ECE446 Project Proposal: Languages and their Acoustic Differences

Oct. 10, 2022

Angelina Zhai and Maisey Perelonia

1. Introduction

Languages are an interesting acoustic topic for their necessity and diversity. There exist tonal languages such as Mandarin, where the spoken words utilizes four different tones; a wrong tone can indicate a completely different meaning. There also exist atonal languages such as English, where tones are not needed to indicate different meanings. As students who are proficient in more than languages, we are especially interested in the acoustic differences between the languages we speak everyday. Through this project, we hope to determine the differences in average pitch between different spoken languages.

2. Background

Previous research has been conducted regarding languages and pitch. A recent study in 2022 focused on how the spoken pitch range could change depending on the language spoken and the social meaning an individual is trying to convey (ie. politeness). This study recruited 41 male and female individuals who were proficient in both Japanese and English, and the participants were asked to read out and record 16 English sentences and their corresponding translations in Japanese given different addressees for each sentence to indicate different social contexts. Acoustic analysis was then conducted on the recorded dataset using autocorrelation pitch tracking using Praat, a linguistic analysis tool and with a custom code script. Noticeable differences in pitch range between the two languages in different social contexts were observed in all participants. Though the results indicated that our pitch range changes depending on the social situations we are in, the results do not shed light on whether language alone alters an individual's spoken pitch. [1]

3. Hypothesis

Based on the background research conducted in this topic, the objective of this research project is to determine if the average pitch of speech can differ depending on the language spoken.

4. Methodology

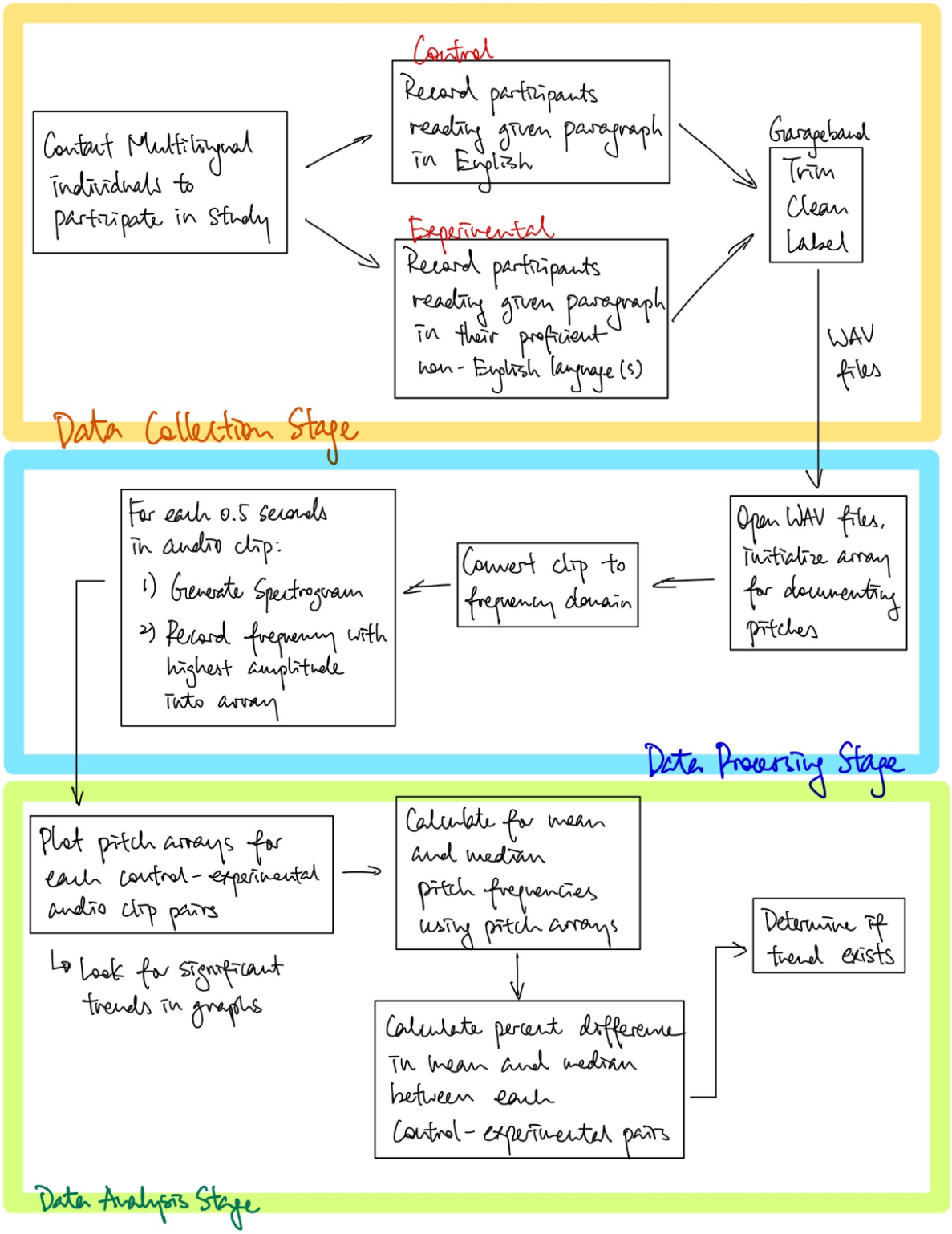

Figure 1: Flow Chart describing Methodology of Data Collection, Data Processing and Data Analysis

Below is the methodology we will use to answer the proposed question. We will first start contacting bilingual and multilingual individuals we are acquainted with and ask for consent to participate in the experiment. Some of the participants will be proficient in tonal languages, and others in atonal languages. Then, appointments will be arranged to record each individual in the same room at different times. A recording microphone with the sampling rate of 44.1 kHz will be used. Each participant will then be asked to read out a 15 second long sentence in English as the control sample, and its translation in the participant's non-English language(s) as the experimental sample(s). The sentence will be in the format of "My name is ___, I am ___ years old, and I like to do___. Today I had ___ for a meal, and I am planning to do ___ later."

After the data collection is complete, all recordings will be trimmed, cleaned, and labeled with participant name and languages using Garageband. The audio clips will then be exported as WAV files, and will be opened in MATLAB using the audioread() function, with a sampling rate of 44.1 kHz. An empty array will be initialized to document the important pitches of the audio clip. The audio will then be processed in half-second batches (22050 frames per batch) in MATLAB using the following procedure. First, the audio clip containing the batch will be converted into the frequency domain; a spectrogram of the batch will be generated using MATLAB's spectrogram function. Then, to ignore the noise, overtones, and undertones of the pitch, the frequency with the highest amplitude will be appended into the pitch array (for it represents the most prominent frequency in the batch). Finally, repeat the process until the entire clip is processed.

After MATLAB analysis is performed on all audio clips, we will then plot the pitch arrays of the control sample clips and the arrays of their corresponding sample clips on the same graph, and look for immediate and significant patterns such as one set of data points always higher than the over, if such trends exist. Using MATLAB, we will then calculate the mean and median frequencies in both pitch arrays and calculate the percent difference between the two. All percent differences between all experimental and control pairs will be marked down, then we will decide if a trend can be observed between tonal and atonal languages, as well as between English and other languages.

5. Resources

- Microphone: Always available
- Garageband (for trimming and cleaning raw audio): Always available
- Participating members: Need to contact and arrange recording sessions
- Computers with MATLAB or Python: Always available
- Phone to connect the microphone to: Always available
- Room used to record audio samples: Need to reserve and coordinate with participating members

6. Timeline

Task	Deadline
Gathering Equipment	Oct. 14
Recruiting Subjects	Oct. 14
Electrical System (code and stuff)	Nov. 1
Recording Data	Oct. 21
Organizing & Examining Data	Nov. 4
(In case: Collect more data & examine them)	Nov. 11
Start Paper Writing	Nov. 11 (if no additional data is needed)
Finish Paper Draft	Nov. 25
Finish Polished Paper	Dec. 2

Table 1 – Deadlines for Project Subtasks

7. References

[1] E. Passoni, E. de Leeuw, and E. Levon, “Bilinguals produce pitch range differently in their two languages to convey social meaning,” Language and Speech, p. 002383092211052, 2022.