

# CS6300: Speech Technology

## Lab Assignment 1: Report

Abhishek Sekar  
EE18B067

February 2021

### 1 Assignment Introduction

- This assignment revolved around a data set comprising 10 different sentences. The ten sentences were recorded and the vowel, syllable and word boundaries were subsequently labelled.
- A histogram pertaining to the duration of vowels and syllables was computed using the information captured in the labels.
- The syllable rate of the speaker was also computed.
- Five different speakers were compared using the mean and standard deviation of each syllable.

### 2 Relevant Plots and Values

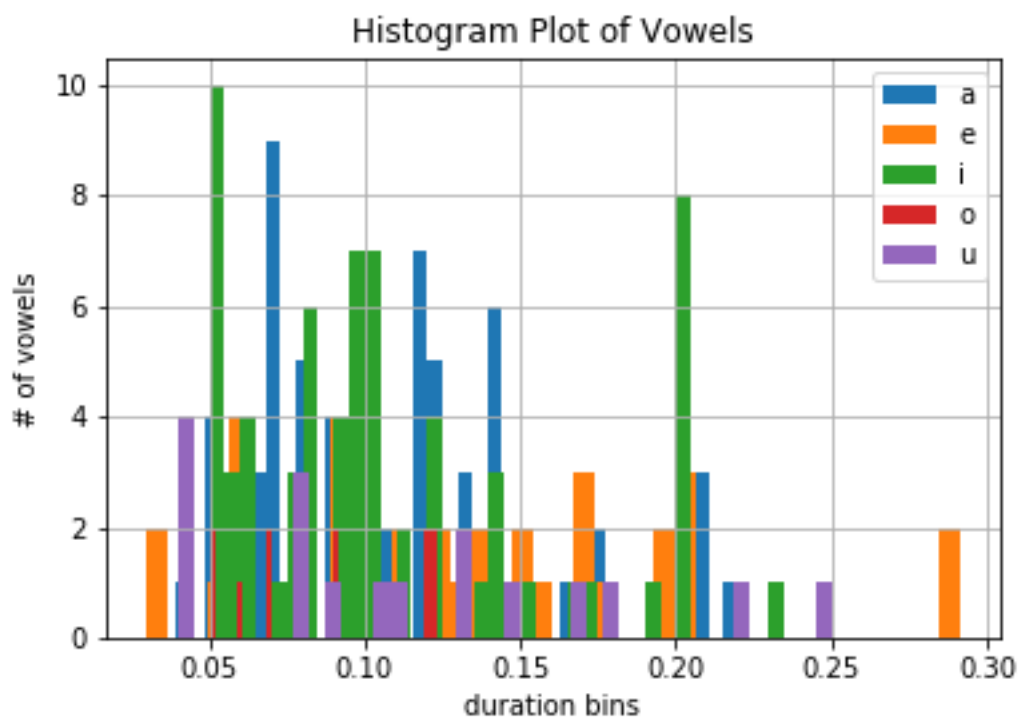


Figure 1: Histogram Plot of Vowels

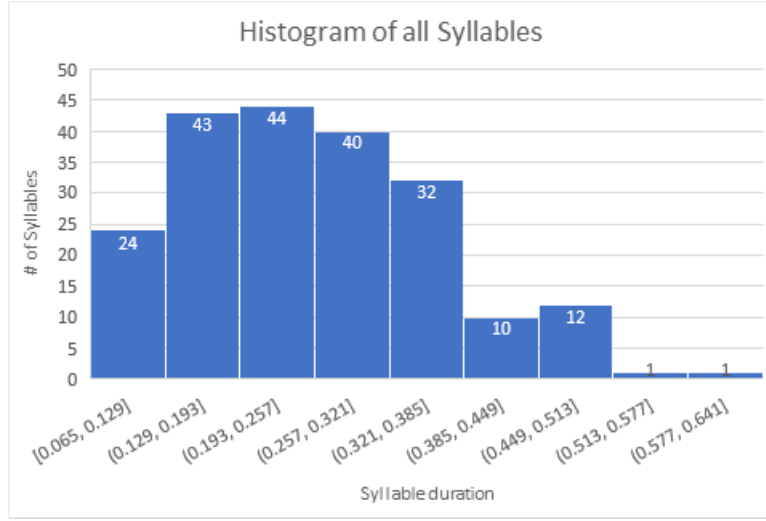


Figure 2: Histogram Plot of all syllables

- My syllable rate was approximately 3.9 syllables per second.
- Mean and standard deviations of an example syllable (wer) is given in the table below for the 4 speakers.

Sr.No.	Mean	Standard deviation
1	0.377324	0.144403064
2	0.2708995	0.035721133
3	0.358116333	0.05366587
4	0.242699	0.034611799

### 3 Observations and Results

- From the histogram plots we can observe that there is no unique duration for any vowel and the context of the vowel in a particular word primarily determines its duration.
- Comparing both the vowel and the syllable histogram plots we can observe that the vowel forms the bulk of the duration of the syllable while most consonants cut off quickly.
- Observing the speech spectrum, we can see that the vowels form magnitude peaks. On further reading, I learnt that this is because they coincide with the resonance regions of the vocal tract. This is why we can easily identify vowels when compared to consonants.
- While observing the histogram plot of the vowels, we see that quite a few of them are too short in duration. Probing further, I found that this happens when the vowel is surrounded by a consonant.
- While probing further into the duration of the syllables I observed that some specific syllables were much longer than the other ones. I then found one thing common with the longer syllables. They ended with a vowel while the shorter ones ended with a consonant.
- From the mean and standard deviation values obtained from multiple speakers, I observe that they are roughly in the same neighbourhood of durations. This could be because the syllable rate is usually limited to a range of approximately 3.3 to 6 syllables per second.