Voice Pitch Control Using a Two-Dimensional Tactile Display

Masatsugu SAKAJIRI[1], Shigeki MIYOSHI[2], Kenryu NAKAMURA[3],
Satoshi FUKUSHIMA[3] and Tohru IFUKUBE[4]

[1]Faculty of Health Sciences,
Tsukuba University of Technology, 4-12-7 Kasuga, Tsukuba, Ibaraki 305-8521, Japan
e-mail: sakajiri@cs.k.tsukuba-tech.ac.jp
[2]Research and Support Center on Higher Education for the Hearing and Visually Impaired,
Tsukuba University of Technology
[3]Research Center for Advanced Science and Technology, the University of Tokyo
[4]Institute of Gerontology, the University of Tokyo

Abstract: It is difficult for the deafblind, or the hearing impaired, to control the pitch of their voice because they cannot perceive it. In particular, when singing, it is very difficult for them to control the pitch of their voice because they need to maintain a stable tone. We have developed a voice pitch control system to assist their singing by means of a tactile display. Using this tactile feedback system, we verified in a previous study that two deafblind subjects were able to control the pitch of their voices with as much accuracy as hearing children. By using this tactile feedback system, the correspondence between musical scale and proprioceptive sensation (muscular sensation, and so on) of the two subjects was returned to pre-hearing loss levels. They sung using not only tactile feedback but also proprioceptive feedback. In this paper, we investigate the ability of hearing subjects to control the pitch of their voice using our tactile feedback system, but without auditory feedback. Seven hearing participants were examined under two conditions—"without tactile feedback" and "with tactile feedback"—to ascertain their abilities to control their pitch while subjected to a masking noise. The results indicated that hearing subjects could not sing with an accurate pitch using only the proprioceptive feedback ("without tactile feedback" condition).
Keywords: deafblind, hearing impaired, singing, tactile display, voice pitch

1. Introduction

A tactile aid that presents vocal information as tactile stimulus is an assistive device for the hearing impaired [1],[2]. For deafblind people, i.e., those who are totally blind and totally deaf and who, therefore, cannot use visual information, a tactile device is the only means by which they can perceive information. Voice pitch and frequency is important information for recognizing a change in accent or intonation. Moreover, voice pitch and frequency is also important for the recognition of a musical scale, which is required to sing a song. Tactile aids that present voice pitch using tactile stimulus are used for understanding context or for assisting lip reading [3],[4]. However, only a few studies have been made on presenting a musical scale using a tactile display, or on voice training to adjust musical interval change using tactile feedback. In light of this, we have developed a

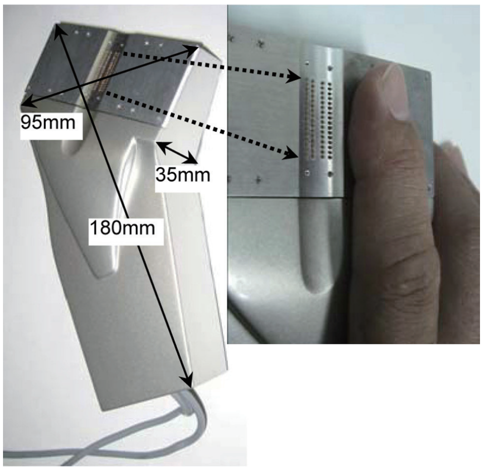

Fig. 1. Main unit of tactile display

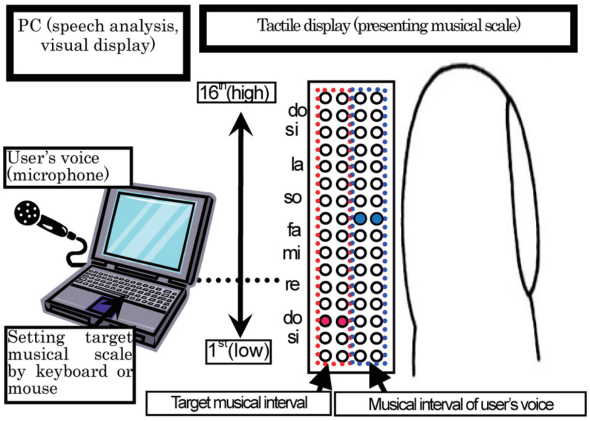

Fig. 2. Schematic diagram of the voice pitch control system

voice pitch control system that uses a two-dimensional tactile display [5]. The results of our research showed that two deafblind subjects could control their voice pitch with as much accuracy as hearing children. Using this tactile feedback system, the correspondence between musical scale and proprioceptive sensation (muscular sensation, and so on) of the two subjects was returned to pre-hearing loss levels. The subjects sung using not only tactile feedback but also proprioceptive feedback. Auditory feedback and proprioceptive feedback are significant cues in acquiring the ability to control pitch [6], and deficiency in the pitch control ability of a deaf person suggests that auditory feedback is important [7]. On the other hand, singers can maintain control of their pitch even when they are not able to hear their own voices. Auditory feedback for singers in a choir is sometimes masked by the sound of other singers [8]. Untrained singers are able to sing to a certain level of accuracy without an auditory feedback condition [9]. Auditory feedback is important for accurate pitch control; however, one can control pitch frequency to a certain level of accuracy without auditory feedback. In other words, without auditory feedback one is able to sing to a certain level using proprioceptive feedback, which is formed by the correspondence between musical scale and proprioceptive sense. In this paper, we investigate the ability of hearing subjects to control the pitch of their voice, without auditory feedback, using our tactile feedback system.

2. Voice Pitch Control System

The voice pitch control system consists of a personal computer (PC) and the main unit of a tactile display shown in Fig. 1. When they use the system, the user places the ball of the first joint of the right index finger on the tactile display to perceive a stimulus. The size of the main unit of the tactile display is 180 mm × 95 mm × 35 mm, it weighs 180 g and it is connected to a USB connector of the PC. Electric supply is only from the USB connector.

A schematic diagram of the voice pitch control system is shown in Fig. 2. For this system, the first step of operation is to set the target musical scale using the PC. Next, a deafblind person vocalizes so that this input can become the same musical scale as the target musical scale. The voice input from a microphone connected to the PC is taken in 10 kHz bands of sampling frequencies (Hanning window length: 51.2 ms.). The voice pitch frequency is calculated by a fast Fourier transform (FFT) cepstrum analysis. The musical scale corresponding to the calculated voice pitch frequency is displayed on the PC. Similarly, a vibrotactile stimulus is presented to the corresponding musical-scale position on the tactile display. The tactile display shown in Fig. 2 consists of 64

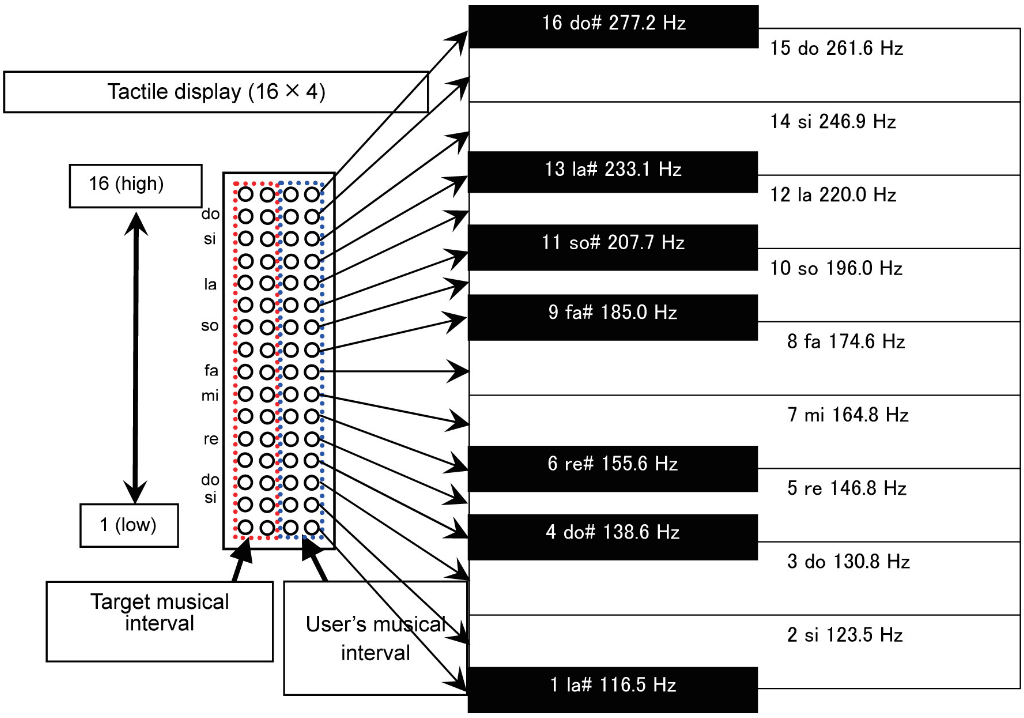

Fig. 3. Correspondence of musical scale and stimulus pins

piezoelectric vibrators in four rows (16 × 4 rows). The diameter of each stimulus pin is 0.6 mm. The displacement of each pin can be changed from 0 mm to 0.1 mm (0 to peak), which corresponds to the sensation level of the vibration in the range of 12–16 dB (S.L.: Sensation Level). The two rows on the left side of the tactile display present the target musical scale, while the two rows on the right side present the musical scale corresponding to the voice pitch frequency of a user. When the target musical interval (left side) and the musical interval of the user's voice (right side) are in the same position, the user's pitch stays on key.

The correspondence of musical scale and stimulus pin is shown in Fig. 3. The pin number is the stimulus pin of the tactile display. The musical scale uses an equal temperament of 12 degrees. One stimulus pin corresponds to a semitone with an equal temperament of 12 degrees. A male voice ranges from C4# to A2#, while the range of a female voice is from C5# to A3#. The user is able to select one of these two modes by using a keyboard or a mouse. The frequency range of each musical scale is shown in Fig. 3. In addition, a semitone is converted into 100 cents that is calculated on the basis of C3 in the male voice and C4 in the female voice.

3. Method

The subjects used in the experiment were one female and six males. All subjects were normal-hearing university students. The mean age of the subjects was 21.9 years, ranging from 20 to 23 years old.

The voice pitch control system shown in Fig. 2 was used in this experiment. During the test, all subjects wore sound-isolating headphones (Sennheiser HD380 pro) and a head-mounted microphone (Audio-technica AT810F) positioned at a constant microphone-to-mouth distance of 1 cm, placed off-center at the left corner of the mouth. The masking noise used was a white noise bandpass filtered (12 dB/octave) at 50 Hz and 2000 Hz. The sound pressure level of the noise was 100 dB. All subjects were tested in a double-walled sound booth. A sound level meter (Bruel & Kajaer Type 2231) was used for visual feedback to assist subjects in controlling the intensity of their vocal output, and was positioned at a constant microphone-to-mouth distance of 0.5 m. The subjects were instructed to maintain the intensity of their voices within 55 dB to 85 dB using the sound level meter.

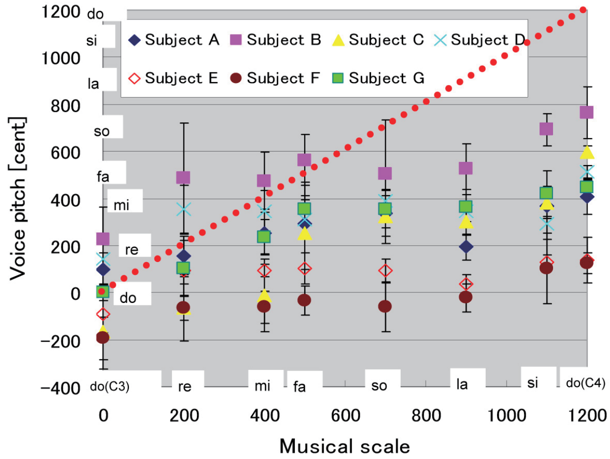

Fig. 4-a Average voice pitch for the first session (without tactile feedback).

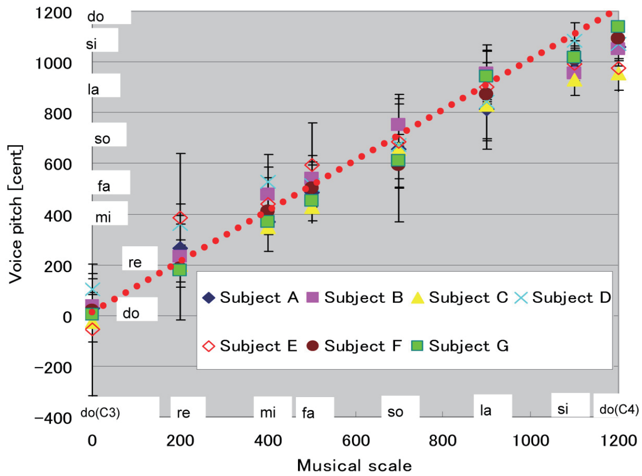

Fig. 4-b Average voice pitch for the second session (with tactile feedback).

Eight musical note cards—do (C3), re, mi, fa, so, la, si, do (C4)—were presented to each subject at random (where the single female subject was one octave higher). Each subject was instructed to sing the presented musical note. The subject phonated the name of the presented musical note card. Each card was presented five times. The masking noise was added in all trials when the subject sung.

These trials were implemented in two sessions. The first session, (1), was without tactile feedback, while the second session, (2), was with tactile feedback. Without tactile feedback means the subject sings using only proprioceptive feedback (auditory feedback is masked). With tactile feedback, the subject sings using tactile feedback and proprioceptive feedback (auditory feedback is masked in an analogous way.) Before implementing the second session, each subject practiced for 20 minutes as follows. First, the subject practiced singing do (C3), re, mi, fa, so, la, si, do (C4) in sequence to match the presented musical scale (pitch frequency). Second, the subject practiced singing descending musical notes [do (C4), si, ra, so, fa, mi, re, do (C3)], the same as above. Third, the subject sung one of the eight musical note cards presented at random. During the second session, the subject sung three times, and the third-sung pitch frequency was treated as the experimental data. The tactile display was covered by a case to prevent the subject from seeing the vibrating pins.

4.Results and Discussions

Figure 4-a shows the average for the first session, without tactile feedback, for all subjects. The horizontal axis of Fig. 4-a is the target musical scale presented by the card. The vertical axis is the musical interval with a cent value that is based on C3 (do). One semitone equals 100 cents. C3 (do) equals 0 cents, D3 (re) equals 200 cents, E3 (mi) equals 400 cents, and F3 (fa) equals 500 cents, and so on. Each set of data is the average of five trials and the error bar shows the standard deviation. If the subject sung with an accurate pitch corresponding to the target musical scale, the data are on the red dotted line. From the results shown in Fig. 4-a, voice pitch data increased as the target musical scale increased; however, the the slope is shallower than the red dotted line. Although subject D reported hearing his own voice, the data showed no significant results compared with the other subjects. Figure 4-b shows the average for the second session, with tactile feedback, for all subjects. From the results shown in Fig. 4-b, voice pitch data increased as the target musical scale increased commensurately, at which point the gradient of the slope is approximately the same as the red dotted line. These results indicate that hearing subjects cannot sing with an accurate pitch using only proprioceptive feedback ("without tactile feedback" condition). Details of our results appeared in our previous study [10].

Acknowledgement

This paper is the excerpt from a paper presented at the 2011 IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2011) in Anchorage, AK, USA [10]. This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (A)21240058 and (C)21500896.

References

[1] I. Summers, "Tactile aids for the hearing impaired;" Whurr Publishers, 1992.

[2] C. Wada., H. Shoji. and T. Ifukube, "Development and evaluation of tactile display for a tactile vocoder;" Technology and Disability, Vol. 11, No. 3, 1999, pp. 151-159.

[3] M. Rothenberg and R. Molitor, "Encoding voice fundamental frequency into vibrotactile frequency;" Journal of Acoustical Society of America, Vol. 66, No. 4, Oct. 1979, pp. 1029-1038.

[4] K. Grant, L. Ardell, P. Kuhi and D. Sparks, "The transmission of prosodic information via an electrotactile speechreading aid;" Ear and Hearing, Vol. 7, No. 5, 1986, pp. 328-335.

[5] M. Sakajiri, S. Miyoshi, K. Nakamura, S. Fukushima and T. Ifukube, "Voice pitch control using tactile feedback for the deafblind or the hearing impaired persons to assist their singing;" IEEE International Conference on Systems, Man, and Cybernetics, 2010, pp. 1483-1487.

[6] L. Rasphael, G. Borden and K. Harris, "Speech science primer - Physiology, acoustics, and perception of speech 5th edition;" Lippincott Williams & Wikins Inc., 2007.

[7] T. Nakata, S. Trehub, C. Mitani and Y. Kanda, "Pitch and timing in the songs of deaf children with cochlear implant;" Music Perception, Vol. 24, No. 2, 2006, pp. 147-154.

[8] S. Temström and J. Soundberg, "Intonation precision of choir singers;" J. Acoustical Society of America, Vol. 84, No. 1, 1988, pp. 59-69.

[9] C.Watts, J. Murphy and K. Barnes-Burroughs, "Pitch matching accuracy of trained singers, untrained subjects with talented singing voices, and untrained subjects with nontalented singing voices in condition of varying feedback;" Journal of Voice, Vol. 17, No. 2, 2003, pp. 185-194.

[10] M. Sakajiri, S. Miyoshi, K. Nakamura, S. Fukushima and T. Ifukube, "Voice pitch control ability of hearing persons with or without tactile feedback using a two-dimensional tactile display system;" IEEE International Conference on Systems, Man, and Cybernetics, 2011, pp. 1069-1073.