

# A Mandarin L2 Learning APP with Mispronunciation Detection and Feedback

Yanlu Xie<sup>1</sup>, Xiaoli Feng<sup>1</sup>, Boxue Li<sup>2</sup>, Jinsong Zhang<sup>1</sup>, Yujia Jin<sup>1</sup>

<sup>1</sup>Advanced Innovation Center for Language Resource and Intelligence, Beijing Language and Culture University, Beijing, China;

<sup>2</sup>Yunfan Hailiang (Beijing) Technology co., LTD, Beijing, China

xieyanlu@blcu.edu.cn, fengxiaoli314@163.com, 1240146049@qq.com, jinsong.zhang@blcu.edu.cn, blcu16wenyijyj@163.com

## Abstract

In this paper, an APP with Mispronunciation Detection and Feedback for Mandarin L2 Learners is shown. The APP could detect the mispronunciation in the words and highlight it with red at the phone level. Also, the score will be shown to evaluate the overall pronunciation. When touching the highlight, the pronunciation of the learner's and the standard's is played. Then the flash animation that describes the movement of the tongue, mouth, and other articulators will be shown to the learner. The learner could repeat the process to improve and excise the pronunciation. The App called 'SAIT 汉语' can be downloaded at App Store.

**Index Terms:** computer-assisted pronunciation training, mispronunciation detection, mispronunciation feedback, Mandarin learning

## 1. Introduction

Pronunciation has always been a challenge in the learning and teaching of Mandarin Chinese as a second language. L2 learners' pronunciation can be improved when providing sufficient meaningful practice opportunities and quick feedback from the instructors[1]. Adult language learners can still improve their perception and production with enough pronunciation training[2]. However, the lack of qualified teachers prevents students from practicing frequently.

In the field of computer-assisted pronunciation teaching (CAPT), it is helpful to provide the mispronunciation position and feedback for L2 learners, just like human teachers do. To improve the pronunciation evaluation, Witt proposed the

Goodness of Pronunciation (GOP) method for phone level scoring[3]. Furthermore, researchers try to detect and diagnosed mispronunciation. Strik studied how to discriminate against the frequently pronounced incorrectly sounds of Dutch[4]. Yinming and Yanlu detected frequent errors in Mandarin Chinese for L2 with DNN and landmark methods[5][6].

With the above technology, a Mandarin pronunciation teaching APP is developed by Beijing Language and Culture University. L2 learners could get the overall score, mistake tendency of their pronunciation with the app. L2 learners can get audio and video feedback for the mispronunciation. A preliminary study shows that after a six-week teaching experiment using the APP at Beijing Language and Culture University, L2 learners can remedy 83.2% of their pronunciation errors. [7]

In the paper, the features and framework of the App will be present in section 2. The performance of mispronunciation detection is described in section 3. Section 4 is the conclusion.

## 2. Features of the App

The architecture of the proposed App is shown in Fig.1. There are different functions for L2 students and teachers in the APP.

### 2.1. Mispronunciation detection

When students pronounce to the APP, it detects the mispronunciation in the initials, vowels, and tones of each syllable and distinguishes the pronunciation problems in red at the phone level. Also, the score evaluation of the overall pronunciation is shown in the APP.

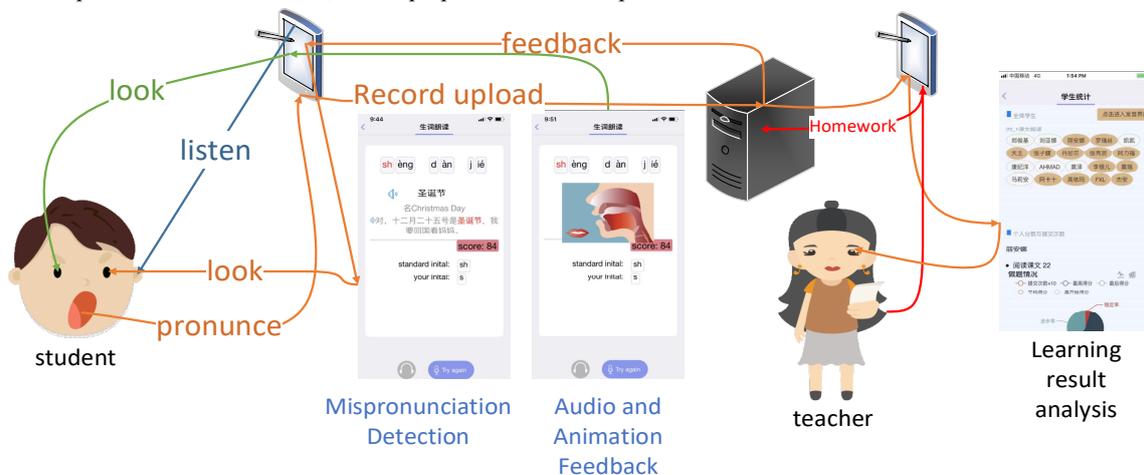


Fig 1: the architecture of the App system

## 2.2. Feedback

Students can check the pronunciation feedback for individual pronunciation. As shown in Figure 1, the student's "sh" of the first syllable has a pronunciation error. According to the feedback block in Figure 1, the standard pronunciation is "s" instead of "sh". Click on the standard initial and your initial, the student can look at the animation of the pronunciation action and listen to the correct sound of the pronunciation. The students can compare correct and wrong pronunciation through repeated listening, visual comparison, and pronunciation imitation, and then learn the correct pronunciation methods to improve the pronunciation quality.

## 2.3. Homework

After applying to join the class, students can use the specific textbook to practice pronunciation. The teacher can assign the homework in the textbook. After students submit the homework, the teacher can check the status of students' completion of homework and check the students' errors. The system will also provide teachers the learning result analysis.

## 3. Performance of Mispronunciation Detection

In order to evaluate the App's performance of detecting the mispronunciation of L2 speech, some experiments were carried on BLCU inter-Chinese speech corpus [8]. 7 speakers from the corpus were selected as the test set. Each speaker has 301 sentences. The total number of phonemes is 26431. The test set has been annotated at the phone level by six annotators who are proficient in phonetics.

Mandarin TDNN chain models developed by CVTE trained on commercial data were used as the baseline model[9]. To improve the robustness of the system, 100 hours of speech data recorded under different environments were used to adapt to the original model. The detecting performance is shown in Fig. 2-3. There are 3 evaluation indicators:

- False Acceptance Rate (FAR): the percentage of mispronunciation phones that are accepted as correct.
- False Rejection Rate (FRR): the percentage of correctly pronounced phones that are rejected as mispronunciation.
- Diagnostic Accuracy (DA): the percentage of correctly detected.

With different thresholds, we can get different FAR, FRR, and DA. It shows that for all thresholds, the performance is improved with more training data. FRR improves more. L2 learners will feel better if FRR is lower. The adaptation system with 79.0%DA is applied to the demonstrated App.

## 4. Conclusions

In this paper, a novel App for Mandarin L2 learners is shown. In the future, we will improve the performance of mispronunciation detection and the interactivity of animation of the App. And we will test it with more learners.

## 5. Acknowledgments

This work is supported by National Social Science Foundation of China (18BYY124), Advanced Innovation Center for Language Resource and Intelligence (KYR17005), Discipline Team Support Program of Beijing Language and Culture

University (GF201906), Education and teaching reform project of Beijing Language and Culture University (XJGZ201904) and Wutong Innovation Platform of Beijing Language and Culture University (19PT04). The corresponding author of the paper is Yanlu Xie.

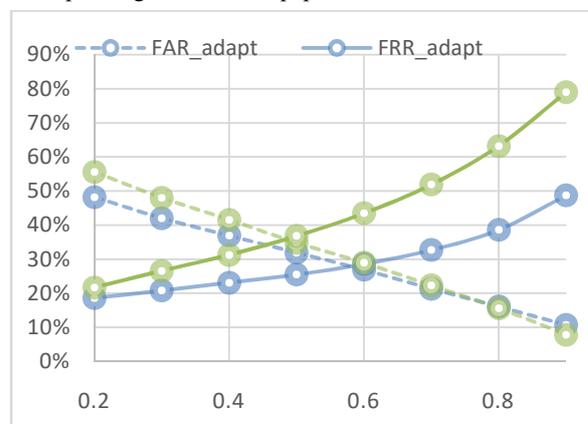


Fig2. FAR and FRR of the Mispronunciation Detection System

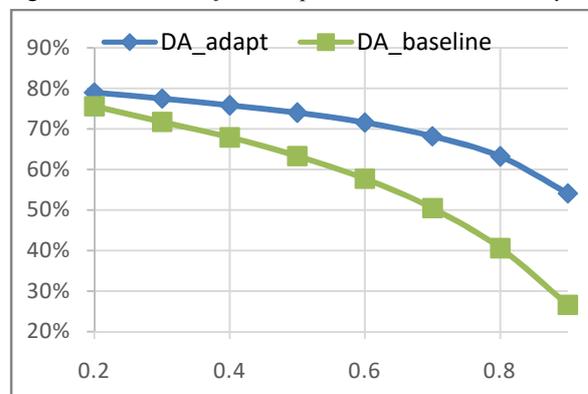


Fig3. DA of the Mispronunciation Detection System

## 6. References

- [1] Krashen, S. The input hypothesis: Issues and implications. London: Longman 1985
- [2] Mei, L. Cross-language speech perception in adults. *Journal of South China Normal University (Social Science Edition)*, 4, 164-169.2014.
- [3] S. M. Witt, and S. J. Young. "Phone-level pronunciation scoring and assessment for interactive language learning." *Speech Communication*, vol. 30, no. 2, pp. 95-108, 2000.
- [4] H. Strik, K. Tru Wong, F. De Wet, et al. "Comparing different approaches for automatic pronunciation error detection." *Speech communication*, 2009, 51(10): 845-852.
- [5] Gao Y, Xie Y, Cao W, et al. A Study on Robust Detection of Pronunciation Erroneous Tendency Based on Deep Neural Network *Interspeech: Conference of the International Speech Communication Association*. 2015.
- [6] Xie Y, Hasegawa-Johnson M, Qu L, et al. Landmark of Mandarin nasal codas and its application in pronunciation error detection *IEEE International Conference on Acoustics, IEEE*, 2016.
- [7] Wei, Wei, Zhang, Jinsong An Intelligent Chinese Pronunciation Teaching App and the Preliminary Result of a Teaching Experiment *Journal of Technology and Chinese Language Teaching* Volume 9 Number 2, December 2018 pp. 83-97
- [8] W. Cao, et al. "Developing a Chinese L2 speech database of Japanese learners with narrow-phonetic labels for computer assisted pronunciation training", in *INTERSPEECH*, 2010.
- [9] <http://kaldi-asr.org/models/m2>