

Design of an Interactive GUI for Pronunciation Evaluation and Training

Naoya Horiguchi and Ian Wilson

University of Aizu, Aizu-wakamatsu City, Fukushima-ken, 965-8580, Japan
wilson@u-aizu.ac.jp

Abstract

Although language learners often desire to improve their pronunciation of a foreign language, the software to help them do so is limited in scope. Most commercial software for pronunciation evaluation and training focuses on the acoustic signal in the evaluation and training of a learner. However, few systems, if any, give visual feedback of the learner's articulators (lips, tongue, jaw). In this paper, we describe the ongoing development of a GUI that is programmed in Objective-C for Mac OS X. Our software uses QtKit framework for video recording and playing, and some open source libraries for audio recording, audio playing, and pitch detection. The GUI incorporates and links together many kinds of phonetic data for the pronunciation learner - for example, real-time frontal video of the learner, recorded frontal and side videos of a native speaker's face during pronunciation, an ultrasound movie of the tongue moving in the mouth, and MRI images of the native speaker's tongue during the production of all the sounds in the training text.

Keywords: Interactive GUI, Pronunciation evaluation/training, Articulatory feedback, Ultrasound, MRI

1. Introduction

The pronunciation ability of second language (L2) learners is one of the most noticeable and influential factors when a native listener makes a snap judgement of the learner's proficiency. Many L2 learners want to improve their pronunciation, but individual feedback from a teacher is often impossible due to time constraints or class sizes. Thus, many L2 learners turn to software to help them with their pronunciation. Unfortunately, most commercial software for pronunciation evaluation and training focuses on the acoustic signal in the evaluation and training of a learner. The acoustics of the learner's speech signal are evaluated and

displayed to the learner, but the learner is left to interpret the link between acoustics and articulation (i.e., what changes he/she should make to his/her tongue, jaw, and lips to produce required changes in the acoustic output). Few systems, if any, give visual feedback of the learner's articulators (lips, tongue, jaw), and this is counterintuitive because it is easier for a learner to interpret articulatory feedback (e.g., the movement of the tongue) than acoustic feedback (e.g., the formant values in the acoustic signal).

Because of the complete lack of this type of pronunciation evaluation software, we decided to create a GUI that incorporates visual and audio information, both native speaker model data and L2 learner feedback. At the University of Aizu, construction was recently completed on two new high-tech classrooms that contain 48 iMac computers each. Each iMac contains a built-in web camera that enables real-time recording and display of the L2 learner as he/she speaks. For this reason, we decided to develop our pronunciation evaluation GUI using Objective-C in the Mac environment.

The rest of this paper proceeds as follows. Section 2 describes some typical pronunciation evaluation software and outlines the problems with these programs. Section 3 lists the phonetic data provided by the native speaker model (the second author) and explains how this data was collected and its use in the GUI. In Section 4, we elucidate the motivation behind the GUI design and give details about its functions. In Section 5, we describe the use of Praat (open source acoustic analysis software) and a rudimentary speech recognition script (written by the first author) that finds and labels syllables for analysis within the speech signal. Finally, Section 6 presents conclusions and future work.

2. Existing Pronunciation Evaluation Software

One type of commercial pronunciation software available is **AmiVoice CALL Lite** [1]