

AN INTERACTIVE SINGING COACHING INTERFACE FOR VOICE THERAPY

**YI-HSIN CHEN
(SONIA)**

TS Voice Club
sonia.yh.chen@gmail.com

**HENG-CHIN YEH
(YERO)**

TS Voice Club
yero1209@gmail.com

**SHU-WEI TSAI
(SUWEI)**

TS Voice Club;
National Cheng Kung
University Hospital
tsaisuwei@gmail.com

ABSTRACT

We are presenting an interactive mobile-based singing coaching and rehabilitation system which is designed from the perspective of the voice therapy. We aim to use this system to help patients and students to remind the “correct” way to use their voice for rehabilitation and singing. This system is developed to help patients and students who cannot use their hearing memory and muscle memory to bring the “correct” voice and singing practice from the clinics and classrooms. By recording and visualizing the correct voice they learned from the clinics and classrooms, the users can revisit the speech-language pathologists or the singing coaches certified “correct” voice from both visual and audio cues so that they can practice with those signs remotely. In the meantime, the gamifying module is also incorporated to increase the practice adherence rate and the communication for the rehabilitation and singing practice. By using this system, patients and students not only take the practice protocols but also the memory of the practice nuances back from the clinics and the classroom.

1. INTRODUCTION

Treatment dropout is a consistent problem among behavior change therapies. According to previous study, the dropout rate of voice therapy is about 65% [1]. Evidence shows that people with voice disorders benefit from voice therapy. [2] However, it is well accepted that behavior-changing therapies, such as voice therapy, are only effective as a patient adhere to the therapy process. [3]

By interviewing the therapy participants (including otolaryngologists, the speech-language pathologists (SLPs), the vocal coaches, and the patients) and reviewing the literatures, we found that there was lack of an efficient way for patients to take out the “correct” practices to do rehabilitation they learned from the SLPs and/or their voice coaches and it was one of the main root causes that patients drop out from the voice therapy and singing prac-

tice.[4] Specifically, it is not easy for most of patients and students to repeat the “correct” way they learned from the clinics and the classroom by their hearing and muscles memory. It made patients to lack of confidence at their practice. It is known that giving proper feedback could increase the rehabilitation rate and the adherence to therapy. [5] Although some IOS-based app/software systems have been developed for SLPs to evaluate the patient’s voice and some of those systems also provided the feature to communicate SLPs and the patients, those interfaces were too complicated for the patients to read and operate. Patients need a more intuitive way to revisit and recall their practices and rehabilitation procedure.

In this LBD, we use audio transcription algorithms to develop an interactive mobile-based singing coaching and rehabilitation system that help patients and students to visually revisit the “correct” voice they learned from the clinics and the classroom.

2. DESIGN CONCEPT

From series of interview with our target users (the patients in voice therapy and the students in vocal class), we figured out that they need more signs and bio-feedbacks to demonstrate how they are using their voice. By enabling the target users to record their “correct” voices certified by the SLPs and/or the voice coach through the mobile device, the patients can take the correct voice cue back to home and practice with the voice in a timely manner. In addition, the system also feedback the visual cues by analyzing the voice real time. Users can see the pattern of their reference voice and their practice voice at the same time on this interface. They can further adjust the way they use their voice to match the pattern reference voice. The more percentage the user can match the pattern, the higher score the user get.

3. FEATURES

This section describes the different feature currently included in this demo. This demo was made by the game engine platform “Unity”. It will be applied to Android Mobile System. The interface is shown as Figure 1.

3.1 I/O



Input and output functionality is implemented in this demo. The audio files, or the “correct” voice files, can be recorded and playback to/from the mobile device, specifically for Android Mobile Device for this demo, by simply pressing “Record Ref” button and “Play Ref”. The user can press “Test” button to start their voice therapy and singing practice

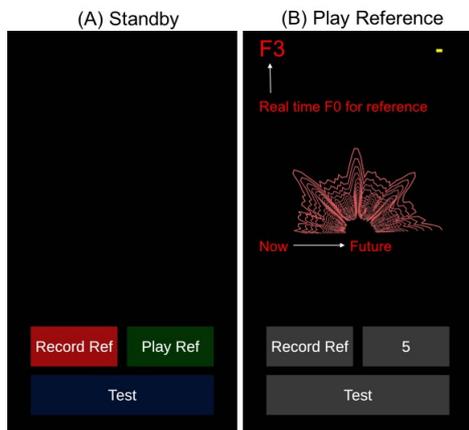


Figure 1. The User interface layout while (A) Standby and (B) Play Reference.

3.2 Time-Frequency Transform

The so-called “correct” or “incorrect” voice indicates different “timber”, which can be translated into the designated setting of the frequency distribution. Therefore, we visualize those audio signal to a new pattern in accordance to their spectrum. The reference and test audio signals were converted into real-time pattern based on Short-Time Fourier Transformation (STFT) with a Hanning Window. The sound frequency between 20Hz to 4000Hz are presented linearly on a radius axis. The peak will be shown while the indicated frequency dB is over certain level. When playback the referenced audio file, the visualized reference audio pattern can be previewed from the center of the semi-circle. The visualized pattern approaches to the outer region of the circle in a timely manner, and the most outer region of the circle is showing the current audio pattern.

3.3 Analysis and Gamifying Module

In the “Test” mode, the user can see the visualized reference audio pattern in the color of red and their real-time input audio pattern in the color of blue. The note for both reference audio and real time test audio will be analyzed. In this interface, the strained voice and the hoarse voice can be distinguished, as shown in Figure 2. The gamifying module are composed with a dynamic voice matching task. The test audio files will be stored for the next visit of clinic and/or class. SPLs and voice coaches may further evaluate their practice outcome through this record.

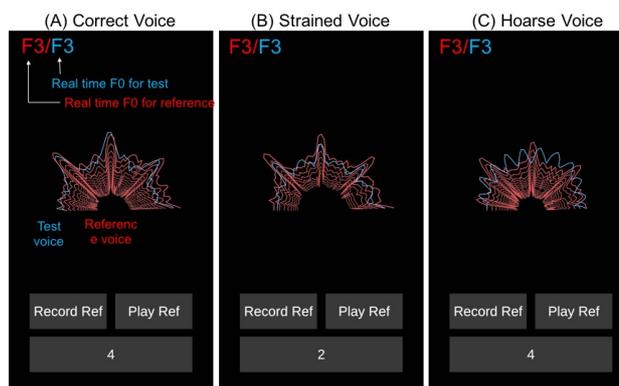


Figure 2. The correct voice, the strained voice, and the hoarse voice can be distinguished.

4. CONCLUSION

We’ve developed an interactive interface for voice therapy patients to practice their “correct” voice. This is the first gamified interactive interface that designed for voice therapy patients. The score system will be implemented for the next version.

5. ACKNOWLEDGEMENT

We acknowledge the support from Women In MIR (Wi-MIR) Grant to this project in the Late Break Demo (LDB) section of ISMIR, 2017, Suzhou, China.

6. REFERENCE

- [1] E. Hapner, *et al.*: “A Study of Voice Therapy Dropout,” *Journal of Voice*, Vol. 23, No. 3, pp. 337–340, 2009.
- [2] L Ramig, and K Verdolini: “Treatment efficacy: voice disorders,” *Journal of Speech, Language, and Hearing Research*, Vol. 41, pp. S101 – S 106
- [3] R Speyer, *et al.*: “Documentation of progress in voice therapy: perceptual, acoustic, and laryngostroboscopic findings pre-therapy and post therapy,” *Journal of Voice*, Vol. 18, pp. 325–340, 2004.
- [4] E van Leer, NP Connor: “Patient perceptions of voice therapy adherence.” *Journal of Voice*, Vol. 24, pp. 458–469, 2009
- [5] E van Leer, NP Connor: “Use of portable digital media players increases patient motivation and practice in voice therapy.” *Journal of Voice*, Vol. 26, pp. 447–453, 2012.