System to quantify the impression of sounds expressed by onomatopoeias

Maki Sakamoto*

The University of Electro-Communications,
West 6 Bldg., 1–5–1 Chofugaoka, Chofu, 182–8585 Japan

Abstract: Sounds can be expressed by onomatopoeias such as tick-tock and ding-dong, where we verbalize the perceived auditory information from environmental sound. Onomatopoeias, i.e., sound symbolic words, indicate linguistic forms closely related to environmental sounds. In recent years, some researchers have reported that onomatopoeias are implicated in affective aspects. Our research group has developed a system to quantify the affective impression or texture of environmental sounds expressed by onomatopoeias. Interestingly, our system can estimate not only sound impressions but also tactile or taste impressions expressed by onomatopoeias.

Keywords: Onomatopoeia, Sounds, Affective, Texture, System

PACS number: 43.71.Sy  [doi:10.1250/ast.41.229]

1. INTRODUCTION

Sounds can be expressed by onomatopoeias such as tick-tock and ding-dong. Japanese people frequently verbalize the perceived auditory information from environmental sounds by onomatopoeias, i.e., sound symbolic words.

In our previous research [1], we asked participants to describe 60 environmental sounds by onomatopoeias and evaluate the affective impression of the 60 environmental sounds using the semantic differential (SD) method with a seven-point scale (very pleasant, +3; pleasant, +2; somewhat pleasant, +1; neither pleasant nor unpleasant, 0; somewhat unpleasant, −1; unpleasant, −2; very unpleasant, −3). The mean result was −0.232 (variance 1.44). We performed pairwise t-test comparisons between the result on the seven-point SD scale for each stimulus and the mean value on the seven-point SD scale of all auditory stimuli as the reference value. We analyzed the relationship between the pleasant/unpleasant evaluation and the phoneme of the onomatopoeia for each stimulus. The result showed that pleasant/unpleasant evaluation of sounds was associated with the phoneme of the onomatopoeia.

A sensory-sound correspondence can be found not only in onomatopoeias expressing environmental sounds but also in words referring to visual shapes. This was demonstrated in landmark studies (e.g., mal/mil and buba/kiki for round and sharp shapes in studies by Sapir [2] and Ramachandran and Hubbard [3], respectively). We have investigated the sound symbolic associations in touch, specifically the association between the phonemes of Japanese onomatopoeias for expressing tactile sensations and subjective evaluations of comfort/discomfort for touched objects.

Compared with other languages, Japanese has a large number of onomatopoeic words for tactile sensations, and associations between the phonemes of Japanese onomatopoic words and typical categories of tactile sensations can be observed [4]. Most Japanese onomatopoeias (SSWs) expressing tactile sensations consist of two-syllable repetitions (C1V1C2V2-C1V1C2V2, where C and V indicate a consonant and vowel, respectively, e.g., "sara-sara"), and the sound of the first syllable (C1V1, e.g., "sa" for "sara-sara") is strongly associated with evaluations of tactile sensations. For instance, "sara-sara" and "zara-zara," which are different only in the first syllable of the repeated unit, denote completely different tactile sensations. Whereas the former is used for expressing a smooth and pleasant sensation, the latter is used for expressing a rough and unpleasant sensation.

In [5], participants touched an object and expressed the tactile sensation using Japanese onomatopoeic words and then rated the pleasantness of touching the object with the SD method. Participants reported an onomatopoeic word to express the tactile feeling while touching each of the 120 materials, and then they rated the comfort/discomfort of touching the material on a seven-point scale (very comfortable, +3; comfortable, +2; slightly comfortable, +1; neither comfortable nor uncomfortable, 0; and three levels −1 to −3, for uncomfortable feelings) while

*e-mail: maki.sakamoto@uec.ac.jp
touched it. We therefore analyzed the relationship between the phonemes of onomatopoeic words in two-mora repeated form and the evaluations of comfort/discomfort using 1,569 cases. The average rating for the 1,569 cases was 0.08. The average ratings across trials in which the same phonemes were used in the first syllable were statistically compared with the average of all 1,569 cases (0.08) (t-test comparing the average with a constant value). Of the vowels, only /u/ had a statistically significant relationship with comfort. The vowels /i/ and /e/ were not used often, but they were strongly related to discomfort (/i/ was marginally significant and /e/ was significant). Among the consonants, /m/, /n/, and /s/ were significantly related to comfort, whereas /g/, /z/, /j/, /b/, and /n/ were significantly related to discomfort.

These results show that the affective impression of various modalities such as sound, vision, and touch by an onomatopoeia is associated with phonological features of the onomatopoeia. Japanese onomatopoeias make up a large body of vocabulary that can express the complex and minute qualities of multiple sensory experiences.

2. System to Quantify Affective Impression of Onomatopoeias

We have proposed a system that can convert an onomatopoeia in Japanese into quantitative ratings in multiple sensory dimensions (26 pairs of adjectives) [5]. We next explain the procedure for constructing the automatic estimation system.

2.1. Experiment to Build Sound-symbolic Database

First, we discuss an experiment that we performed to analyze the sound-symbolic associations of phonemes and 26 pairs of adjectives. In the experiment, participants viewed onomatopoeias displayed on a monitor and rated their impressions in terms of the 26 bipolar adjective scales. Seventy-eight paid participants, aged 20 to 24 years (51 men and 27 women), participated in this experiment. They had no knowledge of linguistics and were all native Japanese speakers. They were unaware of the purpose of the experiments, and informed consent was obtained from all the participants before the experiment started.

To obtain sound-symbolic associations of all Japanese phonemes with the 26 pairs of adjectives, we selected word stimuli that included all varieties of Japanese phonemes. First, we made a list of 14,584 onomatopoeias by combining all sounds in the Japanese syllabary as two-syllable forms (e.g., /sara/). Then, we also prepared repetition expressions of two-syllable forms (i.e., /sara-sara/). In addition, we made another list of expressions by adding words with all types of special phonemes used in Japanese onomatopoeias (syllabic nasals /N/, choked sounds /Q/, long vowels /R/, and adverbs ending in /Li/) to the two-syllable forms (i.e., /saraLi/). Second, from this large word list, 312 words that could be used as tactile onomatopoeia were selected by three experts (including the author) in psychology and linguistics.

The trials started with the presentation of an onomatopoeia on the monitor, and participants were asked to report how they felt about it on a seven-point SD scale (e.g., for the comfortable–uncomfortable scale, participants selected one of the following seven points: 1, very comfortable; 2, comfortable; 3, slightly comfortable; 4, neither comfortable nor uncomfortable; 5, slightly uncomfortable; 6, uncomfortable; 7, very uncomfortable). The participants responded by pressing one of seven buttons. The time allotted for answering was unlimited, but almost all trials took less than 1 minute. The presentation order of onomatopoeias was randomized among the participants. The orders and polarities of the bipolar scales were also randomized in the answer matrix.

The experimental results produced 105,456 items of data (26 rating scales × 312 words × 13 participants). Then, we calculated the averages and standard deviations of the rating values for each scale of all words. The average of the standard deviation was 1.31. Ninety-eight percent of all data (105,181 items of data) were within 2.0 standard deviations of the average.

2.2. Sensory Image Estimation Model

To estimate the sensory impressions of onomatopoeias, we constructed a model in which the following equation was used to predict each rating value:

\[
Y = \sum_{i=1}^{13} X_i + \text{Const.} \quad (1)
\]

where \(Y\) represents the rating values of the 26 scales, and \(X_1-X_{13}\) are the corresponding values defined in Table 1. \(X_1-X_6\) are respectively the mean values of the specific consonant, voiced sound/p-sound, contracted sounds, vowels, semivowels, and special phonemes in the first syllable. \(X_7-X_{12}\) are the same categories for the second syllable, and \(X_{13}\) denotes the presence or absence of repetition in the word.

Using the average of the rating values as the objective variables and the variation of phonemes as the predictor variables, we conducted class I mathematical quantification theory, which is a type of multiple regression analysis. As shown by Eq. (1), the rating values of an onomatopoeia can be calculated by a linear sum of the values (\(X_1-X_{13}\)) of the word. For example, the expression “sara” is composed of the first mora /sa/ (/s/ + /a/) and the second mora /ra/ (/t/ + /a/). Therefore, the value of the “rough–smooth” scale on a seven-point scale (1 = smooth, 7 = rough) is estimated using the Eq. (2) (see Eq. (1) and Table 2).
Table 1  Correspondences between variables and phonemes.

| First syllable | Second syllable | Phonological characteristics | Variation of phonemes |
|----------------|-----------------|------------------------------|-----------------------|
| $X_1$          | $X_1$           | consonants                   |                       |
|                | $X_2$           | voiced sounds/p-sounds       | presence (/g/, /z/, /d/, /b/, /v/) or absence |
| $X_3$          | $X_3$           | contracted sounds            | presence (/ty/, /gy/, /ey/, /ny/, /hy/, /my/) or absence |
| $X_4$          | $X_4$           | vowels                       | /a/, /i/, /u/, /e/, /o/ |
| $X_5$          | $X_5$           | semi-vowels                  |                       |
| $X_6$          | $X_6$           | special sounds               | /N/, /Q/, /R/, /Li/ or absence |
| $X_{13}$       |                 | repetition                   | presence (ex. huwa-huwa) or absence |

Table 2  Examples of category quantities for 26 rating scales.

| Rating scales                  | Consonants | Vowels | $R^2$ |
|-------------------------------|------------|--------|-------|
| smooth–rough                  | $-0.05$    | $0.46$ | 0.88  |
| bumpy–flat                    | $0.56$     | $0.18$ | 0.84  |
| hard–soft                     | $-0.33$    | $-0.35$| 0.91  |
| warm–cold                     | $0.52$     | $0.28$ | 0.88  |
| slippery–sticky               | $-0.50$    | $-0.10$| 0.89  |
| wet–dry                       | $0.49$     | $0.93$ | 0.88  |
| elastic–nonelastic            | $0.80$     | $0.47$ | 0.91  |
| firm–fragile                  | $0.02$     | $0.23$ | 0.85  |
| regular–irregular             | $-0.53$    | $0.01$ | 0.89  |
| repulsive–nonrepulsive        | $0.48$     | $0.24$ | 0.88  |
| sharp–dull                    | $-0.62$    | $-0.12$| 0.92  |
| clean–dirty                   | $-0.51$    | $-0.14$| 0.90  |
| stretchy–nonstretchy          | $0.30$     | $0.57$ | 0.92  |
| thick–thin                    | $0.73$     | $0.37$ | 0.89  |
| heavy–light                   | $0.47$     | $0.37$ | 0.91  |
| strong–weak                   | $0.21$     | $0.23$ | 0.89  |
| comfortable–uncomfortable     | $-0.41$    | $-0.13$| 0.87  |
| relieved–uneasy               | $-0.05$    | $0.02$ | 0.87  |
| good–bad                      | $-0.28$    | $0.01$ | 0.87  |
| impressive–unimpressive       | $0.30$     | $0.15$ | 0.87  |
| luxury–cheap                  | $-0.23$    | $0.13$ | 0.89  |
| pleasant–irritating           | $-0.60$    | $-0.07$| 0.84  |
| familiar–unfamiliar           | $-0.06$    | $0.12$ | 0.86  |
| eccentric–ordinary            | $0.29$     | $0.30$ | 0.85  |
| natural–artificial            | $0.05$     | $0.10$ | 0.84  |
| intense–calm                  | $0.04$     | $0.03$ | 0.90  |

$Y = /s/ (X_1) + \text{absence} (X_2) + \text{absence} (X_3) + /a/ (X_4) + \text{absence} (X_5) + /\text{r/} (X_7) + \text{absence} (X_8) + /a/ (X_{10}) + \text{absence} (X_{11}) + \text{absence} (X_{12}) + \text{absence} (X_{13}) + \text{Const.} = (-0.05) + (-0.32) + (-0.05) + (0.46) + (-0.02) + (-0.03) + (-0.14) + (-0.01) + (0.05) + (-2.19) + (0.2) + (-0.02) + (0.05) + (0.01) + (3.75) = 2.05$ (2)

The estimated value of 2.05 suggests that “sara” is associated with a smooth impression. The multiple correlation coefficients $R^2$ between the predicted values and the mean values of the actual ratings (the values obtained from the participants) were used as indicators of prediction accuracy. For 20 scales, the $R^2$ values were in the range of 0.80 to 0.90, and for the other six scales, they were higher than 0.90 (Table 2). Therefore, we considered our model to be satisfactory for estimating impressions of onomatopoeias by analyzing the phonemes and forms of the words.

2.3. Advantage of our System

In our system, when a word that intuitively expresses
a tactile sensation is input into the text field, information equivalent to evaluations against the 26 pairs of texture related adjectives is obtained from an analysis of the sounds of the word. Figures 1 and 2 show examples of outputs from our system for “sara-sara” and “zara-zara,” respectively. “Sara-sara” shows higher ratings for “slippery,” “dry,” “nonelastic,” “clean,” “thin,” and “light” feelings, while “zara-zara” shows higher ratings for “rough,” “hard,” “dry,” “none-lastic,” “nonstretchy,” “impressive,” and “intense” feelings. This system enables us to analyze, for example, tactile sensations with many criteria by expressing the sensation with only a single word, and this idea can be applied to any combination of phonemes (even to newly created words) in Japanese. To estimate the quantitative information of every possible onomatopoeic word, we built a database of sound-symbolic associations for each phoneme with the 26 pairs of adjectives through psychological experiments. We described how we constructed a system that can estimate multidimensional rating scores from a single onomatopoeia based on sensory-sound associations. Our system can estimate not only onomatopoeias established as Japanese vocabulary, but also newly created novel onomatopoeias.

### 3. CONCLUSION

Onomatopoeias indicate linguistic forms closely related to environmental sound. Sensory-sound correspondences can be found not only in onomatopoeias expressing environmental sounds but also in words referring to various modalities such as visual impressions as well as tactile sensations and affective aspects. Our research group developed a novel system to quantify multimodal impressions expressed by onomatopoeias. In this paper we introduced a system that estimates information equivalent to evaluations against the 26 pairs of impressions and textures when a word that intuitively expresses a tactile sensation is input into the text field. The system is expected to help capture subjective impressions expressed by onomatopoeias.

### ACKNOWLEDGEMENTS

This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas “Shitsukan” (Nos. 23135510 and 25135713) from MEXT Japan and JSPS KAKENHI Grant Number 15H05922 (Grant-in-Aid for Scientific Research on Innovative Areas “Innovative SHITSUKSAN Science and Technology”) from MEXT, Japan.

### REFERENCES

[1] J. Kwon and M. Sakamoto, “Visualization of relation between sound symbolic word and perceptual characteristics of environmental sounds,” American Institute of Physics (AIP) Conf. Proc., pp. 1–6 (University of Electro-Communications, Tokyo, Japan, 2016).

[2] E. Sapir, “A study of phonetic symbolism,” J. Exp. Psychol., 12, 225–239 (1929).

[3] V. S. Ramachandran and E. M. Hubbard, “Hearing colors, tasting shapes,” Sci. Am., 288, 43–49 (2003).

[4] M. Sakamoto and J. Watanabe, “Bouba/Kiki in touch: Associations between tactile perceptual qualities and Japanese phonemes,” Front. Psychol., 9, 1–12 (2018). doi: 10.3389/fpsyg.2018.00295

[5] R. Doizaki, J. Watanabe and M. Sakamoto, “Automatic estimation of multidimensional ratings from a single sound-symbolic word and word-based visualization of tactile perceptual space,” IEEE Trans. Haptics, 10, 173–182 (2017). doi: 10.1109/TOH.2016.2615923