUSSION

In the long history of linguistics, language has been characterized as abstract, amodal, and not directly bound to sensory or bodily experiences. In recent years, a body of studies have demonstrated empirical evidence that challenges this assumption (e.g., Perniss *et al.*, 2010). However, a key question—what mechanism underlies the sense of sound symbolism—has remained inconclusive. The results of the present study have offered new insights into this question from a hitherto untested perspective, one that asked how people with severe hearing loss sense sound symbolism.

Individuals with severe hearing loss sensed sound symbolism very well; thus, hearing ability is not as indispensable for sound symbolism as one might have assumed. We found that the DHH participants, even those with congenital and most profound hearing loss, judged match/mismatch of sound and shape correspondences in a way very similar to that of hearing people, as long as the DHH participants could make articulatory movements freely. However, when articulation was restricted, DHH individuals' sense of sound symbolism deviated away from hearing participants' sense of sound symbolism.

Our results build on the study by Eberhardt (1940) and extended it in an important way. As in Eberhardt's study, we found that deaf participants could make similar sound symbolic judgment as that of hearing participants. Unlike Eberhardt's study, we did not explicitly instruct DHH participants to pronounce the words before judgment, thus, our DHH participants were likely to have implicitly simulated articulatory movement when making sound symbolic judgment. We went beyond Eberhardt (1940) in an important way as we investigated the causal role of articulatory movement more directly by showing that the sense of sound symbolism by DHH people was deteriorated when their articulatory movement was restricted.

Our findings offer great insights into a key question in the literature of sound symbolism, i.e., whether sound symbolism can arise from articulation (Eberhardt, 1940; Sapir,

TABLE II. The fixed effects of the best model. **, $p < 0.01$; **, $p < 0.001$.

	b	SE	z	p
Baseline judgment	6.17	0.256	24.1	$< 2 \times 10^{-16***}$
Group	0.762	0.279	2.73	0.0063^{**}
Baseline judgment:group	–1.78	0.338	–5.28	$1.33 \times 10^{-7***}$

https://doi.org/10.1121/10.0034832

1929; Ramachandran and Hubbard, 2001; Margiotoudi and Pulvermüller, 2020) or acoustic features, such as frequency (Hinton *et al*., 1994; Ohala, 1984), or both. The results of this research provide evidence that articulatory movements play an important role in sound symbolic judgment and suggests that one can sense correspondence between sound and meaning, at least in some (or a large) degree, through articulatory movements. For example, the vowels and bilabial consonants associated with round objects have lip rounding and protrusions, and such features can readily be associated with roundness. Indeed, the DHH participants judged words such as "maruma" and "moma" very accurately (see Fig. 1). It is remarkable that individuals with difficulty in recruiting the auditory channel can sense sound symbolism by mapping oral movements to the meanings.

However, it should be noted that the articulatory account and acoustic account are not mutually exclusive. In fact, restricting articulatory movement did not affect the hearing participants, which suggests that the hearing participants used acoustic image for sound symbolism even in the absence of auditory input. It is likely that hearing people can use articulatory and acoustic bases for sound symbolism judgment (cf. Fort and Schwartz, 2022).

In addition to the articulatory movement itself, the tactile feedback of the movement may also play a role in sound symbolism. When the tongue touches the other parts of the oral tract, it gets tactile feedback, which plays an important role in speech production. The nature of this tactile sensation may also contribute to the sense of sound symbolism.

Could our DHH participants use sources other than articulatory movement in their judgments of sound symbolism? It may be possible. Although DHH individuals were affected by the restriction of articulatory movement more strongly than hearing people, their judgments did not become random in experiment 2. It is possible that DHH individuals first sensed sound symbolism through iconicity between articulatory movement and the referents, and this experience may have developed into an abstract representation of sound-meaning correspondence, which enabled them to image it without actual articulatory movements, but this representation was not as solid as that for hearing people and vulnerable to disturbance of articulatory movement compared to people with typical hearing function.

Some may wonder that the sense of sound symbolism arises through orthographical symbolism, i.e., the iconicity between the shape of the letters in the word and the shape of the figure. However, it is not likely that orthographical symbolism can explain all aspects of our results. In particular, it cannot explain the greater influence of the articulatory restrictions on sound symbolic judgment in DHH participants than in hearing participants. Thus, even if orthographical symbolism was at play in our task, that cannot be the whole story. Furthermore, a recent study by Ćwiek and colleagues (Ćwiek *et al*., 2022) investigated whether and to what extent the bouba-kiki effect could be explained by the letter symbolism across different cultures. They concluded that the contribution of letter symbolism is negligible for the bouba-kiki sound symbolic effect overall but especially for speakers whose first language uses non-Roman scripts. Taken together, we argue that the results of the two experiments cannot be explained by orthographical symbolism. However, it is possible that DHH people are more sensitive to iconicity between the shape of letters and the meaning of the words than hearing people. Future research is required to examine whether the sensitivity to letter-meaning iconicity is different between hearing individuals and those with hearing loss.

The articulatory account of sound symbolism has profound implications for theories of language evolution. Hand gesture has been considered to be a prime candidate for how an open-ended shared lexicon emerged in early stages of language evolution (Stokoe, 2002; Arbib, 2005) because people can move their hands in a way that can iconically map to entities in the world such as events and objects (Goldin-Meadow *et al*., 1996; Ortega and Özyürek, 2020). The hand-gesture origin theories of language evolution maintain that hand movements have unique advantage over the movement of tongue and other articulatory organs when it comes to iconically representing entities in the world (Stokoe, 2002; Arbib, 2005). However, the current results may question such an advantage for hand gestures and suggest that hand movements and articulatory movements can be the basis of iconic meaning, offering a possible account for why language evolved in speech modality as well as in manual modality.

The current research found a possibility that a novel word can get its link to a referent object from articulatory movements. This suggests that articulatory movements may have mediated the link between word form and meaning in the dawn of language. A speaker can give sound to a word by articulatorily mimicking a property of the referent, and a listener who hears the word infers the meaning by articulating the sound and sensing iconicity between the articulatory movement and the referent. This bidirectional process between sound and meaning may have helped our ancestors to quickly build a shared lexicon that can be intuitively understood by members of the community (Kita, 2008; Imai and Kita, 2014). The current result also adds further support to the idea that language is an embodied representational system, where meaning partly originates from how the body mimics events and things in the world (Fort and Schwartz, 2022; Lakoff and Johnson, 1980; Glenberg and Kaschak, 2002; Lupyan and Casasanto, 2015; Stokoe, 2002).

V. LIMITATIONS AND FUTURE DIRECTIONS

Much more research is needed to fully understand the mechanism of sound symbolism sensitivity at the psychological and neural levels. The present research invites future exploration in a variety of exciting directions.

This study relies on the premise that people simulate articulatory movement even when they silently read words. Although there is evidence for such simulation (Yao, 2021; Yan *et al*., 2014), this assumption should be directly tested in future research.

The generalizability of the present results should be investigated. Do DHH people sense sound symbolism equally well in other sensory domains such as tactile, motion, or magnitude? Do they acquire language-specific sound symbolism as hearing people do (cf. Saji *et al.*, 2019)? Do DHH children sense sound symbolism as in adults and, if yes, is their word learning (in spoken language) scaffolded by sound symbolism as in hearing children (Imai *et al.*, 2008; Kantartzis *et al.*, 2011)?

Another important question for future investigation is whether people with hearing loss could detect sound symbolism without training to articulate speech sound. In Japan, oral language has been heavily stressed in schools such that all of the DHH participants in this research had extensive training for understanding and producing spoken language. It is interesting to test DHH individuals who had not received articulation training on this task to determine whether general articulation training is required to detect sound symbolism. This would help us understand not only the nature of sound symbolism sensitivity but also the nature of cross-modal mapping and processing in humans.

SUPPLEMENTARY MATERIAL

See the supplementary material A for the distribution of mismatch judgments in experiment 1 and supplementary material B and C for the background of the DHH participants in experiments 1 and 2, respectively.

ACKNOWLEDGMENTS

This study was supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT)/Japan Society for the Promotion of Science (JSPS) Grants-in-Aids-for Scientific Research (KAKENHI) Grant Nos. 20H00014, 16H01928, and 18H05084 to M.I. We would like to thank Junko Kimura, Daiki Kondo, and Fumika Murayama for their help with data collection.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

The experiments reported in this paper were approved by the research ethics committee at the Tsukuba University of Technology for hearing or visually impaired persons (Approval No. 2022-20) and the Shonan-Fujisawa Campus (SFC) research and experiment ethics committee at the Keio University (Approval No. 2019/262). All participants signed informed consent and agreed on the publication of their anonymized data.

DATA AVAILABILITY

The data that support the findings of this study are openly available in the Open Science Framework at https://osf.io/m64eu/.

[1]As the editor pointed out, the IPA, which most studies on sound symbolism use, is primarily based on articulatory features, such as place and manner of articulation, and this tradition might have biased many researchers toward the articulatory account.

[2]As an anonymous reviewer pointed out, the sound-shape associations used in the stimuli can be found in several ideophones (e.g., *gizagiza*, "serrated" and *kakukaku*, "jagged") and prosaic words (e.g., *maru*, "circle"), and this lexical knowledge might have facilitated the participants' sound symbolic judgments and made the two groups' judgments similar to each other. However, the current experimental result that the on-tongue condition weakened the DHH participants' sound symbolic intuition suggests that they did, in fact, use articulatory information to link sounds and shapes.

Arbib, M. A. (**2005**). "From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics," Behav. Brain Sci. **28**, 105–124.

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., and Thierry, G. (**2015**). "Sound symbolism scaffolds language development in preverbal infants," Cortex **63**, 196–205.

Bartoń, K. (**2020**). "MuMIn: Multi-model inference. R package version 1.43.17," available at https://CRAN.R-project.org/package=MuMIn (Last viewed September 26, 2021).

Bhat, H. S., and Kumar, N. (**2010**). "On the derivation of the Bayesian information criterion," University of California Technical Report, Vol. 99, pp. 1–4.

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., and 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**, 165–172.

Cuskley, C., and Kirby, S. (**2013**). "Synaesthesia, cross-modality and language evolution," in *Oxford Library of Psychology. The Oxford Handbook of Synesthesia*, edited by J. Simner and E. M. Hubbard (Oxford University Press, Oxford, UK), pp. 869–899.

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-Logacev, Ö., Zeller, J., Perlman, M., and Winter, B. (**2022**). "The *bouba/kiki* effect is robust across cultures and writing systems," Philos. Trans. R. Soc. B **377**(1841), 20200390.

de Saussure, F. (**1916**). *Course in General Lingustics* (Columbia University Press, New York).

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., and Monaghan, P. (**2015**). "Arbitrariness, iconicity, and systematicity in language," Trends Cognit. Sci. **19**, 603–615.

Dingemanse, M., Perlman, M., and Perniss, P. (**2020**). "Construals of iconicity: Experimental approaches to form-meaning resemblances in language," Lang. Cognit. **12**, 1–14.

D'Onofrio, A. (**2014**). "Phonetic detail and dimensionality in sound-shape correspondences: Refining the bouba-kiki paradigm," Lang. Speech **57**, 367–393.

Eberhardt, M. (**1940**). "A study of phonetic symbolism of deaf children," Psychol. Monogr. **52**, 23–41.

Fort, M., and Schwartz, J. L. (**2022**). "Resolving the *bouba-kiki* effect enigma by rooting iconic sound symbolism in physical properties of round and spiky objects," Sci. Rep. **12**, 19172.

Glenberg, A. M., and Kaschak, M. P. (**2002**). "Grounding language in action," Psychon. Bull. Rev. **9**, 558–565.

Goldin-Meadow, S., McNeill, D., and Singleton, J. (**1996**). "Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality," Psychol. Rev. **103**, 34–55.

Hillis, C. L., Uchanski, R. M., and Davidson, L. S. (**2023**). "Common sounds audiograms: Quantitative analyses and recommendations," Semin. Hear. **44**, S49–S63.

Hinton, L., Nichols, J., and Ohala, J. J. (**1994**). *Sound Symbolism* (Cambridge University Press, Cambridge, UK).

Hockett, C. F. (**1960**). "The origin of speech," Sci. Am. **203**, 88–96.

Imai, M., and Kita, S. (**2014**). "The sound symbolism bootstrapping hypothesis for language acquisition and language evolution," Philos. Trans. R. Soc. B. **369**, 20130298.

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

Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., and Kita, S. (**2015**). "Sound symbolism facilitates word learning in 14-month-olds," PLoS One **10**, e0116494.

Kantartzis, K., Imai, M., and Kita, S. (**2011**). "Japanese sound-symbolism facilitates word learning in English-speaking children," Cognit. Sci. **35**, 575–586.

Kita, S. (**2008**). "World-view of protolanguage speakers as inferred from semantics of sound symbolic words: A case of Japanese mimetics," in *Origins of Language*, edited by N. Masataka (Springer, Tokyo), pp. 25–38.

Köhler, W. (**1929**). *Gestalt Psychology. An Introduction to New Concepts in Modern Psychology* (Liveright Publishing Corporation, New York).

Lakoff, G., and Johnson, M. (**1980**). *Methapors We Live By* (University of Chicago Press, Chicago, IL).

Ludwig, V. U., and Simner, J. (**2013**). "What colour does that feel? Tactile–visual mapping and the development of cross-modality," Cortex **49**, 1089–1099.

Lupyan, G., and Casasanto, D. (**2015**). "Meaningless words promote meaningful categorization," Lang. Cognit. **7**, 167–193.

Margiotoudi, K., and Pulvermüller, F. (**2020**). "Action sound–shape congruencies explain sound symbolism," Sci. Rep. **10**, 12706.

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

Meng, X. L., Rosenthal, R., and Rubin, D. B. (**1992**). "Comparing correlated correlation coefficients," Psychol. Bull. **111**, 172–175.

Newman, S. S. (**1933**). "Further experiments in phonetic symbolism," J. Psychol. **45**, 53–75.

Newmeyer, F. J. (**1992**). "Iconicity and generative grammar," Language **68**, 756–796.

Nielsen, A., and Rendall, D. (**2011**). "The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon," Can. J. Exp. Psychol. **65**, 115–124.

Ohala, J. J. (**1984**). "An ethological perspective on common cross-language utilization of *F*0 of voice," Phonetica **41**, 1–16.

Ortega, G., and Özyürek, A. (**2020**). "Systematic mappings between semantic categories and types of iconic representations in the manual modality: A normed database of silent gesture," Behav. Res. Methods **52**, 51–67.

Ozturk, O., Krehm, M., and Vouloumanos, A. (**2013**). "Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds," J. Exp. Child Psychol. **114**, 173–186.

Perniss, P., Thompson, R. L., and Vigliocco, G. (**2010**). "Iconicity as a general property of language: Evidence from spoken and signed languages," Front. Psychol. **1**, 227.

Perniss, P., and Vigliocco, G. (**2014**). "The bridge of iconicity: From a world of experience to the experience of language," Philos. Trans. R. Soc. B. **369**, 20130300.

Ramachandran, V. S., and Hubbard, E. M. (**2001**). "Synaesthesia—A window into perception, thought and language," J. Conscious. Stud. **8**, 3–34.

R Core Team (**2024**). "R: A language and environment for statistical computing" (*R* Foundation for Statistical Computing, Vienna, Austria), available at https://www.R-project.org/ (Last viewed September 16, 2024).

Saji, N., Akita, K., Kantartzis, K., Kita, S., and Imai, M. (**2019**). "Cross-linguistically shared and language-specific sound symbolism in novel words elicited by locomotion videos in Japanese and English," PLoS One **14**, e0218707.

Sapir, E. (**1929**). "A study in phonetic symbolism," J. Exp. Psychol. **12**, 225–239.

Sidhu, D. M., and Pexman, P. M. (**2018**). "Lonely sensational icons: Semantic neighbourhood density, sensory experience and iconicity," Lang. Cogn. Neurosci. **33**, 25–31.

Spence, C. (**2011**). "Crossmodal correspondences: A tutorial review," Atten. Percept. Psychophys. **73**, 971–995.

Stokoe, W. (**2002**). *Language in Hand: Why Sign Came Before Speech* (Gallaudet University Press, Washington, DC).

Voeltz, F. K. E., and Kilian-Hatz, C. (**2001**). *Ideophones* (Benjamins, The Netherlands).

Walker, L., and Walker, P. (**2016**). "Cross-sensory mapping of feature values in the size–brightness correspondence can be more relative than absolute," J. Exp. Psychol. Hum. Percept. Perform. **42**, 138–50.

Yan, M., Luo, Y., and Inhoff, A. W. (**2014**). "Syllable articulation influences foveal and parafoveal processing of words during the silent reading of Chinese sentences," J. Mem. Lang. **75**, 93–103.

Yao, B. (**2021**). "Mental simulations of phonological representations are causally linked to silent reading of direct versus indirect speech," J. Cognit. **4**, 1–14.