Dept. of Electronics and Electrical Communications Engineering, Indian Institute of Technology, Kharagpur

Neuronal Coding of Sensory Information (EC60004)

NCSI (Spring 2025) Project I

Prasit Mazumder
22EC10057

Part A:

Question 1: Responses to Tones :

Semilogarithmic tuning curves with logarithmic frequency are shown in Figures 1 and 2. Each ANF fires more when the stimulus matches its Best Frequency. Both curves peak at 400 Hz and 5 kHz.

Both tuning curves broaden from -10 dB SPL to 80 dB SPL, indicating responses to more frequencies at higher intensities. Low intensities make fibers more selective, acting around their optimal frequency.

The rate versus intensity charts for both fibers are given in Figure 3. In sigmoid curves, firing rates are low at low intensities but climb rapidly with intensity and reach saturation. ANFs' dynamic range is when intensity reaches a threshold and stabilizes firing rate.

Figure 1:

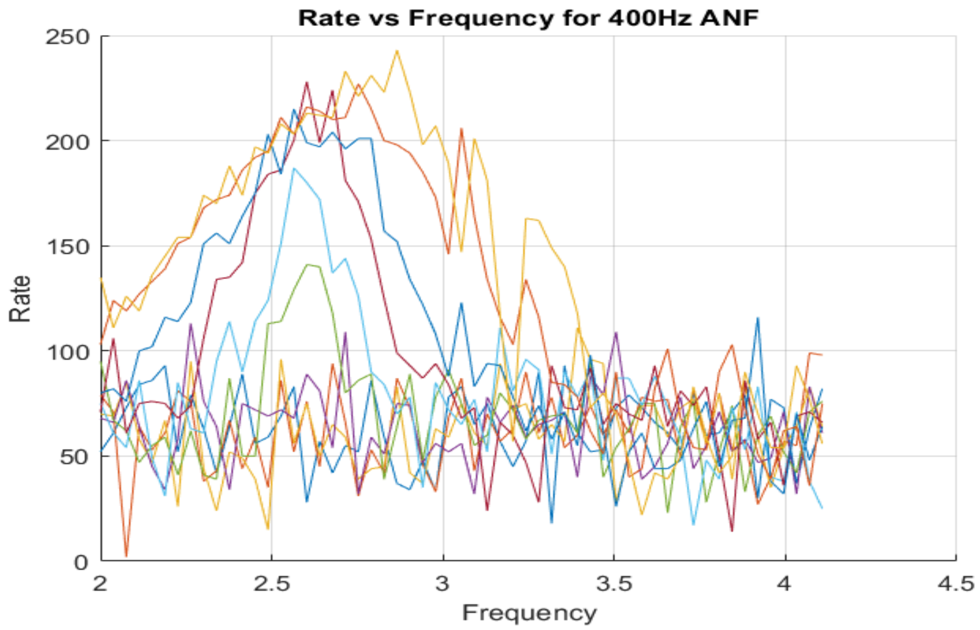


Figure 2:

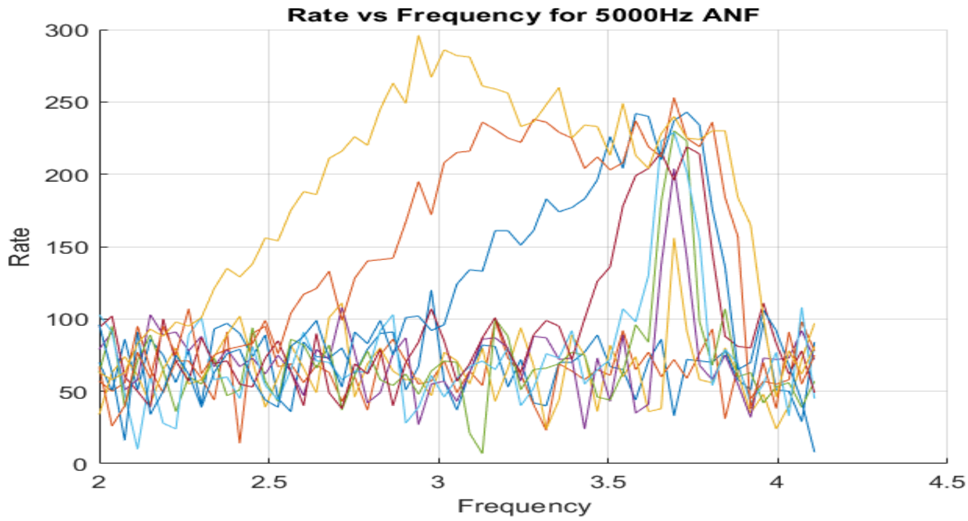
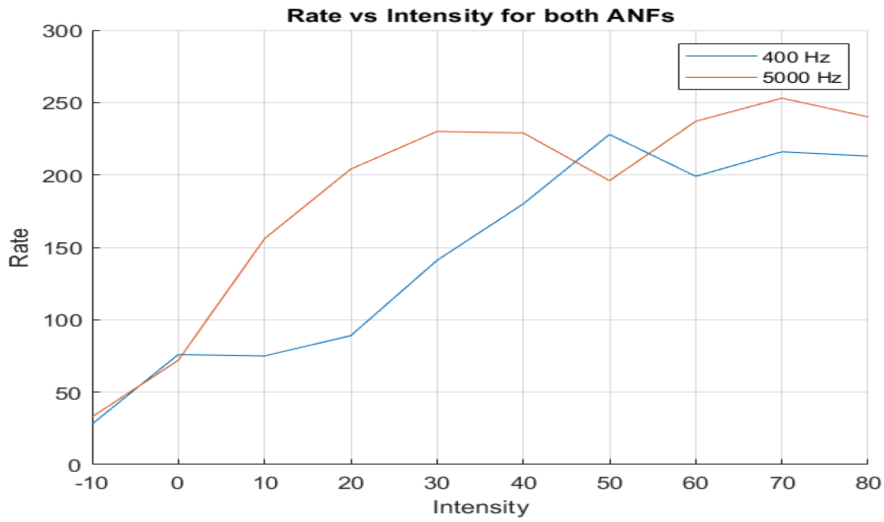


Figure 3:



Question 2: Responses to Speech :

The rate versus intensity function of a 550 Hz auditory nerve fiber creates three sound levels in response to "aaah". Figure 4 shows the method distinguishing firing rate intensities. Levels of 0 dB, 35 dB, and 70 dB indicate threshold, mid-range, and near-saturation.

After then, 25 auditory nerve fibers (ANFs) form spike trains with best frequencies (BFs) from 100 Hz to 6.4 kHz for each of the three sound levels. The stimulus is repeated 80 times to construct spike trains for accurate rate predictions.

Voice signal spectrograms show energy distribution across time in Figure 5. Spike averaging across window widths from 3.2 to 102.4 ms produces comparison ANF rate representations. Firing rate is colored, and time is on the x-axis in Figure 6A-F.

Figure 4:

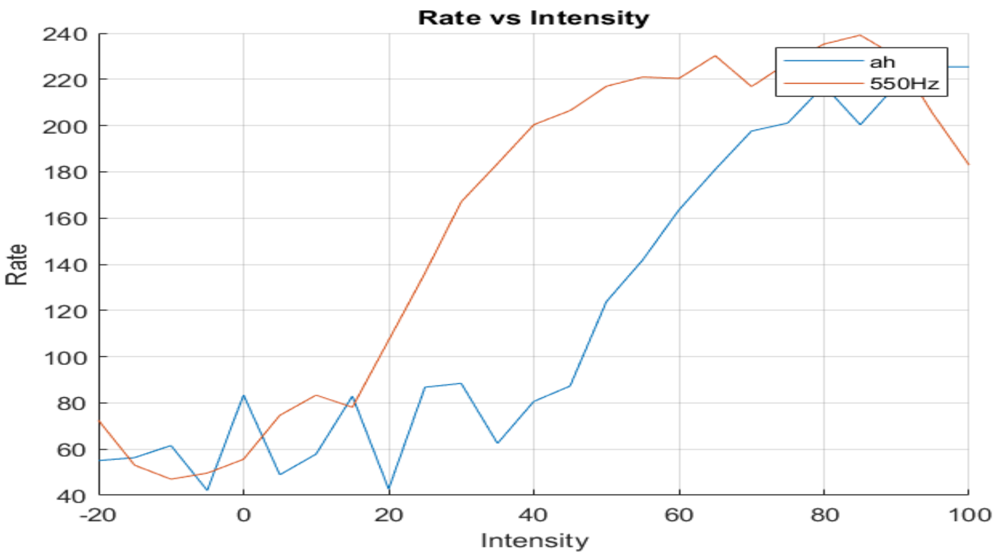


Figure 5:

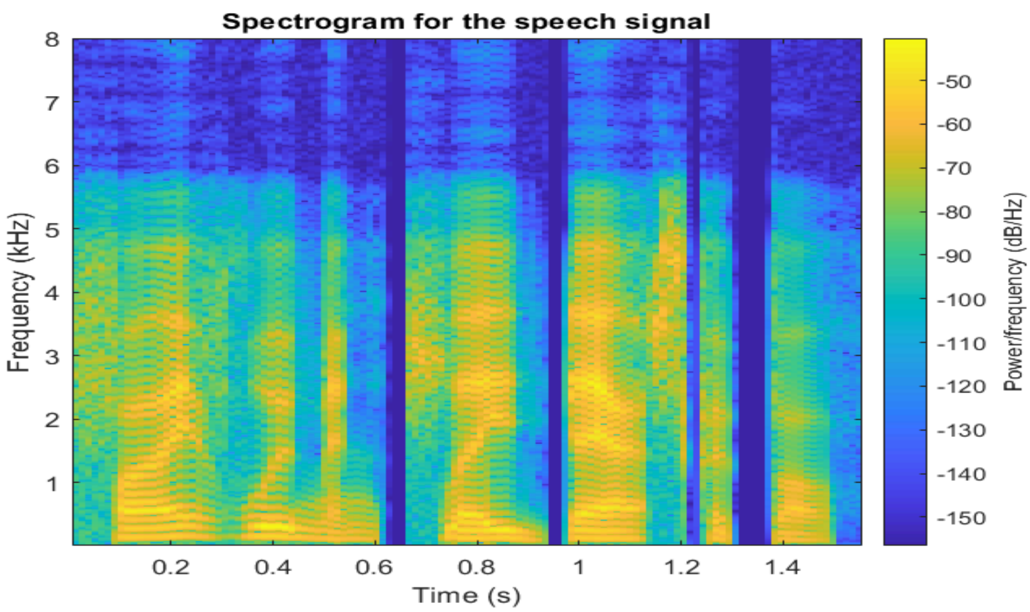
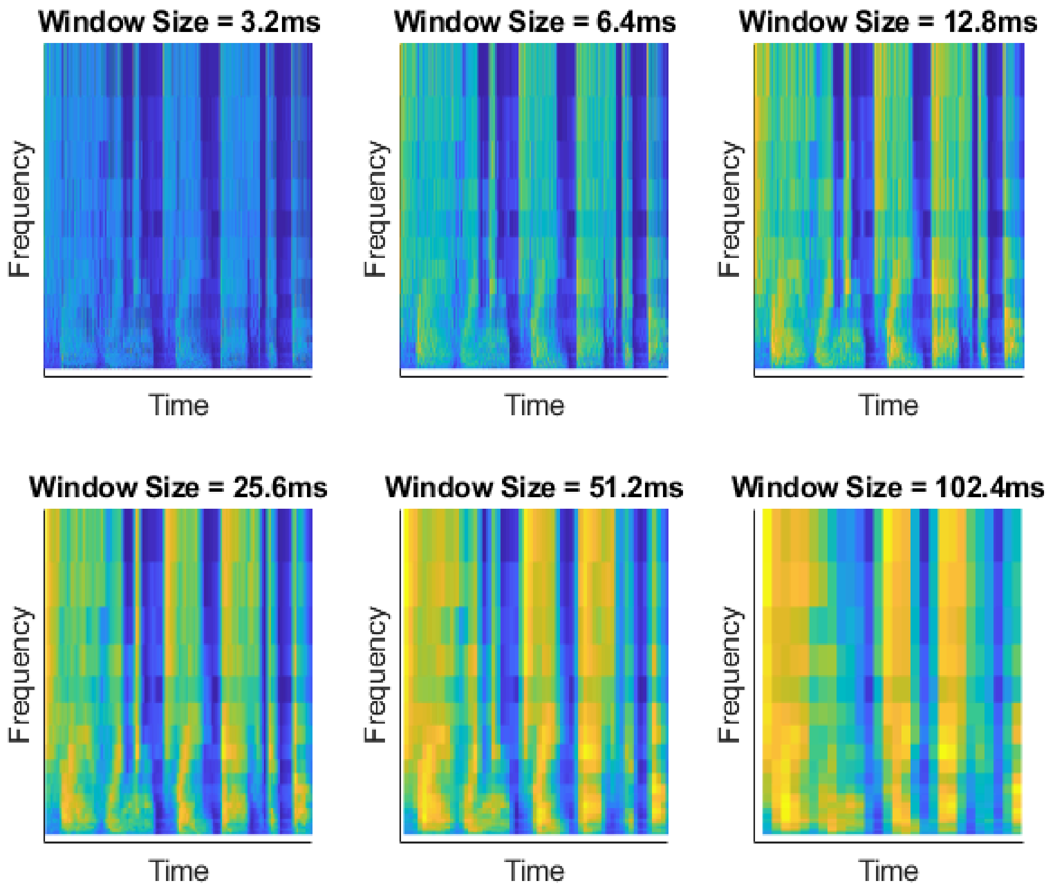


Figure 6:



Question 3: Responses to Speech (Fine Timescale Representation) :

A 12.8 ms window with 50% overlap is applied to the PSTHs derived from 80 repetitions of the stimulus for each ANF. An FFT is calculated for each window after the subtraction of the mean to eliminate the DC component. The dominant frequency is discovered by locating the greatest FFT magnitude and its matching frequency. To guarantee that only substantial phase-locking frequencies are included, a requirement is implemented: the peak FFT magnitude must be a minimum of three (a threshold value) times the mean FFT magnitude of that window; otherwise, no dominating frequency is designated (noted as NaN).

The predominant frequencies are shown as asterisks on the speech spectrogram in Figures 7 and 8 for certain ANFs. The dominating frequencies are often situated around each fiber's Best Frequency (BF) and correspond with areas of high energy in the spectrogram, suggesting that the ANFs are phase-locking to significant frequency components in the spoken signal.

Figure 7:

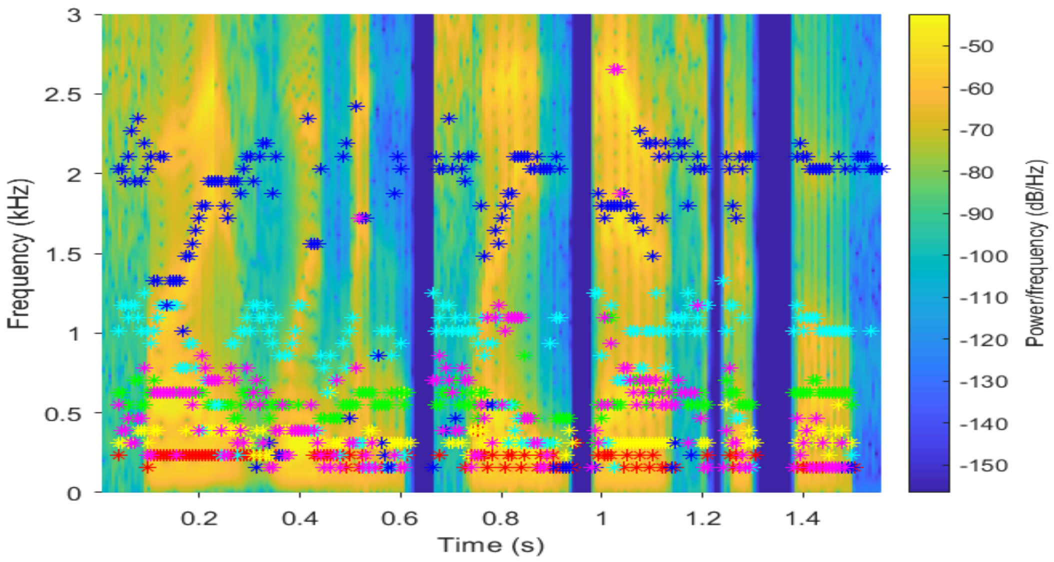
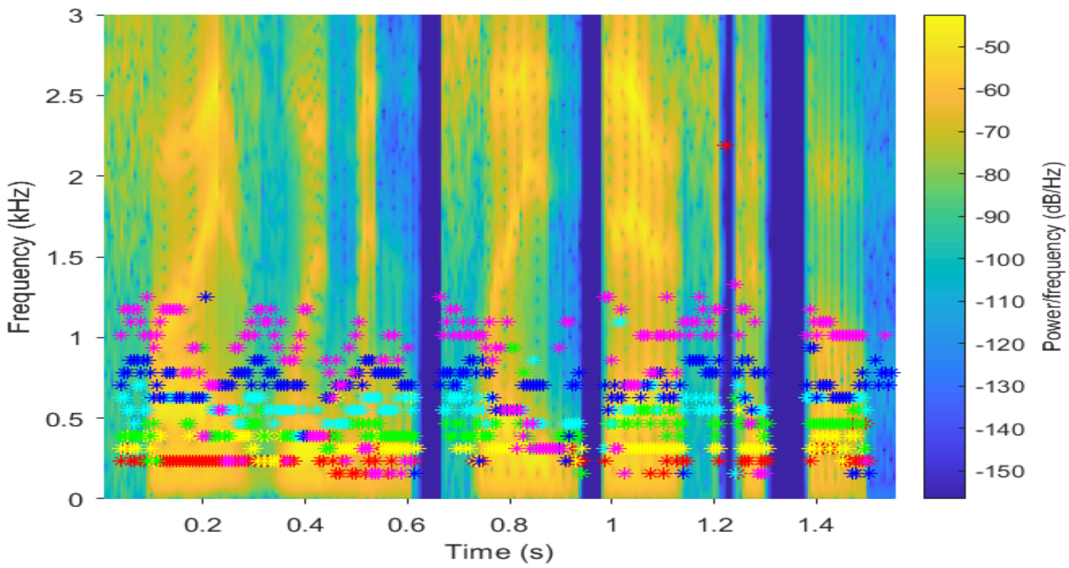


Figure 8:



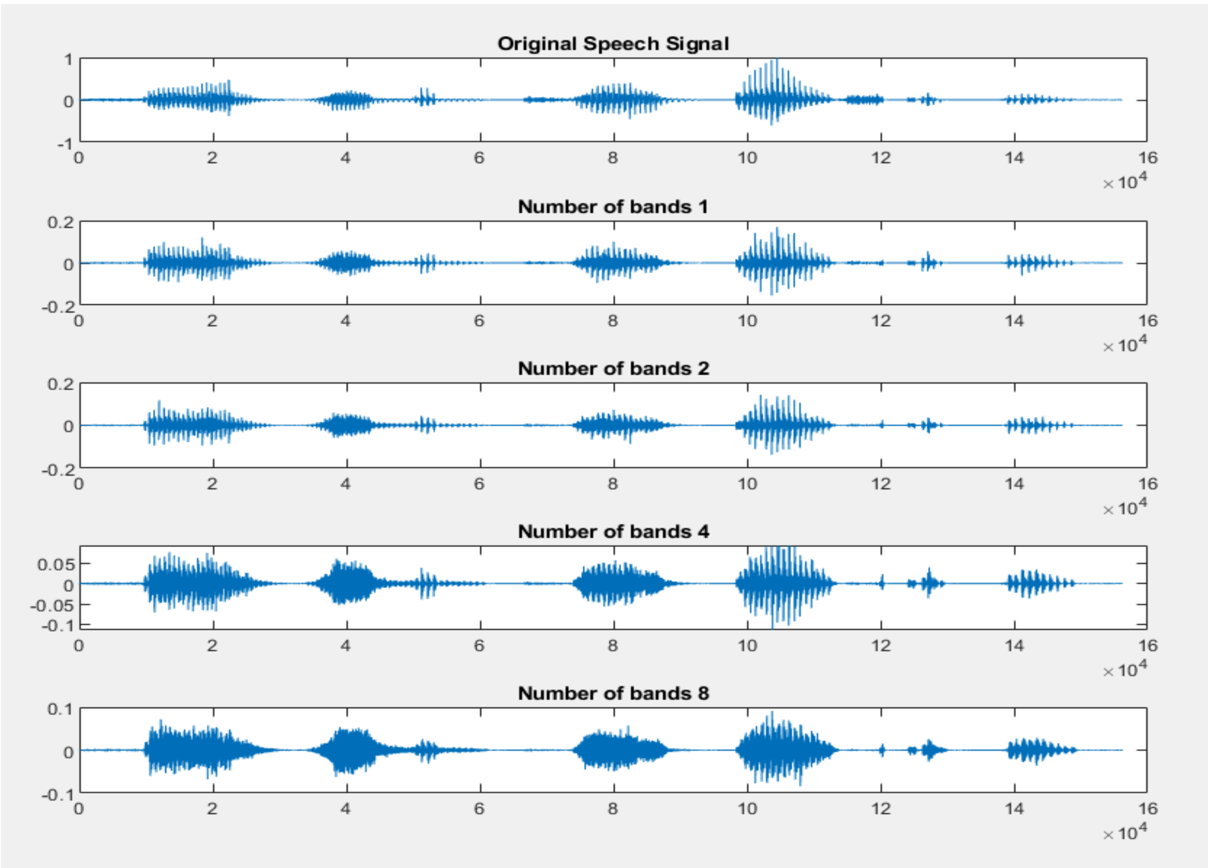
Part B:

In this section, we used a noise vocoder as described by Shannon et al. (1995), producing 1, 2, 4, and 8 band variations of a voice signal using Butterworth bandpass filters throughout the frequency range of 250 Hz to 2000 Hz. The Hilbert envelope of each band was used to modulate band-limited noise, and the resultant signals were presented to a listener unacquainted with the message.

The listener indicated that the 1-band signal was entirely incomprehensible, resembling arbitrary noise. With two bands, some rhythm was discernible; nonetheless, communication remained ambiguous. The 4-band version permitted little word inference but remained difficult to comprehend. The 8-band signal was clearly comprehensible, with the majority of words being recognized, aligning with the results of Shannon et al.

We further examined ANF responses to the 1-band and 8-band stimuli. The 1-band signal produced feeble and indistinct rate responses, but the 8-band signal exhibited pronounced, unique firing patterns among fibers attuned to various frequencies. This aligns with the perceptual observation that enhanced brain representation results in improved speech intelligibility. Consequently, temporal envelopes spanning frequency ranges are essential for speech comprehension.

<u>**Final Audio Spectrum Plot**</u>:



Figures for 1 Band Sound:

Figure 4:

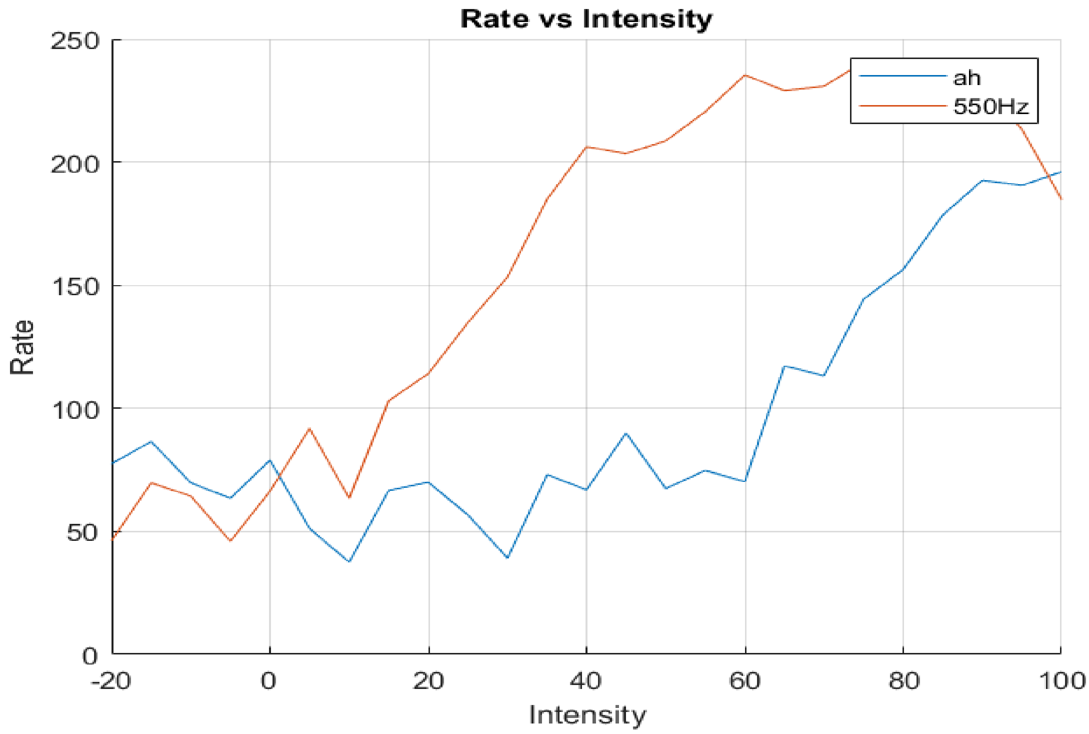


Figure 5:

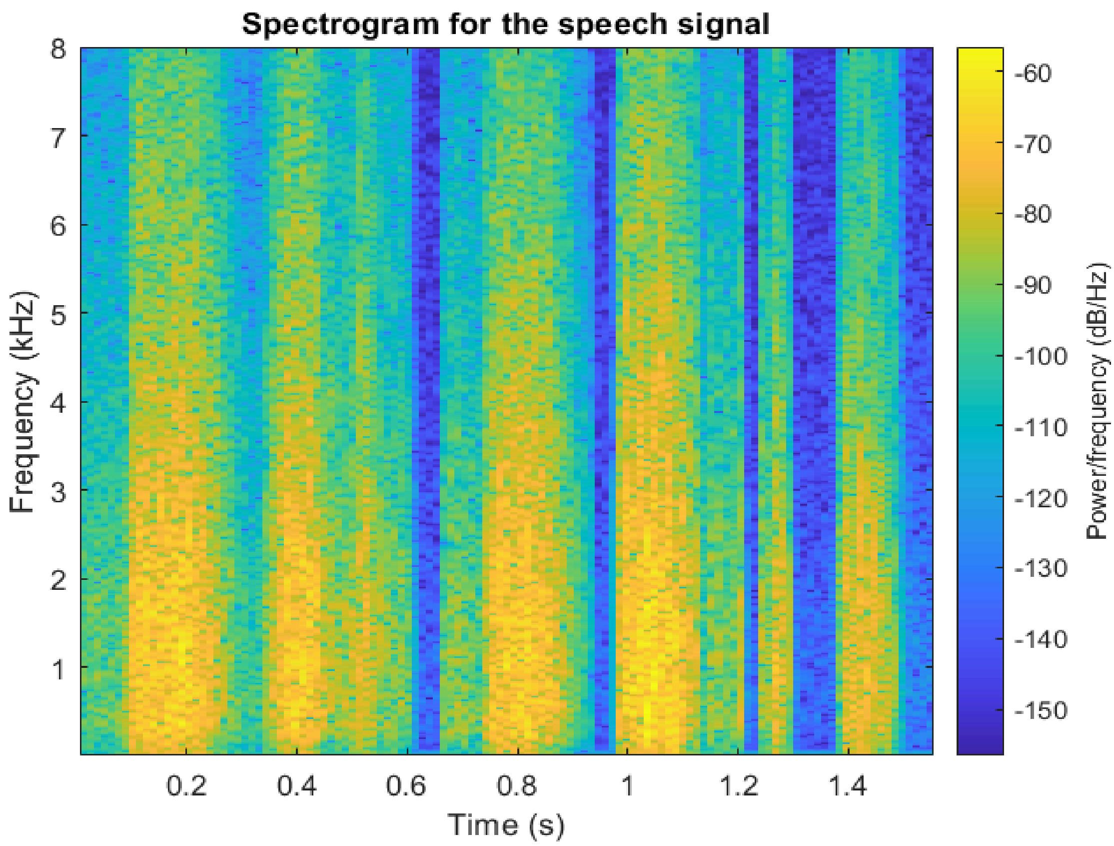


Figure 6:

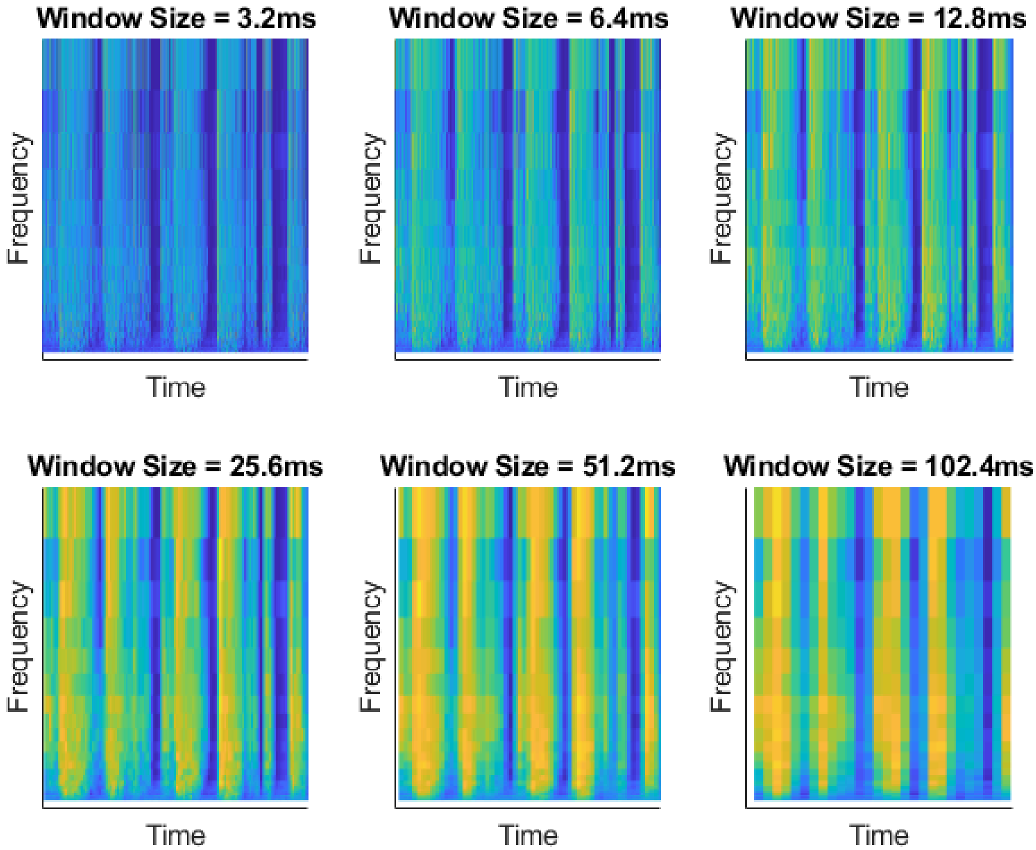


Figure 7:

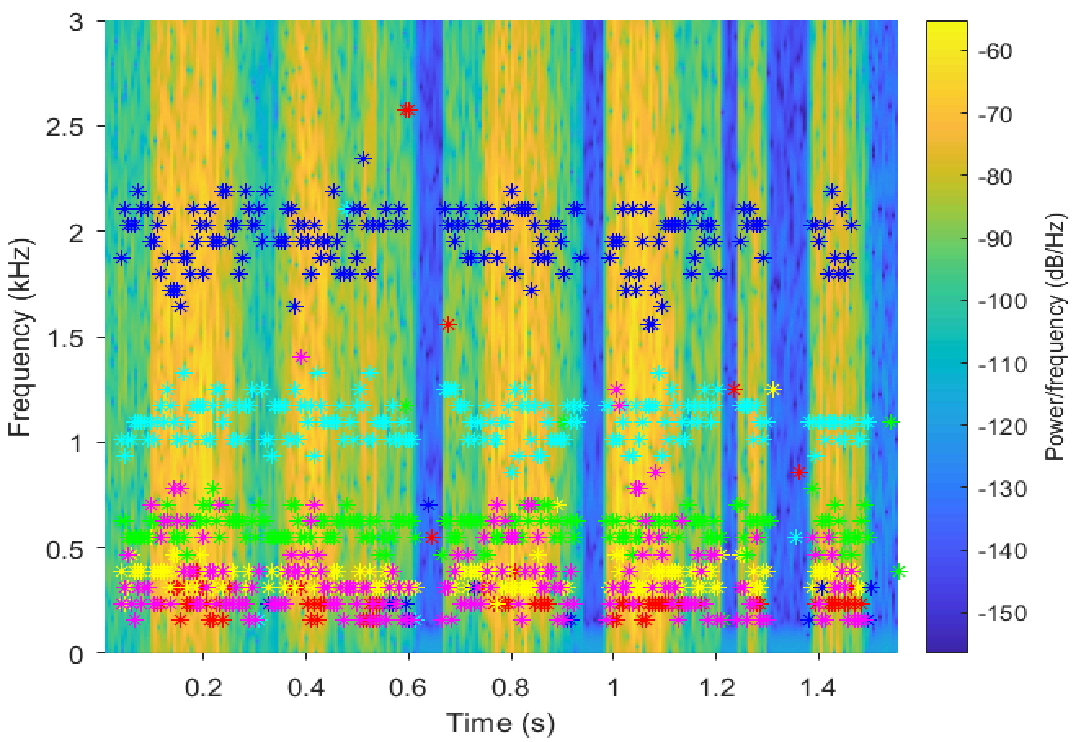
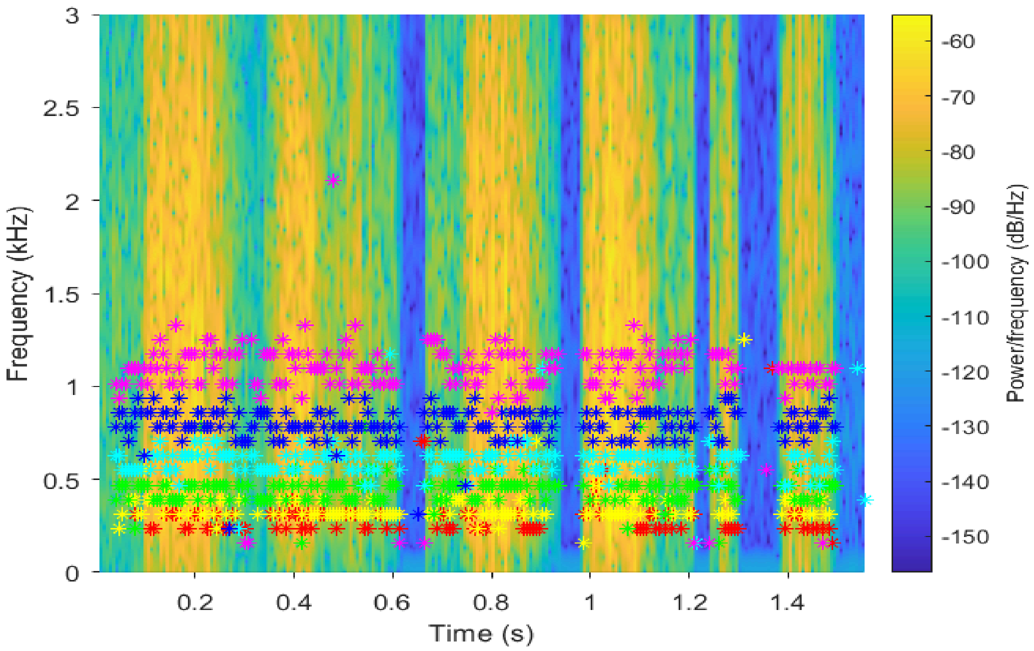


Figure 8:



Figures for 8 Band Sound:

Figure 4:

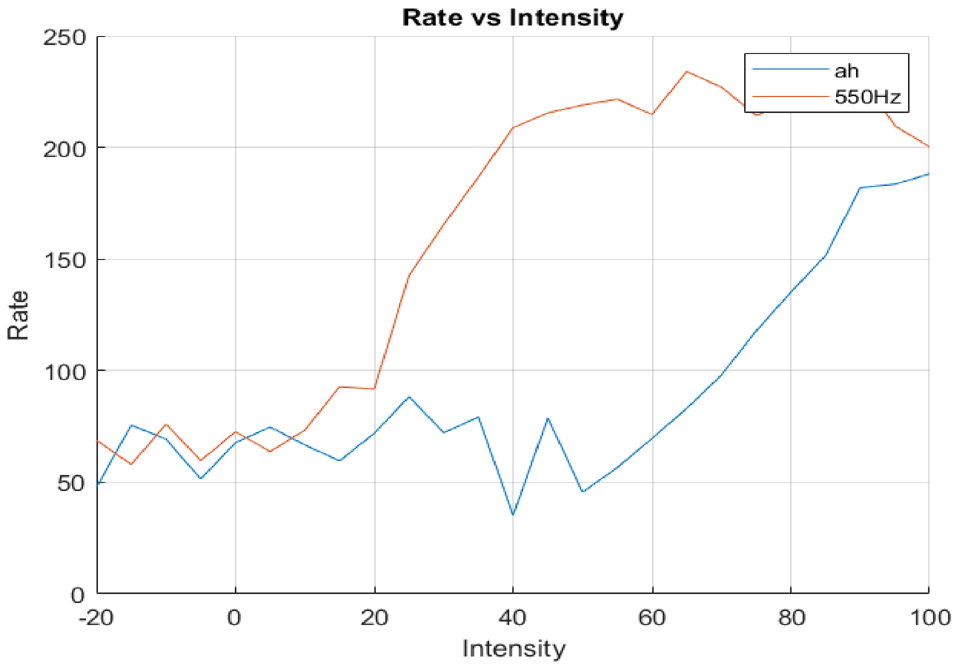


Figure 5:

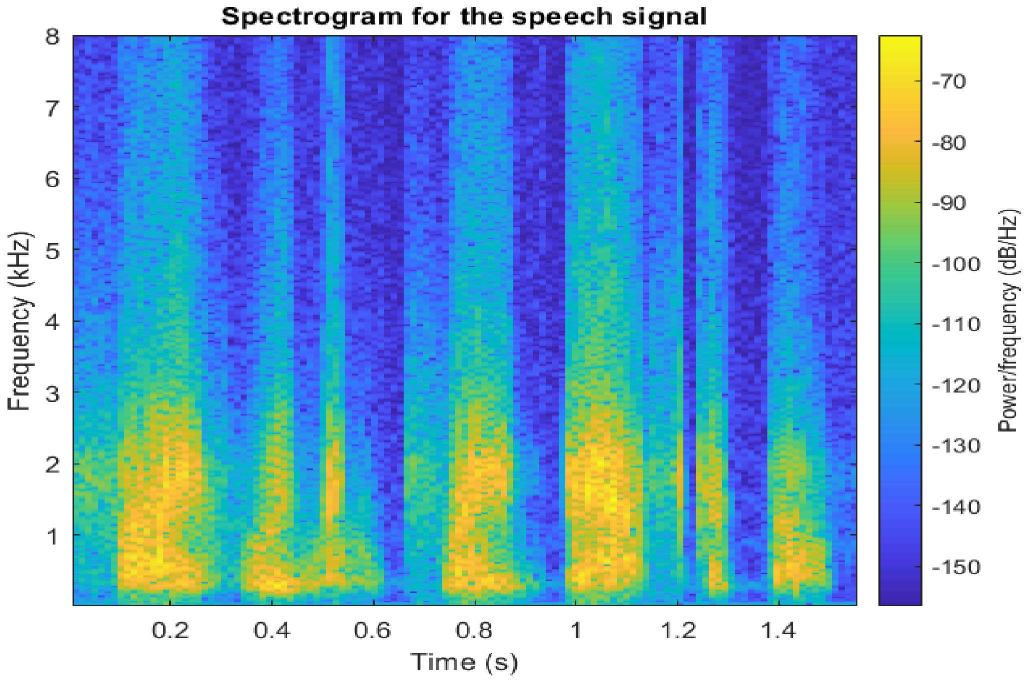


Figure 6:

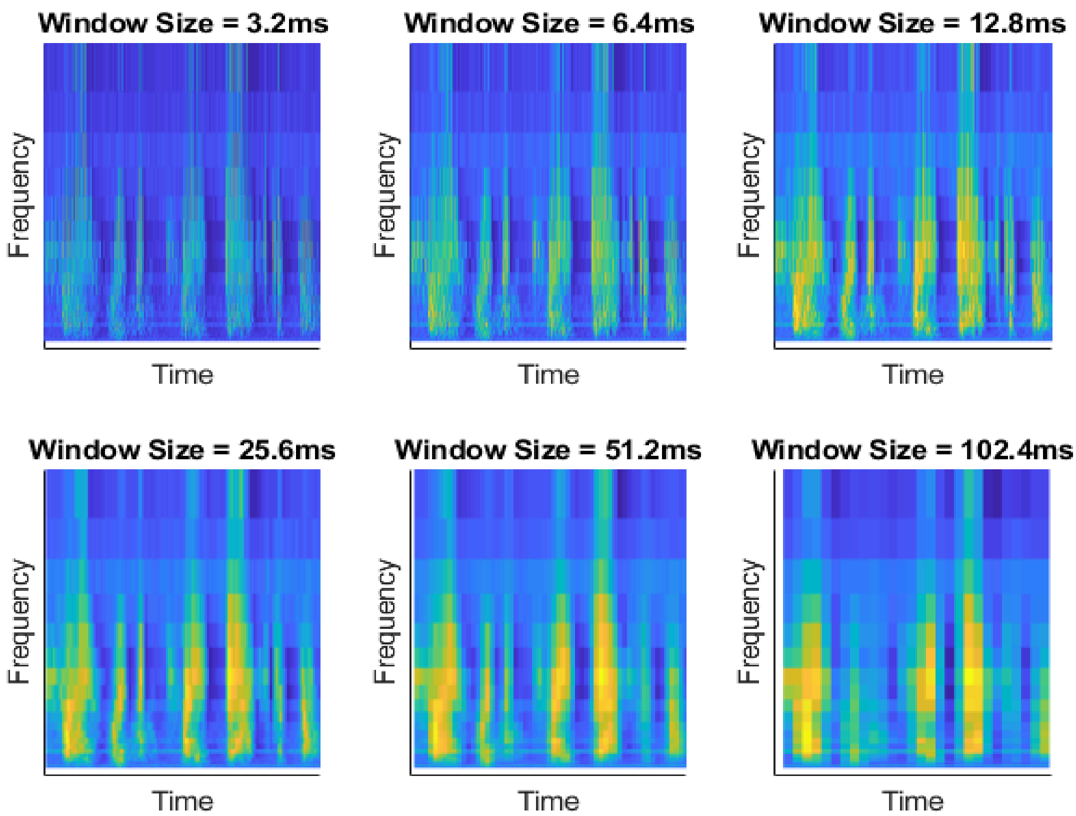


Figure 7:

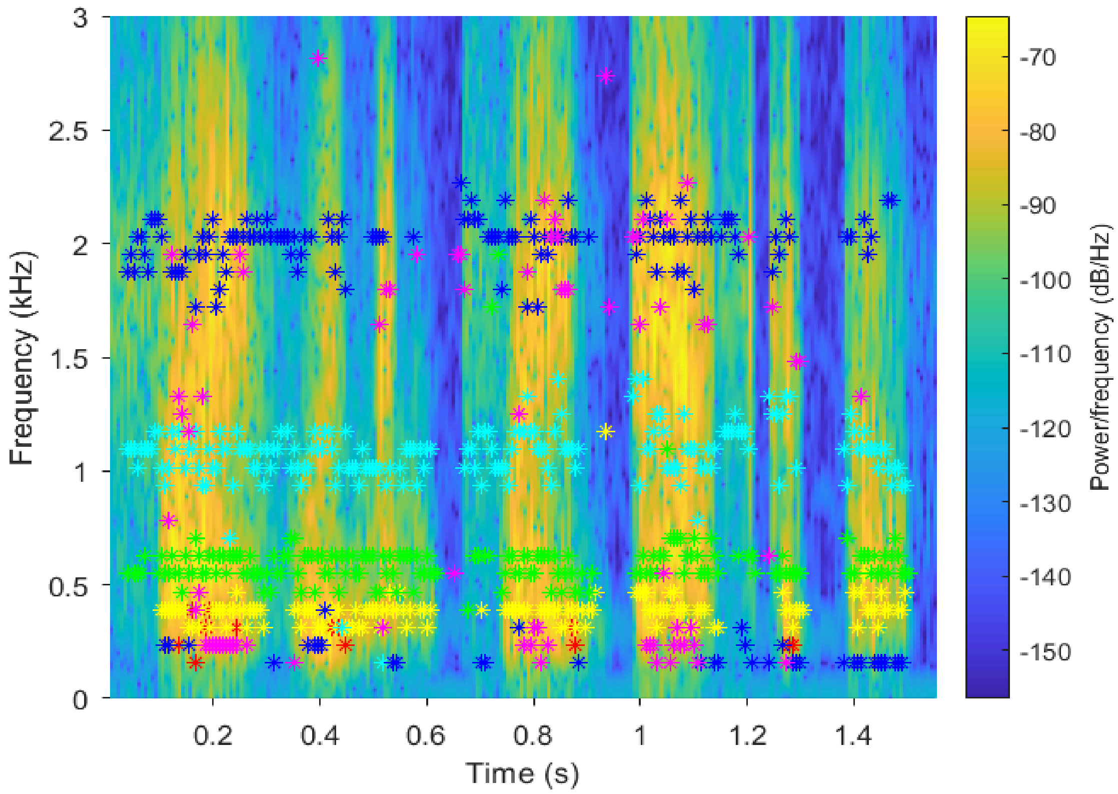


Figure 8: