Technical University of Munich

Neuroprosthetics Exercise 6 Report

Student Name: Abdelrahman Elsissy
Student Number: 03780605

TECHNISCHE
UNIVERSITÄT
MÜNCHEN

1. Audio Signals

Cochlear implants have many challenges to the functionality and effectiveness of the human ear. One of the main challenges for cochlear implants is the reduced number of frequency channels. This exercise is focused on representing the effects of this reduced number of frequency channels. The phrase “effective elephant” was chosen for this exercise as the speech signal as it matches the criteria stated in the handout. The audiorecorder() function was used in MATLAB to obtain the speech signal. The recording was done with a sampling frequency of 21 kHz, at least 16-bit resolution (bits per sample), and a single channel, as instructed in the handout. A threshold of 0.002 was applied to reduce the background noise, meaning that all the values in the recorded signal below this value will be set to 0. An ideal threshold should reduce the background noise but allow the audio to be audible. White noise was generated using the randn() function in Matlab with exact dimensions of the speech signal. The spectrogram of the recorded phrase and the noise signal were plotted with a time window of 15 ms with an overlap of 5 ms, as illustrated in **Figures 1 and 2**, respectively.

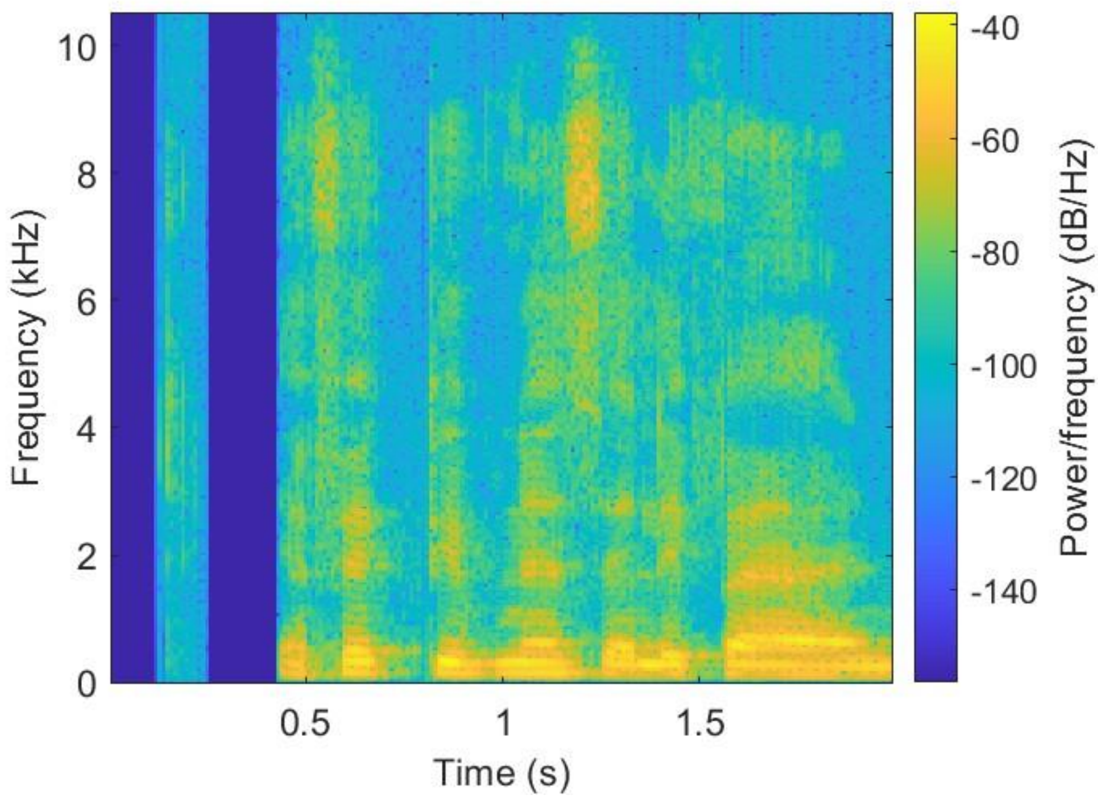

Figure 1. Speech Signal Spectrogram ("Effective Elephant Phrase")

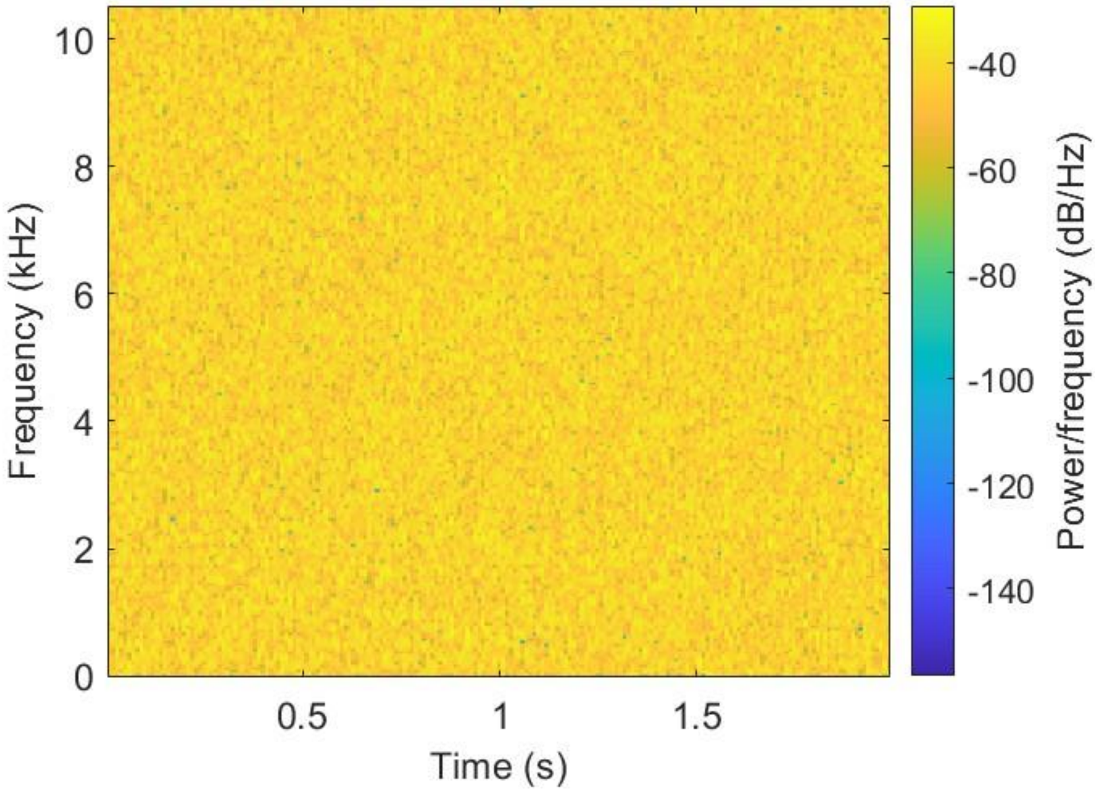

Figure 2. White Noise Signal Spectrogram

2. FILTER BANK

A filter bank with 10 channels was implemented to represent and simulate a CI with 10 stimulating electrodes. **Figure 3** below shows the border frequencies (100 Hz to 8kHz) of a 10-channel filter bank. Each of the two neighboring filters share one border frequency.

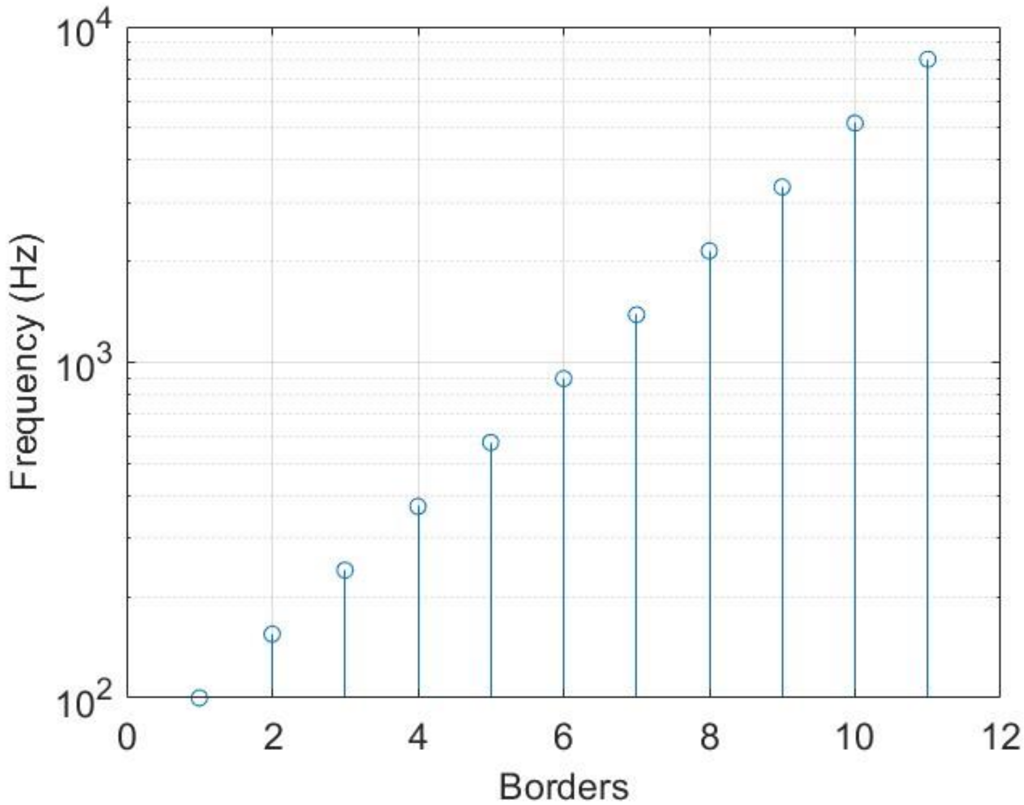

Figure 3. Border frequencies of the 10-channel filter bank

10 fourth-order Butterworth digital filters, with -3 dB at the corresponding border frequencies, were created to implement the filter bank. **Figure 4** below demonstrates the magnitude response of the filter bank.

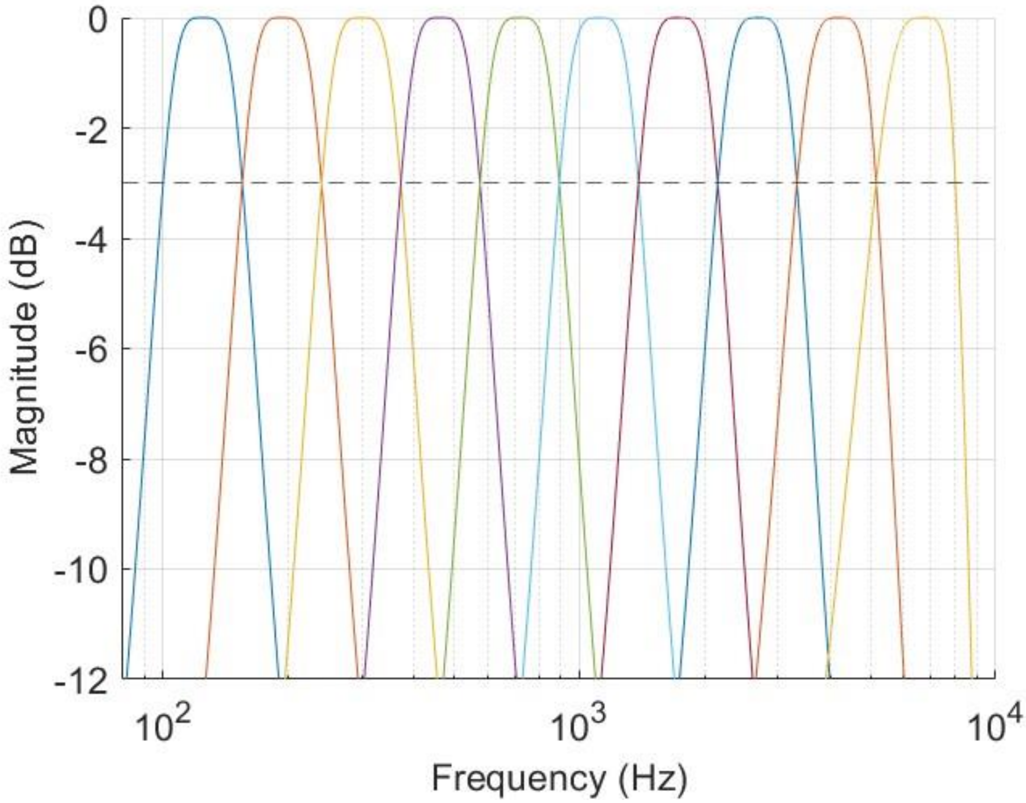

Figure 4. Magnitude response of a 10-channel filter bank

Using our filtered bank, 10 band-limited speech channels and 10 band-limited noise signals can be obtained. After filtering the speech and noise signals with the second highest lowest at channel 2, **figures 4 and 5** were obtained, respectively. Additionally, after filtering the speech and noise signals with the second highest frequency at channel 9, **figures 6 and 7** were obtained, respectively.

Having analyzed the audio, it can be observed that the speech signal subjected to filtering in the second lowest frequency channel (channel 2) exhibits enhanced vowel sounds but with a perceived muffled quality. On the other hand, the speech signal filtered through the highest frequency channel (channel 9) sounds less muffled, somewhat artificial, and features more pronounced consonants. These distinctions in audio characteristics can be attributed to the variations in frequency between the channels.

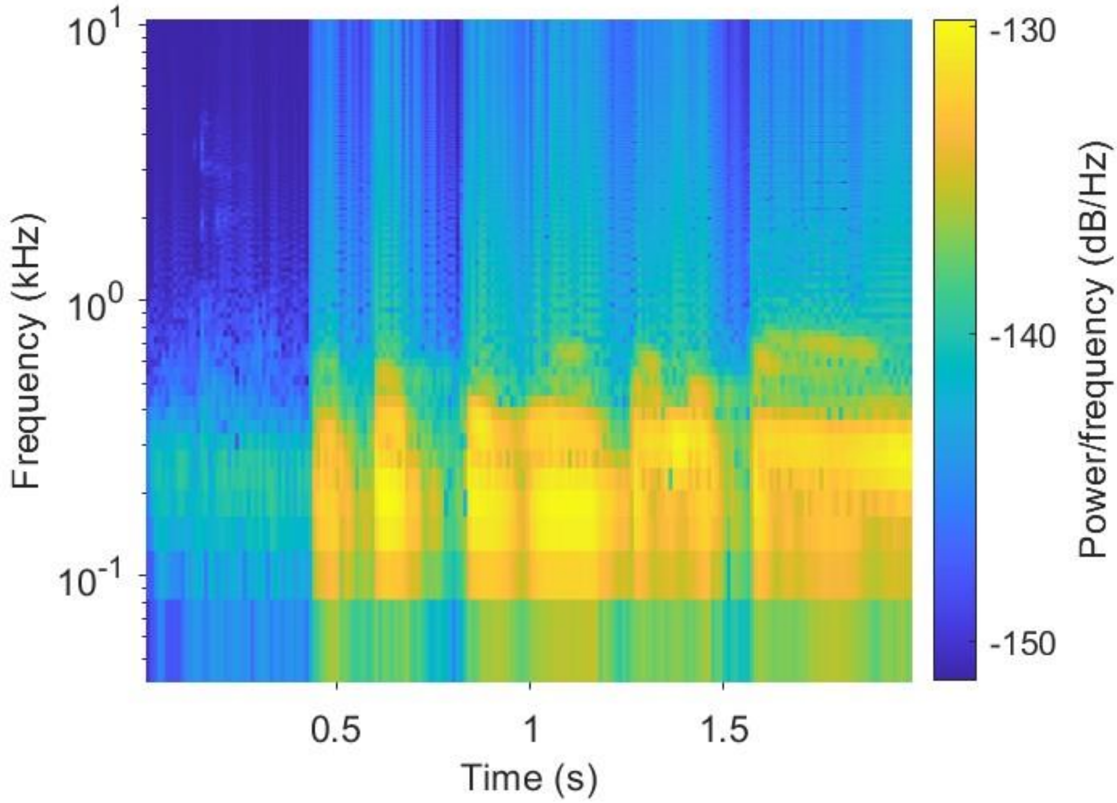

Figure 5.Spectrogram of a speech signal filtered by the second lowest frequency channel.

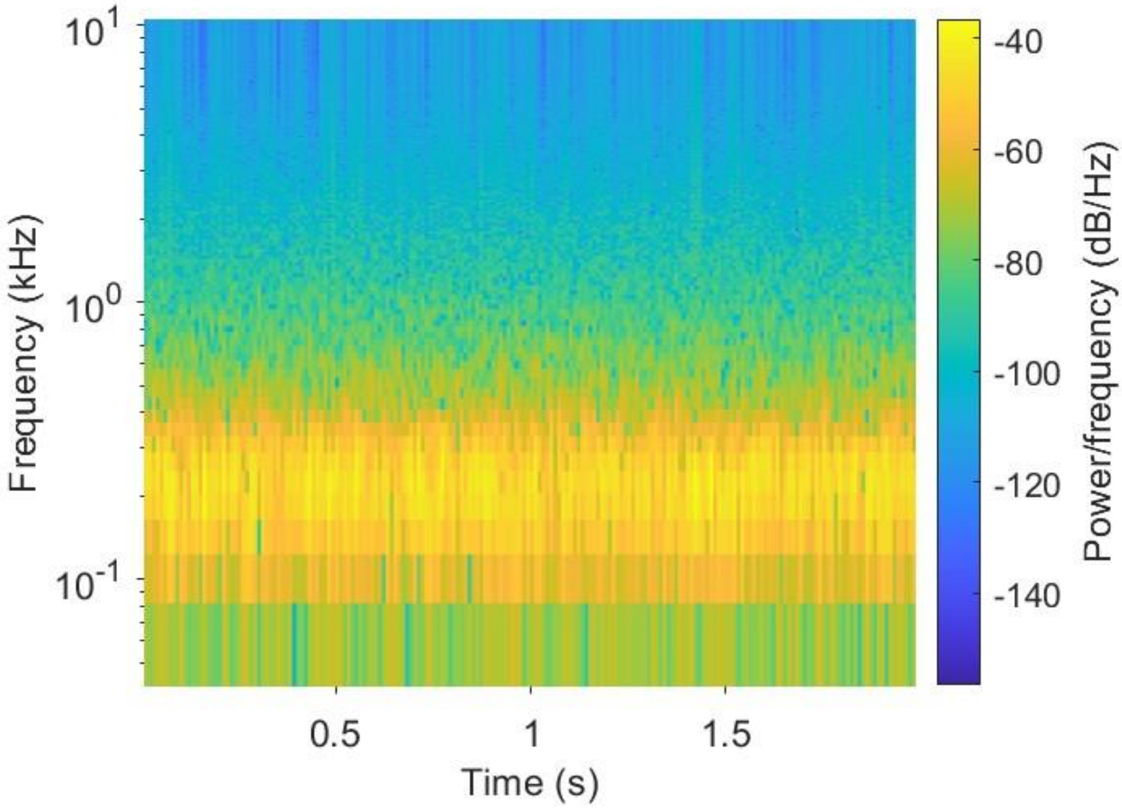

Figure 6. Spectrogram of the noise signal filtered by the second lowest frequency channel

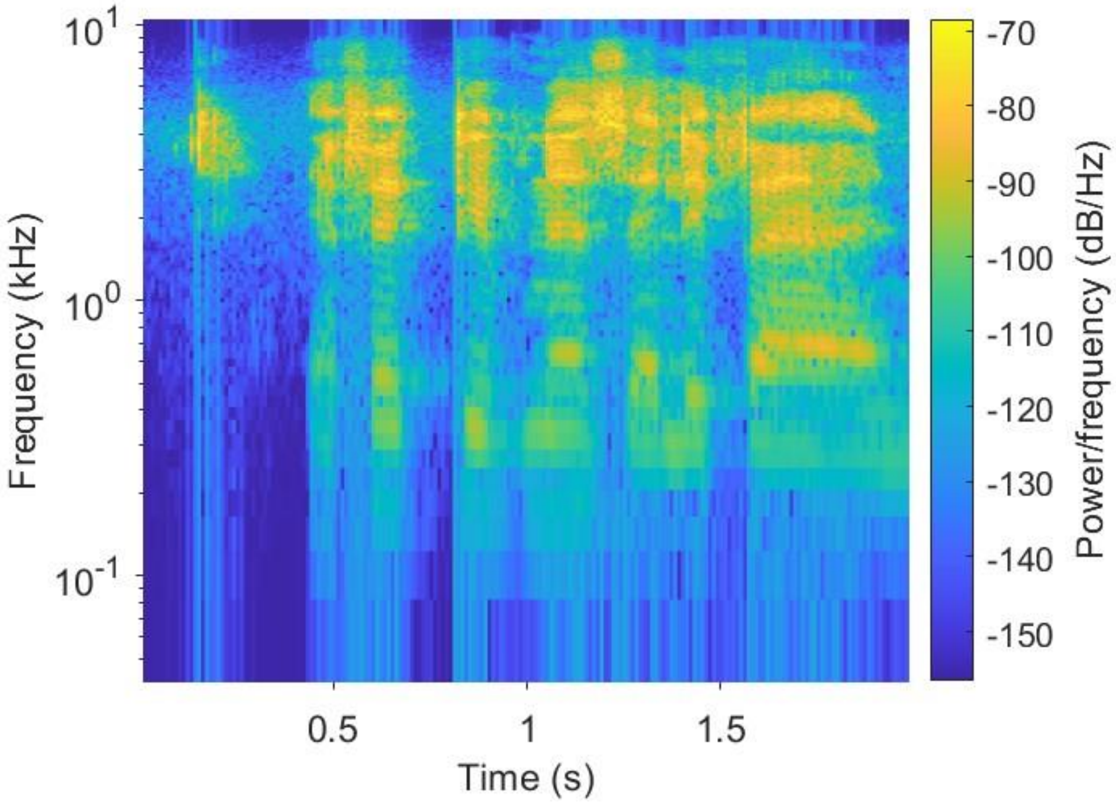

Figure 7. Spectrogram of a speech signal filtered by the second highest frequency channel.

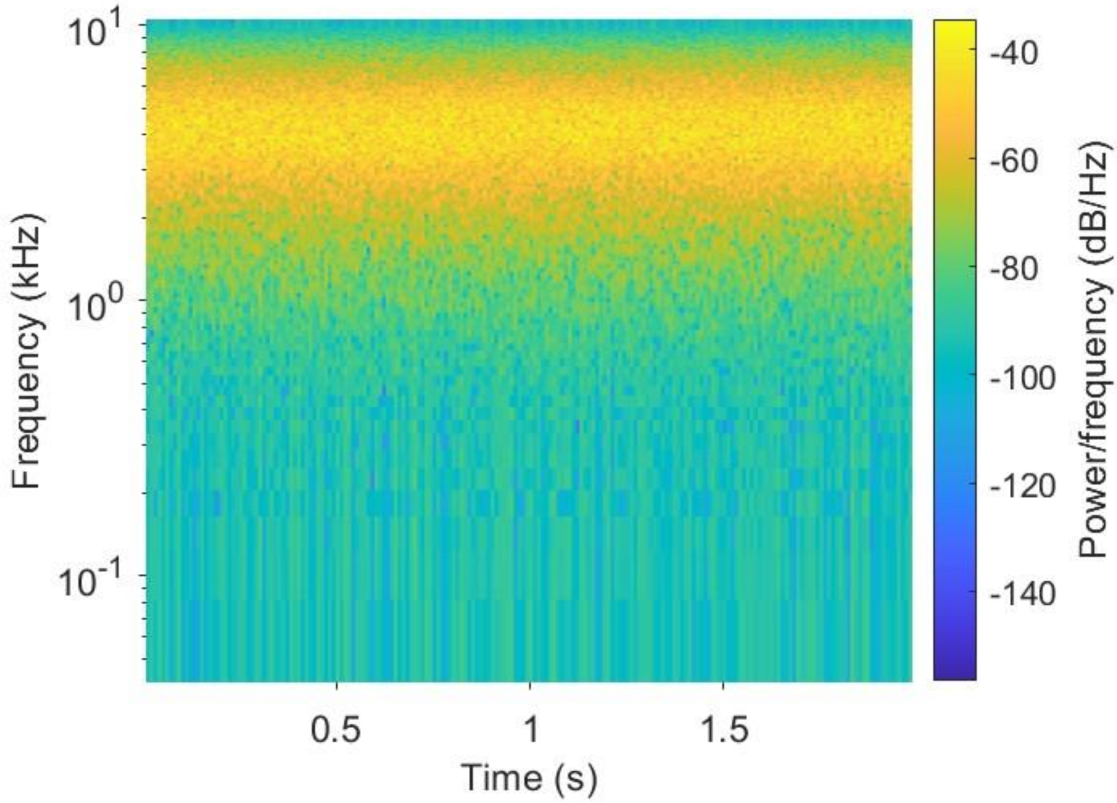

Figure 8. Spectrogram of the noise signal filtered by the second highest frequency channel

3. VOCODER

Each filtered signal's amplitude information (envelopes) over time was obtained using the Hilbert transform to implement the vocoder. As discussed in the lecture and earlier in this report, cochlear implants have many challenges to the functionality and effectiveness of the human ear. One of the main challenges for cochlear implants is the dynamic range of the human ear, as the dynamic range of loudness that humans can distinguish is much larger than the dynamic range of stimulation currents CI electrodes can employ. To represent this effect of the reduced dynamic range on the speech signal, the envelopes were compressed using Equation 1 below. The last step was modulating the band-filtered noise with the respective envelopes obtained from the band-limited speech and summing up all the 10 channels to obtain the final vocoded signal. Figure 9 below shows the vocoded signal, which has a time window of 15 ms with an overlap of 5 ms.

$$env_{compressed} = \frac{log_{10}(1 + 300. env)}{log_{10}(1 + 300)} \tag{8}$$

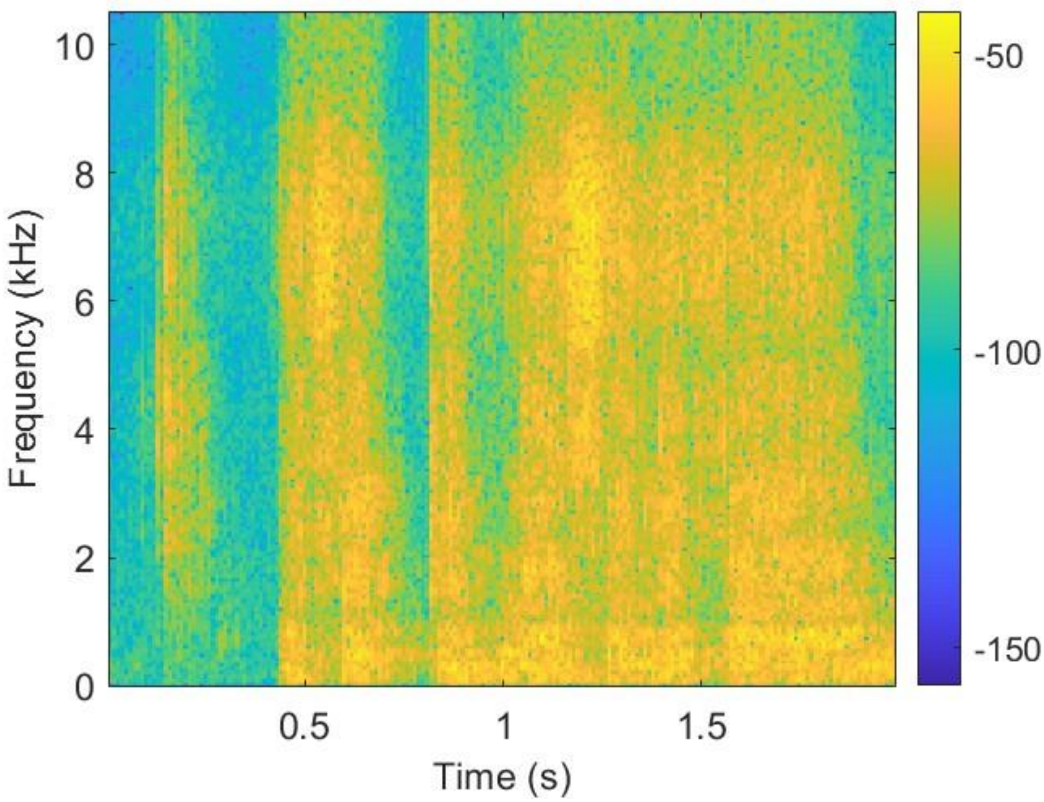

Figure 9. Spectrogram of the vocoded signal

We can see a clear difference when comparing the original speech signal with the vocoded signal. The original speech signal sounds much more comprehensible and intelligible. Even after thresholding the original speech signal, it still sounded perfectly audible. On the other hand, the vocoded signal is very noisy and much less intelligible when compared with the original speech signal, which makes it very difficult for the listener to interpret the phrase. This result illustrates one of the main challenges for cochlear implants.