INTERSPEECH 2009 BRIGHTON

Control of human generating force by use of acoustic information – Study on Onomatopoeic utterances for controlling small lifting-force

Miki Iimura[1], *Taichi Sato*[1], *and Kihachiro Tanaka*[2]

[1] School of Engineering, Tokyo Denki University, Tokyo, Japan
[2] Faculty of Engineering, Saitama University, Saitama, Japan
taichi@n.dendai.ac.jp, tanaka@mech.saitama-u.ac.jp

Abstract

We have conducted basic experiments for applying acoustic information to engineering problems. We asked the subjects to execute lifting actions while listening to sounds and measured the resultant lifting-force.

We used human onomatopoeic utterances as the sounds that are presented to the subjects aiming to make their lifting-force small.

Especially, we focused on the "emotion" or "nuance" contained in humans' utterances, which is a unique characteristic evoked by the utterance' acoustical features. We found that the emotion or nuance can control the lifting-force effectively. We also clarified the acoustical features that are responsible for effective control of lifting-force exerted by human.

Index Terms: onomatopoeia, emotion, kansei, acoustical feature, human sensitivity, lifting action

10.21437/Interspeech.2009-219

1. Introduction

1.1. Focus on human onomatopoeic utterance

In this research, we attempted to induce human to generate large and small force by using acoustical information. Especially here, we have focused on human onomatopoeic utterances as acoustical information to induce human to generate small lifting-force.

Onomatopoeia is a word that concisely expresses human senses, emotion, feelings, body rhythms, or actions in Japanese [1]. For example, there are the Japanese word, onomatopoeia "*suta-suta*" and "*tobo-tobo*". "*Suta-suta*" indicates a situation where a person is walking lively, briskly, or energetic mood. On the other hand, "*tobo-tobo*" indicates a situation where a person is walking slowly, sadly, or trudging. Also Japanese people can imagine these situations easily or intuitively, in other words, we can "sense" the meaning of these words through the filter of human sensitivity, that is, "kansei". Furthermore, these words can be used to intuitively and instinctively convey to others what actions or movements they perform. If actions or movements were described in detailed physical amounts or engineering terms, such actions or movements would be difficult to actually visualize.

These effects of onomatopoeia in Japanese have been already known in some fields by previous studies. For example in rehabilitation, it was shown that when rehabilitants are presented with onomatopoeia as they perform an action, it improves their control and memorization of this action. If young children are given the task of moving a stick while saying the word "*gyu*", then their memory of this action is improved. As trainers gain more experience in the rehabilitation of handicapped children, it can be shown that they make more use of onomatopoeia. These research results suggest that onomatopoeia is through the senses, and that actions presented together with such words are easier to remember. [2]-[4].

As described above, it is effective to use the word "onomatopoeia" to convey actions or movements to others. Furthermore, the word "onomatopoeia" is conveyed to listener via human utterances and has certain acoustical features. The utterances also contain information about intention, emotion, sense, and feeling by itself. For example, the utterance "to shout" means "to say something with loud voice". But it can be said that actual utterances contains joy, fear, or anger. That is, to shout "with joy", or "with fear". The listener perceives an onomatopoeic utterance as a sound that includes such information. It can be said these are caused by some "acoustical features".

The acoustical features are categorized into (1) loudness, (2) pitch, (3) timbre, (4) the speaker's emotion, and (5) nuance or intention [5], [6]. Among these, emotion and nuance are unique to human utterances. Emotion refers to a person's mood, e.g., happy, sad, or angry. Nuance refers to the speaker's implied intention, which may be interrogative, imperative, negative, or affirmative, and so on [7].

We think that there is a point we should focus on in the fact that humans can intuitively and sensitively grasp not only the meaning of onomatopoeia but also the acoustical features contained in human onomatopoeic utterances. This ability is called "kansei" in Japanese [8]. Especially, when human feel the presence of emotion or nuance in humans' utterances via "kansei", human grasps these characteristics by the extra acoustical features.

1.2. Application concept

We have planned to apply these findings, that is, the relationship between humans' lifting-force and sound, to some equipment for training of humans' operation in which the magnitude of force has to be adjusted.

An example of such equipment would be an electrical wheelchair with a joystick that detects force magnitude [10]. Concept of this training-equipment is as follows.

First of all, a trainee puts up a wireless-headphone and sits on the electrical wheelchair with joystick detecting the force magnitude. A trainer only instructs to operate the electrical wheelchair using the joystick while listening to the sounds. Then the trainee operates the electrical wheelchair using the joystick followed by the sounds. Therefore, these sounds that are used for this training-equipment are required not only to be the sound that can induce good-operation, but also to be the sound that can make beginners acquire the operation skill early as well as experts.

As described above, we can say that onomatopoeia is understood easily or intuitively through the human sensitivity without any explanation. Furthermore, we can apprehend the actions that is induced naturally by onomatopoeia or that is presented together with onomatopoeia are easier to remember or memorize. So we can expect an appropriated action as "output" by using the onomatopoeia as "input", aiming to using these relationship for the training-equipment.

Then, we explain what "output" that we expect is, that is, what an operation that needs training for beginners is. There are mainly "to start", "to move", and "to stop" as scene that needs the joystick operation-training of an electrical wheelchair. Especially in "to start", it is required to start slowly and smoothly. That is the cases not only to prevent suddenly-acceleration at starting, but also to move avoiding some obstacles with joystick adjustment, also to overcome low-hurdles.

Therefore, we try to clarify the most effective sound for inducing small force and apply this to the training-equipment.

1.3. Previous work and present study

In a previous work [9], we focused on "*gu*", a Japanese onomatopoeic utterance that expresses the idea of forceful, conveying an emotion or nuance that induces a person to exert a large lifting-force. We call this "*gu* with emotion (forceful)". On the other hand, we prepared also onomatopoeic utterance "*gu*" that was uttered in a monotone or with a lack of any extra emotion or nuance. We call this "*gu* without emotion".

Considering that emotion is a unique feature of human utterances, we investigated the effect of the presence or absence of emotion on lifting-force when the subjects listened to "*gu*" with emotion (forceful) and without emotion. Then, we found that "*gu*" with emotion can make people exert a large lifting-force. We also clarified that the acoustical features of emotion (forceful) that elicit the large lifting-force response are its sweeping high-frequency components [9].

In this work, we focused on an onomatopoeic utterance that can control small lifting-forces exerted by humans.

For this investigation, we used the "*su*", a Japanese onomatopoeia that expresses the meaning of forceless [11], conveying an emotion that induces a person to exert a small lifting-force. We call this "*su* with emotion (forceless)". As we did with "*gu*", we prepared also "*su*" without emotion, and investigated the effect of emotion (forceless) on lifting-force by comparing these two utterances.

More specifically, we explain the Japanese onomatopoeia "*su*" according to the results that we have already reported [11]. Based on these previous results, an unvoiced-consonant (such as "*ku*", the unvoiced-consonant corresponding to the voiced consonant "*gu*") is superior to a voiced-consonant for generating a small force in people. Because there is a tendency that unvoiced-consonant utterances such as "*ku*" have longer duration of consonant part than voiced-consonant utterances such as "*gu*". Based on this, we selected onomatopoeia "*su*" as the most effective one for inducing small force, because onomatopoeia "*su*" has a tendency to have the longest duration of consonant part when human utters some onomatopoeia naturally.

2. Lifting Action Experiments

The test was performed in a semi-anechoic room using the apparatus depicted in Fig. 1(a). Subjects, five males in their 20s with normal hearing, were instructed to lift the handle of a force gauge with one hand while listening to the sounds (This was the only instruction given).

Each test sound was presented to the subjects ten times in random order. That is, when we prepared five sounds for this experiment, a total of fifty sounds are presented to a subject. We set the interval of one minute between each sound. The sound was recorded simultaneously with the above action. The magnitude of each sound was 60 dB(A), and the duration was 2.5 seconds. An example of the sounds and force waveforms is shown in Fig. 1(b).

Before this experiment, the only instruction given to the subjects was "Please perform the lifting action while listening to the sound." The reason why we apply this method for experiment is because we concentrate on the "output" obtained via only "human intuition" or "human sensitivity", that is kansei, when the sounds are input to human. This means that it can indicate that even if there is no explanation but the effective sound as the instruction, it can control human lifting-force.

Furthermore, as we mentioned above, we can expect that it becomes earlier for a person to get used to the operation in the case that the person acquired the operation skill using onomatopoeia that reaches directly human intuition. Therefore, we evaluate the magnitude of lifting force exerted by test subjects by only listening to the sounds.

3. Effect of emotion on lifting-force

3.1. Effect of emotion in human utterances

Focusing on the extra features of human onomatopoeic utterances we analyzed how lifting-force is affected by the presence or absence of emotion in the onomatopoeic utterance "*su*". As shown in Fig. 2, the lifting-force for "*su*" with emotion is smaller than that for "*su*" without emotion. This means the subjects perceive emotion in "*su*" with emotion as a signal to reduce lifting-force.

In all experiments, the tendencies in the results were, unless specifically mentioned otherwise, almost exactly the same for all subjects. In this report, we therefore show the results for only one subject due to space limitations.

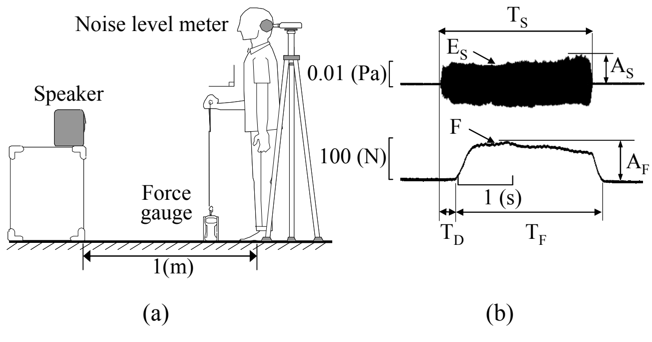


Figure 1: (a) *Experimental setup in a semi-anechoic room,* (b) *Example of the sound and force waveform.*

3.2. Investigation of acoustical features

Figures 3(a) and (b) show the sound waveform and frequency characteristics of "*su*" with and without emotion We investigated the vowel part and the consonant part and discovered five acoustical features of "*su*" with emotion:

(1) Absence of high-frequency components in the vowel part
(2) Presence of periodic amplitude variation in the sound waveform of the vowel part
(3) A wide distribution of frequency components in the consonant part
(4) Smaller amplitude in the consonant part than in the vowel part
(5) Long utterance duration in the consonant part

There is a possibility that these acoustical features make humans feel the presence of emotion in "*su*" with emotion; that is, they are what cause the lifting-force to be reduced.

4. Analysis of emotion (forceless)

4.1. Effect of absence of high-frequency components

As mentioned above, the acoustical features seem to have a large effect on lifting-force. Here, we examine the effect of the absence of high-frequency components in the vowel part.

Five onomatopoeic utterances were prepared and presented to the subjects: "*su*" with and without emotion and three other sounds processed from "*su*" without emotion by cutting off high-frequency components with 500-Hz, 800-Hz, and 2k-Hz low-pass filters (LPFs).

As shown in Fig. 4, the lifting-force was largest for "*su*" without emotion, and gradually decreased with the 2-, 800-, 500-Hz LPF sound. The lifting-force was smallest with "*su*" with emotion. These results indicate that the absence of high-frequency components in the vowel part affects the lifting-force, and it is a one of the acoustical features that makes humans exert a small lifting-force. It can be also said that the subjects perceive it as the signal to reduce lifting-force smaller.

Notably, the lifting-force for "*su*" with emotion was even smaller than that for the500-Hz LPF sound, which is same condition as "*su*" with emotion with respect to the absence of high-frequency components. This suggests that the other four acoustical features mentioned in 3.2 affect the difference in the magnitude of lifting-force.

4.2. Condition of acoustical features included in utterance for small lifting-force

Next, we investigated the effect of the other four above-mentioned acoustical features on the lifting-force.

We prepared seven sounds: the 500-Hz LPF sound (sound A), "*su*" with emotion (sound B), three sounds processed to have one acoustical feature in sound A, and two sounds processed to have three or four acoustical features in sound A simultaneously. The five additional sounds are composed of sound A with and are some acoustical features based on sound B added. The details of these sounds are as follows;

(a) Sound A: The same 500-Hz LPF sound used in the previous analysis.
(b) A*Wf(V) (sound [i]): Sound A processed to have periodic amplitude variation in the sound waveform of the vowel part. (Note that process of amplitude variation is faithfully imitated to the waveform of Sound A
(c) A*Fr(C) (sound [ii]): Sound A processed to have a consonant part of white noise with a wide distribution of frequency components

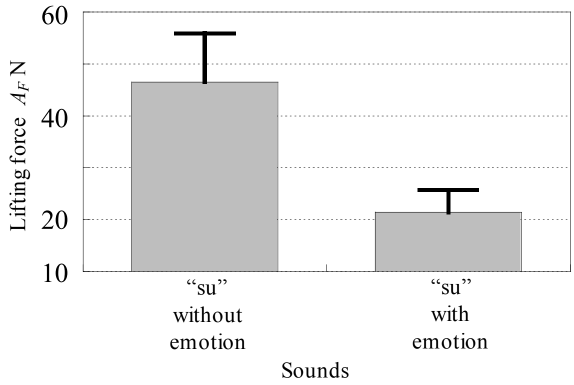
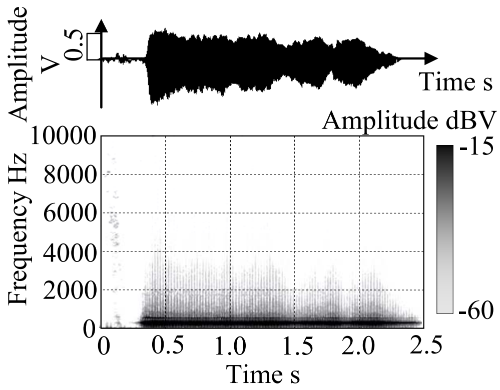
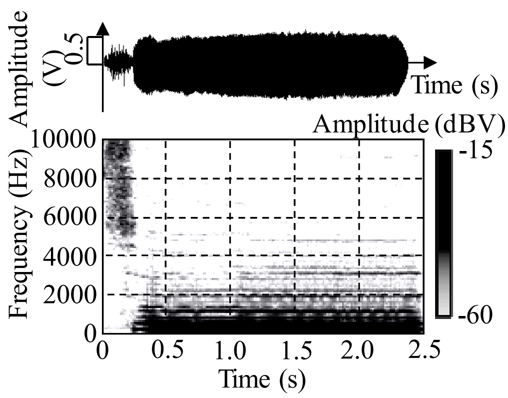


Figure 2: *Test result showing effect of emotion included in onomatopoeic utterance.*



(a)



(b)

Figure 3: (a) *Sound waveform and frequency characteristic of onomatopoeic utterance "su" with emotion,* (b) *Sound waveform and frequency characteristic of onomatopoeic utterance "su" without emotion*

(d) A*Wf(C) (sound [iii]): Sound A processed so that consonant part has smaller amplitude than the vowel part
(e) [i]+[ii]+[iii]: Sound A processed to have the [i], [ii], and [iii] simultaneously.
(f) [i]+[ii]+[iii]*Du(C): Processed [i]+[ii]+[iii] to have the long duration of sound in the consonant part simultaneously
(g) Sound B: "*su*" with emotion

The results are shown in Fig. 5. At first, we focused on the result with sounds [i], [ii], [iii] and with sound A. Compared with the lifting-force with sound A, the lifting-force for [i] and [iii] decreased; however, the lifting-force for [ii] increased. This means that the presence of a periodic amplitude variation in the sound waveform of the vowel part and having smaller amplitude of the consonant part are effective for reducing lifting-force. However, a wide distribution of frequency components in the consonant part was not effective for this subject, indicating that he had difficulty perceiving the difference in sounds that have one acoustical feature as the signal to reduce lifting-force small, especially for [ii].

On the other hand, the lifting-force for [i]+[ii]+[iii] was a little smaller than that for [i], [ii], and [iii], and the lifting-force for [i]+[ii]+[iii]*Du(C) was much smaller than that for [i]+[ii]+[iii]. It can be also seen that the lifting-force for [i]+[ii]+[iii]*Du(C) was as small as that for sound B. This means that it was easier for this subject to perceive the sounds that had multiple acoustical features simultaneously as the signal to reduce lifting-force.

To summarize, the more acoustical features the sound has simultaneously, the more the lifting-force tends to decrease. That is, to elicit a small lifting-force, it is necessary for "*su*" to have all five acoustical features simultaneously.

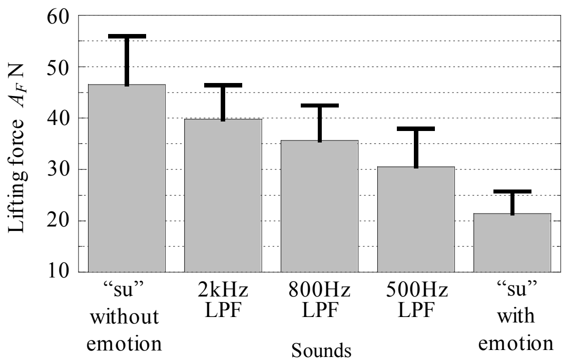
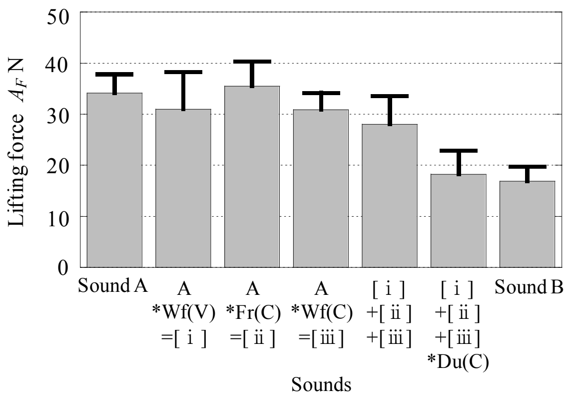


Figure 4: *Effect of absence of high-frequency components in vowel part.*



*Wf: Sound waveform of "Sound A" is processed
*Fr: Frequency characteristic of "Sound A" is processed
*Du: Duration of Consonant of "Sound A" is processed
(V): Vowel, (C): Consonant

Figure 5: *Test result when acoustical features were included in utterance for small lifting-force.*

5. Conclusions

We attempted to induce human to generate small lifting-force using the onomatopoeic utterances. We clarified that the emotion or nuance in human utterances can elicit a small lifting-force. We also clarified the acoustical features and the conditions that are responsible for inducing human to generate small lifting-force.

We will attempt to control much smaller force magnitudes by using an artificial sound.

6. Acknowledgements

This research was supported by a Grand-in-Aid for the COE Research Project of the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

7. References

[1] N. Osaka, Study on words of sensitivity, (in Japanese), Shin-yo-sha, Japan, 1999.

[2] Y. Fujino, K. Inoue, M. Kikkawa, S. Horie, E. Nishina, T. Yamada, Y. Sagisaka, Actual condition survey of the usage of the sport onomatopoeias in Japanese, Tokai J. of Sports Medical Science (in Japanese), No.17, (2005), p.28-38.

[3] K. Tohya, Effects of verbalization of mimetic phoneme on a motor-memory in infants, J. of Education Psychology (in Japanese), Vol.40, No.2, 1992, pp. 148-156.

[4] K. Tohya, Developmental changes in effect of verbalization strategy of onomatopoeia on the motor-memory, J. of Education Psychology (in Japanese), Vol.40, No.4, 1992, pp. 436-444.

[5] S. Murata, Encyclopedia of the sounds (in Japanese), Maruzen-sha, 2006, pp. 46-47, 94, 178-181, 729-731, 799-803.

[6] I. Tamori, Enjoy Onomatopoeia (in Japanese), Iwanami-shoten, 2004, pp.134-151.

[7] A. Matsumura, Daijisen (in Japanese), Shogaku-kan, 1998, p.2029.

[8] S. Inokuchi, Kansei Information Processing, The Journal of the Institute of Electronics, Information, and Communication Engineers (in Japanese), Vol.80, No.10, 1997, pp. 1007-1012.

[9] T. Sato, K. Oyama, M. Iimura, H. Kobayashi, K. Tanaka, Control of human generating force by use of acoustic information―Utilization of onomatopoeic utterance, JSME International Journal, Series C, Vol. 49, No. 3, 2006, pp.687-694.

[10] T. Hayama, T. Sato, Y. Noguchi, A. Funakubo, K. Tanaka, Supporting for operation of electric wheelchair by use of acoustic information, JSME annual meeting 2007(7) (in Japanese), 2007, pp. 259-260

[11] M. Iimura, T. Sato, K. Tanaka, Control of Human-generating Force by Use of Onomatopoeia ―Method for Controlling Light Force, JSME Conference on Information, Intelligence and Precision Equipment (in Japanese), No. 07-7, 2007, pp.161-166.