

A Dictionary-based Quantitative Analysis of the Sound-symbolic System of Japanese Ideophones

Hinano Iida (iida.hinano.i1@s.mail.nagoya-u.ac.jp)

Department of English Linguistics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya
Aichi 464-8601 Japan

Ryo Hori (hor.ryo.w9@s.mail.nagoya-u.ac.jp)

Department of Social Informatics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya
Aichi 464-8601 Japan

Abstract

Ideophones are characterized by sound symbolism, a non-arbitrary relationship between speech sounds and meaning, which is observed across languages worldwide. Despite the increasing interest in sound symbolism and ideophones in cognitive science, Japanese ideophone research has lacked a robust quantitative approach. This paper proposes a dictionary-based method to quantify the sound-symbolic system of Japanese ideophones and presents statistical results that support and extend previous findings. The proposed method aims to advance the objective and comprehensive analysis of ideophones, sound symbolism, and iconicity, offering valuable contributions to ongoing discussions in the cognitive sciences.

Keywords: sound symbolism; ideophones; Japanese; dictionary; quantitative approach; iconicity; systematicity

Introduction

Ideophones, also known as “sound-symbolic words” (e.g., Tamamura, 1984), are words characterized by sound symbolism as a key feature, which abound in numerous languages of the world (Dingemanse, 2012). Sound symbolism refers to the non-arbitrary relationship between speech sounds and meaning (Hinton et al., 1994). As a form of iconicity or non-arbitrariness of language, sound symbolism has been extensively studied in both psychology and linguistics (e.g., Köhler, 1947; Ramachandran & Hubbard, 2001; Sapir, 1929), and it continues to be an important area of interest in the cognitive sciences (Winter, Woodin, & Perlman, 2026).

Despite the sustained and growing interest in sound symbolism and ideophones across linguistics, psychology, and cognitive science more broadly, Japanese ideophone research—particularly concerning the sound-symbolic system—often lacks a robust quantitative, objective, and statistical perspective. Regarding the analysis of the sound-symbolic system, Hamano’s (1986, 1998, 2014) introspective generalization is well-known and has been regarded as a “bible” for Japanese ideophone researchers (Akita, 2020). For example, the sound /k/ is associated with ‘hard’, as exemplified by the ideophones such as *kachikachi* ‘hard’. In fact, recent experimental studies have posited that language-specific iconicity may derive from the lexical patterns or sound-symbolic systems of Japanese ideophones based on Hamano’s observations (e.g., Akita & McLean, 2021; Akita et al., 2024; Iida & Funakura, 2024).

Considering the current state of research in this field, the purpose of this paper is to propose a dictionary-based quantification method for the sound-symbolic system of Japanese ideophones and provide quantitative results that support the claims made in previous research. The method presented here will enable us to further refine and expand the objective and quantitative analysis of ideophones, sound symbolism, and iconicity, offering valuable insights for advancing the field.

Previous Research

The most influential analysis of the sound-symbolic system of Japanese ideophones is Hamano’s generalization (1986, 1998, 2014) (Akita, 2009, 2020). Hamano’s research identifies a set of semantic features for each consonant and vowel (i.e., sound symbolism) that constitute ideophones, based on dictionary entries, actual data from novels, and a diachronic perspective. For example, velar stops /k, g/ are associated with ‘hard surface’, and the vowel /a/ is linked to meanings like ‘large’ (Hamano, 1986, 1998), ‘flat’, and ‘wide’ (Hamano, 2014). This method can be considered one of the most comprehensive sound-symbolic descriptions based on introspection (Akita, 2020). However, the study lacks a quantitative perspective, such as providing a specific number of ideophones associated with ‘hard’. In fact, there are many ideophones, such as *kirakira* ‘shining’ and *giragira* ‘glittering’, that contain /k, g/ as initial consonants but have no connection to ‘hard’ or even convey the opposite meaning, such as *kunyakunya* ‘soft’.

Hamano’s studies have had significant influence in not only linguistic but also psycholinguistic experimental studies, which investigate the perception of sound symbolism by language speakers—an area that has attracted increasing attention (e.g., Akita & McLean, 2021; Akita et al., 2024; Kumagai, Uno, & Shinohara, 2021). In this field, the language-specific perception of sound symbolism is often assumed to derive from the lexical patterns (i.e., sound-symbolic systems) of ideophones, such as the contrast between voicing based on Hamano’s observations. Now, as interest in this field continues to increase, it is crucial to rigorously quantify the sound-symbolic structure within ideophones and to solidify the discussions.

Some previous studies have attempted to quantify the sound-symbolic system. For example, Akita (2020) investigated how representative morphological and

phonological properties of Japanese ideophones together constitute their sound-symbolic system, using dictionary data and multi-factorial analysis to reveal the crucial role of language-specific features. However, the study employed broad categories, such as sound, movement, visual patterns, and inner feelings, for its analysis (cf. Dingemanse, 2012). A different approach is needed to discuss sound symbolism in line with psycholinguistic research.

Another quantitative approach to sensory meanings of Japanese ideophones involves semantic coding methods, such as sensory ratings and semantic differential ratings by native speakers (Doizaki, Watanabe, & Sakamoto, 2016; Iida & Akita, 2023; Yaguchi, 2021). For example, Yaguchi (2011) collected perceptual strength norms (auditory, visual, haptic, olfactory, gustatory) for 47 ideophones in Japanese. A more fine-grained semantic coding method is semantic differentials (e.g., Doizaki et al., 2016), where native Japanese speakers rate ideophones using 26 pairs of adjectives related to tactile qualities (e.g., smooth–rough, warm–cold, thick–thin, good–bad) to estimate impressions of tactile-related ideophones. However, since their quantitative approach focuses on words belonging to particular categories (e.g., tactile, gustatory (Sakamoto & Watanabe, 2018), or personality (Sakamoto, Watanabe, & Yamagata, 2021) and did not address the features discussed in psycholinguistic studies (e.g., place of articulation), it is not possible to capture the entire sound-symbolic system. We could use the same methodology as theirs, but it may be challenging due to the need for a large number of Japanese native speakers, as Japanese has thousands of ideophones (e.g., Kakehi, Tamori, & Schourup, 1996; Ono, 2007).

In this paper, we propose a novel dictionary-based quantitative analysis of Japanese ideophones. Unlike previous approaches, our method allows for the quantification of the entire sound-symbolic system of Japanese ideophones, integrating both phonological and semantic data in a comprehensive manner. By using a structured, data-driven approach, we offer a more systematic and scalable method for analyzing the relationships between sound and meaning (see also Westbury et al. (2018) for a similar systematic approach to English, using nonwords). This quantification method not only contributes to ongoing discussions of language-specific aspects of iconicity but also enriches cross-linguistic and diachronic studies of ideophones based on dictionary data (see also the Discussion).

Method

Dataset

To quantify the sound-symbolic system of Japanese ideophones, we first developed an automated data generation system using Python (version 3.11.11) and its libraries (Iida & Hori, 2024). Our original dataset is based on Kadooka

(2007), which includes a list of 1,648 ideophones, primarily derived from Kakehi et al. (1996). For semantic definitions, we mainly relied on the English definitions found in Kakehi et al. (1996). In cases where certain ideophones were not listed in the dictionary, we sourced definitions from Ono (2007) and translated them into English. 438 of the ideophones (27.6%) are polysemous, meaning they have multiple meanings (e.g., *tsuntsun* can refer to “the state in which something small is sharp”, “the state of feeling a sharp sensation in the nose”, or “the state of appearing proud, affected, or sullen”). As a result, the dataset contains 2,279 rows, with two¹ columns: the Japanese ideophones (written in Japanese characters) and their corresponding English definitions.

The automated data generating system decomposes each ideophone into smaller meaningful units identified in the previous research. We first automatically romanize each word using a romanization python library, *pykakasi* (version 2.3.0). We also enable the automatic extraction of morphological features (e.g., Akita, 2009; Hamano, 1998; Tamori & Schourup, 1999) based on the Japanese characters by pattern matching. The extraction includes identifying templates, such as reduplication (e.g., *kirakira* ‘glittering’), suffixation (/Q²/ (e.g., *kiraQ* ‘glistening’), /N/ (e.g., *kiraN* ‘glistening’), /ri/ (e.g., *kirari* ‘glistening’), vowel lengthening /R/ (usually in combination with /N/ or /Q/) (e.g., *kiraRN* ‘glistening’), CVCCVri form (e.g., *fuNwari* ‘fluffy’, *saQpari* ‘refreshing’), and altering form (e.g., *chirahora* ‘here and there’). After extracting templates, the system identifies roots (e.g., “kira” for *kirakira*, *kirari*; “po” for *poN*, *poRQ*) and root type (e.g., “kira” is bimoraic; “po” is monomoraic (Hamano, 1998)).

Then, it identifies phonemes by dividing monomoraic roots into (C₁)V₁ and bimoraic roots into (C₁)V₁C₂V₂. Considering that certain phonological properties are suggested to play a special role in the sound-symbolic system (Akita, 2009) and discussed in the previous experimental literature on sound symbolism (Akita & Mclean, 2021; Akita et al., 2024; Lockwood & Dingemanse, 2015; Saji et al., 2019; Shinohara & Kawahara, 2010), the phonemes are coded for phonetic and phonological features such as voicing (e.g., voiced /b, d, g/ vs voiceless /p, t, k/), obstruency (obstruent (e.g., /b, p, t, d, k, g, s, z/) vs sonorant (e.g., /m, n, j, w/)), manner of articulation (e.g., fricative /s, z/ , stop /p, t, k/), and place of articulation for consonants (e.g., bilabial /b, p/, alveolar /d, t/, velar /g, k/), and frontness (front /i, e/ and nonfront /u, o, a/), height (high /i, u/, mid /e, o/, and low /a/, and lip rounding (rounded /u, o/ and unrounded /i, e, a/) for vowels.

After the systematic automated coding processes, the system decomposes the English definitions word by word to establish correspondences between each sound and meaning. For example, the definition for *korokoro* “The state of being round and fat” is broken down into individual words: “The”, “state”, “of”, “being”, “round”, “and”, and “fat”. Each word

¹ Our dataset contains one additional column: semantic type (sound or non-sound) from Kakehi et al. (1996) which is not investigated here (see Iida & Hori (2024) for this discussion).

² The symbol *Q* represents a moraic segment that is realized as the first half of a geminate cluster in the middle of a word and as a glottal stop at the end of a word.

(i.e., meaning component; e.g., “round”) is paired with the information of the ideophone (e.g., $C_1 = k$ (voiceless, obstruent, stop, velar), $V_1 = o$ (nonfront, mid, rounded), template = reduplication)).

We remove inflections from words using the spaCy library (version 3.7.5) with the `en_core_web_sm` pipeline (version 3.7.1) in Python. This process ensures that variations like *lose*, *loses*, *losing*, and *lost* are all regarded as the base form “lose”. Additionally, when the definition contains a negation word (e.g., *not*, *no*, *without*), as in the definition “The manner of repeatedly cutting down or defeating **without** difficulty (*baQtabaQta*)” we manually add the prefix “NO_” to the word that is negated (in this case, NO_difficulty) to distinguish it from the ideophone that contains *difficulty*. These processes enable us to quantitatively analyze the sound-symbolic system of Japanese ideophones based on dictionary data.

Statistical Analysis

All statistical analyses reported in this paper used the automatically generated dataset described above and were conducted using R version 4.4.0 (R Core Team, 2024). All plots were generated using the `ggplot2` package version 3.4.2 (Wickham, 2016). For the current analysis, we employed logistic regression models and multinomial logistic regression models, a statistical approach designed for modeling categorical dependent variables. Logistic regression was used when the dependent variable has two categories (e.g., voicing (voiced vs. voiceless)), while multinomial logistic regression is applied when the dependent variable has more than two categories (e.g., place of articulation (e.g., alveolar, bilabial, glottal, palatal, postalveolar, or velar)). These models allow us to assess the relationship between various factors, such as phonetic features of consonants and vowels, morphological features (template and root), and meanings³.

Results

The obtained form–meaning correspondences in the present study provide quantitative evidence that supports the observations made in previous research. Moreover, the current approach and dataset allow us to go beyond a mere replication of previous studies, enabling us to uncover additional statistically significant form–meaning correspondences. In this study, we primarily focus on bimoraic ideophones ($N = 1,729$; 76%), which can be more systematically analyzed within the Japanese ideophone system in terms of sound symbolism⁴ (Hamano, 1998). Here, we concentrate on segmental sound symbolism, such as initial consonants and vowels, as well as morphophonological sound symbolism, particularly in relation to templates.

³ The whole dataset, R scripts, and Python scripts are available at: <https://osf.io/uwx2g/>

Voicing of Initial Consonants (C_1)

Voicing of initial obstruents is suggested to be the key sound-symbolic feature in Japanese ideophones. Voiced and voiceless obstruents are systematically associated with “heavy; large; coarse” and “light; small; fine” objects, respectively (Hamano, 1998). The following are the results in Figure 1.

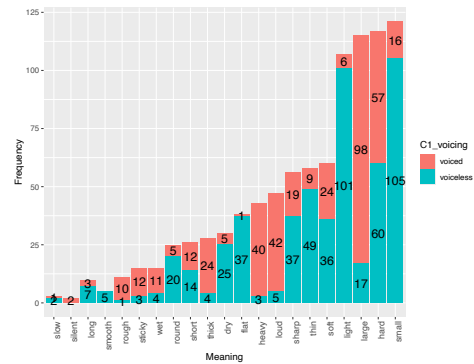


Figure 1: The frequency of voicing types by meaning among bimoraic ideophones with initial obstruents. Only the meanings that are related to voicing contrasts are shown in the figure.

Heavy vs. Light In line with Hamano’s (1998) observation, ideophones associated with the meaning ‘heavy’ (e.g., *zuQshiri* ‘heavy’) contain more initial voiced obstruents ($N = 40$) than voiceless obstruents (e.g., *pochaN* ‘the sound of a small and relatively heavy object falling into somewhat deep water’) ($N = 3$). In contrast, ‘light’ is more strongly associated with initial voiceless (e.g., *fuwafuwa* ‘light and fluffy’) ($N = 101$) than voiced obstruents (e.g., *gachigachi* ‘rustling of movement of light, thin, dry objects’) ($N = 6$). We used a logistic regression model to predict voicing types from meaning and the model revealed a statistically reliable difference between the two groups, indicating that ‘light’ is more strongly associated with voiceless obstruents than with voiced obstruents compared with ‘heavy’ ($b = -5.41$, $SE = 0.73$, $z = -7.40$, $p < .001$).

Large vs. Small Ideophones associated with the meaning ‘large’ (e.g., *doshiQ* ‘large, calm, and impressive’) contain more initial voiced obstruents ($N = 98$) than voiceless obstruents (e.g., *pokaN* ‘having a large, conspicuous hole’) ($N = 17$). In contrast, ‘small’ ideophones are more strongly associated with initial voiceless (e.g., *potsupotsu* ‘covered with small spots’) ($N = 105$) than with voiced obstruents (e.g., *botsubotsu* ‘having many small holes’) ($N = 16$). The model revealed a statistically reliable difference between the two groups, indicating that ‘small’ is more strongly associated with voiceless obstruents than with voiced compared with ‘large’ ($b = -3.63$, $SE = 0.38$, $z = -9.67$, $p < .001$).

⁴ According to Hamano (1998), both consonants and vowels in bimoraic ideophones make positional distinction of functions (e.g., C_1 (tactile information) vs C_2 (movement information)).

Coarse (Rough) vs. Fine (Thin or Sharp) To test Hamano’s (1998) claim on voicing and coarse vs. fine, we used the meanings of ‘rough’, ‘thin’, and ‘sharp’ used in the English definitions in our dataset. Ideophones associated with the meaning ‘rough’ (e.g., *zarazara* ‘rough to the touch’) contain more initial voiced obstruents (N = 10) than voiceless obstruents (e.g., *kasakasa* ‘rough, dry, and often brittle’) (N = 1). In contrast, ‘thin’ associated with initial voiceless (e.g., *perapera* ‘thin and cheap’) (N = 49) more than voiced obstruents (e.g., *gasagasa* ‘loud rustling sound made by light, thin, dry object’) (N = 9). Also, ‘sharp’ associated with initial voiceless (e.g., *togetoge* ‘sharp’) (N = 37) more than voiced obstruents (e.g., *zukiQ* ‘feeling a sharp pain’) (N = 19). The model also revealed a statistically reliable difference between the two groups, indicating that ‘thin’ and ‘sharp’ more strongly associated with voiceless obstruents than with voiced compared with ‘rough’ (thin: $b = -4.00$, $SE = 1.11$, $z = -3.60$, $p < .001$; sharp: $b = -2.97$, $SE = 1.09$, $z = -2.73$, $p < .01$).

Other Semantic Contrasts All the results reported above provide quantitative evidence supporting the introspective observations on voicing contrast in Hamano (1998). Additionally, we identified previously unreported correspondences between voiced sounds and ‘wet’ and between voiceless sounds and ‘dry’ in Japanese ideophone research (Winter et al., 2022). Ideophones associated with the meaning ‘wet’ (e.g., *bishobisho* ‘thoroughly wet’) contain more initial voiced obstruents (N = 11) than voiceless obstruents (e.g., *kasakasa* ‘rough, dry, and often brittle’) (N = 4). Similarly, ‘sticky’ associated with initial voiced (e.g., *betobeto* ‘sticky’) (N = 24) more than voiceless obstruents (e.g., *toroQ* ‘sticky and glutinous’) (N = 9). In contrast, ‘dry’ associated with initial voiceless (e.g., *kasakasa* ‘rough, dry, and often brittle’) (N = 25) more than voiced obstruents (e.g., *basabasa* ‘dry’) (N = 1). The logistic regression model revealed that ‘dry’ is more strongly associated with voiceless obstruents than with voiced compared with ‘wet’ and ‘sticky’ ($b = 2.62$, $SE = 7.62$, $z = 3.44$, $p < .001$; $b = 3.40$, $SE = 7.29$, $z = 4.30$, $p < .001$). Another correspondence is between voicing and thickness: ‘thick’ is more strongly associated with voiced than with voiceless compared to ‘thin’ ($b = 3.49$, $SE = 0.65$, $z = 5.36$, $p < .001$).

Importantly, there was no statistically reliable difference in voicing based on the contrast between hard and soft ($b = -0.39$, $SE = 0.32$, $z = -1.21$, $p > 0.1$) (cf. Kumagai et al., 2021), or between long and short ($b = -0.07$, $SE = 0.65$, $z = -0.11$, $p > 0.1$), which aligns with Hamano’s observations.

Place of Articulation of C₁ of bimoraic ideophones

According to the previous research, place of articulation of C₁ of bimoraic ideophones also plays an important role in the sound-symbolic system (Akita, 2009). Here, we focus on the

role of velar and bilabial consonants, which have been frequently discussed in previous studies.

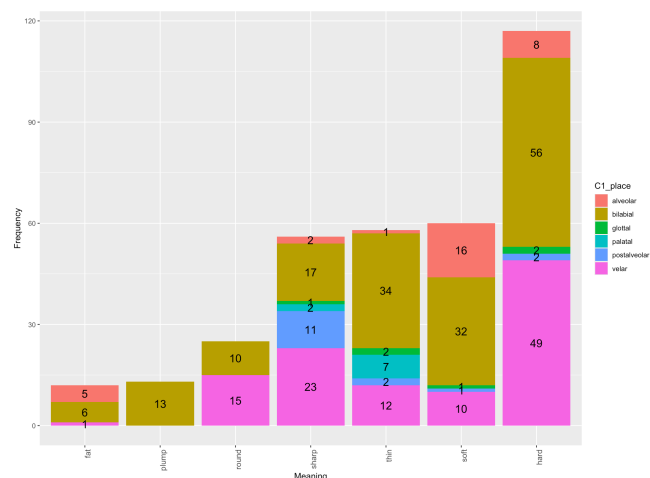


Figure 2: The frequency of each place of articulation by meaning among bimoraic ideophones with initial obstruents. The meanings discussed in the previous research in relation to place of articulation are shown.

Hard vs. Soft It is suggested that velar stops /g, k/ are associated with hardness (Hamano, 1998). In our dataset, ideophones associated with ‘hard’ contain more velar sounds C₁ (N = 49; 41.9%) (e.g., *kachikachi* ‘very hard’; *gachigachi* ‘hard or solid’) compared to those with ‘soft’ (N = 10; 16.7%) (e.g., *kunyakunya* ‘soft and relaxed’). We used a multinomial logistic regression model to predict place of articulation of C₁ from meaning. The model revealed that compared to ‘soft’, the velar strongly increases the likelihood of the meaning ‘hard’ ($b = 2.28$, $SE = 0.55$, $z = 4.11$, $p < .001$).

Thin vs. Fat According to Hamano (1998, 2014), bilabial /b, p/ is associated with ‘fat’. We find twelve ideophones associated with ‘fat’ and half of them contain bilabials (e.g., *bukubuku* ‘swollen with water or fat’; *pukupuku* ‘extremely fat’). However, we also find bilabial in ideophones associated with ‘thin’ (e.g., *perapera* ‘thin and cheap’, *pekapeka* ‘very thin, easily bent’, *berabera* ‘thin and usually cheap’) (N = 34; 58%). The model showed a statistically reliable difference between the proportion of bilabial between ‘thin’ and ‘fat’, indicating that ‘thin’ ideophones are more strongly associated with bilabial ($b = 3.34$, $SE = 1.18$, $z = 2.83$, $p < .01$). When we group ‘fat’ and ‘plump’ as a one category as FAT and construct another model, it did not show a statistical difference on bilabial based on ‘thin’ and ‘fat’ ($b = 2.19$, $SE = 1.13$, $z = 1.93$, $p > .05$). This finding may be taken as evidence against Hamano’s generalization on bilabial and fatness⁵⁶.

⁵ It can be possible that ‘thin’ is also related to ‘fat’ in terms of ‘soft’, ‘bendable,’ or ‘flexible,’ however, further discussion on this point is needed.

⁶ It is also possible that the semantic category of ‘fat’ is too narrowly defined and should instead be considered more broadly as representing ‘softness’—as an anonymous reviewer suggested.

Vowels

Next, we will look at the sound symbolism of vowels, which is well discussed in previous research on Japanese ideophones (Hamano, 1998) and experimental studies (e.g., Akita et al., 2024). Here, we focus on V_1 in both monomoraic and bimoraic ideophones (Figure 3).

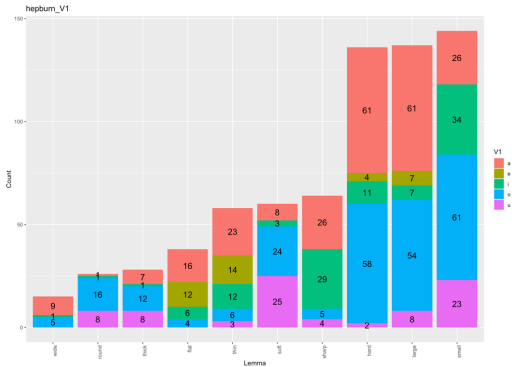


Figure 3: The frequency of each vowel by meaning among all ideophones. The meanings discussed in the previous research in relation to vowels are shown.

Large vs. Small Vowel size symbolism (/a/ is big; /i/ is small) is well-known in psychological studies (Sapir, 1929) and it is suggested to be true in the Japanese ideophones. Hamano (1998) associates /a/ with big and /u/ is small. As Figure 3 shows, ‘large’ ideophones have more /a/ as V_1 (e.g., *bachaQ* ‘large splashing sound’; *gabugabu* ‘drinking large quantities of liquid’) compared to ‘small’ ideophones (*kaQpokaQpo* ‘sound made by a small drum’). In contrast, ‘small’ ideophones contain /i/ (*piyopiyo* ‘small, shrill peeping sounds’; *chikachika* ‘small light flashing’). As for /u/, ‘small’ ideophones contain more /u/ than ‘large’ ideophones. We used a multinomial logistic model to predict V_1 from meaning. The model revealed that /i/ is more strongly associated with ‘small’ ($b = 1.46, SE = 0.46, z = 3.20, p < .05$), and /u/ is more strongly associated with ‘small’ ($b = 0.93, SE = 0.45, z = 2.07, p < .05$). It is also confirmed that /a/ is more strongly associated with ‘large’ ($b = 0.97, SE = 0.30, z = 3.25, p < .05$).⁷

Thick vs. Thin (Sharp) Hamano (2014) also suggests that /i/ is associated with ‘sharp’ or ‘thin’. We can see that ‘sharp’ and ‘thin’ ideophones have a large amount of /i/ (e.g., *chikuchiku* ‘sharp stinging sensation’; *pirapira* ‘thin and dry objects fluttering’ compared to other meanings. Strikingly, we can find only one example of ‘thick’ ideophones containing /i/ in our dataset (e.g., *biriQ* ‘sound of relatively thick being torn’). The differences are also statistically reliable: /i/ is more strongly associated with ‘sharp’ ($b = 4.24, SE = 1.14, z = 3.70, p < .001$) and ‘thin’ ($b = 3.18, SE = 1.15, z = 2.75, p < .01$).

⁷ A logistic regression model also revealed that ‘small’ is more strongly associated with the front vowels than with back vowels ($b = 1.00, SE = 0.34, z = 2.91, p < .005$).

⁸ The meanings ‘repeat’ and ‘repeatedly’ are included.

Hard vs. Soft Another property discussed in relation to vowel sound symbolism is ‘soft’ and /u/ (Akita et al., 2024). As Figure 3 shows, ‘soft’ ideophones have a larger amount of /u/ compared with other meanings (e.g., *fukafuka* ‘soft and fluffy’; *kunyaQ* ‘soft and easily bendable’). In contrast, we can find only two ideophones associated with ‘hard’ (e.g., *zuboN* ‘loud, resonant sound of heavy object striking a hard surface’). The model showed a statistically reliable difference between the two, suggesting that /u/ is more strongly associated with ‘soft’ than ‘hard’ ($b = 3.41, SE = 0.77, z = 4.40, p < .001$).

Template and Aspectual Meanings

It is suggested that morphological (especially template) features correlate to the aspectual meaning of ideophones. For instance, reduplication is assumed to be associated with repetition and continuation, whereas suffixed ideophones with punctual eventualities (e.g., Hamano, 1998).

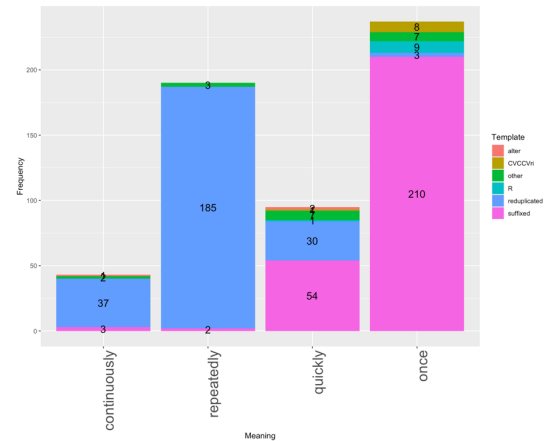


Figure 4: The frequency of each template by meaning among all ideophones. The meanings discussed in the previous research in relation to template are shown.

As Figure 4 shows, ideophones associated with repetition⁸ and continuity⁹ have more reduplicated forms (*pokapoka* ‘striking something repeatedly’; *nitanita* ‘grinning continuously’) compared with ‘once’¹⁰ ideophones (*waiwai* ‘many shouting at once’). In contrast, ‘once’ and ‘quickly’¹¹ ideophones have more suffixed forms (e.g., *piriQ* ‘pricking once’) than reduplicated forms (e.g., *batabata* ‘sound of quick footsteps’). We used a multinomial logistic model to predict template from meaning. The model revealed a statistically reliable result that ‘repeatedly’ ideophones are more strongly with reduplication compared to ‘once’ ($b = 5.00, SE = 0.91, z = 5.50, p < .001$) and ‘continuously’ ideophones are more strongly associated with reduplication compared to ‘quickly’ ($b = 1.46, SE = 0.84, z = 1.74, p < .01$).

⁹ The meanings ‘continue’, ‘continuous’, ‘continuously’, and ‘continual’ are included.

¹⁰ The meanings ‘once’ and ‘single’ are included.

¹¹ The meanings ‘quick’ and ‘quickly’ are included.

General Discussion

The current study provided a dictionary-based quantitative analysis of the sound-symbolic system of Japanese ideophones. We began by systematically generating form–meaning mappings based on dictionary data, which include a range of features, such as phonological, morphological, and semantic properties. This automated process enables a comprehensive analysis of several thousand ideophones, allowing us to move beyond reliance on intuition and conduct more systematic and objective research.

We further validated the effectiveness of our approach by presenting results based on the generated dataset. We investigated the form–meaning correspondences discussed in the previous research, including voicing of initial consonants, vowels, and morphological templates. Some of our findings provide quantitative and statistical evidence supporting prior discussions, while others contradicted previous claims. Furthermore, we found other sound symbolism, such as the association of voicing with dryness and with thickness. These associations might be difficult to uncover through traditional methods or with limited human capacity. This quantitative approach contributes to the ongoing discussion of language-specific sound symbolism, which is often attributed to individual languages’ lexical patterns.

Nonetheless, future research needs to extend the current study. For example, we did not examine the positional effects of consonants within the sound-symbolic system (Hamano, 1998). Additionally, we were unable to delve deeply into the potential division of labor between these and other features in the sound-symbolic system (Akita, 2020; Hamano, 1998). These aspects may be explored in future studies by applying multifactorial approaches, such as random forest (Kilpatrick, Ćwiek, & Kawahara, 2023; Winter & Perlman, 2021). This will enable us to uncover a more detailed and comprehensive understanding of the sound-symbolic system in Japanese ideophones.

There is also room for improvement regarding semantic granularity. First, we did not account for synonyms such as ‘hard’ and ‘firm’ because we could not find an objective way to group them at this stage. One possible solution is to use resources such as thesauri to address this issue. Second, we did not consider the semantic relatedness of meanings. For example, we treated gradable concepts in a dichotomous fashion, such as ‘small’ vs. ‘large’ and ‘heavy’ vs. ‘light’ whereas assuming a continuum may be more appropriate. The English definitions in our dataset contain adverbs that express degree (e.g., “relatively”, “very”, “completely”) which could be incorporated into the analysis of gradability. Moreover, we did not explore the relationships among meanings—for instance, the relationship between ‘heavy’ and ‘large’, or the polysemy of ideophones. Previous studies have suggested that Japanese voiced stops are associated with both ‘large’ and ‘dark’, with the two concepts being related (Akita & McLean, 2021; Akita et al., 2024). These aspects remain as areas for future research.

Another limitation of our approach is the reliance on a primarily single English-based dictionary dataset, which may

be heavily influenced by the language and the subjective descriptions of the dictionary compilers. This could lead to potential biases or limitations in the semantics, so it might be more beneficial to consult multiple dictionaries. Furthermore, it may be helpful to combine our results with those based on a dictionary and subjective ratings study by multiple native speakers (e.g., Doizaki et al., 2016) to determine whether we can achieve convergent results on the sound-symbolic system of Japanese ideophones.

As a direction for future work, an important aspect is the comparison between ideophones and non-ideophonic words. Ideophones are traditionally assumed to possess a highly systematic and cohesive sound-symbolic system (Hamano, 1998). However, recent extensive cross-linguistic comparisons of languages worldwide have also uncovered systematic form–meaning correspondences in non-ideophonic basic vocabulary items (Blasi et al., 2016; Erben Johansson et al., 2020; Joo, 2020), English size adjectives (Winter & Perlman, 2021), and English shape nouns (Sidhu et al., 2021). Approaches like ours can ultimately provide a quantitative answer to the question of whether the sound-symbolic system of ideophones is more systematic than that of non-ideophonic words.

Our approach can further be extended to contrastive or cross-linguistic analyses (including dialects and languages from around the world), as well as diachronic studies. While other semantic coding methods, such as semantic differentials and sensory ratings by native speakers, are also fine-grained and powerful, they are often difficult to apply to languages with a limited number of speakers or to historical languages (cf. McLean, Dunn, & Dingemans, 2023). Therefore, our dictionary-based quantification approach has the potential to make a significant contribution to the ongoing debate on the universality and specificity of sound symbolism, helping to identify which aspects of form–meaning correspondences tend to be universal or language-specific, and whether these associations are more easily changeable or more stable over time.

Conclusion

In this paper, we proposed a new dictionary-based approach to quantify the sound-symbolic system of Japanese ideophones. While there is room for improvement, we have demonstrated that our approach can make a meaningful contribution to the ongoing discourse on sound symbolism. We hope that future research will further refine and expand the objective and quantitative analysis of ideophones, sound symbolism, and iconicity.

Acknowledgments

We thank Kimi Akita and four anonymous reviewers of CogSci 2025 for their thought-provoking comments on earlier versions of this paper. Any remaining inadequacies are our own. This work was supported by the Graduate Program for Lifestyle Revolution based on Transdisciplinary Mobility Innovation.

References

- Akita, K. (2009). *A grammar of sound-symbolic words in Japanese: Theoretical approaches to iconic and lexical properties of mimetics*, Doctoral dissertation, Kobe University, Kobe.
- Akita, K., & McLean, B. (2021). Onsyootyooo chikaku no nichiei taisyoo: Imichizu ni yoru kentoo [A contrastive study of Japanese and English speakers' sound-symbolic perceptions: A semantic map approach]. In K. Shinohara & R. Uno (Eds.), *Jikken ninchi gengogaku no shinka* [Progress in experimental cognitive linguistics]. Tokyo: Hituzi Syobo.
- Akita, K., McLean, B., Park, J., & Thompson, A. L. (2024). Iconicity mediates semantic networks of sound symbolism. *The Journal of the Acoustical Society of America*, 155(4), 2687–2697.
- Akita, K. (2020). Onomatopoe no onsyootyoosei saihoo [Sound symbolism of ideophones revisited]. In *Proceedings of the 20th Annual Meeting of the Japanese Cognitive Linguistics Association*, 424–438.
- Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, 113(39), 10818–10823.
- Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, 6(10), 654–672
- Doizaki, R., Watanabe, J., & Sakamoto, M. (2016). Automatic estimation of multidimensional ratings from a single sound-symbolic word and word-based visualization of tactile perceptual space. *IEEE transactions on haptics*, 10(2), 173–182.
- Erben Johansson, N., Anikin, A., Carling, G., & Holmer, A. (2020). The typology of sound symbolism: Defining macro-concepts via their semantic and phonetic features. *Linguistic Typology*, 24(2), 253–310.
- Hamano, S. (1986). *The sound-symbolic system of Japanese*, Doctoral dissertation, University of Florida, Gainesville, FL.
- Hamano, S. (1998). *The sound-symbolic system of Japanese*. Stanford, CA: CSLI.
- Hamano, S. (2014). *Nihongo no onomatopoe: Onsyootyoo to koozoo* [Japanese ideophones: the sound symbolism and structure]. Tokyo: Kuroshio Syuppan.
- Hinton, L., Nichols, J., & Ohala, J. J. (Eds.) (2006). *Sound symbolism*. Cambridge: Cambridge University Press.
- Iida, H., & Akita, K. (2023). Perceptual strength norms for 510 Japanese words, including ideophones: A comparative study with English. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 45, 2201–2207.
- Iida, H., & Funakura, H. (2024). Investigating iconicity in vision-and-language models: A case study of the bouba/kiki effect in Japanese models. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46, 5273–5279.
- Iida, H., & Ryo, H. (2024). Towards a comprehensive and quantitative data analysis of Japanese ideophones. *Proceedings of the 2nd International Conference on ICT Application Research*, 2, 1–6.
- Joo, I. (2023). The sound symbolism of food: the frequency of initial /PA-/ in words for (staple) food. *Linguistics*, 61(1), 33–46.
- Kadooka, K. (2007). *Nihongo onomatopoe goi ni okeru keitaiteki/on'teiki taikesei nitsuite* [On the morphological and phonological systematicity of the Japanese ideophone lexicon]. Tokyo: Kurosio Publishers.
- Takehi, H., Tamori, I., & Schourup, L. (1996). *Dictionary of iconic expressions in Japanese*. Berlin & New York: Mouton de Gruyter.
- Kilpatrick, A. J., Ćwiek, A., & Kawahara, S. (2023). Random forests, sound symbolism and Pokémon evolution. *PLOS ONE*, 18(1), e0279350.
- Köhler, W. (1929). *Gestalt Psychology*. New York: Horace Liveright.
- Kumagai, G., Uno, R., & Shinohara, K. (2022). The sound-symbolic effects of consonants on food texture: An experimental study of snack names in Japanese. In *The language of food in Japanese: Cognitive perspectives and beyond*. Amsterdam: John Benjamins.
- Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: A review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in psychology*, 6, 1246.
- McLean, B., Dunn, M., & Dingemanse, M. (2023). Two measures are better than one: combining iconicity ratings and guessing experiments for a more nuanced picture of iconicity in the lexicon. *Language and Cognition*, 15(4), 716–739.
- Ono, M. (Ed.). (2007). *Giongo/gitaigo 4500: Nihongo onomatopoe jiten* [4500 mimetics: A dictionary of Japanese ideophones]. Tokyo: Shogakukan.
- R Core Team (2024). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia -- A window into perception, thought and language. *Journal of consciousness studies*, 8(12), 3–34
- Saji, N., Akita, K., Kantartzis, K., Kita, S., & 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(7), e0218707.
- Sakamoto, M., & Watanabe, J. (2016). Cross-modal associations between sounds and drink tastes/textures: a study with spontaneous production of sound-symbolic words. *Chemical Senses*, 41(3), 197–203.
- Sakamoto, M., Watanabe, J., & Yamagata, K. (2021). Automatic estimation of multidimensional personality from a single sound-symbolic word. *Frontiers in Psychology*, 12, 595986.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12(3), 225–239.
- Shinohara, K., & Kawahara, S. (2010). A cross-linguistic study of sound symbolism: The images of size. In *Annual meeting of the Berkeley linguistics society* (pp. 3965–410).

- Sidhu, D. M., Westbury, C., Hollis, G., & Pexman, P. M. (2021). Sound symbolism shapes the English language: The maluma/takete effect in English nouns. *Psychonomic Bulletin & Review*, 28(4), 1390–1398.
- Tamamura, F. (1984). Onsyootyoogo no gokei (sono 1) [Forms of sound-symbolic words, Part 1]. *Doosisya kokubungaku*, 24.
- Tamori, I. & Schourup, L. (1999). Onomatope: Keitai to imi [Ideophones: Form and Meaning]. Tokyo: Kurosiso Publishers.
- Westbury, C., Hollis, G., Sidhu, D. M., & Pexman, P. M. (2018). Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of memory and language*, 99, 122–150.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.
- Winter, B., & Perlman, M. (2021). Size sound symbolism in the English lexicon. *Glossa: a journal of general linguistics*, 6(1), 79.
- Winter, B., Sósokuthy, M., Perlman, M., & Dingemanse, M. (2022). Trilled /r/ is associated with roughness, linking sound and touch across spoken languages. *Scientific Reports*, 12(1), 1035.
- Winter, B., Woodin, G., & Perlman, M. (2026). Defining iconicity for the cognitive sciences. In O. Fischer, K. Akita, & P. Perniss (Eds.), *The Oxford handbook of iconicity in language*. Oxford: Oxford University Press.