

Proposal of Adaptive CALL system for Japanese Learners of Mandarin*

©Minyi Ma, Hiroko Hirano, Keikichi Hirose, Nobuaki Minematsu

(University of Tokyo)

1 Introduction

As China continues to be one of the world's fastest growing economies, many foreigners are eager to learn Mandarin Chinese to enhance their international business. Most people believe Chinese pronunciation is difficult. There are 39 vowels and 21 consonants in Chinese. The number of vowels and consonants implies the difficulty of Chinese pronunciation. There are four patterns of tones in a Chinese (Mandarin) syllable, and, therefore, F0 movements are much more complicated than non-tonal languages like Japanese and English. There is an increased urgency to find ways to help foreign speakers to learn Chinese pronunciation.

Our research objective is to develop a system specifically for helping Japanese people to improve their Chinese pronunciation. In this paper, we first present speech analysis results of Japanese students' utterances, then propose an adaptive CALL system for effective Mandarin pronunciation education.

2 Speech Analysis

2.1 Speech Data

We collected Mandarin speech data from 12 Japanese speakers (7 beginners and 5 advanced learners). Participants read a script containing 24 words and 28 sentences. All the Chinese are labeled with Pinyin and Japanese translations.

2.2 Experimental Results

A native Chinese trained in phonetic labeling labeled the utterances at the phone level using Praat^[1]. Several subjective listening tests were held and the features of Japanese students' utterances we found are shown as follows.

1. Phone

We cut students' utterances into syllables and asked natives to listen to these syllables and write down what did they hear. From the test result, we

concluded a list of easy-to-confuse phones of Japanese students shown in Table 1.

Table 1 Easy to confuse phones of students

Vowels	an / ang
	ou / ao
	lao / ou
	ua / uan / uang
	wu / ' ǔ '
Consonants	sh / ' し '
	f / h
	b / p
	t / d
	l / r

In addition, students perform vowels unstably: the same vowel uttered differently when pronounced in different words. Furthermore, lip movements for Japanese utterances are generally smaller than those for Chinese. This makes, for instance, pronunciation of vowel "a" in Japanese and Chinese is totally different, but Japanese students tend to use Japanese version of "a".

2. Tone

Our reading script was prepared so that all tone combinations were included. Through a listening test focused on tone pronunciation, the combination of [tone3+tone2] was found especially difficult for Japanese students. In addition, they failed to pronounce tone3+tone3 sequence as tone2 and tone3. Utterances with this type of error are hard to be understood by native.

We also compared pronunciations of the same word in word level and in sentence level. It was obvious that the performance in word level was better than that in sentence level.

3. Duration

To check how students' utterances are degraded by duration errors, we randomly selected 14 sentences and modified their phone and pause durations to those of native utterances. After modification, native speakers scored

* 日本人学習者を対象とした適応型中国語発音教育システムの提案, 馬敏懿, 平野宏子, 広瀬啓吉, 峯松信明 (東京大学)

naturalness of the original utterances and duration-modified utterances. The result showed that after modification, the average score rose obviously. This result indicates that telling students their mispronunciation on duration is surely necessary for the CALL system of Mandarin.

2.3 Experimental Summary

From our speech analysis, we found that the problems of beginners and advanced learners are different. Beginners are usually poor at phones and tones, which make their utterances hard to be understood. On the other hand, advanced learners may manage to pronounce phones well, but when speak sentences, there are problems in tones and durations that make their utterances sounds unnatural.

3 System Proposal

From the analysis results, we propose an adaptive CALL system that can judge students' pronunciation level by speech recognition, suggest practice emphases and provide corresponding training. Also, the system can provide corrective audio feedback in students' own voice using speech modification^[2].

The system structure is shown in Figure 1. Before starting pronunciation practice, system will ask a learner to record several sentences. System will recognize learner's utterance (mainly phone and tone recognition), compare the recognition result with the correct script, and do level-judgment to suggest learner which course to enter: beginner course or advanced course.

Beginner course puts emphases on phone training and tone training on word level. Its aim is to help learner pronounce clearly and correctly to make their pronunciation easy to be understood. Advanced course puts emphases on tone training and duration training on sentences level. Its aim is to help learner to pronounce naturally and thus gradually reach a native level.

The system will give learner information about what kinds of errors they make, and how to correct them. Also it modifies learner's speech. The system will use phone insertion, reduction and replacement technologies to modify phones;

use TD-PSOLA to modify pitch and duration. The synthesized speech keeps learner's own voice quality, and is played back to the learner as corrective audio feedback. According to system's advice, the learner can practice several times and record again. Then system will give feedback again. Through the exchange between learner and system, effective training and improvement of learner's pronunciation is expected.

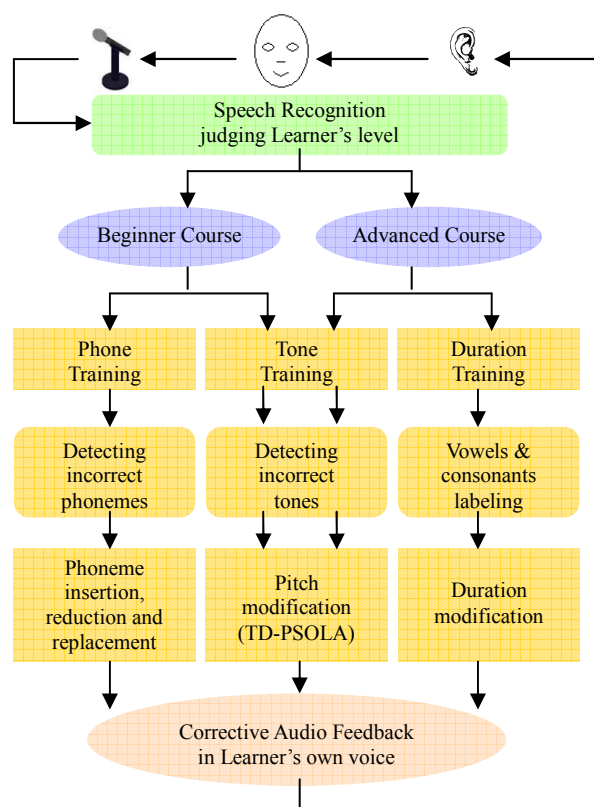


Fig. 1 System Structure

4 Conclusion and Future Work

Through analysis of Japanese learners' utterances, we proposed a CALL system for Chinese. It first finds out learners' level, and then applies different training method depending on the level.

Our future work includes CALL system development and system evaluation.

Reference

- [1] Boersma, P. and Weenink, D. "Praat", <http://www.praat.org/>
- [2] Frederic Gendrin, Keikichi Hirose and Nobuaki Minematsu, "Corrective feedback for accent pattern CALL systems using speech modification", IEICE Technical Report, SP2002-161, pp.1-6 (2003-1).