Auditory-Centered Vocal Feedback System Using Solmization for Training Absolute Pitch Without GUI

Nozomu Yoshida[1](✉), Kosaku Namikawa[1], Yusuke Koroyasu[1],
Yoshiki Nagatani[2], and Yoichi Ochiai[1,2]

[1] University of Tsukuba, Tsukuba 305–8577, Japan
{nozo,namikawa,koroyu,wizard}@digitalnature.slis.tsukuba.ac.jp
[2] Pixie Dust Technologies, inc., Tokyo 101-0061, Japan
yoshiki.nagatani@pixiedusttech.com
https://digitalnature.slis.tsukuba.ac.jp
https://pixiedusttech.com

Abstract. This study proposes an auditory-centered training system using solmization to artificially acquire and train absolute pitch by providing vocal feedback of musical notes through a musician's voice. Most current training systems and applications for absolute pitch acquisition have focused on providing visual feedback. However, many people having perfect pitch describe that they hear music as words rather than envisioning visual notes. Therefore, we propose a training system that does not require a graphical user interface. In an experiment with 10 participants, our system's training with vocal feedback improved six non-potential absolute pitch users' absolute pitch by approximately 25%, although extant system with visual feedback didn't make an improvement.

Keywords: Auditory feedback · Learning method · Pitch training

1 Introduction

This study proposes a auditory-centered training system using solmization to artificially acquire and train absolute pitch by providing vocal feedback of musical notes by musician's voice. Absolute pitch is the ability to name or produce a note of a given musical tone without an external reference tone [24]. People with this ability are rare, and its prevalence in the general population in North America and Europe is estimated to be less than 10,000 people [2,20,24]. Absolute pitch is not an essential skill; however, having absolute pitch can sometimes be beneficial, especially for professional musicians. For example, it is difficult to detect minor pitch errors by determining pitch relations in atonal music [5].

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-85610-6_1) contains supplementary material, which is available to authorized users.

Published by Springer Nature Switzerland AG 2021
C. Ardito et al. (Eds.): INTERACT 2021, LNCS 12935, pp. 3–19, 2021.
https://doi.org/10.1007/978-3-030-85610-6_1

Additionally, a composer with absolute pitch can sometimes work efficiently with one note at a time, compared to a composer without absolute pitch who has to work in two-note increments [15].

Most current absolute pitch training systems provide musical note names and pitches via the graphical display as visual feedback. Here, visual feedback refers to the feedback of information via the user's vision. However, most individuals having absolute pitch state that they hear music as words rather than envisioning it as notes [12,17]. Therefore, the proposed method adopts a musical note-singing method called solmization; a method of singing while calling out each sound by its note name (e.g., *do, re, mi*). Solmization provides singers with vocal feedback of notes in the form of words, facilitating pitch training. Here, vocal feedback refers to the feedback of information via the user's hearing by speech. Furthermore, our proposed system does not require a Graphical User Interface (GUI), in contrast to traditional training systems, such as tuners and the visualized guide melody of Karaoke.

Our contributions are as follows. First, we constructed a system that allows users to recognize note names aurally via solmization. Second, we compared our vocal feedback system with existing visual feedback systems and identified which type is more effective.

2 Related Work

2.1 Pitch Training

Various training methods have been developed to improve the musical sense of pitch [9,10,16,21,23]. Solfege, which applies solmization at the scale of C, uses the *sol-fa* syllables to name or represent the tones of a melody or scale. It refers to training that involves writing down the music heard by ear onto a score and singing the melody while looking at it. Additionally, musical-instrument and singing instruction can be regarded as sound training that requires a teacher. It is known that paying attention to the pitch of both voice and instrument improves the vocal and aural senses.

Karaoke provides a visualized training tool for similar purposes. For example, the Live Dam Stadium DX-G[1] is an analytical scoring system for Karaoke, and it feeds back the singer's pitch in real-time as a visualized melody. There are several such pitch-training applications. In the work of [13], a self-training mobile application was built that improved the sense of pitch as the user sang into an interactive user interface. Additionally, [6] provided a singing and pitch-training application for children with cochlear implants. With the Ear Training Myu-Tre[2] and OtoAte[3] applications, users train their absolute pitch by listening to a specific pitch and guessing the note. The C-major scale (i.e., CDEFGAB) is displayed on the screen to assist the user visually. The Pitch[4] application provides information on note names and pitches of 3–7 scales, and by guessing the presented

---

[1] https://www.clubdam.com/app/dam/seimitsusaiten/, Accessed on Jan 18, 2021.

[2] https://apps.apple.com/jp/app/id1015269208, Accessed on Jan 18, 2021.

[3] https://apps.apple.com/jp/app/id763325064, Accessed on Jan 18, 2021.

[4] https://apps.apple.com/jp/app/id942373213, Accessed on Jan 18, 2021.

notes, users train their relative pitch. This tool requires a GUI. The Singing Assessment and Development [11,26] tool was a pioneer project that incorporated real-time visual feedback into educational software for singing. It provided a pedagogical tool for assessing and developing the singing voices of UK primary school children using real-time visual feedback technology. The ALBERT tool [22] provides acoustic output while monitoring laryngeal action. Thus, the system provides a greater diversity of feedback displays and parameters, such as F0, larynx closed quotient, spectral ratio, sound-pressure level (amplitude), shimmer, and jitter. The SING and SEE [3] tools distinguished three parameters (i.e., pitch, vowel identification, and timbre). Although SING and SEE focused on maximizing visual feedback, the VOXed project [25] involved psychologists, vocalists, singing instructors, and students to investigate the usefulness of direct feedback using commercially available visual feedback software. MiruSinger [18] proposed a method of automatically assessing a user's singing ability by providing visual feedback of musical information without using musical notation.

2.2 HCI Study Using Auditory Feedback

Kuber et al. demonstrated the potential for non-visual speech-based audio and haptic feedback systems in the design of accessible memory games to aid visually impaired users who would otherwise remain excluded from most mainstream gaming applications [14]. They recommended that speech-based feedback be presented in conjunction with haptic cues when developing a non-visual game. Stina et al. proposed a running-technique training system that combined real-time visual and audio feedback [19]. In their proposed system, the visual feedback displayed acceleration, and the audio feedback conveyed the runner's rhythm using beeps. This study suggested that, if the users were in a situation where the screen could not be seen, the audio feedback would provide suitable feedback. Therefore, when inducing user movement, voice feedback may be suitable if the amount of data transmitted is succinct, such as that of rhythm or note names. Guardati et al. proposed a method to assist with writing rehabilitation [8]. The system was designed for human-in-the-loop operations, and it could analyze handwriting in real-time while providing vocal feedback to guide the patient during exercise. The advantage of the proposed method is that it could be easily used at home without a trainer. Christiansen et al. investigated in-vehicle input and output techniques to compare their effects on driving behavior and attention, finding that using audio resulted in significantly fewer eye glances but longer task completion times with inferior primary driving task performance compared with visual cues [4]. Gaver et al. proposed Auditory Icon, an alternative approach to the use of sound in computer interfaces that emphasized the role of sound in conveying information about the world to the listener [7]. By combining auditory and visual feedback, all information could be accessed more efficiently. For example, when an error occurred on a computer, a beeping sound in addition to a warning screen display would instantly inform the user of the situation.

2.3 Vocal Feedback as Cybernetics

Cybernetics is a research field that attempts to study the structure of control and communication/information transmission in a wide range of mechanical, biological, and social organizations from basically the same methodological perspective [27]. The core concept of the discipline is circular causality or feedback, i.e., where the outcomes of actions are taken as inputs for further action [1]. The vocal feedback system also includes the concept of Cybernetics in terms of performing a feedback loop. The user communicates with the vocal feedback system by actions (such as one's voice or UI operation), receives feedback by listening to the system output, and performs self-regulation after referring to it.

2.4 Position of This Study

The research question for this study is "can absolute pitch be trained by a vocal feedback system?" Extant pitch-training methods provide visual feedback containing the note information. However, individuals with absolute pitch describe their experience as audible instead of visual. Furthermore, previous studies indicate that auditory feedback is useful for providing simple information. Therefore, we construct a system that allows users to aurally recognize note names. Then, we compare the auditory feedback system with the extant method to identify which one is more effective.

3 Implementation

3.1 Design Principles

Figure 1 shows the design of our proposed system. It consists of five modules.

Random Test Generator: A module that determines the correct pitch. The fundamental tone is 60 (C4, 261.6 Hz), and it randomly selects the correct pitch from the C-major scale range at one octave from the fundamental tone.

User Input: A module in which the user inputs a note number via a slider. The slider's pitch change rate per width is random for each quiz to prevent the user from memorizing the note according to slider position.

Oscillator: A module that sounds input and correct pitches. The triangular-wave oscillator reproduces the pitch corresponding to the input. The user interface has a fundamental tone button, an input tone button, and correct tone button, but they are not played simultaneously to prevent delivering hints for guessing.

Vocal Feedback: A module that outputs vocal feedback corresponding to the input pitch. From the input sounds, the microtonal music is rounded off to the nearest semitone unit. The sound source material used for vocal feedback is sampled from a vocalist who had mastered vocal music.

Display: A module that returns the user's input pitch in text form. The displayed pitch entered with the slide bar is rounded off to the nearest semitone unit.

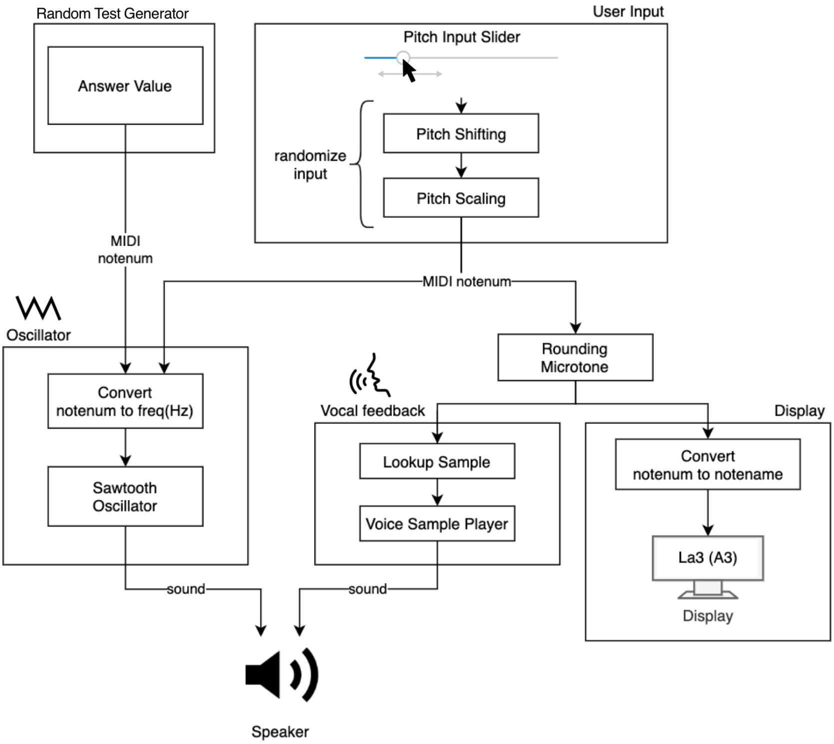

Fig. 1. Design of absolute pitch trainer consisting of five modules: random test generator, user input, oscillator, vocal feedback, and display. In our system, the note name is given auditory feedback in real-time, unlike the existing software learning method that visually feeds back the note name.

3.2 User Interface

Figure 2 provides a screenshot of the interface, and Table 1 describes each button of the display interface. The interface provides a random pitch test for training, and it is used as the proposed system for absolute pitch training in an experiment. In addition to the vocal feedback feature, our application has a visual feedback feature for comparative experiments, as shown in Nos. 12 and 13 in Table 1. The visual feedback feature displays staff notation with note names. In the proposed vocal feedback system, GUI-free pitch training is possible. However, because the experiment requires the user to provide the pitch input using a slide bar, a GUI was adopted. The only difference between vocal and visual feedback is the type of note-name fed back, and the other specifications are the same.

Table 1. User interface description. Each number in Table 1 corresponds to the number in Fig. 2.

No	Name	UI type	Description
1	Random test	Display	The designated note name as random test is displayed
2	Slide bar	Drag	Manipulate the pitch by dragging the slide bar
3	Keynote[a]	Button	Listen to the fundamental triangle wave by pressing a button
4	Start	Button	The bar's pitch's triangular wave at the current position will be fed back by pressing the button. Also, it resumes oscillator feedback from Pause
5	Enter	Button	Confirm the pitch of the bar at the current position as an answer
6	Pause	Button	Stop playing the oscillator by pressing a button
7	Input note	Display	The input note name is displayed
8	Correct pitch	Button	The note name corresponding to the entered pitch is displayed
9	Next question	Button	Move on to the next question by pressing a button
10	Mode selection[a]	Switch	Switch between absolute pitch mode and relative pitch mode
11	Vocal feedback	Toggle	Turn on/off the vocal feedback of the note name
12	Visual feedback	Toggle	Turn on/off the visual feedback of the note name
13	Visual feedback	Display	The staff notation and the note name are fed back when the visual feedback is on

[a] These features are for relative pitch training. In this study experiment, we focused only on absolute pitch, so we did not use them.

4 Experiment

We anticipated that the vocal feedback of the note name presented by the proposed system would be more effective in improving the absolute pitch than the visual feedback of the existing system. Thus, we conducted an experiment to determine the same. The experiment consisted of a preliminary test and a main test. Additionally, a pre-questionnaire was conducted to determine attributes, such as the participants' musical experience prior to the experiment. A post-survey was also conducted to evaluate the experiment process and the usability of the system. The experiment was conducted with the approval of our institution's ethics review committee.

4.1 Participants

We recruited 10 men with no reported physical or mental disability ranging in age from 20 to 24 years old as participants (M: 22; SD: 1.3). We accepted participants regardless of their musical experience. All participants have never received singing instruction or music theory instruction, though four people (P2, P8, P9, P10) had experience playing the piano or electric piano, and three people (P2, P8, P10) had ever taken the piano lesson.

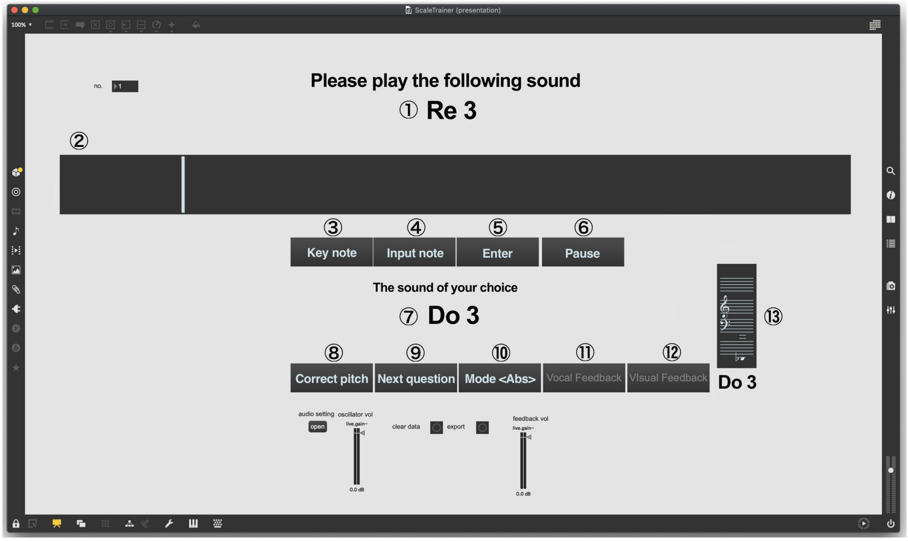

Fig. 2. User interfaces for training and learning absolute pitch. Each number in Fig. 2 corresponds to the number in Table 1. There are two types of training modes: the auditory feedback mode (proposed system) and the visual feedback mode (existing system). There is also a test mode for absolute pitch test, which excludes the training mode's feedback function.

4.2 Preliminary Test

Using an existing pitch-training application (Ear Training Myu-Tre), the participants attempted the absolute pitch quiz 10 times and recorded their scores. The contents of the test was as follows. First, for each trial, one of the seven C-major scale notes was given as a piano note sound: *Do*/C, *Re*/D, *Mi*/E, *Fa*/F, *Sol*/G, *La*/A, and *Si*/B. Next, the participants attempted to identify the note name at the same pitch as the given piano note sound. The score was recorded as "correct" or "incorrect," but the pitch distance between the correct pitch and the input pitch was not recorded.

4.3 Main Test

We conducted the main test to verify the hypothesis: vocal feedback of the note name by the proposed system is more effective in improving the absolute pitch than the visual feedback of the note name by the existing system-based counterpart. For the convenience of explanation, we defined the training provided by the proposed system as A. The training provided by the existing system is B. The absolute pitch test, C, is shown in Table 2. In this experiment, the participants were trained in absolute pitch using both systems (A and B). Additionally, we quantified the improvement of the absolute pitch tests conducted before and after the training. They took test C before and after training A and B. To reduce possible bias caused by the training order, five participants trained on the proposed system first, and the other five trained on the existing system first. The order of the experiments is shown in Table 3.

Table 2. Definition of the experimental process

Process	Definition	Time
A	Absolute pitch training with the vocal feedback of the proposed system	5 min
B	Absolute pitch training with the visual feedback of the existing system	5 min
C	An Absolute pitch test with 10 questions	Until finish

Training Task: A and B. The procedure of the training tasks A and B is shown in Table 4. The only difference between A and B is the note name's type of feedback; The note name is fed back auditory in the A and visually in the B.

Absolute Pitch Test: C. the procedure of the absolute pitch test is shown in Table 5. Two values were recorded in the system: an integer value as the midi-note number of the correct note name and a real value as the midi-note number of the input note name. Therefore, it could measure the distance between the correct and input pitches. For example, when the correct pitch was C4 (midi-note #60) and the input pitch was C#4 (midi-note #61). A distance of 1.0 was recorded as an error. Both the correct and input pitches could be octaves, and the distance was calculated by the difference between the input pitch and the nearest correct pitch.

5 Result

5.1 Preliminary Test

For grouping purposes, in the preliminary test, the four people who recorded perfect scores were defined as "potential absolute pitch," and the other six were defined as "non-potential absolute pitch." The scores from the preliminary test for potential and non-potential absolute pitch are shown in the second columns of Tables 6 and 7, respectively. All potential absolute pitches had experience playing the piano or electric piano, and three (P2, P8, P10) had ever taken the piano lesson for several years between 4 and 13 years old. All non-potential absolute pitches had neither experience playing the instruments nor lessons of any musical instruments (except for the class in the school).

Table 3. Experiment pattern

Pattern	# of Participants	Order
1	5	C A C Interval C B C
2	5	C B C Interval C A C

Table 4. Procedure of training task: A, B

Step	Description
1	A randomly specified note name is displayed; participants imagine the pitch and note name
2	Operate the slide bar to approach the pitch you imagined while listening to the oscillator sound
3	Stop the bar at the position where the specified note name is auditory (A) or visually (B) fed back, and press the enter button. At that time, be sure that the pitch name and pitch are associated and memorized
4	Listen to the correct pitch. At that time, be sure that the floor name and pitch are associated and memorized
5	Click the "next question button" to move to the next question
6	Repeat a series of steps 1–5 for 5 min

Table 5. Procedure of absolute pitch test: C

Step	Description
1	A randomly specified note name is displayed; participants imagine the pitch of that note name
2	Operate the slide bar to approach the pitch you imagined while listening to the oscillator sound
3	Stop the bar at the position that is thought to be correct and press the enter button
4	The note name corresponding to the input pitch is displayed, and the real value of the midi note number is recorded in the system
5	Click "the next question button" to move to the next question
6	Repeat a series of steps 1–5 10 times

5.2 Main Test

The results were divided into two categories: potential and non-potential absolute pitch. The average error is the average difference between the midi number of the correct pitch and that of the input pitch from the 10 questions of test C (c.f. Sect. 4.3). The improvement value refers to the difference between the average error before and after training. The improvement rate refers to the value obtained by dividing the average error before training by that after training.

Potential Absolute Pitch. Table 6 shows the results of the experiment of potential absolute pitch. Generally, neither the proposed nor the existing system showed significant improvement. The average improvement value was slightly negative (−0.09), and the average improvement rate was positive (+11%) in the proposed system. The average improvement value was slightly negative (−0.07), and the average improvement rate was almost 0 in the existing system (Fig. 3).

Table 6. Result of potential absolute pitch (PS: Pretest Score, BT: Before Training, AT: After Training, IV: Improvement Value, IR: Improvement Rate).

		Proposed system				Existing system			
		Average error		Improvement		Average error		Improvement	
ID	PS	BT	AT	IV	IR	BT	AT	IV	IR
2	10	0.29	0.16	0.13	+82%	0.23	0.21	0.02	+12%
8	10	0.34	0.49	−0.16	−32%	0.50	0.52	−0.02	−4%
9	10	0.41	0.39	0.02	+4%	0.33	0.32	0.01	+2%
10	10	2.97	3.30	−0.33	−10%	2.87	3.16	−0.30	−9%
Ave	10	1.00	1.09	−0.09	+11%	0.98	1.05	−0.07	+0%

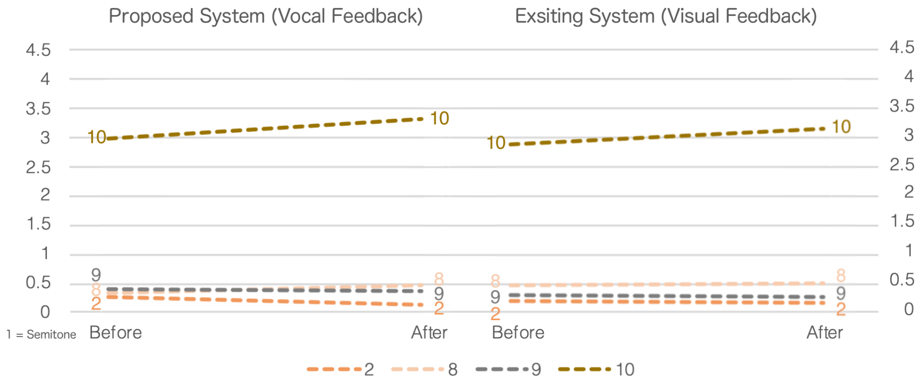

Fig. 3. Average error transition of potential absolute pitch.

Non-potential Absolute Pitch. Table 7 shows the result of non-potential absolute pitch. Generally, the proposed system showed some improvement, whereas the existing system did not. In the proposed system, both the average improvement value and the average improvement rate were positive (0.65 and +25%, respectively). In particular, P4 had an improvement value of 1.40, implying that the absolute pitch accuracy improved by more than a semitone. In contrast, in the existing system, both the average improvement value and the average improvement rate were negative (−0.49 and −12%, respectively). The existing system did not show any improvement, except for one person on P7 (Fig. 4).

Statistical Hypothesis Testing. We conducted a Wilcoxon signed-rank test at a significance level of 5% to compare the average error before and after training in the proposed/existing system. The null hypothesis stated that there would be no difference in the average error before and after training with the proposed/existing system, and the alternative hypothesis stated that there would

Table 7. Result of non-potential absolute pitch (PS: Pretest Score, BT: Before Training, AT: After Training, IV: Improvement Value, IR: Improvement Rate).

		Proposed system				Existing system			
		Average error		Improvement		Average error		Improvement	
ID	PS	BT	AT	IV	IR	BT	AT	IV	IR
1	6	1.91	1.88	0.02	+1%	1.66	2.20	−0.54	−25%
3	4	3.47	3.04	0.43	+14%	1.91	3.53	−1.63	−46%
4	4	4.15	2.76	1.40	+51%	2.74	3.76	−1.02	−27%
5	2	4.17	3.30	0.87	+26%	2.35	3.41	−1.07	−31%
6	5	3.11	2.47	0.64	+26%	2.47	2.45	0.02	+1%
7	5	2.28	1.73	0.56	+32%	3.58	2.27	1.31	+58%
Ave	4.33	3.18	2.53	0.65	+25%	2.45	2.94	−0.49	−12%

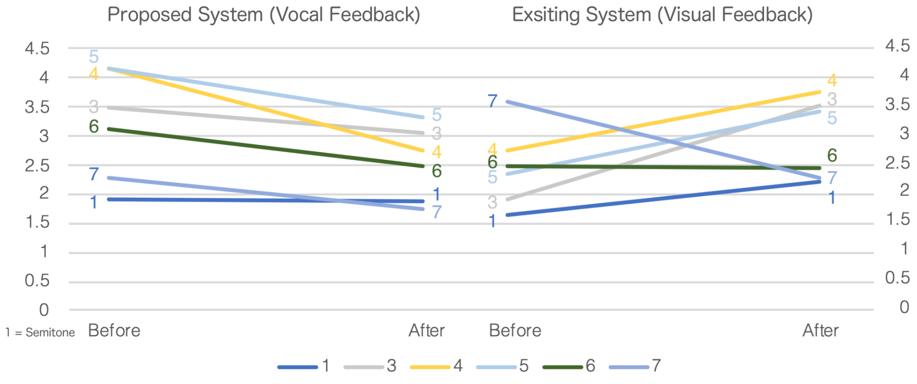

Fig. 4. Average error transition of non-potential absolute pitch.

be a difference in the average error before and after training with the proposed/existing system.

- Potential absolute pitch with the proposed system: We confirmed that there was no significant difference (N = 4, p-value = 0.625, Z = −0.730).
- Potential absolute pitch with the existing system: We confirmed that there was no significant difference (N=4, p-value = 0.875, Z = −0.365).
- Non-Potential absolute pitch with the proposed system: We confirmed a significant difference (N = 6, p-value = 0.0313, Z = 2.20).
- Non-Potential absolute pitch with the existing system: We confirmed that there was no significant difference (N = 6, p-value = 0.438, Z = −0.943).

In summary, it showed significance only when people with non-potential absolute pitch trained with the proposed system. No significance was shown when the non-potential absolute pitch trained with the existing system or when the potential absolute pitch trained with either the proposed system or the existing system.

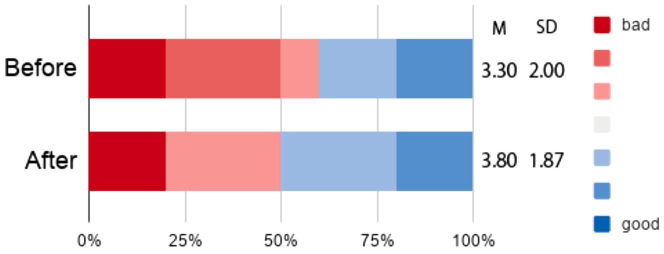

Fig. 5. Self-evaluation of the pitch of the participants

5.3 User Study

Self-evaluation of the Pitch. Figure 5 shows the results of the self-evaluation of the participants' sense of pitch before and after the experiment. After the training task, some participants responded that their sense of pitch improved slightly, but there was no significant difference. The participants were asked the reason for the evaluation after the experiment. Positive opinions included "I learned how to get to the scale (P3)" and "I felt that there was less deviation (P6)". However, some participants reviewed their self-evaluation of the sense of pitch by saying, "I noticed that there was no sense of pitch (P4, P10)."

Usability. After completing all the experimental tasks, we asked the questions about the system's usability as shown in Fig. 6 and 7. The following is a summary of the subjective feedback. More than 50% of the participants answered that they were satisfied with the usability. Although the difficulty level of the sound test and training task was not appropriate for P8, who had with musical experience, 70% of the participants in the experiment answered that it was appropriate. 90% said they were less tired overall, and the degree of fatigue was only 50% for the eyes and 30% for the ears. Additionally, we asked whether the existing system or the proposed system could be used for long-term sound training (e.g., 1 h/day for a month). 80% said that they would like to use the proposed system. When asked about improvements, they mentioned some related to the user interface and feedback system. One of the user interface improvements was to change the button layout. Regarding vocal feedback, it was recommended that the voice of the vocal feedback should be selectable. There was also a desire for multiple choices of feedback voice and for shortening the long word, "sharp".

User Interview. Fig. 8 displays the impressions of the users of the proposed system. 90% of the respondents displayed a lack of confidence regarding listening to music and singing with others. Most said that the proposed system did not affect their motivation for music activities: "music is fun even if you do not know the scale (P4)," and "music is naturally fun (P9)." However, more than 50% said that they became more aware of pitch (Third question in Fig. 8) and more than

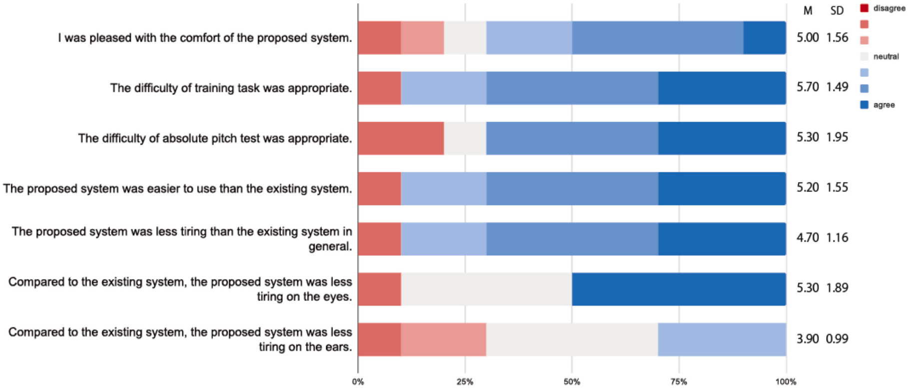

Fig. 6. Questionnaire about Usability

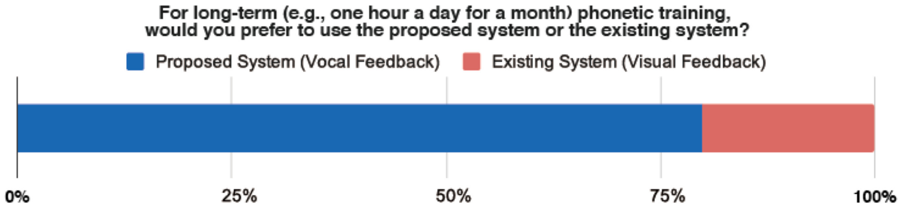

Fig. 7. Questionnaire about Usability2

80% said they thought the proposed system is a great way to improve their pitch (Fourth question in Fig. 8).

6 Discussion

For the "potential absolute pitch" participants, as discussed in Sect. 5.2, neither the proposed system nor the existing one showed significant training effect. One possible reason is that the error recorded in test C before training for potential absolute pitch participants was initially small. Thus, it was difficult to show improvement. However, as a result of the discussion in Sect. 5.2, in the case of non-potential absolute pitch participants, the average improvement rate was 25% higher than before training and 37% higher with the existing system. In addition, Wilcoxon signed-rank test comparing the average errors before and after the proposed system's training confirmed a significant difference (Sect. 5.2). Thus, the proposed system may show beneficial results regarding the absolute pitch training of non-potential absolute pitch participants.

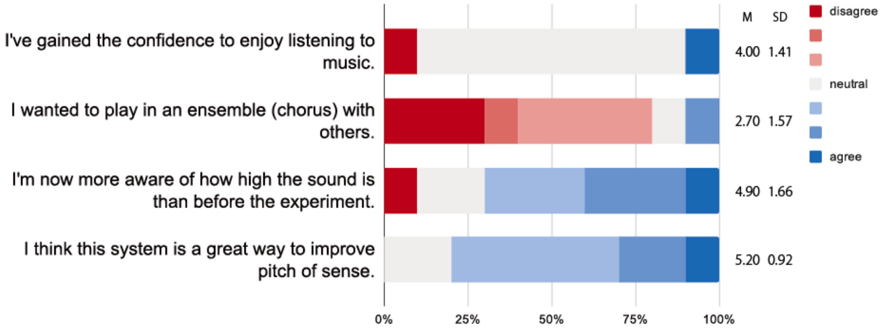

Fig. 8. User interview

Figure 9 shows the distribution of pitch improvement values in the proposed system between non-potential and potential absolute pitch. From this figure, it is clear that the non-potential absolute pitch had a higher improvement value than did the potential absolute pitch. This is because the potential absolute pitch score before training was already high. Thus, the score after training was unlikely to change. Therefore, as shown in Sect. 5.2, we investigated the improvement value by limiting it to the non-potential absolute pitch. As a result, there was a significant improvement in the absolute pitch. Thus, it is considered that the proposed system has certain advantages over the existing system.

7 Future Work

7.1 Relative Pitch Training

In this study, the experiment was designed to train absolute pitch. However, initially, it could also train relative pitch. When training the relative pitch, the idea of movable *Do* is applied. For example, with the voice that feeds back the pitch name of *Si*/B when *La*/A is used as the fundamental tone, the pitch remains *Si*/B, but the pitch name becomes *Re*/D. Because it is possible to test the relative pitch using this principle, future verification is needed.

7.2 Non-GUI Application

In this research experiment, a GUI-based slide bar was used as the system input to unify the experimental conditions. However, by limiting the system input system to auditory system only, ubiquitous sound training becomes possible, as opposed to the GUI-based system.

Voice Input. In the training task, by inputting the user's voice into the system input, it is possible to perform sound training by vocalization. Therefore, the users can train their pitch sense by simply wearing earphones with a microphone. Additionally, users can improve their ability to utter their voice at the height of the pitch they imagined. As a case study, users could walk around outside while training their pitch sense and vocalization.

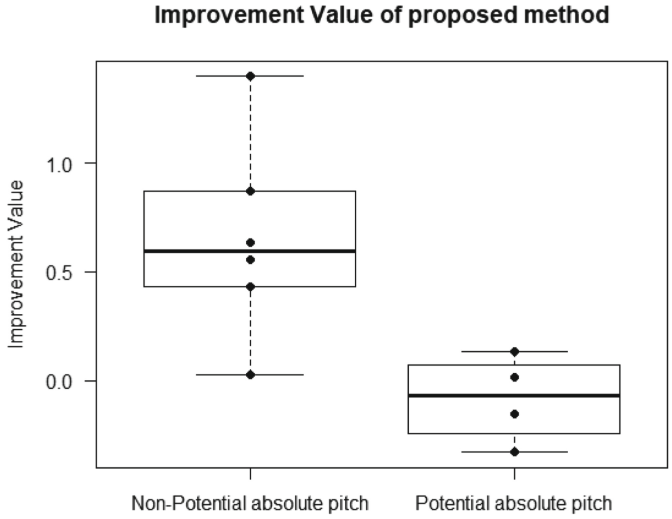

Fig. 9. Distribution of pitch improvement values in the proposed system

Gesture Input. By using a pressure sensor-type slide bar attached to the hand as a system input, the users can perform a training task by indicating pitch by gesturing the sliding of the finger. As a case study, users can employ gesturing when in crowds or when it would otherwise be difficult to sing.

8 Conclusion

This study proposes a novel auditory-centered training system using solmization to artificially acquire and train absolute pitch by providing vocal feedback to a musician's musical notes. Our contributions are as follows. First, we constructed a system that allows users to recognize note names aurally via solmization. Second, we compared our vocal feedback system with existing visual feedback systems and identified which type is more effective. In an experiment with 10 participants, our system's training improved six non-potential absolute pitch users' absolute pitch by approximately 25%. We also proposed GUI-free pitch training by vocalization or reflecting pitch with finger motion, which is beneficial in situations where it is difficult to hear or speak. Future verification is needed on relative pitch training and GUI-free application with vocal feedback.

References

1. Ashby, W.R.: An Introduction to Cybernetics. Chapman & Hall Ltd., Boca Raton (1961)
2. Bachem, A.: Absolute pitch. Physcol. Bull. **161**(1948), 1180–1185 (1948)
3. Callaghan, J., Thorpe, W., van Doorn, J.: The science of singing and seeing, pp. 15–18 (2004)

4. Christiansen, L.H., Frederiksen, N.Y., Jensen, B.S., Ranch, A., Skov, M.B., Thiruravichandran, N.: Don't look at me, i'm talking to you: investigating input and output modalities for in-vehicle systems. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011. LNCS, vol. 6947, pp. 675–691. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23771-3_49
5. Crutchfield, W.: There may be more to music than meets a typical ear. The New York Times (1990)
6. Duan, Z., Gupta, C., Percival, G., Grunberg, D., Wang, Y.: SECCIMA: singing and ear training for children with cochlear implants via a mobile application (2017)
7. Gaver, W.W.: Auditory icons: using sound in computer interfaces. Hum. Comput. Interact. **2**(2), 167–177 (1986). https://doi.org/10.1207/s15327051hci0202_3
8. Guardati, L., Casamassima, F., Farella, E., Benini, L.: Paper, pen and ink: an innovative system and software framework to assist writing rehabilitation. Presented at the (2015). https://doi.org/10.7873/DATE.2015.0453
9. Hoppe, D., Sadakata, M., Desain, P.: Development of real-time visual feedback assistance in singing training: a review. J. Comput. Assist. Learn. **22**(4), 308–316 (2006). https://doi.org/10.1111/j.1365-2729.2006.00178.x
10. Howard, D.M.: Technology for real-time visual feedback in singing lessons. Res. Stud. Music Educ. **24**(1), 40–57 (2005). https://doi.org/10.1177/1321103X050240010401
11. Howard, D.M., Welch, G.F.: Microcomputer-based singing ability assessment and development. Appl. Acoust. **27**(2), 89–102 (1989). https://doi.org/10.1016/0003-682X(89)90002-9
12. Itoh, K., Suwazono, S., Arao, H., Miyazaki, K., Nakada, T.: Electrophysiological correlates of absolute pitch and relative pitch. Cerebral cortex (New York, N.Y. : 1991) 15, 760–769 (2005). https://doi.org/10.1093/cercor/bhh177
13. Kin Wah Edward, L., Anderson, H., Hamzeen, H., Lui, S.: Implementation and evaluation of real-time interactive user interface design in self-learning singing pitch training apps (2014). https://doi.org/10.13140/RG.2.1.1757
14. Kuber, R., Tretter, M., Murphy, E.: Developing and evaluating a non-visual memory game. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011. LNCS, vol. 6947, pp. 541–553. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23771-3_41
15. Langfeld, H.S.: Psychological monographs. Psychological Review. Publication (1932)
16. Leong, S., Cheng, L.: Effects of real-time visual feedback on pre-service teachers' singing. J. Comput. Assist. Learn. **30**(3), 285–296 (2014). https://doi.org/10.1111/jcal.12046
17. Matsuda, M., Igarashi, H., Itoh, K.: Auditory t-complex reveals reduced neural activities in the right auditory cortex in musicians with absolute pitch. Front. Neurosci. **13**, 809 (2019)
18. Nakano, T., Goto, M., Hiraga, Y.: MiruSinger: a singing skill visualization interface using real-time feedback and music CD recordings as referential data. In: Ninth IEEE International Symposium on Multimedia Workshops (ISMW 2007), pp. 75–76 (2007). https://doi.org/10.1109/ISM.Workshops.2007.19
19. Nylander, S., Jacobsson, M., Tholander, J.: Runright - real-time visual and audio feedback on running (2014). https://doi.org/10.1145/2559206.2574806
20. Profita, J., Bidder, T.G.: Perfect pitch. Am. J. Med. Genet. **29**(4), 763–771 (1988). https://doi.org/10.1002/ajmg.1320290405

21. Richardson, P., Kim, Y.: Beyond fun and games: a framework for quantifying music skill developments from video game play. J. New Music Res. **40**(4), 277–291 (2011). https://doi.org/10.1080/09298215.2011.565350
22. Rossiter, D., Howard, D.: Albert: a real-time visual feedback computer tool for professional vocal development. J. Voice Official J. Voice Found. **10**(4), 321–336 (1996). http://europepmc.org/abstract/MED/8943135
23. Shiraishi, M., Ogasawara, K., Kitahara, T.: HamoKara: a system that enables amateur singers to practice backing vocals for karaoke. J. Inf. Process. **27**, 683–692 (2019). https://doi.org/10.2197/ipsjjip.27.683
24. Takeuchi, A.H., Hulse, S.H.: Absolute pitch. Psychol. Bull. **113**(2), 345–361 (1993). https://doi.org/10.1037/0033-2909.113.2.345
25. Welch, G., Himonides, E., Howard, D., Brereton, J.: VOXed: Technology as a meaningful teaching aid in the singing studio, pp. 166–167 (2004)
26. Welch, G., Howard, D., C, R.: The SINGAD (singing assessment and development) system: A classroom-based pilot study with poor pitch singers. Canad. J. Music Educ. Res. Edn. (1988)
27. Wiener, N.: Cybernetics: or Control and Communication in the Animal and the Machine, 2nd edn. MIT Press, Cambridge (1948)