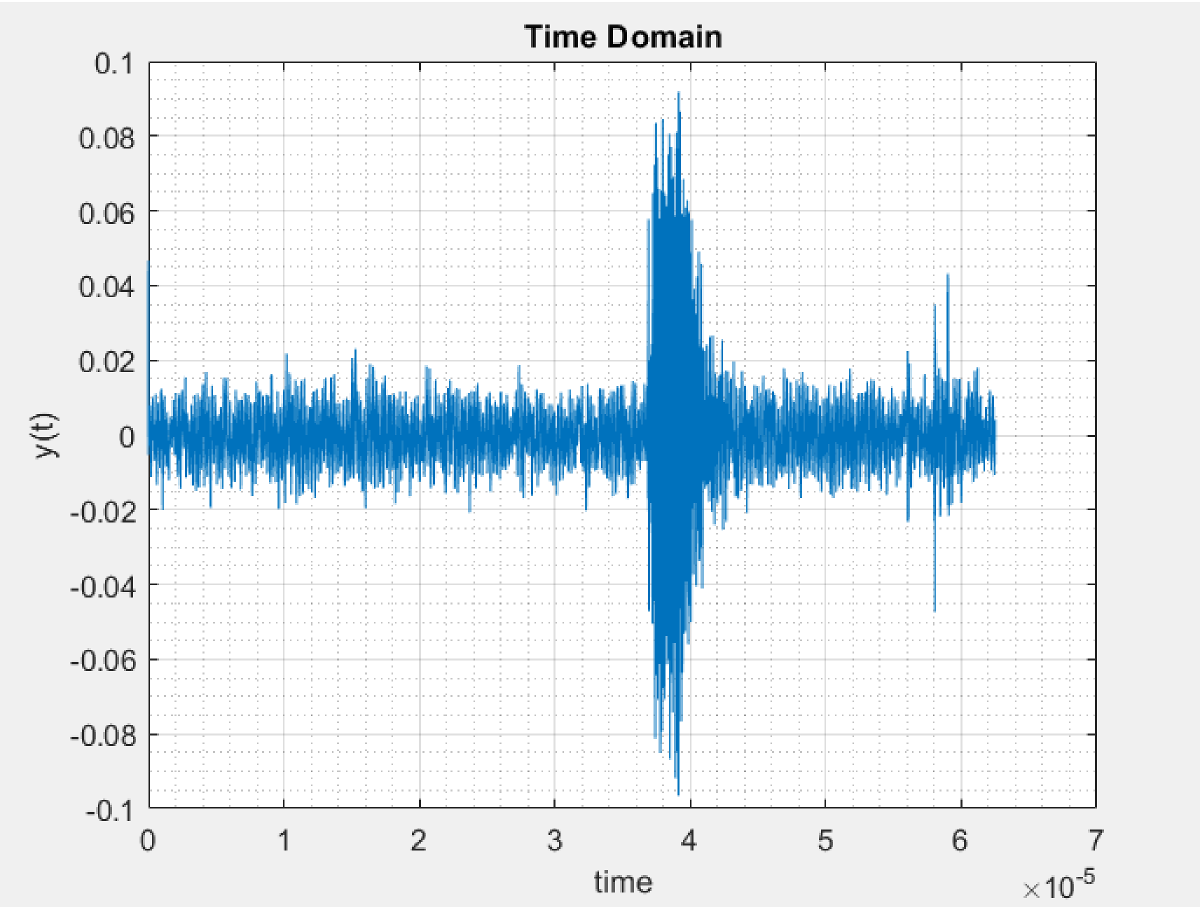
Assignment-3

Speech Signal Processing

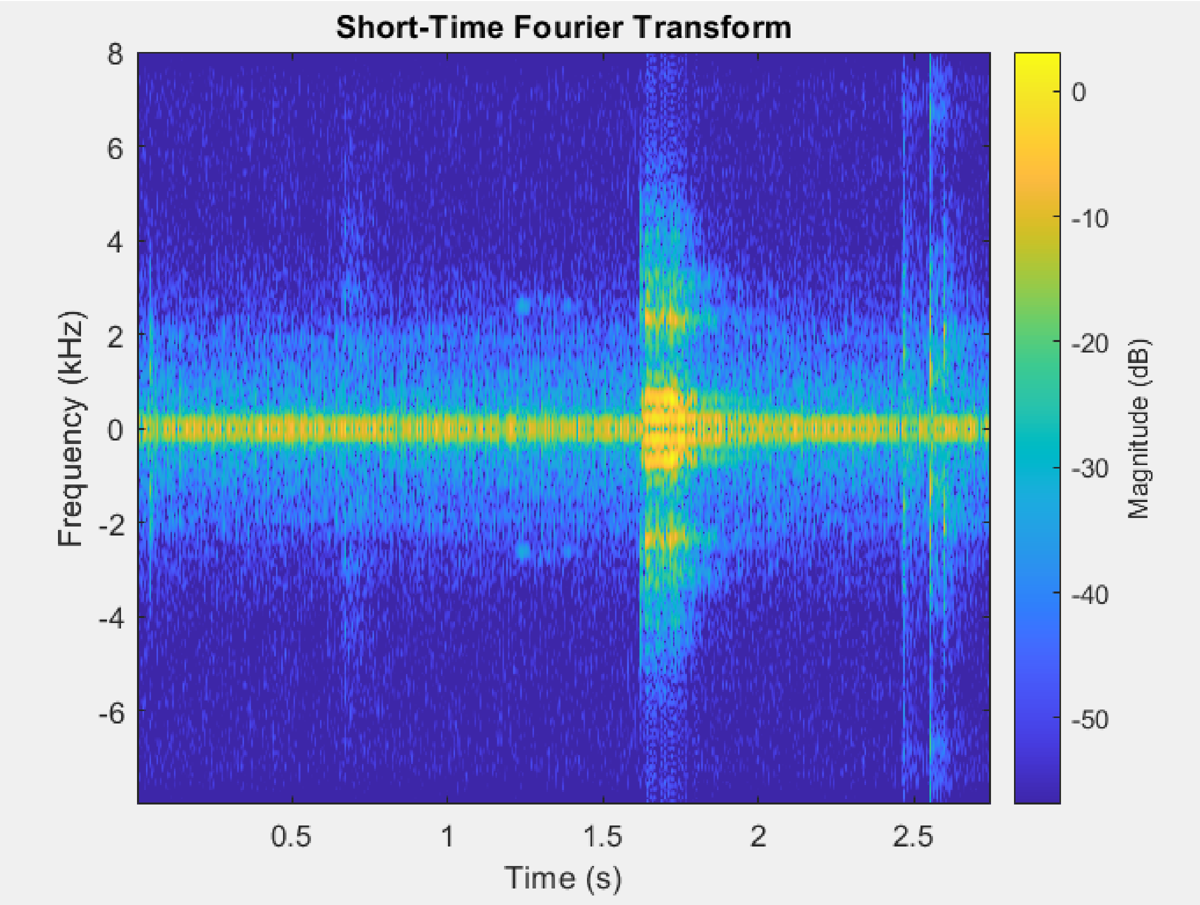
Anjali Singh, 2020102004

$Q1.$ *Record a vowel utterance of your choice and perform the following operations:*

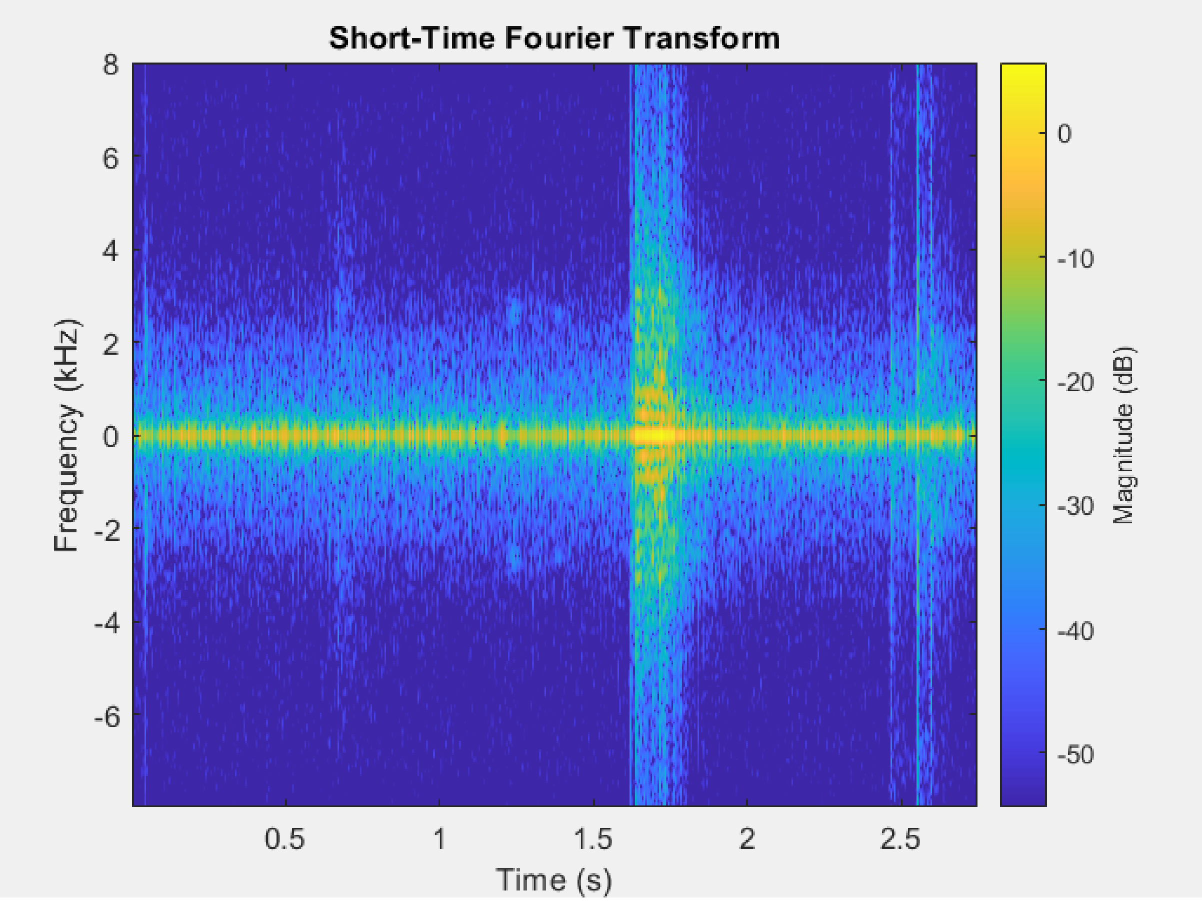
- The vowel used is: /A



1. *Compute the short term autocorrelation function for a hamming window of N=512 and plot.*



2. *Compute the N=512 point magnitude spectrum of the waveform based on hamming window and STFT and plot magnitude spectrum*

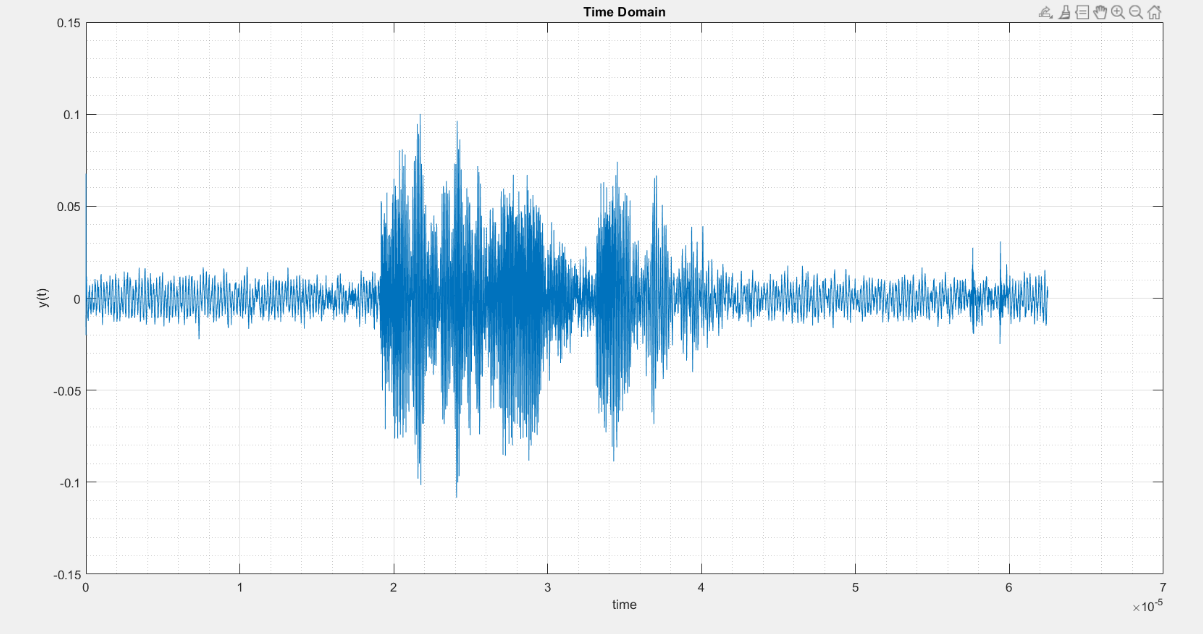


3. *Comment the changes in both the autocorrelation and the spectrum. What do these changes indicate about the effects of the clipping operations on the waveform?*

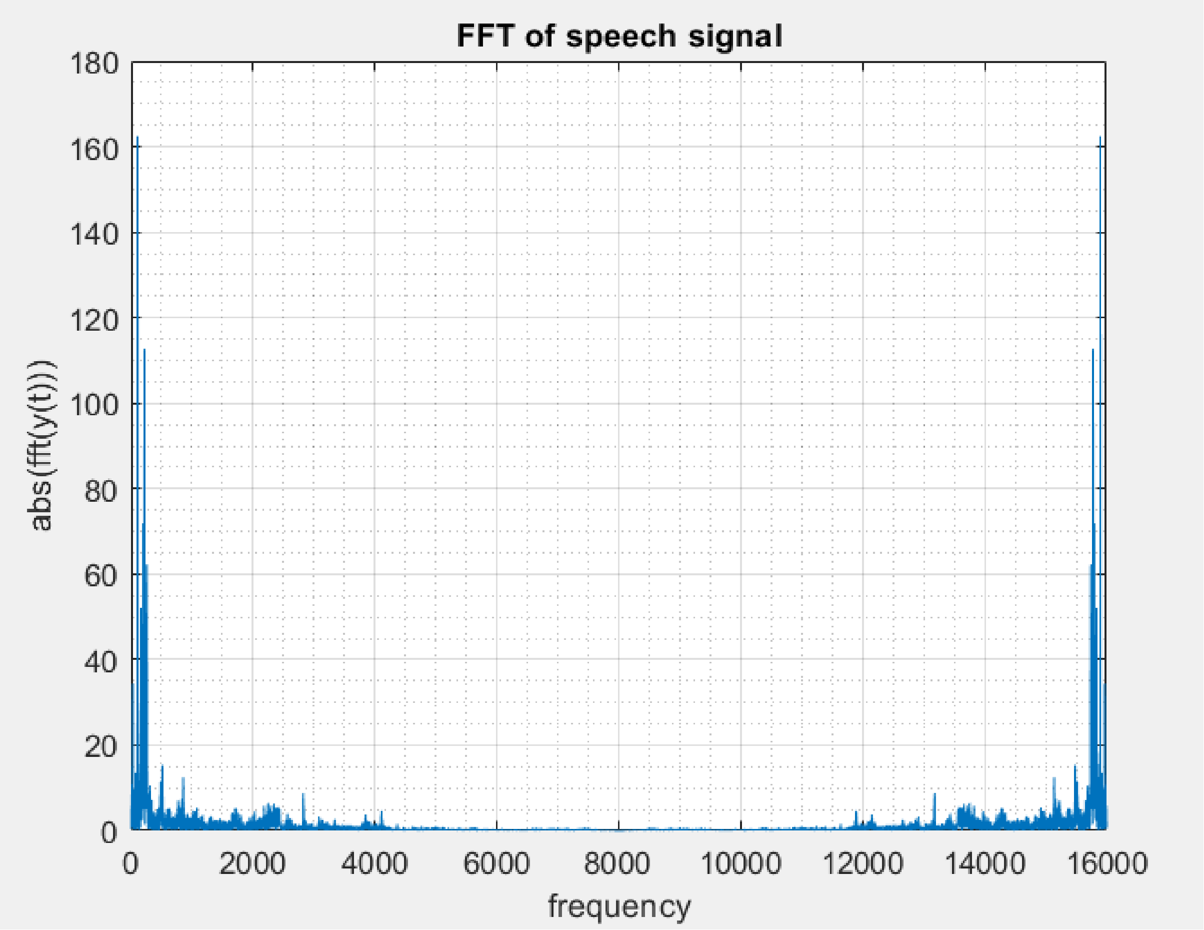
    - From the above two, we can observe that the width of the plot has reduced when magnitude of the audio signal is taken while plotting the STFTs.
    - The intensities of frequency measurement is more in case of plotting the STFT of the audio signal than that of the magnitude of the signal.

$Q2.$ *Record your name and Compute MFCC, and take first 13 coefficients of each frame and plot it. Comment on the plots.*

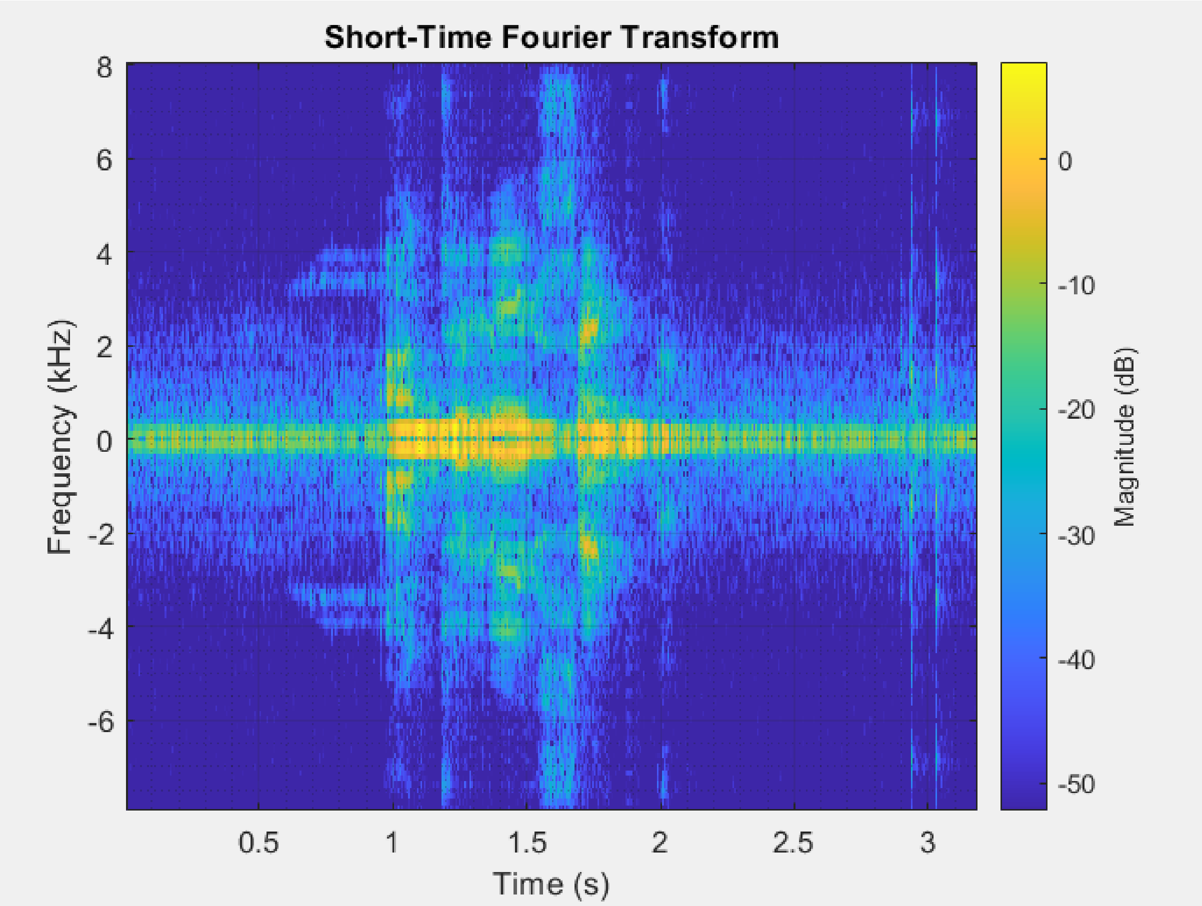
- Time domain of speech signal



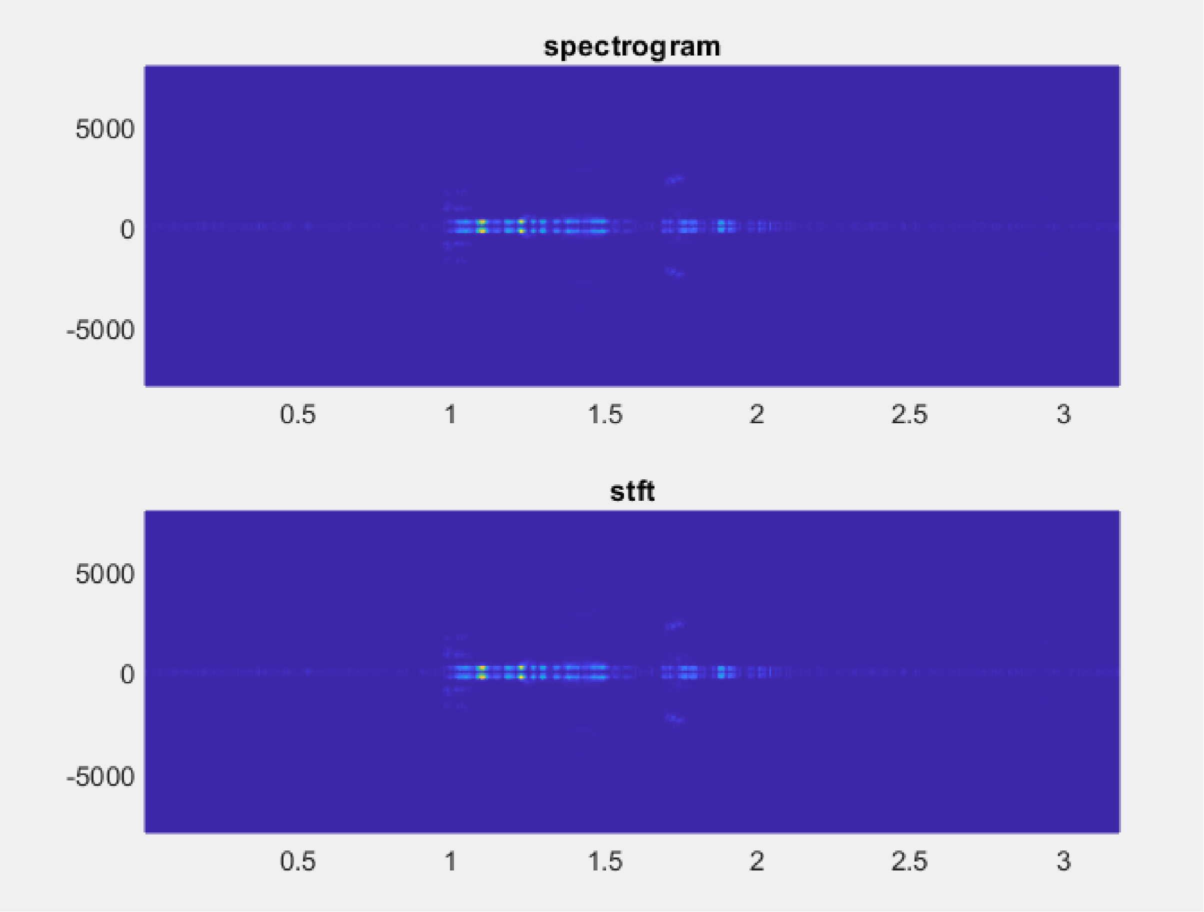
- Fast Fourier transform of speech signal, taken absolute to get rid of imaginary parts



- Finding the short-time Fourier Transform of the signal, y.



- Comparing the STFT and Spectrogram of the signal function

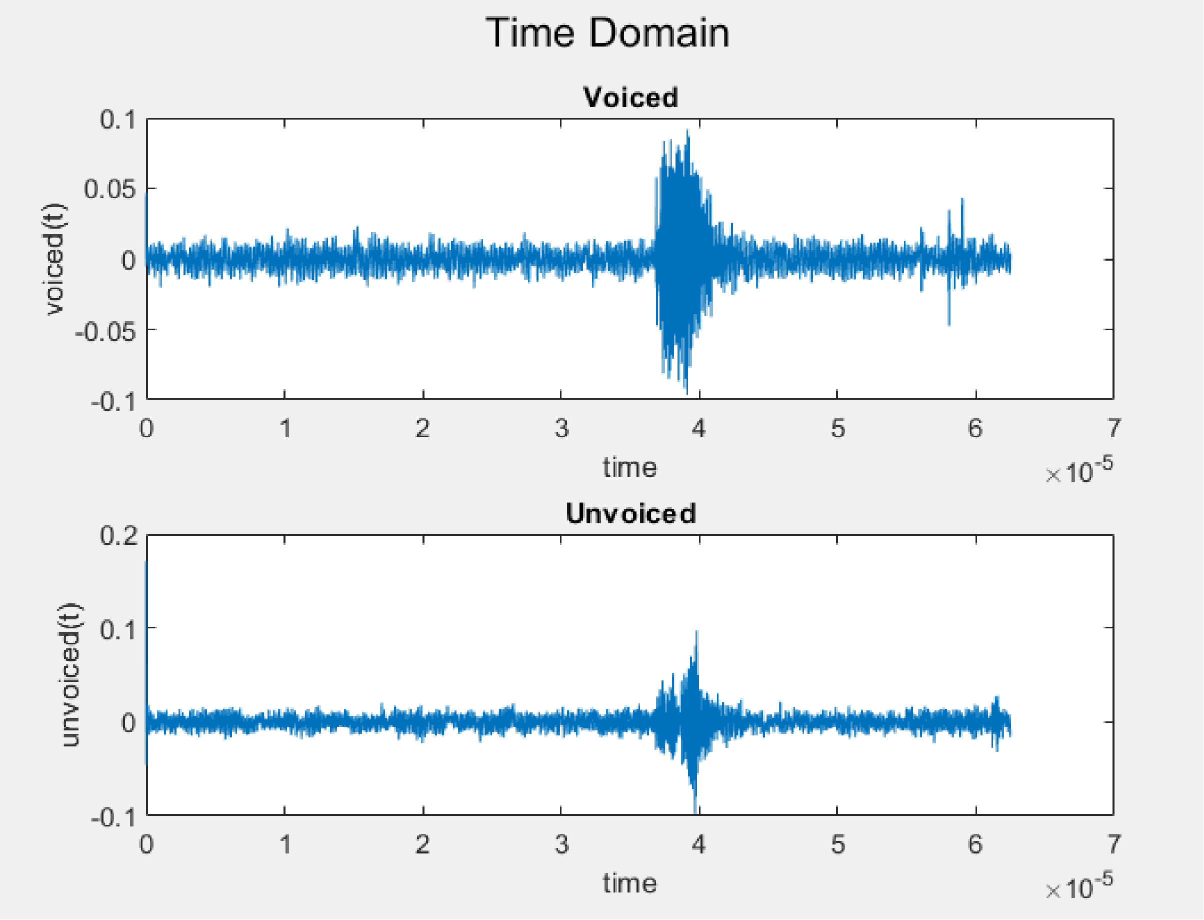

spectrogram
5000
0
-5000
0.5
1
1.5
2
2.5
3
stft
5000
0
-5000
0.5
1
1.5
2
2.5
3

- For MFCC

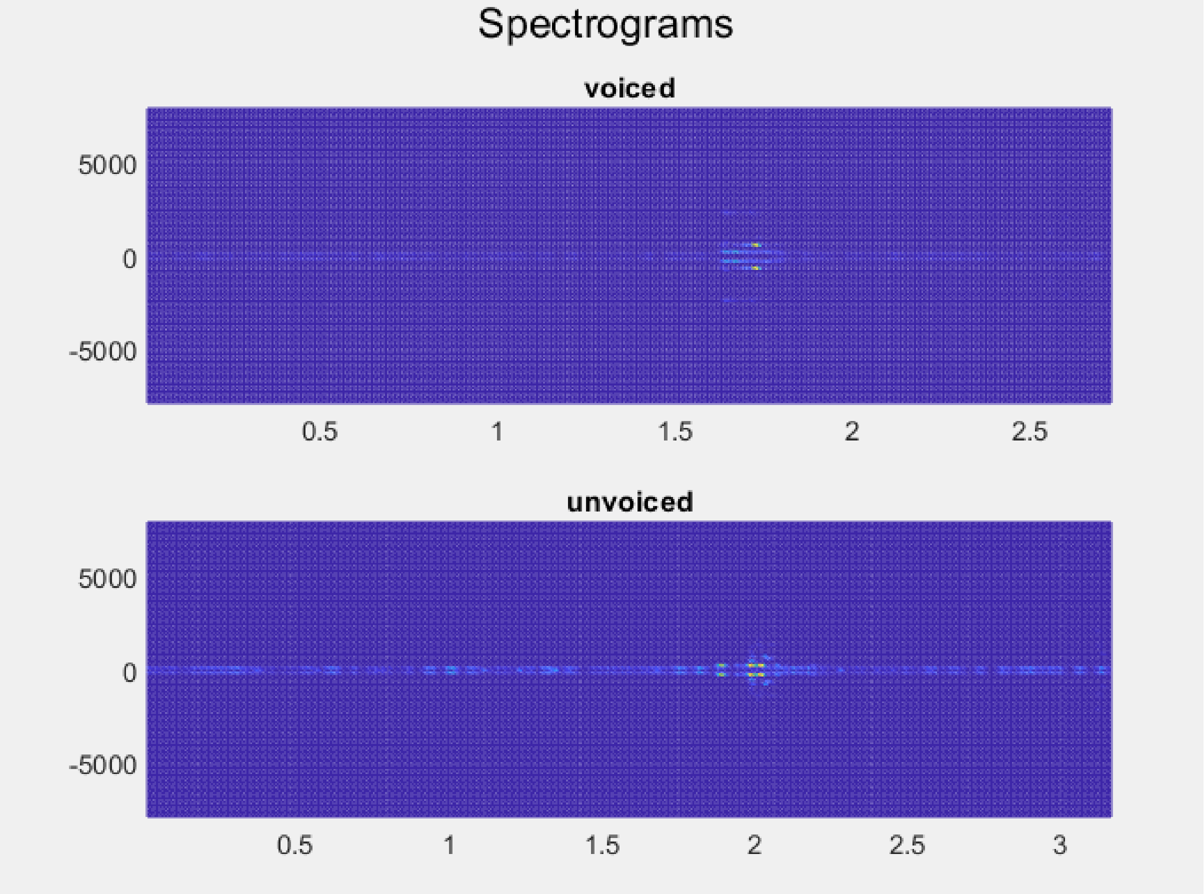
```
[y, fs] = audioread('my_name.wav');
Ts = 1/fs;

n = length(y)-1;
t = 0:Ts/n:Ts;
f = 0:fs/n:fs;

% sound(y, fs);
figure;
plot(t, y);
xlabel('time');  ylabel('y(t)');
title('Time Domain');
grid on
grid minor


y_coeff = mfcc(y, fs);

n_coeff = size(y_coeff, 1);
coeff = y_coeff(:, 1: 13);
for i = 1:10:val
    figure;
    plot(coeff(i, 1:13));
    xlabel('MFCC Coeff');
    ylabel('Value of MFCC coeff');
end
```

$Q3.$ *Consider a voiced and unvoiced sound and plot spectrogram of it. Comment on the formant structures.*

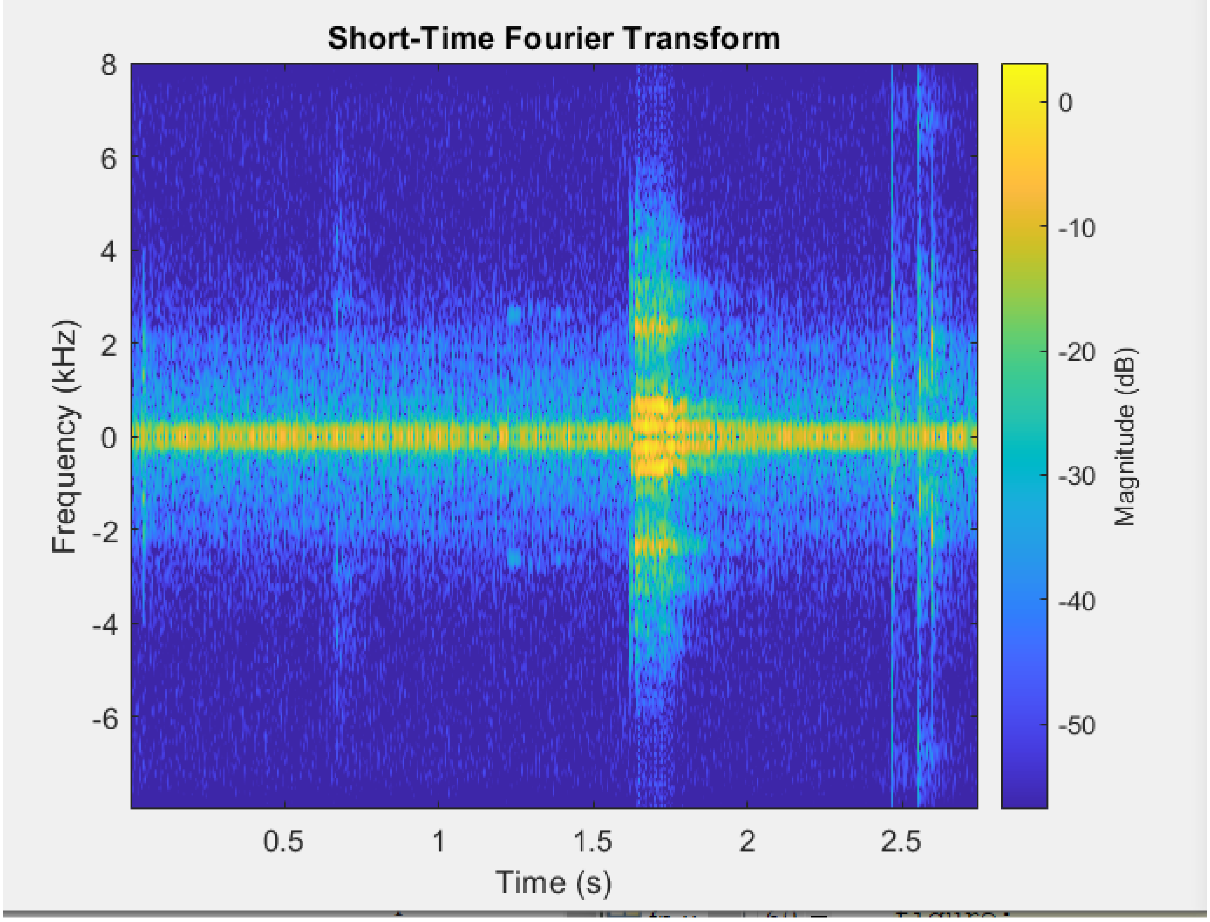
- Here, for voiced speech, I'm using the vowel "/A" and for unvoiced speech, I'm using "/Ba".
- Time Domain plots of voiced and unvoiced signals



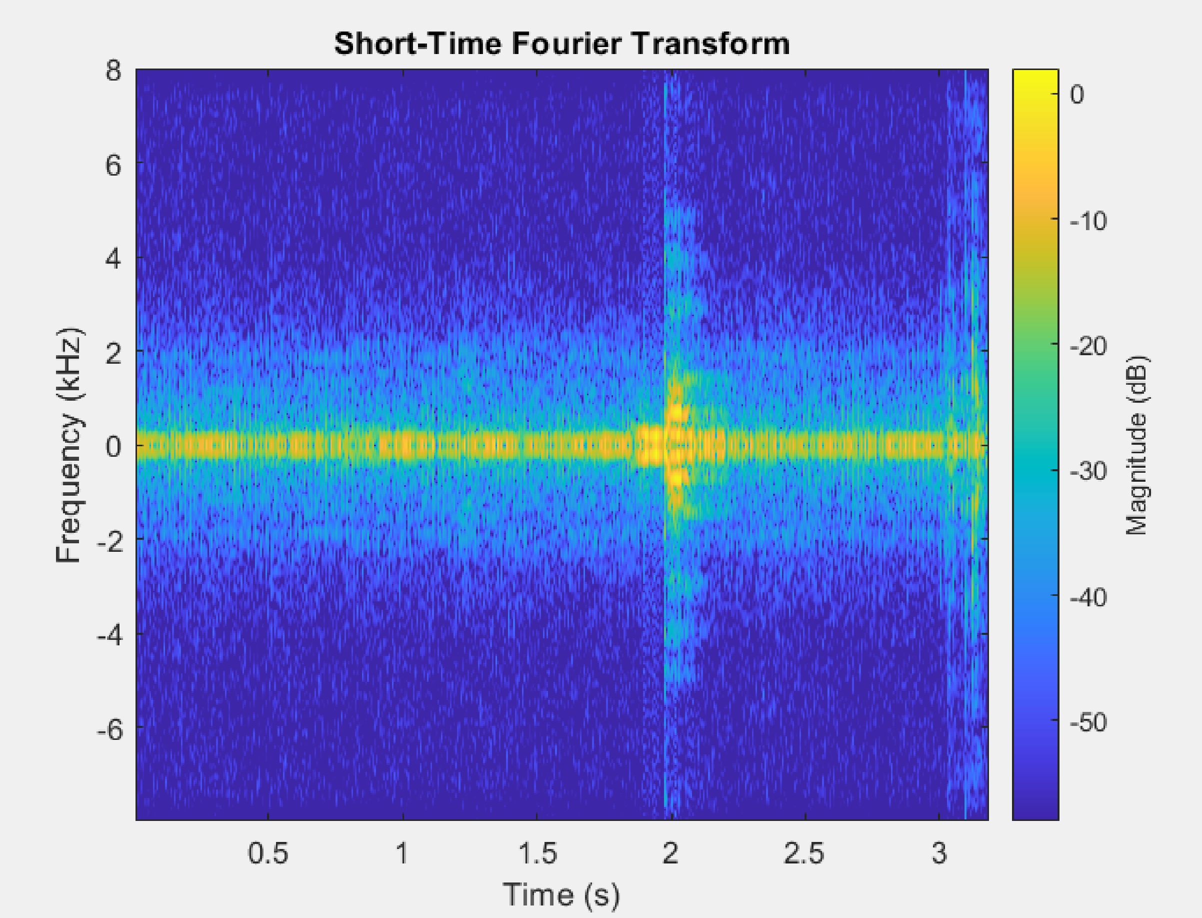
- Spectrogram plots of both unvoiced and voiced signals



- STFT for Voiced signal



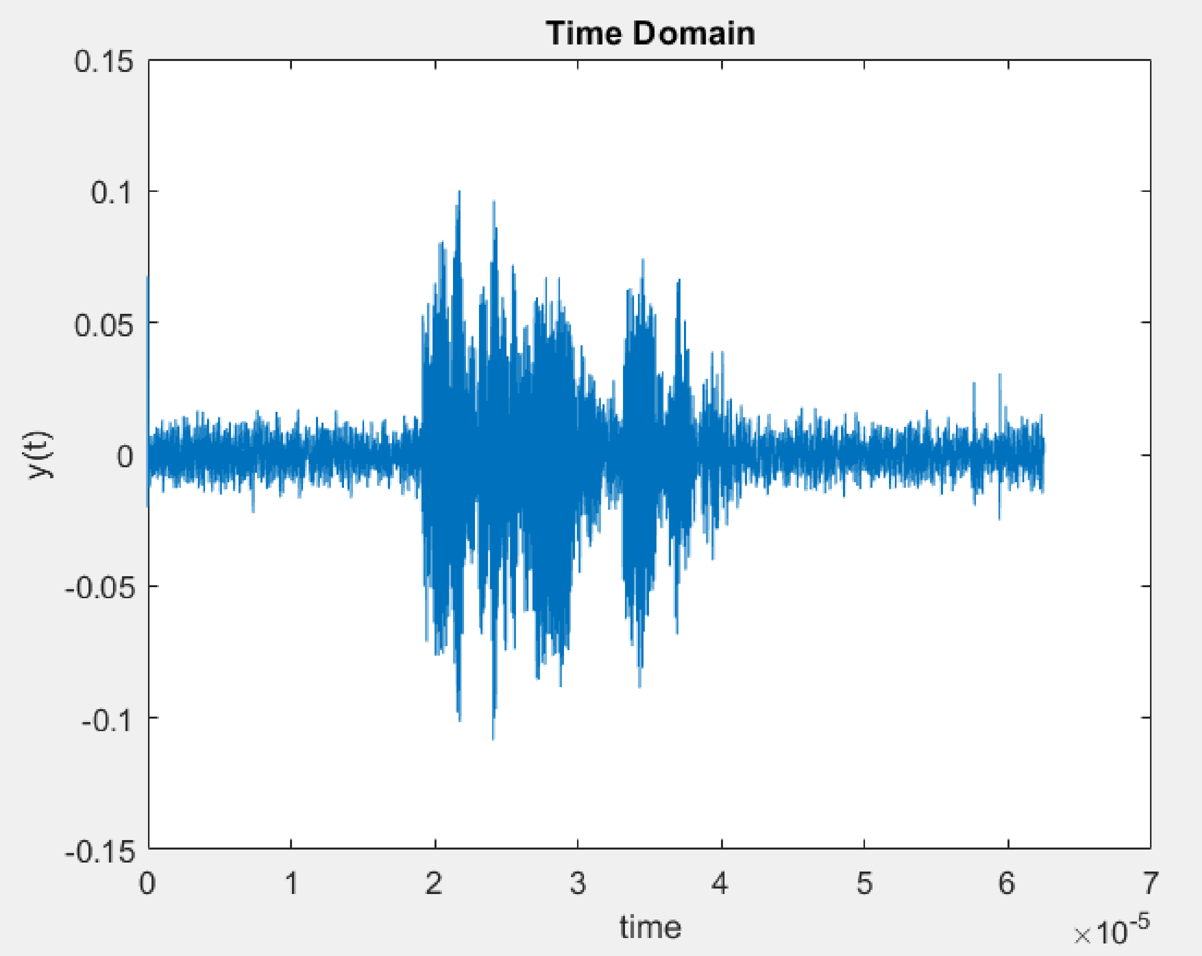
- STFT for Unvoiced signal



- The formant structure for voiced signal has higher amplitude frequency than the formant structure for unvoiced signal, with the width of the visible higher amplitude for voiced signal is more than that of unvoiced signal.

$Q4.$ *Record a voice and enhance the speech using spectral subtraction method from voice-box library and compute the SNR before and after the speech enhancement.*

- Audio Signal



- After spectral subtraction [I don't have voice box in MATLAB, so the below code isn't working]

```
[y, fs] = audioread('my_name.wav');
Ts = 1/fs;
n = length(y)-1;
t = 0:Ts/n:Ts;
f = 0:fs/n:fs;

figure;
plot(t, y);
xlabel('time'); ylabel('y(t)');
title('Time Domain');

snr_before = snr(y);

[s_new, g, tt, f_new, z] = v_specsub(y, fs);
snr_after = snr(s_new);

n_new = length(s_new)-1;
t_new = 0:Ts/n_new:Ts;

figure;
plot(t_new, s_new);
xlabel('time'); ylabel('s_new(t)');
title('Audio Signal after Spectral Subtraction');
```