Comparisons of auditory impressions and auditory imagery associated with onomatopoeic representation for environmental sounds

Masayuki Takada (1), Nozomu Fujisawa (2), Fumino Obata (3)

and Shin-ichiro Iwamiya (1)

(1) Department of Communication Design Science, Faculty of Design, Kyushu University, 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan
(2) Department of Information and Media Studies, Faculty of Global Communication, University of Nagasaki, 1-1-1 Manabino, Na-gayo-cho, Nishi-Sonogi-gun, Nagasaki 851-2195, Japan
(3) Nippon Telegraph and Telephone East Corp., 3-19-2, Nishi-shinjuku, Shinjuku, Tokyo 163-8019, Japan

ABSTRACT

Humans represent sounds to others and receive information about sounds from others using onomatopoeia. Such representation is useful for obtaining and reporting the acoustic features and impressions of actual sounds without having to hear or emit them. But how accurately can we obtain such sound information from onomatopoeic representations? To examine the validity and applicability of using verbal representations to obtain sound information, experiments were carried out in which the participants evaluated auditory imagery associated with onomatopoeic representations created by listeners of various environmental sounds. Furthermore, participants provided answers to questions asking about the sound sources themselves or the phenomena that create the sounds associated with the onomatopoeic stimuli. Comparisons of impressions between real sounds and onomatopoeic stimuli revealed that impressions of sharpness and brightness for both real sounds and onomatopoeic stimuli were similar, as were emotional impressions such as "pleasantness" for real sounds and major (typical) onomatopoeic stimuli. The auditory imagery of powerfulness associated with onomatopoeia was different from the same impression of real sounds. Furthermore, recognition of the sound source from onomatopoeic stimuli affected the emotional impression similarity between real sounds and onomatopoeic representations.

INTRODUCTION

When we describe sounds to others in our daily lives, we often use onomatopoeic representations related to the actual acoustic properties of the sounds we listen. Moreover, because the acoustic properties of sounds induce auditory impressions in listeners, onomatopoeic representations and the auditory impressions associated with actual sounds may be related.

In a number of previous studies, the relationships between the temporal and spectral acoustic properties of sounds and onomatopoeic features have been discussed [1–4]. We have also conducted psychoacoustical experiments to ascertain the validity of using onomatopoeic representations to identify the acoustic properties of operating sounds emitted from copy machines and audio signals emitted from domestic electronic appliances [5,6]. As a result, relationships between subjective impressions, such as product imagery and functional imagery evoked by machine operation sounds, audio signals, and the onomatopoeic features were found. Furthermore, we also investigated the validity of using onomatopoeic representations to identify the acoustic properties and auditory impressions of various kinds of environmental sounds [7].

Knowledge concerning the relationship between the onomatopoeic features and the acoustic properties or auditory impressions of sounds is useful since it would allow one to more accurately obtain or describe the auditory imagery of sounds without actually hearing or emitting them. Practical applications of such knowledge may include situations in which electronic home appliances such as vacuum cleaners and hair dryers break down and customers contact customer service representatives and use onomatopoeic representations of the mechanical problems they are experiencing; engineers who listen or read accounts of such complaints may be able to obtain more accurate information about the problems being experienced by customers and better analyze the cause of the problem through the obtained representations. Wake and Asahi [8] conducted psychoacoustical experiments to clarify how people communicate sound information to others. Sound stimuli were presented to subjects, and they were asked to freely describe the presented sounds to others. Their results showed that verbal descriptions including onomatopoeic representations, mental impressions expressed using adjectives, sound sources, and situations were frequently used by subjects. Thus it is possible to obtain sound information such

as acoustic properties of sounds and auditory impressions for sounds from representations created by listers of sounds.

Table 1. "Major" and "minor" onomatopoeic representations for each sound source.

No	Sound source	"Major (1)" and "minor (2)" onomatopoeic representations
1	whizzing sound	(1) /hyuN/, (2) /pyaN/
2	idling sound of a diesel engine	(1) /burorororo/, (2) /karakarakarakarakara-korokorokorokorokoro/
3	sound of water dripping	(1) /potyaN/, (2) /pikori/
4	bark of a dog	(1) /waN/, (2) /wauQ/
5	ring of a telephone	(1) /pirororororo/, (2) /piririririririririri/
6	owl hooting	(1) /kurururu/, (2) /fororoo/
7	vehicle starter sound	(1) /bururuuN/, (2) /tyeQ baQ aaN/
8	hand clap	(1) /paN/, (2) /tsuiN/
9	vehicle horn	(1) /puu/, (2) /faaQ/
10	baby crying	(1) /Ngyaa/, (2) /buyaaaN/
11	sound of a flowing stream	(1) /zyorororo/, (2) /tyupotyupoyan/
12	jackhammer noise	(1) /gagagagagagagagagagaga/, (2) /gyururururururururu/
13	sound of fireworks	(1) /patsuQ/, (2) /putiiiN/
14	sweeping tone	(1) /puiQ/, (2) /poi/
15	knock	(1) /koNkoN/, (2) /taQtoQ/
16	chirping of a cricket	(1) /ziizii/, (2) /kyuririririiii/
17	twittering of a sparrow	(1) /piyo/, (2) /tyui/
18	harmonic complex tone	(1) /pii/, (2) /piiQ/
19	sound like a wooden gong	(1) /pokaQ/, (2) /NkaQ/
20	sound of a trumpet	(1) /puuuuuuN/, (2) /waaN/
21	sound of a stone mill	(1) /gorogorogoro/, (2) /gaiaiai/
22	siren of an ambulance	(1) /uuuu/, (2) /uwaaaaa/
23	shutter sound of a camera	(1) /kasyaa/, (2) /syagiiN/
24	white noise	(1) /zaa/, (2) /suuuuuu/
25	sound of a temple bell	(1) /goon/, (2) /gaaaaaaaaaaN/
26	thunderclap	(1) /baaN/, (2) /bababooNbaboonbooN/
27	bell of a microwave oven	(1) /tiiN/, (2) /kiNQ/
28	sound of a passing train	(1) /gataNgotoN/, (2) /gararatataNtataN/
29	typing sound	(1) /katakoto/, (2) /tamutamu/
30	beach sound	(1) /zazaaN/, (2) /syapapukupusyaapaaN/
31	sound of wind blowing	(1) /hyuuhyuu/, (2) /haaaououou ohaaa ouo-haaao/
32	sound of wooden clappers	(1) /taN/, (2) /kiQ/
33	sound of someone slurping noodles	(1) /zuzuu/, (2) /tyurororo/
34	sound of a wind chime	(1) /riN/, (2) /kiriiN/
35	sound of a waterfall	(1) /goo/, (2) /zaaaaa/
36	footsteps	(1) /katsukotsu/, (2) /kotoQ kotoQ/

In practical situations in which people communicate sound information to others using onomatopoeic representation, it is necessary that the receivers of onomatopoeic representations (in the above-mentioned case, for example, engineers) be able to identify the acoustic properties and auditory impressions of the sounds that onomatopoeia represent. The present study examines this issue. Experiments were carried out in which participants evaluated the auditory imagery associated with onomatopoeic representations. The auditory imagery of onomatopoeic representations was compared with the auditory impressions for their corresponding actual sound stimuli, which were obtained in our previous study [7].

Furthermore, one of the most primitive behaviors humans engage in related to sounds is the identification of the sound source [9]. If we recognize events related to everyday sounds using acoustic cues [10-12], therefore, is it possible to also recognize sound sources from onomatopoeic features instead of acoustic cues? Moreover, such recognition of the source may affect the auditory imagery evoked by onomatopoeic representation. Although Fujisawa et al. [13] examined the auditory imagery evoked by simple onomatopoeia with two morae such as /don/ and /pan/ ("mora" is a standard unit of rhythm in Japanese speech), the effect of sound source recognition on the auditory imagery evoked by onomatopoeia was not discussed in their study. In the present study, therefore, we took sound source recognition into consideration while comparing the auditory imagery of onomatopoeic representations to the auditory impressions induced by their corresponding real sounds.

EXPERIMENTS

Stimuli

In our previous study [7], 8 participants were aurally presented with 36 environmental sounds, and evaluated their auditory impressions of sound stimuli. The sound stimuli were selected based on their relatively high frequency of occurrence both outdoors and indoors in our daily lives. Additionally, participants expressed sound stimuli using onomatopoeic representations, as shown in Table 1.

For each sound stimulus, 8 onomatopoeic representations which were described by participants in our previous psycho-acoustical experiment [7] were classified into 2 groups based on the similarities of onomatopoeic features. First, the onomatopoeic representations were encoded using 24 phonetic parameters, consisting of combinations of 7 places of articulation (labio-dental, bilabial, alveolar, post-alveolar, palatal, velar, and glottal), 6 manners of articulation (plosive, fricative, nasal, affricate, approximant, and flap) [14], the 5 Japanese vowels (/a/, /i/, /u/, /e/, /o/), voiced and voiceless consonants, syllabic nasals, geminate obstruents, palatalized consonants, and long vowels. Furthermore, for each sound, onomatopoeic representations were classified based on the similarities of the abovementioned phonetic parameters using a hierarchical cluster analysis in which the Ward method of using Euclidean distance as a measure of similarity was employed. For the two groups obtained from cluster analysis, two onomatopoeic representations were selected for each sound. One was selected from the larger group (described as the "major" representation), and the other from the smaller group (the "minor" representation). A "major" onomatopoeic representation is regarded as being frequently described by many listeners of the sound, that is, a "typical" onomatopoeia, whereas a "minor" onomatopoeic representation is regarded as a unique representation for which there is a relative smaller possibility that a listener of the sound would actually use the representation to describe it. In selecting the "major" onomatopoeic stimuli, a Japanese onomatopoeia dictionary

[15] was also referenced. Consequently, 72 onomatopoeic representations were used as stimuli, as shown in Table 1.

Procedure

Seventy-two onomatopoeic representations printed in random order on sheets were presented to the 20 participants. They were asked to rate their impressions of the sounds associated with the onomatopoeic stimuli. The impressions of the auditory imagery evoked by the onomatopoeic stimuli were measured using the semantic differential method [16]. The 13 adjective pairs shown in Table 2 were used as the SD scales, which were also used in our previous listening experiments (i.e., in measurements of auditory impressions for environmental sounds) [7]. Each SD scale had 7 Likert-type scale categories (1 to 7). For example, for the scale "pleasant/unpleasant," the categories "1" and "7" corresponded to "extremely pleasant" and "extremely unpleasant," respectively. The participants selected a number from 1 to 7 for each scale for each onomatopoeic stimulus.

Participants were also requested to provide answers to questions asking about the sound sources themselves or the phenomena that create the sounds associated with the onomatopoeic stimuli by free description.

RESULTS

Analysis of subjective ratings

The rating scores were averaged for each scale and for each onomatopoeic representation. To compare impressions between actual sound stimuli and onomatopoeic representations, factor analysis was applied to the averaged scores for onomatopoeic representations together with those for the sound stimuli (i.e., the rating results of auditory impressions) obtained in our previous experiment [7].

By taking into account the factors for which the eigenvalues were more than 1, a three-factor solution was obtained. Finally, the factor loadings for each factor on each scale were obtained using a varimax algorithm, as shown in Table 2.

Table 2. Factor loading of each adjective scale for each factor.

Pair of adjectives			Fac. 1	Fac. 2	Fac. 3
tasteful	-	tasteless	0.905	0.055	0.154
desirous of hearing	-	not desirous of hearing	0.848	0.292	0.214
pleasant	-	unpleasant	0.788	0.458	0.254
rural	-	Urban	0.693	-0.210	0.294
soft	-	Hard	0.381	-0.101	0.327
muddy	-	Clear	-0.165	-0.901	-0.288
bright	-	Dark	-0.007	0.830	-0.018
smooth	-	Rough	0.190	0.726	0.356
sharp	-	Dull	-0.393	0.712	-0.323
strong	-	Weak	-0.259	-0.391	-0.860
modest	-	Loud	0.391	-0.020	0.805
powerful	-	powerless	-0.153	-0.486	-0.805
slow	-	Fast	0.504	-0.208	0.538

The first factor is interpreted as the emotion factor because adjective pairs such as "tasteful/tasteless" and "pleasant/unpleasant" have high loadings for this factor. The second factor is interpreted as the clearness factor because adjective pairs such as "muddy/clear" and "bright/dark" have high factor loadings. The third factor is interpreted as the powerfulness factor because the adjective pairs "strong/weak," "modest/loud," and "powerful/powerless" have high factor loadings. Similar factors were also obtained in our previous psychoacoustical study [7].

Furthermore, the factor scores for each stimulus for each factor were computed. Figure 1(a) to (c) shows the factor scores for the sound stimuli and the "major" and "minor" onomatopoeic representations on the emotion, clearness, and powerfulness factors, respectively.

Analysis of free description answers of sound source recognition questions

From the free descriptions regarding sound sources associated with onomatopoeic representation, the percentage of participants who correctly recognized the sound source or the phenomenon creating the sound was calculated for each onomatopoeic stimulus. In Gaver's study on the ecological approach to auditory perception [17], sound-producing events were divided into three general categories: vibrating solids, gasses, and liquids. Considering these categories, participants' descriptions in which keywords related to sound sources or similar phenomena were contained were regarded as being correct. For example, for "whizzing sound (No.1)", descriptions such as "sound of an arrow shooting through the air" and "sound of a small object slicing the air" were counted as a correct answer. The percentages of correct answers for sound sources associated with "major" and "minor" onomatopoeic stimuli are shown in Fig. 2.

The percentage of correct answers averaged across all "major" onomatopoeic stimuli was 64.3%, whereas the same percentage for "minor" onomatopoeic stimuli was 24.3%. "Major" onomatopoeic stimuli seemed to allow participants to better recall the corresponding sound sources. These results suggest that sound source information might be communicated by "major" onomatopoeic stimuli more correctly than by "minor" stimuli.

DISCUSSION

Comparison between onomatopoeic representations and real sound stimuli factor scores

Fig. 1(a) shows that sound stimuli such as "owl hooting (No.6)," "vehicle horn (No.9)," "sound of a flowing stream (No.11)," "sound of a noisy construction site (No.12)," and "sound of a wind chime (No.34)" displayed highly positive or negative emotion factor scores (e.g., inducing strong impressions of tastefulness or tastelessness and pleasantness or unpleasantness). However, the factor scores for the onomatopoeic representations of the same sound stimuli were not as positively or negatively high. On the other hand, the factor scores for the "major" onomatopoeic representations of stimuli such as "sound of water dripping (No.3)," "sound of a temple bell (No.25)," and "beach sound (No.30)" were nearly equal to those of the corresponding real sound stimuli.

To compare between the auditory impressions of sounds and the auditory imagery evoked by the corresponding onomatopoeia, the absolute differences in factor scores between the sound stimuli and the "major" or "minor" onomatopoeic representations were averaged across all sound sources in each of the three factors (see Table 3).

For the emotion factor, the factor scores for the real sound stimuli were closer to those for the "major" onomatopoeic representations than to those for the "minor" onomatopoeic representations (see Fig. 1(a) and Table 3). The correlation coefficient of the emotion factor scores between the real sound stimuli and the "major" onomatopoeic stimuli was statistically significant at $p<0.01$ ($r=0.682$), while the same scores of the "minor" onomatopoeic stimuli were not correlated with those of their real sounds.
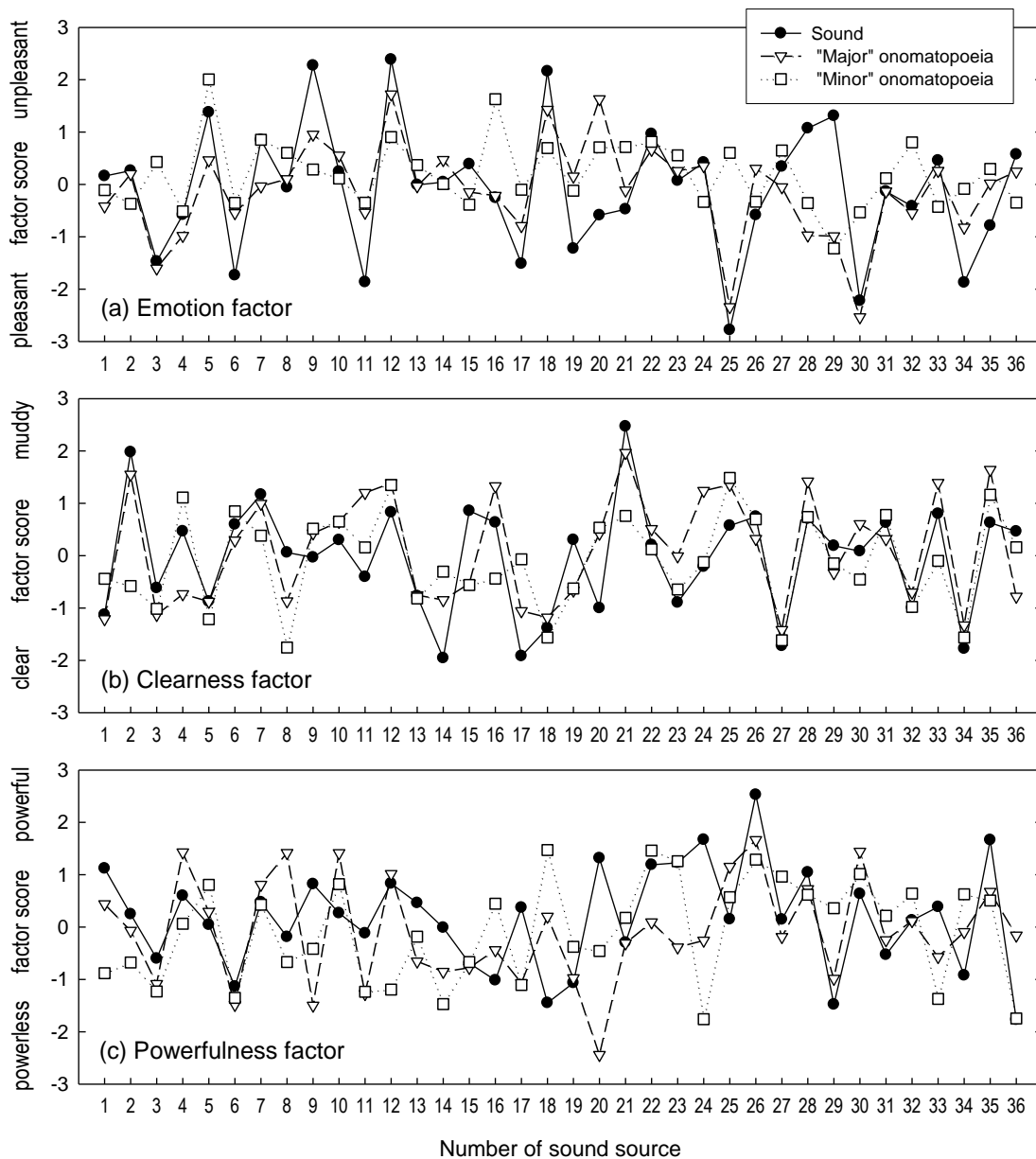
Figure 1. Factor scores for real sound stimuli and "major" and "minor" onomatopoeic representations on the (a) emotion factor, (b) clearness factor, and (c) powerfulness factor
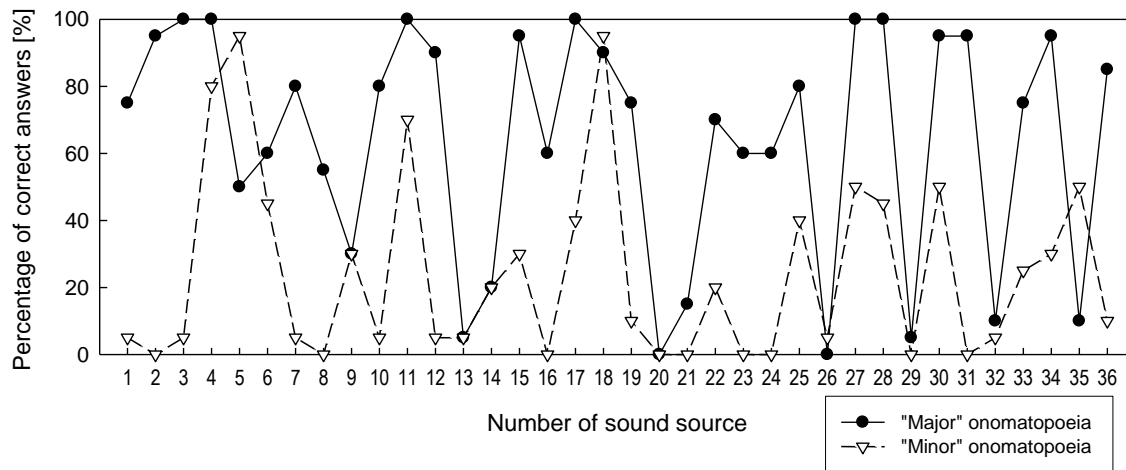

Figure 2. Percentage of correct sound source answers associated with "major" and "minor" onomatopoeic stimuli

Table 3. Averaged absolute differences of factor scores between real sound stimuli and "major" or "minor" onomatopoeic representations (standard deviations shown in parentheses).

	Onomatopoeic representation	
	"Major"	"Minor"
emotion factor	0.66 (±0.61)	1.04 (±0.77)
clearness factor	0.65 (±0.43)	0.68 (±0.64)
powerfulness factor	0.90 (±0.76)	1.00 (±0.80)

As shown in Fig. 1(b), for the clearness factor, the factor scores for the "major" and "minor" onomatopoeic representations were close to those for the real sound stimuli as a whole.

Table 3 also shows that the averaged differences of the clearness factor score between the real sound stimuli and both the "major" and "minor" onomatopoeia were the smallest among the three factors. The correlation coefficients of the clearness factor scores between the real sound stimuli and the "major" or "minor" onomatopoeic stimuli were both statistically significant at $p<0.01$ (sound vs. "major" onomatopoeia: $r=0.724$; sound vs. "minor" onomatopoeia: $r=0.544$). The impressions of muddiness (or clearness) and brightness (or darkness) for the onomatopoeic representations were similar to those for the corresponding real sound stimuli.

For the powerfulness factor, factor scores for the "major" and "minor" onomatopoeia were different from those for the corresponding sound stimuli as a whole, as shown in Fig. 1(c) and Table 3. Moreover, no correlation of the powerfulness factor scores between the real sound stimuli and the onomatopoeic stimuli was found.

These results suggest that the receiver of onomatopoeic representations can more accurately guess auditory impressions of muddiness, brightness and sharpness (or clearness, darkness and dullness) for real sounds from their heard onomatopoeic representations. Conversely, it seems difficult for listeners to report impressions of strength and powerfulness for sounds using onomatopoeic representations.

In the present study, while onomatopoeic stimuli with highly positive clearness factor scores included the Japanese vowel /o/ (e.g., the "major" onomatopoeic stimuli Nos. 2 and 21), those with highly negative clearness factor scores contained vowel /i/ (e.g., the "major" and "minor" onomatopoeic stimuli Nos. 27 and 34). According to our previous study [7], the Japanese vowel /i/ was frequently used to represent sounds with spectral centroids at approximately 5 kHz, which induced impressions of sharpness and brightness. Conversely, vowel /o/ was frequently used to represent sounds with spectral centroids at approximately 1.5 kHz, which induced impressions of dullness and darkness. From a spectral analysis of the five Japanese vowels produced by speakers, the spectral centroids of vowels /i/ and /o/ were actually the highest and lowest, respectively, among all the five vowels [7]. Thus it can be said that these vowels are at least useful in communicating information about the rough spectral characteristics of sounds.

As mentioned above, a relatively small difference in addition to a significant correlation of emotion factor scores between the real sound stimuli and the "major" onomatopoeic stimuli were found. Participants could identify the sound source or the phenomenon creating the sound more accurately from the "major" onomatopoeic stimuli (see Fig.2 and Table 3).

Preis et al. have pointed out that sound source recognition influences differences in annoyance ratings between bus recordings and "bus-like" noises, which were generated from white noise to have spectral and temporal characteristics similar to those of original bus sounds [18]. Similarly, in case of the present study, good recognition of sound sources may be the reason why the emotional impressions of the "major" onomatopoeic stimuli were similar to those for the real sound stimuli. This point was discussed in the latter section.

Our previous study reported that the powerfulness impressions of sounds were significantly correlated with the number of voiced consonants [7]. However, as shown in Fig. 1(c), the auditory imagery of onomatopoeic stimuli containing voiced consonants (i.e., Nos. 26 and 35) was different from the auditory impressions evoked by real sounds. Thus, we can conclude that it is difficult to communicate the powerfulness impression of sounds by voiced consonants alone.

Effects of sound source recognition on the differences between the impressions associated with onomatopoeic representations and those for real sounds

As mentioned regarding the emotion factor in the previous section, there is some possibility that differences in impressions between real sound stimuli and onomatopoeic representations may be affected by sound source recognition. That is, impressions of onomatopoeic representations may be similar to those for real sound stimuli when the sound source can be correctly recognized from the onomatopoeic representations. To investigate this point for each of the three factors, the absolute differences between the factor scores for the onomatopoeic representations and those for the corresponding sound stimuli were averaged for each of two groups of onomatopoeic representations, that is, one group comprised of onomatopoeic stimuli for which more than 50% of the participants correctly identified the sound source question, and another group comprised of those for which less than 50% of the participants correctly answered the sound source question. These two groups comprised 30 and 42 representations, respectively, from the 72 total onomatopoeic representations (See Fig. 2). The averaged differences of factor scores for both groups mentioned above for each factor were shown in Table 4.

Table 4. Absolute differences between factor scores for onomatopoeic representations and those for real sound stimuli, averaged for each of the two groups of onomatopoeic representations: those for which more than 50% of the participants had correct sound source identifications, and those for which less than 50% of the participants had correct identifications (standard deviations shown in parentheses).

	Groups	
	Above 50 %	Below 50 %
emotion factor	0.60 (±0.53)	1.02 (±0.78)
clearness factor	0.65 (±0.41)	0.68 (±0.62)
powerfulness factor	0.90 (±0.64)	0.99 (±0.86)

The difference in the group of onomatopoeic representations in which participants had higher sound source recognition was slightly smaller than that in the other group for each factor. In particular, regarding the emotion factor, the difference between the averaged differences in both groups was statistically significant at $p<0.05$. For the other two factors, no significant differences were found. These results revealed that the recognition of a sound source from an onomatopoeic representation may affect the difference between the emotional impressions associated with an onomatopoeic representation and those evoked by the real sound that it represents. Furthermore, it can be concluded that impressions of the clearness, brightness and sharpness of both the sound and onomatopoeic stimuli were similar, regardless of sound source recognition.

CONCLUSION

The auditory imagery of sounds evoked by "major" and "minor" onomatopoeic stimuli was measured using the semantic differential method. From a comparison of impressions made by real sounds and their onomatopoeic stimuli counterparts, the clearness impressions for both sounds and "major" and "minor" onomatopoeic stimuli were found to be similar, as were the emotional impressions for the real sounds and the "major" onomatopoeic stimuli. Furthermore, the recognition of a sound source from an onomatopoeic stimulus was found to influence the similarity between the emotional impressions evoked by such onomatopoeic representations and their corresponding real sound stimuli. However, this effect was not found for the factors of clearness and powerfulness. From these results, it can be said that it was relatively easy to communicate information about impressions of clearness, including the muddiness, brightness and sharpness of sounds, to others using onomatopoeic representations, regardless of sound source recognition. These impressions were mainly related to the spectral characteristics of the sounds [19]. The present results also suggested that we could communicate emotional impressions through onomatopoeic representations, enabling listeners to imagine the sound source correctly.

ACKNOWLEDGMENTS

REFERENCES

1   K. Tanaka, K. Matsubara, and T. Sato, "Onomatopoeia expression for strange noise of machines," *Journal of Acoustical Society of Japan*, **53**(6), 477–482 (1997) (In Japanese).

2   S. Iwamiya, and M. Nakagawa, "Classification of audio signals using onomatopoeia," *Soundscape*, **2**, 23–30 (2000) (In Japanese).

3   K. Hiyane, N. Sawabe, and J. Iio, "Study of spectrum structure of short-time sounds and its onomatopoeia expression," *Technical report of IEICE.*, SP97-125, 65–72 (1998) (In Japanese).

4   T. Sato, M. Ohno, and K. Tanaka, "Extraction of physical characteristics from onomatopoeia: Relationship between actual sounds, uttered sounds and their corresponding onomatopoeia," *Proceedings of Forum Acusticum Budapest 2005*, 1763–1768 (2005).

5   M. Takada, K. Tanaka, S. Iwamiya, K. Kawahara, A. Takanashi, and A. Mori, "Onomatopoeic features of sounds emitted from laser printers and copy machines and their contribution to product image," *Proceedings of 17th International Congress on Acoustics*, 3C.16.01 (2001).

6   K. Yamauchi, M. Takada, and S. Iwamiya, "Functional imagery and onomatopoeic representation of auditory signals," *Journal of Acoustical Society of Japan*, **59**(4), 192–202 (2003) (In Japanese).

7   M. Takada, K. Tanaka, and S. Iwamiya, "Relationships between auditory impressions and onomatopoeic features for environmental sounds," *Acoustical Science and Technology*, **27**(2), 67–79 (2006).

8   S. H. Wake and T. Asahi, "Sound retrieval with intuitive verbal descriptions," *IEICE Transactions on Information and Systems*, **E84**(11), 1568–1576 (2001).

9   R. Guski, "Psychological methods for evaluating sound quality and assessing acoustic information," *Acta Acustica united with Acustica*, **83**, 765–774 (1997).

10 W. H. Warren and R. R. Verbrugge, "Auditory perception of breaking and bouncing events: A case study in ecological acoustics," *Journal of experimental psychology: Human perception and performance..* **10**(5), 704–712 (1984).

11 J. A. Ballas, "Common factors in the identification of an assortment of brief every day sounds," *Journal of experimental psychology: Human perception and performance..* **19**(2), 250–267 (1993).

12 L. D. Rosenblum, "Perceiving articulatory events: Lessons for an ecological psychoacoustics," in *Ecological psychoacoustics* ed. G. Neuhoff (Elsevier Academic Press, San Diego, London, 2004), Chap.9, pp.219–248.

13 N. Fujisawa, F. Obata, M. Takada, and S. Iwamiya, "Impression of auditory imagery associated with Japanese 2-mora onomatopoeic representation," *Journal of Acoustical Society of Japan*, **62**(11), 774–783 (2006) (In Japanese).

14 International Phonetic Association, *Handbook of the international phonetic association* (International Phonetic Association, Cambridge, 1999).

15 T. Asano, *The dictionary of onomatopoeia* (Kadokawa Books, Tokyo, 1978).

16 C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, *The measurement of meaning* (University of Illinois Press, Urbana, Chicago and London, 1967).

17 W. W. Gaver, ''What in the world do we hear?: An ecological approach to auditory event perception,'' *Ecological Psychology*, **5**, 1–29 (1993).

18 A. Preis, H. Hafke, and T. Kaczmarek, "Influence of sound source recognition on annoyance judgment," *Noise Control Engineering Journal*, **56**(4), 288–293 (2008).

19 G. von Bismarck, ''Timbre of steady sounds: A factorial investigation of its verbal attributes,'' *Acustica*, **30**, 146–159 (1974).