CPSC 490 Proposal: Improved Second Language Acquisition with Phoneme Biofeedback

Acshi Haggenmiller

Faculty Adviser Steven Zucker

January 26, 2016

1   Introduction

Second language acquisition is known for being difficult, especially with languages that differ significantly from the learner's native language. For an English speaker, romantic languages like Spanish and French are considered easier to learn because they both use Latin alphabets and relatively similar sounds. Languages like Mandarin are considered far harder to learn because they are character based, use tone to distinguish meaning, and contain different sounds. Romanization systems alleviate only the character problem. The conventional wisdom of language learning states that if you don't begin learning a language as a child, you have passed your cognitive prime and can never learn it as well as a native speaker.

Although fluency and comprehension may improve arbitrarily given enough time, it can be very hard for an advanced learner to make their speech indistinguishable from a native speaker's. Many different languages have similar phonemes. Although they may both have an "a" vowel, close analysis of the Fourier transformations will reveal different formant characteristics. For vowels that are similar enough, the learner may not be able to hear these differences because they cognitively "correct" unknown phonemes to the closest matches. A

native speaker would be able to indicate that their pronunciation is not quite correct, but would have trouble describing what exactly is wrong. Tone in tonal languages and intonation present a similar problem. It is hard for the learner to know when their usage is natural and awkward for a native speaker to explain. As these language features are inherently auditory, they are difficult to communicate precisely in words.

2   Problem Statement

I propose a computer-based biofeedback system which would both understand the characteristics of correct pronunciation in a given language and be able to evaluate the pronunciation of the language learner. Through this feedback system, the learner would be able to "see" the difference between correct and incorrect pronunciations that they cannot hear and be able learn this difference. As vowels are significantly simpler and fewer in number than consonants in most languages, I would mainly focus on correcting the pronunciation of vowel sounds. I will start by using the Fast Fourier Transform (FFT) to characterize the pronunciation of my own vowels and attempt to confirm published information about the vowel formants in American English. I will then develop a method for determining the formants of speech in real-time. I will then make a user interface that visualizes both the correct targeted pronunciation as well as the current pronunciation of the user. Towards the end of the project, I will test the efficacy of the biofeedback system by conducting a small-size preliminary study. If time permits, I will also analyze and visualize tone and intonation. The program will be made general for use with different languages. To enhance the educative value of the project, especially to future students, I will write blog posts to detail the progress of the project on an ongoing basis.

Preliminary work has already revealed the complexity of speech analysis to me. Directly taking the FFT of a vowel waveform sample results in a complicated spectrum mainly showing peaks at multiples of the fundamental frequency (the harmonics) of the voice. The

formants can be viewed as active filters of certain bandwidth, frequency, and amplitude that modify the source vocal waveform. Determining the formant characteristics then is not as simple as finding the peaks of the vowel FFT, but involves an implicit comparison of the vowel waveform to the unfiltered waveform produced by the larynx. As this source waveform differs from individual to individual, determining a proper method to characterize vowel formants for any individual will be my first task. I will start by attempting to characterize the voice of an individual as generally as I can, determining the fundamental frequency and and other traits of their voice.

3 Deliverables

1. A pronunciation visualization program and source code providing biofeedback on vowel formants and possibly tones and intonation.

2. Blog posts illustrating the problems encountered and solutions attempted.

3. Final report detailing the extent to which the problem goals where achieved as well as a preliminary study on the efficacy of the biofeedback software on improving pronunciation in second language learners.

4 Timeline

1. Research speech and formant characterization. (~3 weeks)

2. Develop application to visualize and correct vowel pronunciation. (~4 weeks)

3. Extend application to visualize tone and tonality (time permitting). (~3 weeks)

4. Conduct efficacy study, refine application, and write final report. (~3 weeks)