Phone Detection Notes

Maybe meet every other week. Once everything’s settled down, possibly twice a week.

By April we need a written report. The deliverable for phone recognition needs 80% accuracy.

There are resources for us to use.

Our task for this month is to read and understand what our Client is doing/trying to do. Materials will probably be given to us this week.

Speech recognition used to be done solely by humans, but now with improved technology, we can detect accents, prosody, tones, etc. and give a score based on these qualities.

Right now, there isn’t enough accuracy with this technology, and this needs to be improved.

This technology may be used to give speakers a score, based on how well they generally speak. This includes pauses, use of the word ‘uh’, and other factors mentioned above.

ETS relies heavily on speech rate, so its score accuracy is closer to 60%. We need to add additional features and improve accuracy.

One additional feature would be phone detection.

Phone means any sound you might make. Pronunciation split down to each syllable.

The scoring is split into separate steps, and our goal is to make the first step, listening for speech/phones, much better to improve accuracy overall.

We need to focus on replicating a complex program to recognize only phones, rather than words, and then improving on it.

MATLAB was chosen because it’s a language often used for AI and other complicated learning programs. Very robust, has a lot of built in functions coded especially for AI.

The audio files given to use will be sentences. 637 speakers, of different American-English dialects, Californian, Army brat, etc.

Corpuses – look this up

Deliverable:

We need a program to do this: Audio file -> List of phones