

Week 4 - Speech Breathing Empathy Project

Nicolò Loddo

What happened

- Made structure of thesis proposal
- Examined other models and approaches to speech
- Extended speech synthesis literature
- Wrote scripts for segmenting audio to feed to flowtron
- Wrote scripts to use Montreal Forced Aligner
- Wrote script to use AssemblyAI Speech to text API
- Evaluated results the Montreal Forced Aligner against the previously used Gentle
- Evaluated results of AssemblyAI against the previously used IBM Watson
- Wrote script that systematically identifies breathing pauses in the audio from the aligned transcriptions

- Looked more into HiFi-GAN which turned out to be a vocoder, not a full text to speech model. Still it can be useful for later processes
- Searched a bit for a spontaneous speech emotional dataset in english, with not great results

Next week to-dos

- Search better for emotional datasets in english
- Start putting papers into each section of the thesis proposal
- Read and comment better the speech synthesis literature with its evaluation methods
- Better exploration of INTERSPEECH
- Make script to label the breathing pauses on the dataset and possibly make a go at style transfer with INTERSPEECH (finally!)

