Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Cross-Linguistically Shared and Language-Specific Sound Symbolism for Motion: An Exploratory Data Mining Approach

Permalink

https://escholarship.org/uc/item/2s01d8pf

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN

1069-7977

Authors

Saji, Noburo
Akita, Kimi
Imai, Mutsumi
et al.

Publication Date

2013

Cross-Linguistically Shared and Language-Specific Sound Symbolism for Motion: An Exploratory Data Mining Approach

Noburo Saji (nons@sfc.keio.ac.jp)
Japan Society for the Promotion of Science/Keio University, 4-1-1 Hiyoshi, Kohoku
Yokohama, Kanagawa 223-8521 Japan

Kimi Akita (akitambo@lang.osaka-u.ac.jp)
Studies in Language and Culture, Osaka University, 1-8 Machikaneyama-cho
Toyonaka, Osaka 560-0043 Japan

Mutsumi Imai (imai@sfc.keio.ac.jp)
Faculty of Environment and Information Studies, Keio University, 5322 Endo
Fujisawa, Kanagawa 252-0882 Japan

Katerina Kantartzis (k.f.kantartzis@bham.ac.uk), Sotaro Kita (s.kita@bham.ac.uk)
School of Psychology, University of Birmingham, Edgbaston
Birmingham B15 2TT UK

Abstract

This paper demonstrates a new quantitative approach to identify what is behind universally sensed sound symbolism and sound symbolism detected only by speakers of a particular language. We presented 70 locomotion videos to Japanese and English speakers and asked them to create a word that would sound-symbolically match each action, then to rate the action on five semantic dimensions. Multivariate analyses revealed that certain sound-meaning links (e.g., voicing and speed) were more consistent than others within and across languages. Language-specific sound symbolism was also found for some sound-meaning links (e.g., the affricate manner of articulation was associated with light motions in Japanese, but with heavy motions in English). This implies that cross-linguistically shared and language-specific parts of sound symbolism are intricately intertwined within each language. This research underscores the importance of a bottom-up approach which can exploratorily investigate the complex sound-symbolic systems as a whole.

Keywords: sound symbolism; mimetics; canonical correlation analysis

Introduction

Traditional linguistics has long assumed that the relationship between the form and meaning of a word is arbitrary (de Saussure, 1916/1983). However, words whose forms are motivated by their meanings (i.e., sound-symbolic words) are widely found across languages. For example, *bump* and *thump* sound like what they mean: events with an abrupt end (Firth, 1935/1957). Some languages have a large lexical class of sound-symbolic words called "ideophones," "expressives," or "mimetics" (Voeltz & Kilian-Hatz, 2001; Kita, 1997). For example, Japanese is rich in not only onomatopoeic (e.g., *piyopiyo* 'tweet-tweet') but also non-onomatopoeic mimetic words (e.g., *tobotobo* 'plodding').

Sound symbolism is not limited to ideophones and mimetics. Sapir (1929) points out that English speakers associate novel words containing the vowel /i/ with smallness more frequently than words containing /a/. Another celebrated example of sound symbolism is the association between sonorancy and roundness (Köhler, 1929/1947). It has been repeatedly observed that speakers of many languages prefer a round shape for *maluma* and an angular shape for *takete* (Brenner, Caparos, Davidoff, Fockert, Linnell, & Spence, 2013; Davis, 1961; Holland & Wertheimer, 1964).

Thus, there has been accumulating evidence that language does contain some non-arbitrary sound-meaning correspondences and people are sensitive to them. However, the exact nature of sound symbolism has not been fully clarified and one of the most important questions about sound symbolism is still open: what sound-meaning associations are shared by speakers of different languages, and why? In fact, researchers have recognized that not every case of sound symbolism may be detected as commonly as *maluma/bouba* vs. *takete/kiki*.

For example, Iwasaki, Vinson, & Vigliocco (2007) examined whether English speakers can detect the meanings of some Japanese mimetics depicting motion events, by asking them to rate the mimetics on a set of semantic-differential scales (e.g., energetic vs. non-energetic; fast vs. slow). Iwasaki et al. demonstrated that English and Japanese speakers' ratings agreed on some dimensions but not others. Specifically, Japanese speakers associated mimetics starting with a voiced consonant with the meaning component of ''big person," and the mimetics with voiceless consonants with "feminine" and "formal" walking. English speakers agreed only with the former association (see also Haryu & Zhao, 2007 for the language-specific nature of magnitude-voicing symbolism).

Limitations of Comprehension Tasks

The question of universal and language-specific *facets* of sound symbolism has not been properly addressed or pursued in previous studies, mainly due to the nature of their experimental method. Most experimental studies on sound symbolism have aimed at detecting the universality of sound symbolism and mainly employed comprehension tasks, such as forced-choice and semantic-differential rating tasks. These experiments were designed to examine whether subjects can detect "correct" sound-meaning correspondences, or how they rate each sound or word on a predetermined set of semantic scales, such as size and brightness. These tasks are effective in the examination of particular sound-meaning associations. However, no one knows how many such associations—how many sound patterns, how many meaning dimensions, and how many combinations of sounds and meanings—we have to examine before we reach the whole picture of the sound-symbolic system of a language, let alone its universality.

The Present Study

The goal of the present research was to extract cross-linguistically shared and language-specific parts of sound symbolism and to give phonological or phonosemantic explanations to them. We approach this issue by examining intuitions for sound symbolism in Japanese speakers and English speakers. To rectify the above mentioned limitations in using comprehension tasks, we employed a production method in which participants were asked to make mimetic words that matched human locomotions in short video clips. This method would reveal an unlimited set of phonologically and phonotactically possible phoneme sequences available to the subjects. We then conducted a multivariate analysis which detects underling correlations between sounds and sounds, meanings and meaning, and sounds and meanings, and evaluates what sound-meaning correlations are more significant than others in Japanese and English. The comparison of the detected sound-meaning pairs in each language shows us the shared and language-specific sound symbolism.

We will present the Japanese and English speakers' data separately in Experiments 1 and 2, respectively.

Experiment 1: Japanese

Method

Materials We created 70 short video clips of various types of human locomotion (*M* = 7.3 sec, *SD* = 2.7). In each video, a person appeared from the left side of the monitor and moved to the right out of the frame in a certain manner of walking or running.

Participants and Procedure Ninety-three native Japanese speakers, all undergraduate students, participated in the experiment. They went through both an attribute rating task and a word creation task. The participants were first presented with the 70 video clips on a computer screen in random order and asked to evaluate them on five 11-point semantic-differential scales (from 1 to 11): "size" (big-small), "speed" (slow-fast), "weight" (heavy-light), "energeticity" (energetic-non-energetic), and "jerkiness" (jerky-smooth). After the rating task, the videos were shown again to the participants in a random order. They were asked to generate sound-symbolic words and type them on a keyboard.

Data Preparation For analysis, we excluded sound-symbolic words that were obviously made on the analogy of existent nouns and verbs (e.g., *robo-robo*, cf. *robotto* 'robot'). We also excluded the data obtained for the videos whose most common semantic rating was "6" (neutral), which we assumed to blur the rest of the data. A total of 1,442 mimetics were submitted for analysis. They were phonetically coded and listed with the rating scores. For phonetic coding, we limited ourselves to the first moras (C_1V_1) of the obtained mimetics, as they have been discussed to have particular sound-symbolic significance (Kawahara, Shinohara, & Uchimoto, 2008; see also Hamano, 1998). The coding scheme for consonants, shown with the one for vowels in Table 1 (the coding for English will be used in Experiment 2), is based on Bailey & Hahn (2005). The data is thus a 1,442 × 13 matrix, consisting of five semantic ratings and eight phonetic values for each mimetic.

Table 1: The coding scheme for phonetic features

	Japanese	English
C_1 place of articulation	labial (Lab), velar (Vel), alveolar (Alv), glottal (Glot), palate (Pal), dental (Dent)	labial, velar, alveolar, glottal, palate, dental
C_1 sonorancy	sonorant (Son), obstruent (Obs)	sonorant, obstruent
C_1 manner of articulation	stop (Stop), affricate (Aff), fricative (Fric), glide (Gld), flap (Flap)	stop, affricate, fricative, glide, lateral (Lat), nasal (Nas), rhotic (Rhot)
C_1 voicing	voiced, voiceless	voiced, voiceless
C_1 palatalization	palatalized, not palatalized	n/a
C_1 nasality	nasal, not nasal	n/a
V_1 height	high, mid, low	high, mid-high, mid-low, low
V_1 backness	front, central, back	front, central, back

Note: The abbreviations in parentheses will be used in Figure1 and Figure2.

Analysis and Results

Canonical Correlation Analysis We conducted a variant of canonical correlation analysis (CCA) designed for categorical variables (see Thompson, 2005 for its detailed algorism) developed by Van der Burg (1988). Generally speaking, CCA enables us to visualize an implicit structure underlying multiple datasets. In common with other multivariate analyses, such as principle component analysis, CCA attempts to explain all possible correlations in a low-dimensional space. While principle component analysis is applied to only one dataset, CCA investigates relationships among two or more different variable sets and derives estimates by applying weights to the variables.

In the current context, CCA examines all possible correlations both within and across the two sets of variables (i.e., the sound and meaning datasets). This means that we can explore not only sound-meaning associations but also sound-sound or meaning-meaning correlations simultaneously, not limiting ourselves to a predetermined set of sound-meaning pairs. Notice that this analytical method is meaningful due to the very nature of sound symbolism, in which sound and meaning are intertwined with each other.

The Consistency of Sound-Meaning Associations The data matrix was fed into the program for canonical correlation analysis packaged in IBM SPSS Statistics 20 (IBM, 2012). We employed a two-dimensional solution, as the canonical correlation values of the first and second dimensions, which represent the latent correlations between the canonical variable of sounds and that of meanings, were significantly high (*r*s = .56 (first dimension) and .25 (second dimension), *p*s < .001). These values guarantee consistent sound-meaning associations in the two dimensions, indicating systematic sound symbolism in the present free production experiment.

The Focal Sound-Meaning Associations in the Sound-Symbolic System of Japanese To examine how sound and meaning are correlated in the present dataset, we considered the component loadings of each variable (see Table 2). As in principle component analysis, component loadings represent the correlation between the data and the extracted dimensions; each absolute value approximates the importance of the variables on each dimension.

Table 2: Component loadings of canonical correlation analysis in Japanese

Dataset	Variable (positive – negative)	Dimension 1	Dimension 2
Meaning	Size (large – small)	.40	−.36
	Speed (slow – fast)	.56	.39
	Weight (heavy – light)	.85	.07
	Energeticity (energetic – non-energetic)	−.21	−.56
	Jerkiness (jerky – smooth)	.31	−.35
Sound	C_1_place	.05	.17
	C_1_sonorancy	.36	.26
	C_1_manner	.05	−.42
	C_1_voicing	.74	−.24
	C_1_palatalization	−.43	−.05
	C_1_nasality	.38	.29
	V_1_height	.05	−.40
	V_1_backness	.28	.33

Table 2 shows that the semantic attribute “weight” in the meaning group and the phonetic feature “C_1 voicing” in the sound group obtained high positive loadings on Dimension 1 (.85 and .74, respectively). This suggests that the voicing-weight association was critically important in Japanese sound symbolism for motion. Noteworthy contributions were also observed for “speed” (.56), “size” (.40), and “jerkiness” (.31) among the meaning features and “C_1 palatalization” (−.43), “C_1 nasality” (.38), and “C_1 sonorancy” (.36) among the sound features. The four semantic variables were positively correlated. Heavy, large, slow, and jerky (or light, small, fast, and smooth) manners of motion tended to appear together, corresponding to the four consonantal features. On the other hand, in Dimension 2, “speed” (.39) and “V_1 backness” (.33) obtained high positive absolute values, while “size” (−.36), “energeticity” (−.56), “jerkiness” (−.35), “C_1 manner” (−.42), and “V_1 height” (−.40) obtained high negative absolute values. This suggests that the correspondences between this set of consonantal and vocalic features and slow, small, non-energetic, and smooth (or large, fast, energetic, and jerky) manners have the second most important status in Japanese sound symbolism for motion.

Details of the Sound-Meaning Associations The loading scores tell us which variables (e.g., manner of articulation) play a primary role in the discrimination of the dimensions, but it does not specify how much individual values in each variable (e.g., “affricate” in manner of articulation) contribute to those dimensions. We therefore computed the centroids of object scores for the semantic and phonetic values (see Van der Burg, 1988 for the details of this algorism). Specifically, each point in Figure 1 represents the weight of each value on the two dimensions. Note that the figure only shows sound variables for the sake of clarity; relevant meaning variables are indicated in the dimension labels, based on their loading scores above. The all abbreviations in Figure 1 is corresponds to those in Table1.

First, it is evident that the “voiced” and “voiceless” points are contrastively located in the positive and negative sides of Dimension 1, respectively. This is consistent with the large contribution of the voicing feature to this dimension in component loading. Moreover, Figure 1 reveals large positive contributions of the two phonetic values, “nasal” and “sonorant,” to the same dimension, although the component loadings of the “C_1 nasality” and “C_1 sonorancy” variables were not as large as that of “C_1 voicing.” These coordinates indicate that voiced consonants that are nasal and sonorant (i.e., [m], [n], as in *moji* and *noro*) have particular significance in Dimension 1. In contrast, the negative half of Dimension 1 features the voiceless obstruent that is “palatalized” and “affricate” (i.e., /ty/, realized as [tʃ]) as a sound that is strongly associated with a small, fast, light, smooth motion (e.g., *tyoko*).

Dimension 2 also shows clear contrasts for the variables which received high loading scores in Table 2: “fricative” and “affricate” (C_1 manner), “high” and “low” (V_1 height), and “back” and “central” (V_1 backness). Each of these contrasts is paired with a set of positive (slow, small, non-energetic, smooth) or negative semantic values (fast, large, energetic, jerky) in Figure 1. The positive half of the same figure also contains “glottal,” “palatal,” “sonorant,” and "nasal." These results allow us to think of particular phones to be relevant to the present case of sound symbolism, such as [h(j)] (fricative, glottal, (palatal)), [m] and [n] (nasal), and [ɯ] (high, back) (e.g., *heto*, *hura*, *moso*). Similarly, in the negative half of Dimension 2, [tʃ] and [ts] (affricate) and [a] (central, low) are associated with large, fast, energetic, and jerky manners of motion (e.g., *tyaki*).

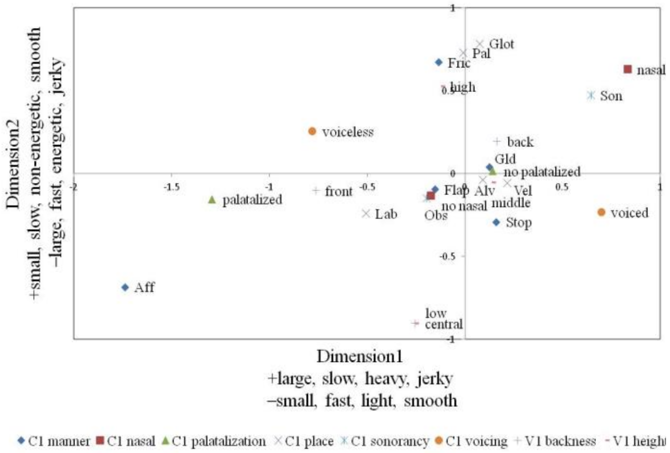

Figure 1: Category centroids for individual phonetic values in Japanese (See Table 1 for the explanations for the abbreviations)

Experiment 2: English

Method

Participants and Materials Twenty-seven English speakers at University of Birmingham, UK, participated in the experiment. The same 70 videos as we used in Experiment 1 were used as stimuli.

Procedure As in Experiment 1, English participants first saw randomly presented videos and were asked to rate them on five semantic dimensions. After the rating task, they watched the videos again and produced sound-symbolic words to depict the human motions shown in the video clips. Unlike Experiment 1, however, the participants were instructed to create $C_1V_1C_2V_2$ words that intuitively (or "sound-symbolically") matched the motions. This change was made because English speakers were not likely to be familiar with the notion of mimetics.

Data Preparation The data went through the same noise exclusion procedure as in Experiment 1. 1,227 "mimetic" words were retained for analysis. The C_1V_1 of each mimetic was phonetically coded according to the scheme in Table 1. Thus, the resultant data matrix consisted of 1,227 rows of mimetics and 8 columns of phonetic/evaluative features.

Analysis and Results

The Consistency of Sound-Meaning Associations Non-linear canonical correlation analysis was conducted with the English data matrix. We adopted a two-dimensional solution. The canonical correlation values for Dimensions 1 and 2 were .17 and .15, respectively ($ps < .01$). These values were substantially lower than their Japanese equivalents, indicating that the associations between the sound and meaning datasets in English are relatively weaker than those in Japanese. This may suggest that Japanese speakers have a better established sound-symbolic sense than English speakers due to the existence of the sound-symbolically systematized lexical class of mimetics in Japanese.

The Focal Sound-Meaning Associations in Sound-Symbolic System of English The component loadings of each variable are listed in Table 3. It shows that "size" (−.40), "speed" (.56), "energeticity" (−.62), "C_1 voicing" (.58), and "V_1 height" (−.39) obtained high absolute values in Dimension 1, while "weight" (.47), "energeticity" (−.32), "jerkiness" (−.46), and "C_1 place" (−.70) were heavily weighted in Dimension 2. Thus, Dimension 1 is associated with small, slow, non-energetic motions, and Dimension 2 with heavy, non-energetic, smooth motions.

Table 3: Component loadings of canonical correlation analysis in English

Dataset	Variable (positive – negative)	Dimension 1	Dimension 2
Meaning	Size (large – small)	−.40	.12
	Speed (slow – fast)	.56	.06
	Weight (heavy – light)	−.11	.47
	Energeticity (energetic – non-energetic)	−.62	−.32
	Jerkiness (jerky – smooth)	−.09	−.46
Sound	C_1_place	.04	−.70
	C_1_sonorancy	.27	.07
	C_1_manner	.18	−.19
	C_1_voicing	.58	−.03
	V_1_height	−.39	−.15
	V_1_backness	.07	.12

Details of the Sound-Meaning Associations Figure 2 plots the centroids of object points, which indicate how each value of the sound/meaning categories was weighted. Dimension 1 is clearly divided by the two phonetic features "C_1 voicing" and "V_1 height," with "voiced" and "mid-low" being positive and "voiceless" and "high" being negative. The figure also contains "nasal," "lateral," "rhotic," and "sonorant" in the positive area, suggesting that [n], [l], and [r], as in *medi*, *lela*, and *reso*, are strongly connected with small, slow, non-energetic motion. Likewise, the negative domain contains a voiceless glottal fricative (i.e., [h], as in *hali*), which was associated with large, fast, energetic motion.

Dimension 2 exhibits a wide distribution of the places and manners of articulation. A marked contrast is observed between the two positive phonetic features "glottal" and "affricate" and the three negative ones "palatal," "velar," and "glide." Among these features, "glottal" and "affricate" can be unambiguously identified as [h] (e.g., *hopi*) and [tʃ], respectively, which are linked with heavy, non-energetic, smooth motion. Similarly, the combination of “palatal” and “glide” is synonymous with [j].

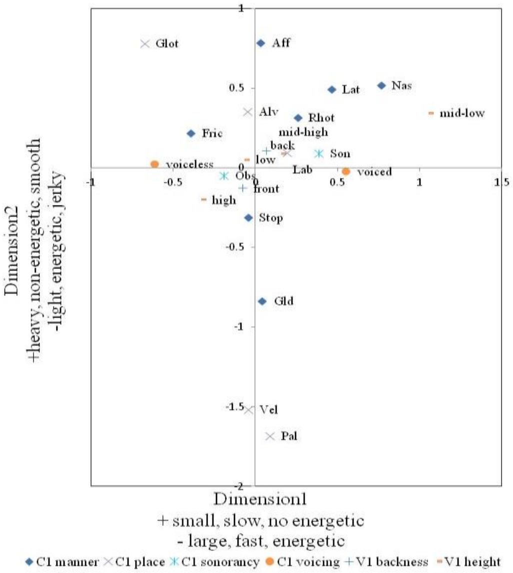

Figure 2: Category centroids for individual phonetic values in English (See Table 1 for the explanations for the abbreviations)

General Discussion

Comparing the detected sound-meaning correlations in Japanese and English shows us the shared and language-specific sound symbolism. Table 4 summarizes the sound-symbolic mappings found in the two languages, which shows what sound and meaning components have priority in motion sound symbolism of the two languages.

First, the results suggest that the two languages share in a large part a set of “sound-symbolically relevant” phonetic features. For example, both languages utilized the phonetic features “sonorancy,” “voicing,” “nasality,” and “vowel height,” and specific phonetic values “glottal,” “palatal,” “affricate,” and “fricative.” It should be noted here that some phonetic values, such as “alveolar,” “labial,” and “stop,” did not make a large contribution in the present data. This might reflect the unmarked nature of these sounds in the phonological systems of the two languages. In the present data, alveolar, labial, and stop consonants were found in 66%, 16%, and 61% of all Japanese mimetics and 51%, 28%, and 37% of all English mimetics, respectively.

Second, the two languages share many semantic features in their primary sound symbolism. Most notably, both of them use “weight” and “energeticity” as the most significant semantic features in sound symbolism of manner of motion. The two features are correlated with “size” and “speed” (see Tables 2 and 3).

Thus, speakers of Japanese and English use a similar set of phonetic and semantic features in sound symbolism of locomotion. However, these similarities do not directly mean that English and Japanese speakers mapped these sounds and meanings in the same way. They shared the most important sound-symbolic mapping: the voicing-speed mapping in the primary dimension. This can be accounted for by the long VOT (voice onset time) of voiced consonants, which appears to be readily mapped to the long duration of slow motion. Phonosemantic descriptions in the literature support this interpretation (Hamano, 1998; Tamori & Schourup, 1999). Further, the present study revealed that this sound-symbolic effect is especially strong in nasals (i.e., [m], [n]).

Table 4: Sound-meaning associations obtained in the two experiments

Dimension	Japanese	English
Dimension 1	heavy, **slow**, jerky, large **voiced** , **nasal + sonorant** ,	non-energetic, **slow**, small **voiced** , **nasal + sonorant**, lateral , rhotics , mid-low
	light, **fast**, smooth, small **Voiceless**, palatalized + affricate	energetic, **fast,** large, **voiceless** , glottal + fricative, high
Dimension 2	small, slow, non-energetic, smooth glottal + fricative, palatal + fricative, nasal + sonorant, high + back	heavy, non-energetic, smooth glottal, affricate
	large, fast, energetic, jerky central, low vowel, affricate	light, energetic, jerky palatal + glide, velar

Note: Sound-meaning associations shared by the two languages are given in **boldface**.

The present results also established the presence of language-specific sound symbolism. Most strikingly, Japanese and English speakers mapped some sounds to opposite meanings. For example, Japanese speakers associated the palato-alveolar affricate [tʃ] with light, fast, smooth, small motion in the primary dimension, whereas English speakers linked it to heavy, non-energetic, smooth motion in the secondary dimension. Likewise, the high back vowel /u/ was connected to slow, non-energetic motion in Japanese, but to fast, energetic motion in English.

These disagreements may be explained by the cross-linguistic differences in the phonological status of these sounds. First, in Japanese, the phone [tʃ] often appears secondarily, in a palatalized environment (i.e., /ty/), whereas this is not the case in English. Moreover, another affricate in Japanese (i.e., [ts]) is analyzed into [t] and [s] in English. Second, /u/ is realized as [ɯ] (unrounded) in Japanese, but as [ʊ] (rounded) or [u(:)] (rounded) in English. This cross-linguistic contrast in the roundedness of high back vowels suggests an articulation-based link between roundedness and energetic (hence, fast) motion. Thus, the present comparative observation illustrates the possibility that at least some parts of language-specific sound symbolism may be accounted for in terms of phonological typology. This possibility has been assumed widely in the literature, yet has not been much investigated.

Our study provides some important insights for theories of sound symbolism. We revealed that the sense of sound symbolism is realized in a complex system, which involves both universality and language-specificity. Sound symbolism is often alluded as "phonetic iconicity," but despite the name, this linguistic phenomenon is subject to a certain degree of arbitrariness, which originates from our language experience (Brenner et al., 2013 for a similar discussion). Our holistic and exploratory approach greatly contributes to the clarification of the complexity of iconic and arbitrary mappings in sound symbolism.

Acknowledgements

We are grateful to Shigeto Kawahara for his insightful and helpful comments from the perspective of phonetics and phonology. We also thank Farzana Bhaiyat, Alicia Griffiths, Junko Kanero, and Yayoi Tajima, who participated in the current project as annotators. Remaining inadequacies are ours. This research was supported by MEXT/JSPS KAKENHI to Saji (#232913), Akita (#24720179), Imai (#15300088, #22243043, #23120003), and Biological Sciences Research Council'[curly apostrophe]s Research Development Fellowship [BB/G023069/1] to Kita..

References

Bailey, T., & Hahn, U. (2005). Phoneme similarity and confusability. *Journal of Memory and Language,* 52, 339–362.

Brenner, A. J., Caparos, S., Davidoff, J., Fockert, J., Linnell, K. J., & Spence, C. (2013). "Bouba" and "kiki" in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition*, 126, 165–172.

Davis, R. (1961). The fitness of names to drawings. *British Journal of Psychology,* 52, 259–268.

de Saussure, F. (1916/1983). *Course in General Linguistics* (trans. R. Harris). London: Duckworth.

Firth, J. R. (1935/1957). The use and distribution of certain English sounds. In *Papers in Linguistics 1934–1951,* 34–46. London: Oxford University Press.

Hamano, S. (1998). *The Sound-Symbolic System of Japanese*. Stanford: CSLI Publications.

Haryu, E., & Zhao, L. (2007). Understanding the symbolic values of Japanese onomatopoeia: Comparison of Japanese and Chinese speakers. *Japanese Psychological Research*, 78, 424–432.

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

IBM Corp. (2012). IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007). What do English speakers know about *gera-gera* and *yota-yota*? A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese-Language Education around the Globe,* 17, 53–78.

Kawahara, S., Kazuko S., & Uchimoto, Y. (2008). A positional effect in sound symbolism: An experimental study. *Proceedings of the Eighth Annual Meeting of the Japanese Cognitive Linguistics Association,* 8, 417–427.

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

Köhler, W. (1929/1947). *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*. New York: Liveright.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12, 225–239.

Tamori, I., & Schourup, L. (1999). *Onomatopoeia: Its form and meaning*. Tokyo: Kurosio Publishers.

Thompson, B. (2000). Canonical correlation analysis. In L. Grimm & P. Yarnold (eds), *Reading and Understanding More Multivariate Statistics,* 285–316. Washington, DC: American Psychological Association.

Van der Burg, E. (1988). Homogeneity analysis with *k* sets of variables: An alternating least squares method with optimal scaling features. *Psychometrika,* 53, 177–197.

Voeltz, F. K. E., & Kilian-Hatz, C. (eds) (2001). *Ideophones*. Amsterdam/Philadelphia: John Benjamins.