Automatic extraction of vowel formant values

Master Natural Language Processing

TABLE OF CONTENT

Record a speech signal in French

Segment the speech using Astali

Extract the F1 and F2 values of the vowels

Praat Script to visualize the F1 & F2 values

Comparison of vowel F1 & F2

French speech signal

For the speech signal I read a wikipedia page on “fenêtrage”, a technique used on speech analysis.

I used Audacity to record and manipulate the speech signal

Segment speech using Astali

After recording 4 minutes of speech signal I had to create a text file for the Astali annotation but actually Astali do not authorize speech signal > 180 secondes so I cut it at 3 minutes, but still Astali could not take the file in charge, so I cut it in 3 parts and did the annotation separately.

Extraction of the F1 & F2

After the annotation by Astali I had 3 inputs files, and 3 TextGrid files. So I used Praatio and specially tgio to get the TextGrid and starting-ending time of the different segments, take the middle of each vowel segment and with Parselmouth, a library that allow Praat function on Python to get the first and second formant from a point, I get the formants values in the middle point of each vowel. Where the modification of vowels due to coarticulation is smaller.

The last step was to combined the 3 outputs files.

Praat Script to visualize the F1 & F2 values

The vowel with its corresponding Formant 1 and 2, saved in a text file.

I used Praat command to get a representation of those formants in a graph with F1 and F2 as axes.

Actually it’s doesn’t seem like a triangle but, it’s because of the fact that I’m sick my voiced is nasalized and, I spoke fast and could not articulate properly, those facts neutralised vowels formant that tend to be closer to the neutral vowel shwa.

Comparison of short and long vowel : F1 & F2

Unfortunately, on my corpus, I don't have enough data on long vowel to do a comparison.

To know if it's a short vowel or not, I used Praatio, thanks to the born of the segment. I subtract the time point of the end of the segment by the time point at the start of the segment. If it's < 60 ms it's considered 'Short' if it's > 90 ms it's consider 'Long'.

Segment between 60 and 90 ms was not taken into account.

Thank you !