Week 6 - Speech Breathing Empathy Project

Nicolò Loddo

What happened

- Decided to not choose a specific aligner but rather merge two of them to get the best parts of both.

- Developed a script that merges Gentle and Montreal aligners putting priority on Gentle's alignment, and that manages inconsistencies in timing of close words.

- Generalized and polished all the scripts to be used on datasets of audio even without the breathing signal

- Finished the breath labeling script, added parameters to be used from command line

- Qualitative and quantitative evaluation of the parameters to use for the breath detection and labeling

- Ran the breath detection and labeling: I now have transcriptions with breathing tags and their alignments

- Wrote script to segment big audios into smaller sentences with breathing instances separating them. It features upper and lower limits of time that cuts the audio even if there is no breathing.

- Qualitative evaluation of the audio-segmenting script: everything seems fine. The preprocessing pipeline is done

- Wrote script that encapsulates the preprocessing pipeline

Next week to-dos

- Report the evaluations on overleaf

- Run the segmentation script

- Make a go at style transfer with INTERSPEECH

- Make a go at training some model with the data

- Put papers into each section of the thesis proposal

- Start the writing of the literature review

- Read and comment better the speech synthesis literature with its evaluation methods

- Search for more emotional datasets in english

```
ndt.search_breath_analysis('devel_00.wav', gentle_mfa_json, nonstop = True, printonlyfinal = True)

Final total breath sections: 62
*****************************************************
```
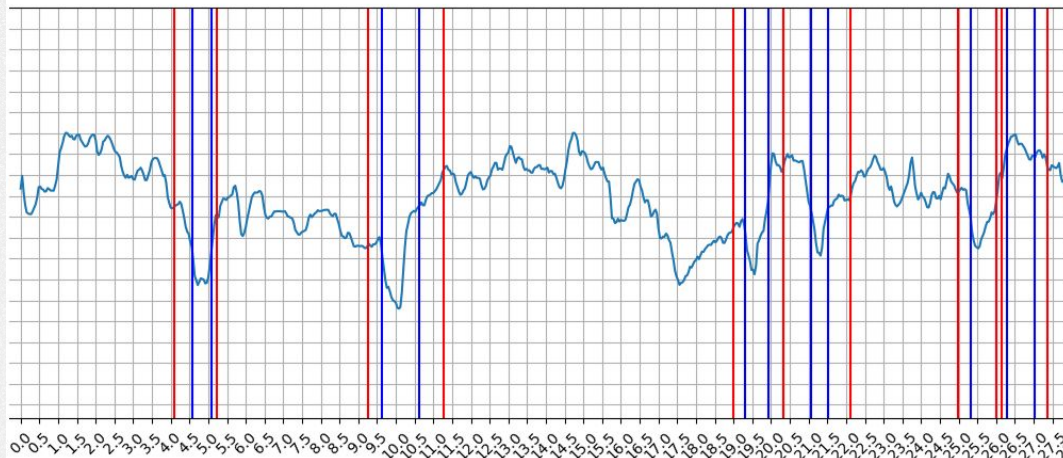
Breath detection with visual plots on the breathing signal

This is the detection output of the script, with statistics on the breaths of all files:

```
    FINAL STATS:
    Average number of breaths: 62.43181818181818
    Max breaths in a file: 101
    Min breaths in a file: 29
    Average breath duration: 0.5777072092602468
    Max breath duration: 4.88
    Min breath duration: 0.27
```

Since all the recordings have a length of 4 minutes we have the following respirations per minutes (rpm) stats:
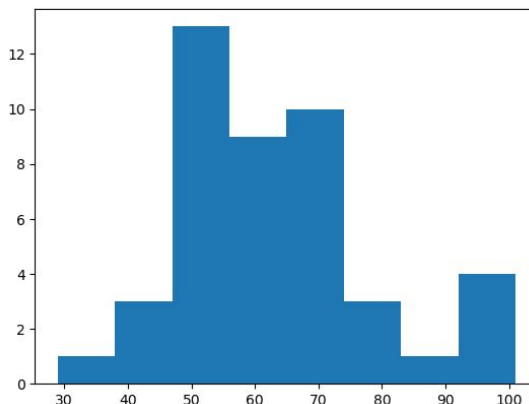
```
    Average rpm: 15.5
    Max rpm: 25
    Min rpm: 7
```

Which perfectly checks out with the known average rpm of a resting person (between 12 and 16)

mean: 62.43
std: 14.78

Breath labeling stats (on the left); number of breaths histogram (on the right)