



Web Application System for Pronunciation Practice by Children with Disabilities and to Support Cooperation of Teachers and Medical Workers

Ikuyo Masuda-Katsuse

Faculty of Humanity-oriented Science and Engineering, Kindai University, Fukuoka, Japan

katsuse@fuk.kindai.ac.jp

Abstract

We developed a Web application system for children with pronunciation difficulties to practice pronunciation. The system users are assigned one of the following authorizations: student, teacher, medical worker, or speech evaluator. The teachers, medical workers, and speech evaluators are grouped by the student to whom they are linked and can access exercise records and student speech sounds by the Internet. The teachers can individually tailor practice words to each child's pronunciation needs. The medical workers and speech evaluators confirm the accuracy of the student's pronunciation and can share information with teachers. Thus, our system will encourage students to practice their pronunciation and promote the cooperation of teachers and medical workers for more effective instruction.

Index Terms: speech training, speech disorders, medical-educational-engineering collaboration

1. Introduction

In Japan, special education classes are provided in elementary schools to support children whose language development lags behind their peers due to hearing or speaking disorders or developmental disabilities. Sometimes the teachers of such special education classes have not been formally trained in speech education techniques. In some cases, they learned the techniques on their own or participated in training sessions or cooperate with such medical workers as speech-language-hearing therapists (STs).

In our previous research, we developed a stand-alone system for children with pronunciation difficulties to practice their pronunciation [1]. Our system allows teachers to individually set practice words based on each child's pronunciation needs. In addition to registered words, inspection words [2] are presented so that the speech sounds are used for subsequent articulation tests.

Nine elementary school students practiced their pronunciation using our previous system for three or four months. Because several sets of inspection words were recorded

in the system, we observed the improvement process. Table 1 shows the improvement process of one participant's /s/ sound.

The table suggest that since the pronunciation changed every month, frequent speech evaluations are needed so that the articulation test results contribute to training. Our previous stand-alone system was inconvenient for such frequent speech evaluations, because the teachers had to copy the data from the system into media and provide them to STs to evaluate student speech sounds.

In this research, we introduce a new Web application system in which data are managed in a database, and teacher and medical workers can access the data through the Internet at any time.

2. System outline

2.1 Main technical components

- Audio processing on client side: WebAudio API
- Client-server communication: WebSocket and WebRTC
- Authorization and data management: MySQL Database
- Audio processing in servlet: automatic speech recognition (SPHINX4 [3]), phonemic discrimination [1], and voice activity detection [4]

2.2 User authorizations and interrelation among users of system usages

IDs, passwords, and user authorizations are allocated to all users. User authorizations include students, teachers, medical workers, and speech evaluators. We grouped the users by the student to whom they are related. The users with teacher, medical worker, and speech evaluator authorizations can access the permitted data of the students who belong to the group same as they do.

Figure 1 shows a typical interrelation of the system users. Here a medical worker also works as a speech evaluator.

■ Teachers (or STs)

The teachers usually instruct their students to work on their pronunciation and give them homework to practice pronouncing words repeatedly.

Table 1. Process of pronunciation improvement

Correct pronunciation	Changes of pronunciations of a student			
	Oct-13	Dec-13	Jan-14	Feb-14
/sakana/	/teakana/	/teakana/	/eakana/	/sakana/ (distorted s)
/sora/	/teora/	/teora/	/eora/	/sora/ (distorted s)
/semi/	/temi/	/temi/	/teemi/	/semi/ (distorted s)
/suika/		/teuika/	/euika/	/euika/
/teisai/		/teiteai/	/teiteai/	/teisai/ (distorted s)

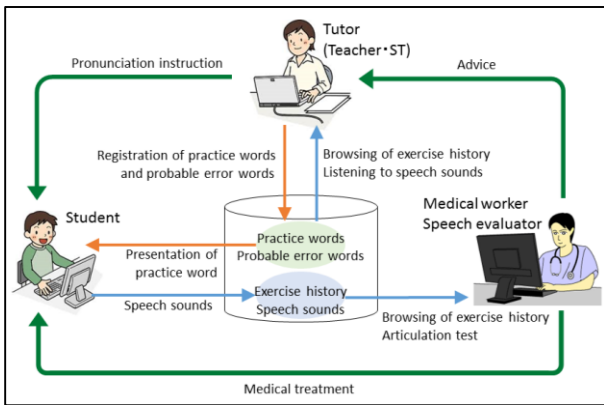


Figure 1: Typical interrelation of system users

To help students with their homework, teachers can use the system by registering practice words selected from the dictionary and registering probable error pronunciations. They can also register new words in the dictionary and check the exercise records of their students and accordingly change the practice words.

■ Students

After students login to the website, a practice screen is presented that contains a volume bar, a level meter, a block for the presentation of a practice word, three buttons (next word, again, quit), a picture of the speech evaluation result, and a graph of the practice amount (Fig. 2).

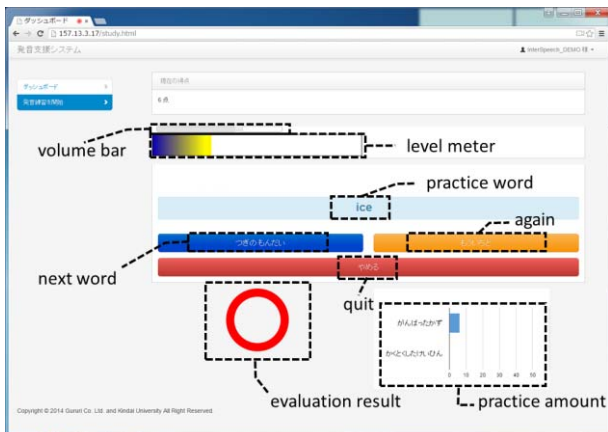


Figure 2: Web screen for pronunciation practice

In a block that presents the practice word, practice and inspection words are randomly presented. The students are encouraged to say them aloud. The speech sounds are sent to the server and evaluated, and the evaluation results are presented on the screen. When a probable error pronunciation is registered for the presented word, the possible scores for the correct and probable error phonemes are represented in an animation graph. Practice logs and speech sounds are recorded in the database.

Because a teacher observed that some students could notice their own mispronunciations by listening to their speech sounds, we will soon add to our system a new ability to listen to recorded speech sounds.

■ Medical workers (doctors or STs)

They can listen to recorded speech to perform articulation tests. They can also browse their students' exercise records at any time and advise teachers.

■ Speech evaluators (teachers or STs)

This authorization investigates and records each student's specific pronunciation errors in detail. To describe student mispronunciations, speech evaluators listen to the speech sounds of the words for which the probable error pronunciation was registered. Fig. 3 shows part of the screen for playing speech sounds and evaluating them. The records can be downloaded as a csv file.

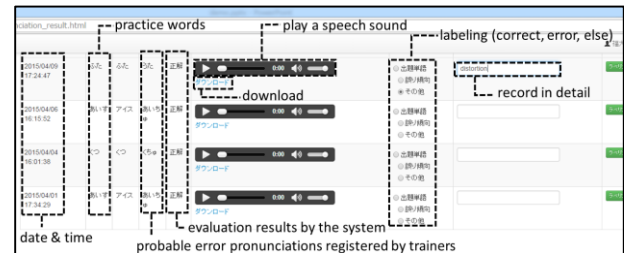


Figure 3: Web screen for speech evaluation

3. Demonstration and practical field test

Demonstration videos are linked to the following website: <http://www.joho.fuk.kindai.ac.jp/~katsuse>

Access the following URL of our application's website to try our system's student mode:

<http://mm.joho.sv.fuk.kindai.ac.jp/login.html>

user ID: InterSpeech2015, password: InterSpeech2015.

The application works with Google Chrome browser. Audio input should be set as 48k Hz sampling and 16 bit.

We are performing practical field tests with informed consent at hospitals and special support education classes in elementary schools. We will report the results in future works.

4. Conclusion

We developed a Web application system for students who suffer from pronunciation difficulties to practice their pronunciation. Our system not only encourages students to practice their pronunciation but also promotes the cooperation of teachers and medical workers for more effective instruction.

5. Acknowledgements

Part of this work was supported by JSPS KAKENHI Grant Number 25350692 and the Project of Humanity-Oriented Science and Engineering by Kindai University.

6. References

- [1] I. Masuda-Katsuse, "Pronunciation Practice Support System for Children who Have Difficulty Correctly Pronouncing Words," in INTERSPEECH 2014 Proceedings, 2014, pp. 2144-2145.
- [2] Articulation Clinics Workshop [Ed], Japanese Articulation Test, Second Edition, Chiba Test Center, 2010. (in Japanese)
- [3] Sphinx-4, A Speech Recognizer Written Entirely in the Java TM-programming Language, Online: <http://cmusphinx.sourceforge.net/sphinx4>, accessed on March 18, 2013.
- [4] K. Ishizuka, T. Nakatani, M. Fujimoto, and N. Miyazaki, "Noise robust voice activity detection based on periodic to aperiodic component ratio," Speech Communication, 52, 41-60, 2010.