Sound Processing with pyaudio (microphone_streaming_with_spectrum.py)

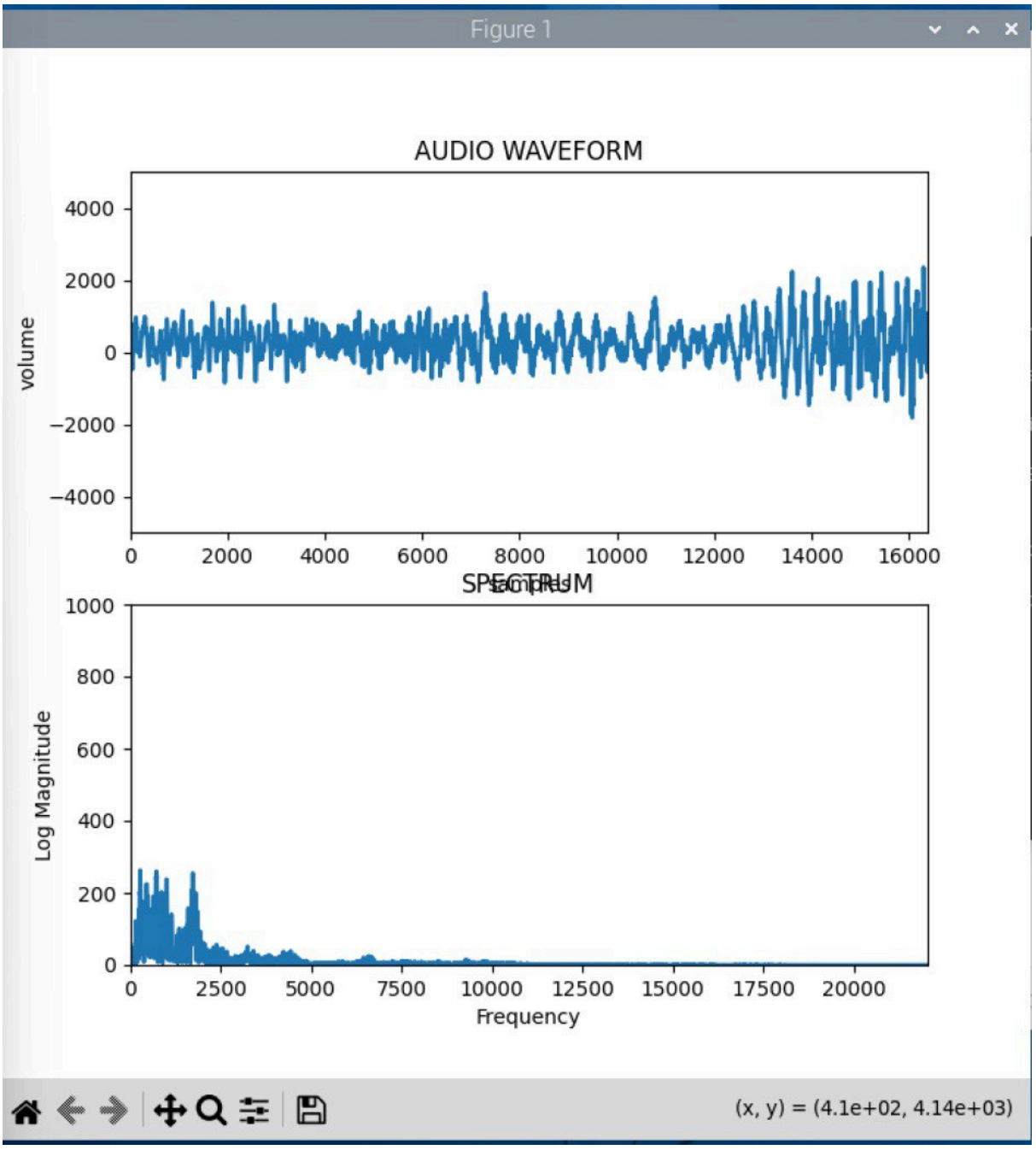

Fig 1. Sample Captured Speech using pyaudio

Speech will be captured through the microphone of the webcam, and visualised onto the audio waveform spectrum as seen in Figure 1.

Basic Sound Analytics with pyaudio (filtering_audio.py)

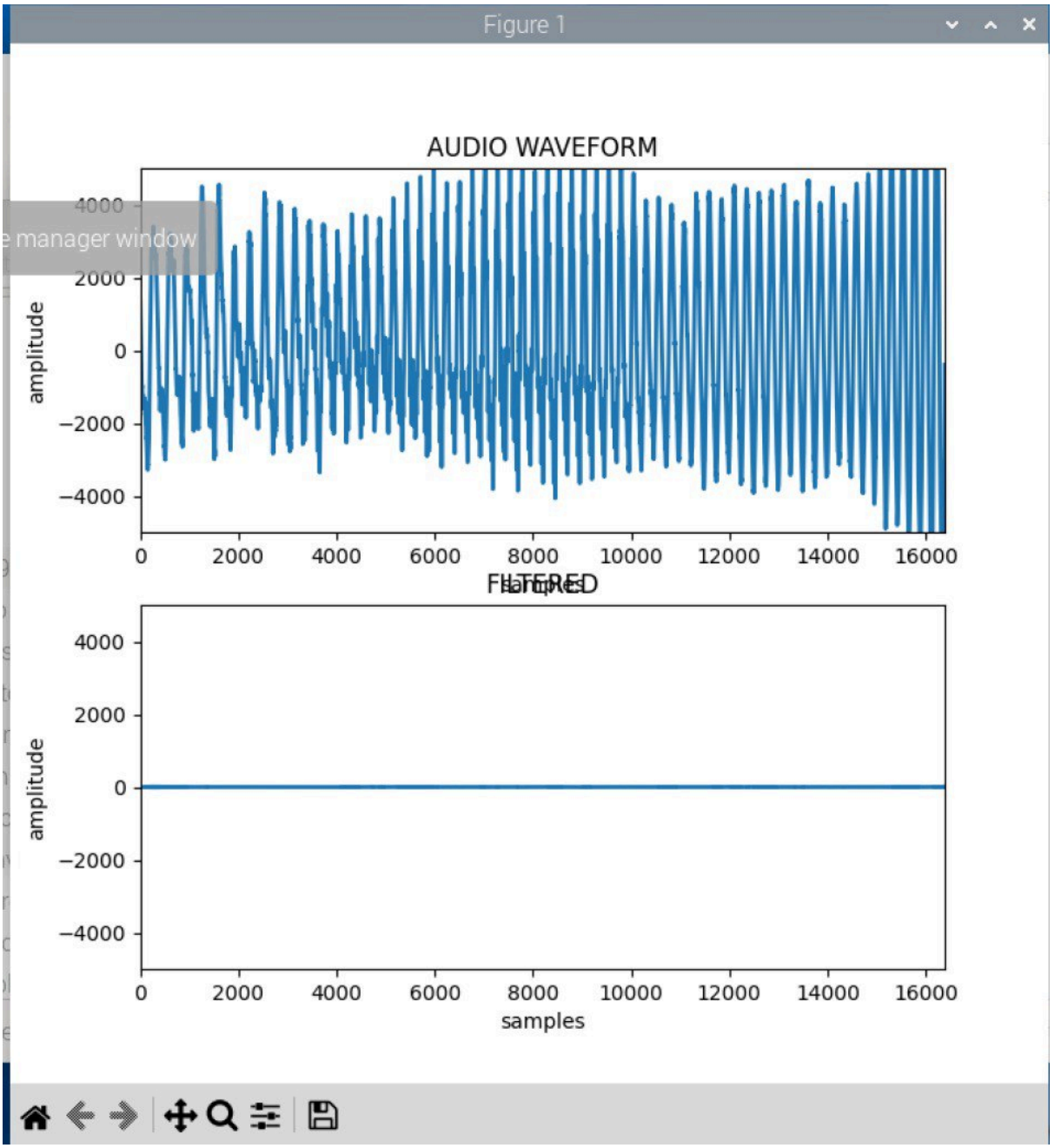

Fig 2. Filtered audio spectrum visualisation with pyaudio

The filtering removes noise and unwanted frequencies, which can be seen in Figure 2 whereby the filtered audio spectrum of our voice is at a amplitude of 0 due to the code specifying the frequency range of 19400Hz to 19600Hz.

Librosa Spectrogram (audio_features.py)

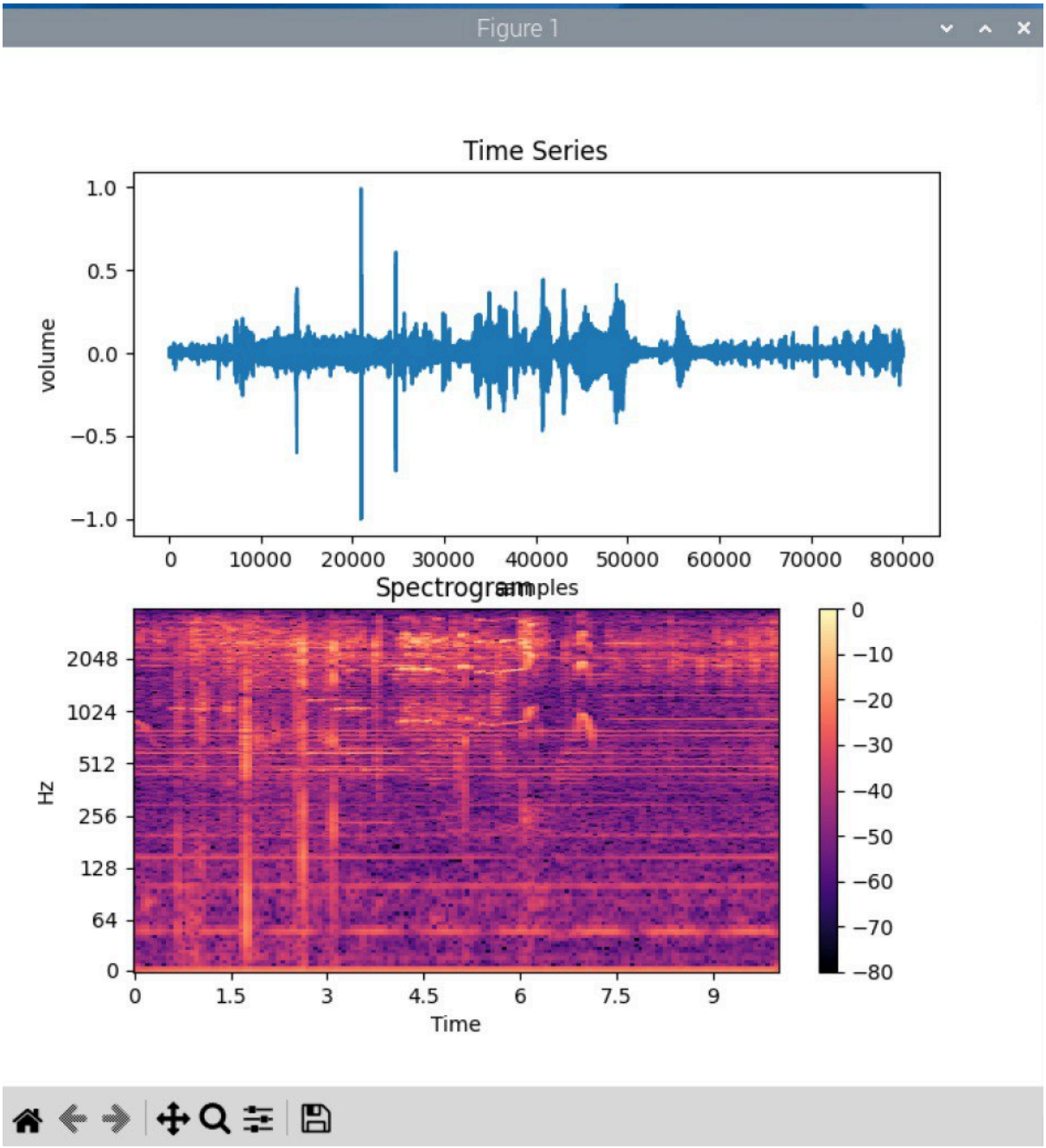

Fig 3. Spectrogram visualisation

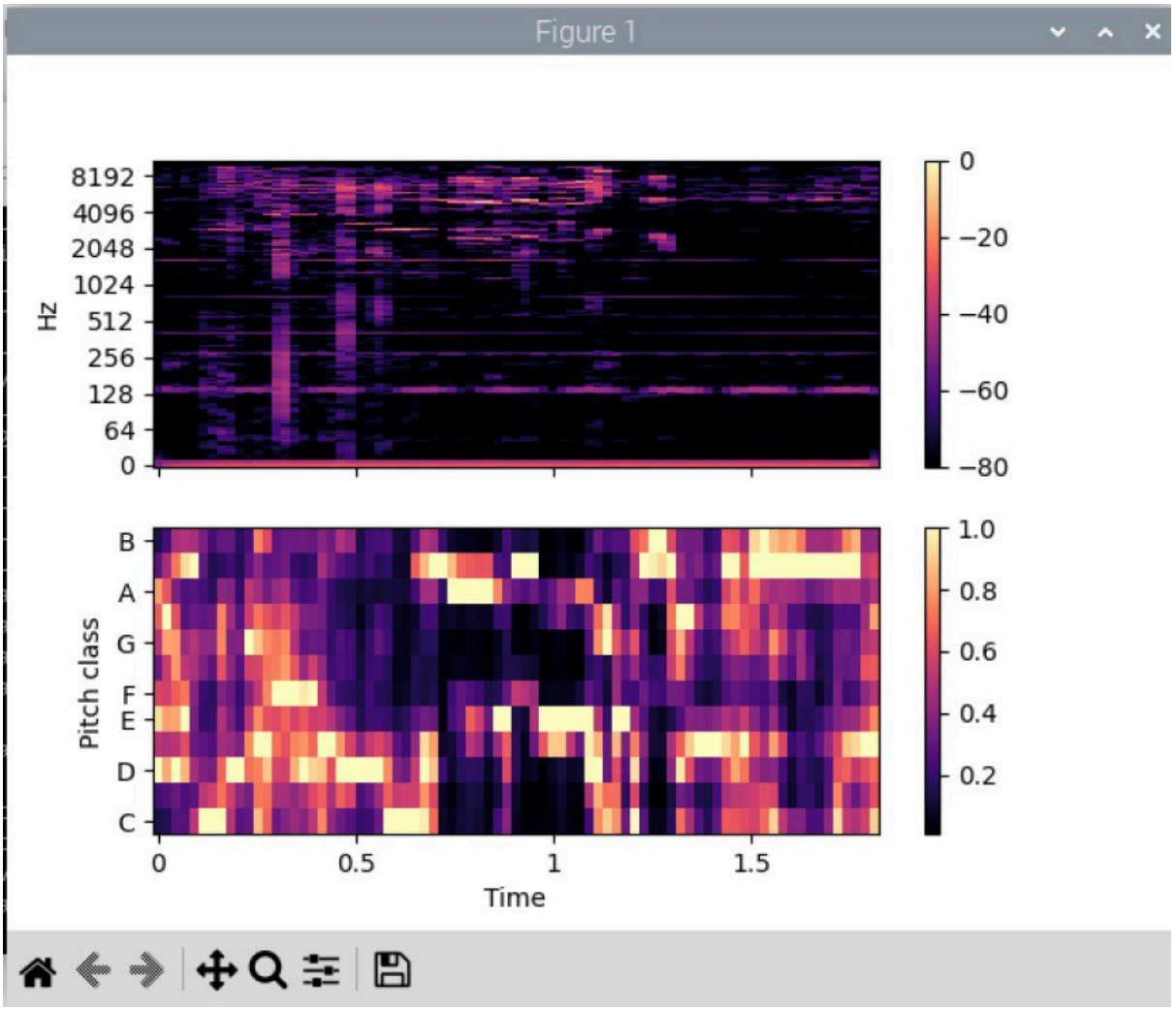

Fig 4. Chromogram visualisation

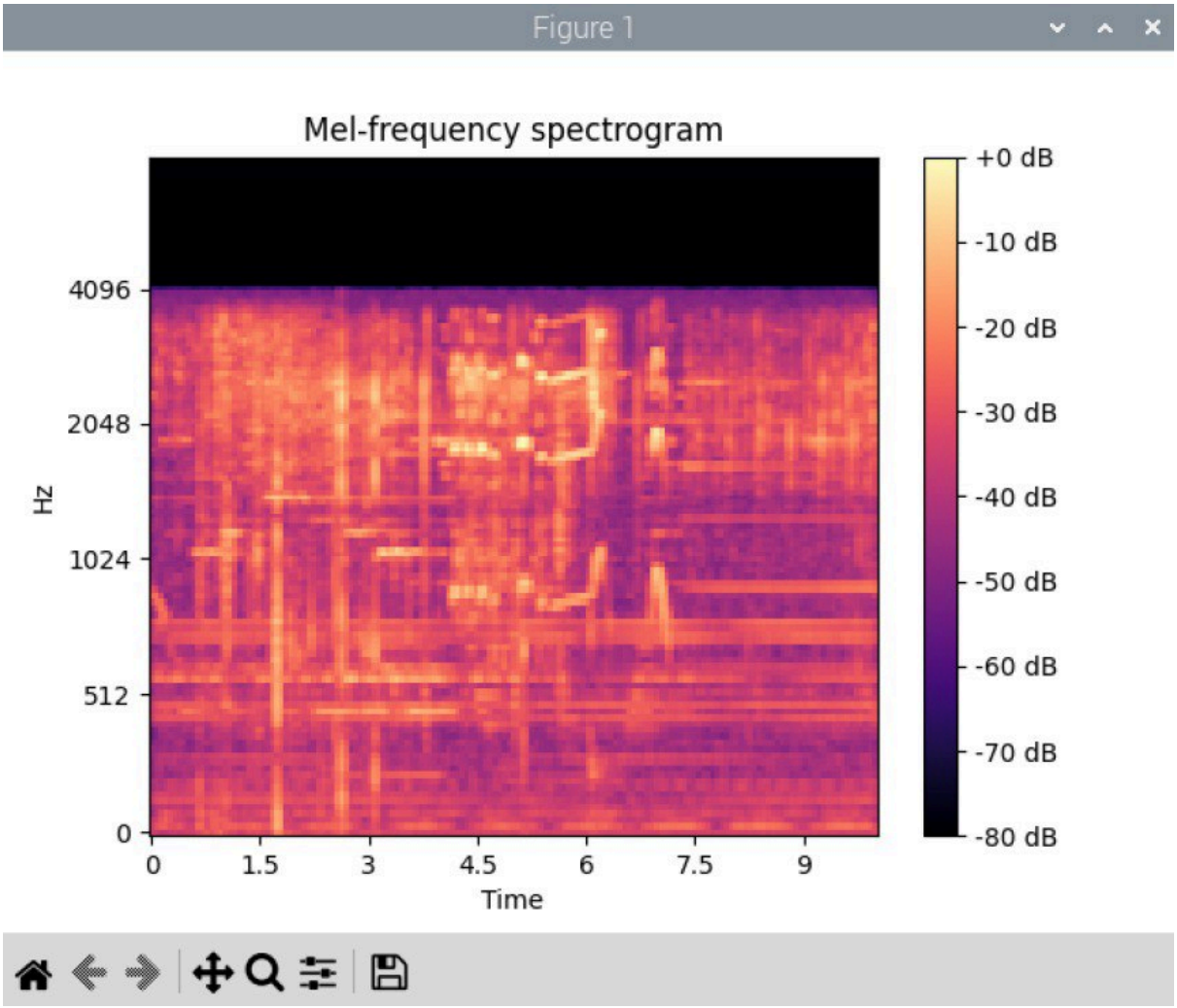

Fig 5. Mel-frequency spectrogram

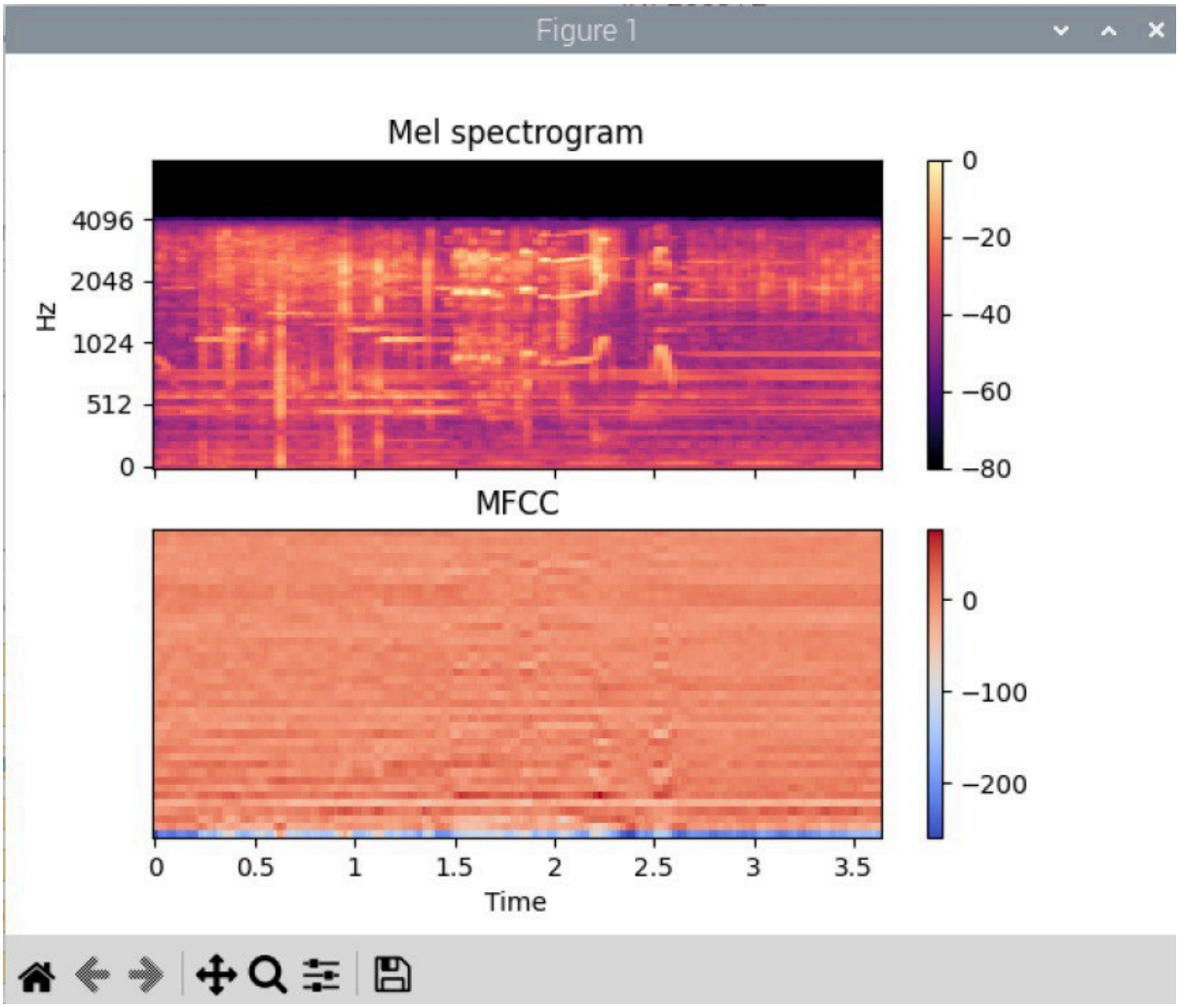

Fig 6. Mel Frequency Cepstral Coefficients (MFCC) and Mel Spectrogram comparison

Advanced Sound Analytics (microphone_recognition.py)

The performance of the OpenAI Whisper speech recognition API is compared to the Google and Sphinx speech recognition API as seen in Figure 7.

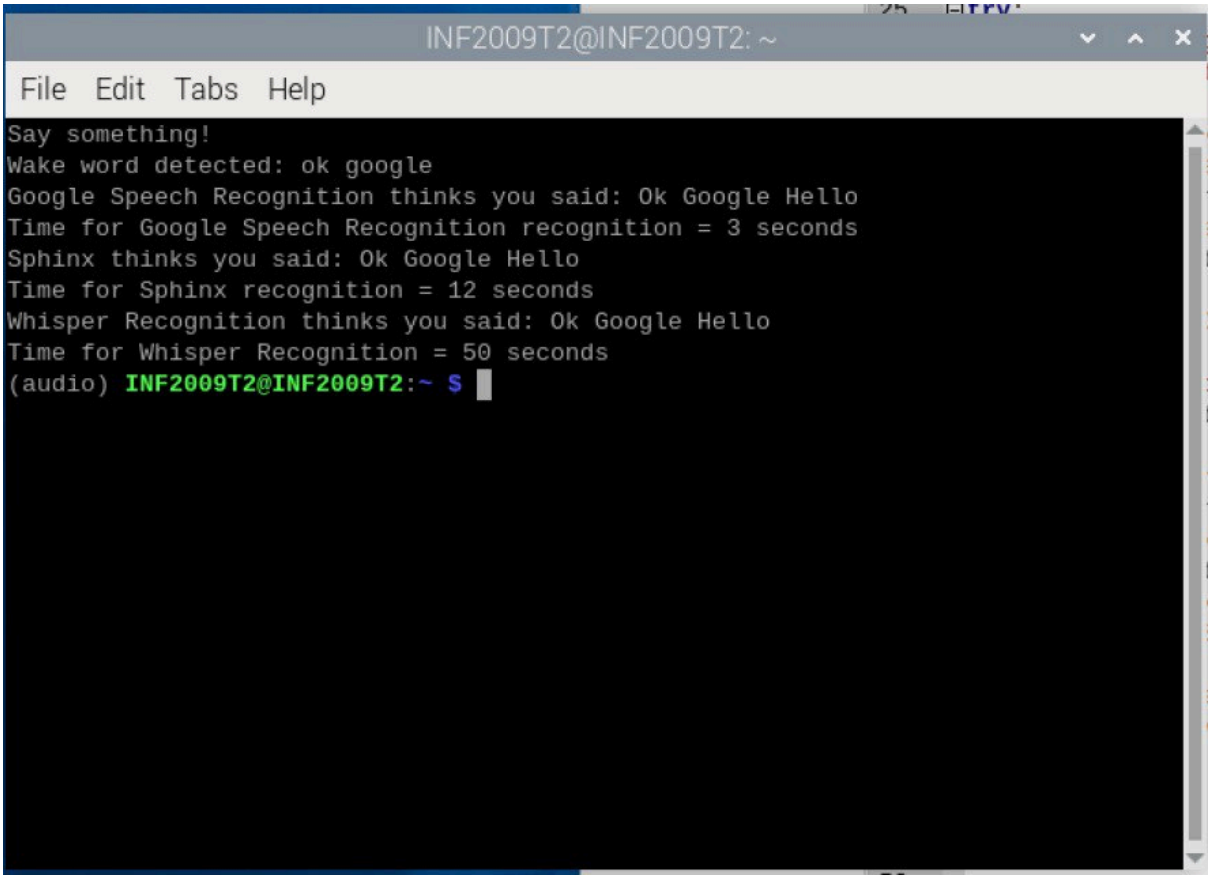

Fig 7. Performance comparison between the speech recognition APIs