EE-414 Speech Processing Lab

Lab-5-Part-2 - 180020002

Aim:

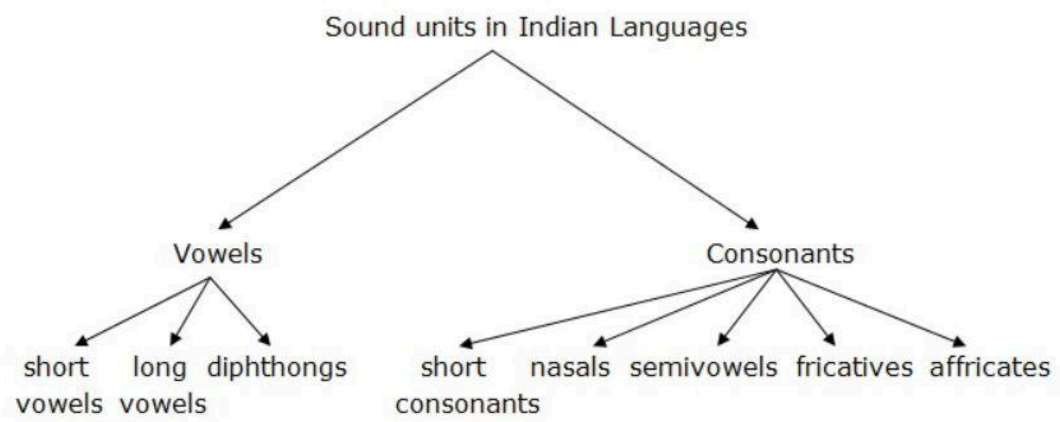
● To study different sound units present in majority of Indian languages.

● To understand the production mechanism of each sound unit.

● To learn the time domain and frequency domain characteristics of different sound units.

Theory:

Classification of Sound units in Indian Languages:



Vowels & Consonants:

The sound units of most languages in India are broadly classified into two categories, namely, vowels and consonants. These two broad categories are mainly based on the shape of the vocal tract. In case of vowels, the vocal tract shape is wide open without any constriction along its length starting from the glottis till the lips and is excited by voiced excitation source. Alternatively, in case of consonants, their may be constriction in vocal tract shape some where along its length and is excited by either voiced, unvoiced and both types of excitation.

Nasals:

Nasal sounds are similar to vowels having lower formant energy compared to vowels. Nasal sounds are produced with the help of air flow in nasal cavity. The examples of nasal sound units found in Indian languages are */ng/*, */nj/*, */n/*, */N/* and */m/*

Semivowels:

The set of semivowels in Indian language include /y/, /r/, /l/ and /v/. The semivowels are weakly periodic as compared to the vowels and having lower energy as compared to vowels.

Fricatives:

The fricatives are the consonants produced by a narrow constriction somewhere along the length of the vocal tract. The basic difference between fricatives and stop consonants is that the closure will be partial & narrow in case of fricatives & is complete in case of stop consonants. Depending on the place of narrow constrictions, we have different fricatives.

Affricates:

The affricates are the consonants where the production involves combination of stop and fricative consonant production. Initially, the vocal tract will be completely closed somewhere all the length to create a total constrictions. After this, the constriction will be partially released to create a fricative excitations.

Procedure:

C. Nasals

a. Record the sounds of any two nasal sounds and plot their time domain waveform, the magnitude spectrum and the spectrogram.

b. Inspect the above plots and write your observations. Also, comment on how they compare to vowel sounds.

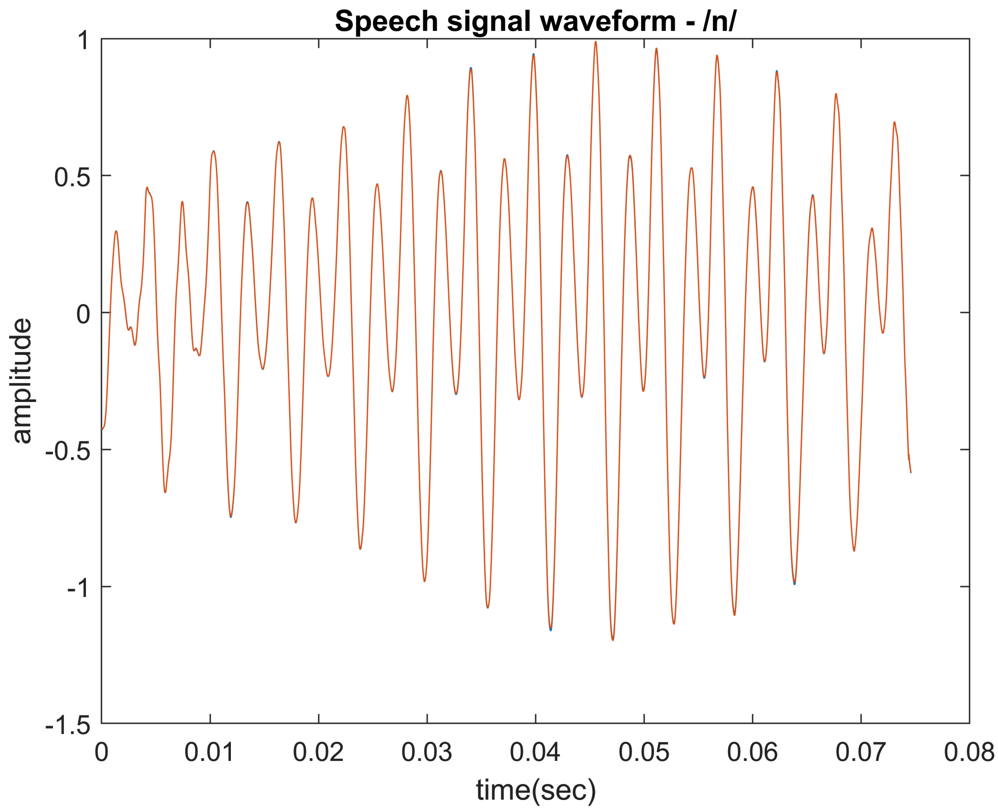
PART : C

We can use wavesurfer to record the sounds and convert into sampling frequency of 16kHz and bit resolution as 16bits/sample. Then here we plot the waveforms.

Time domain waveforms:

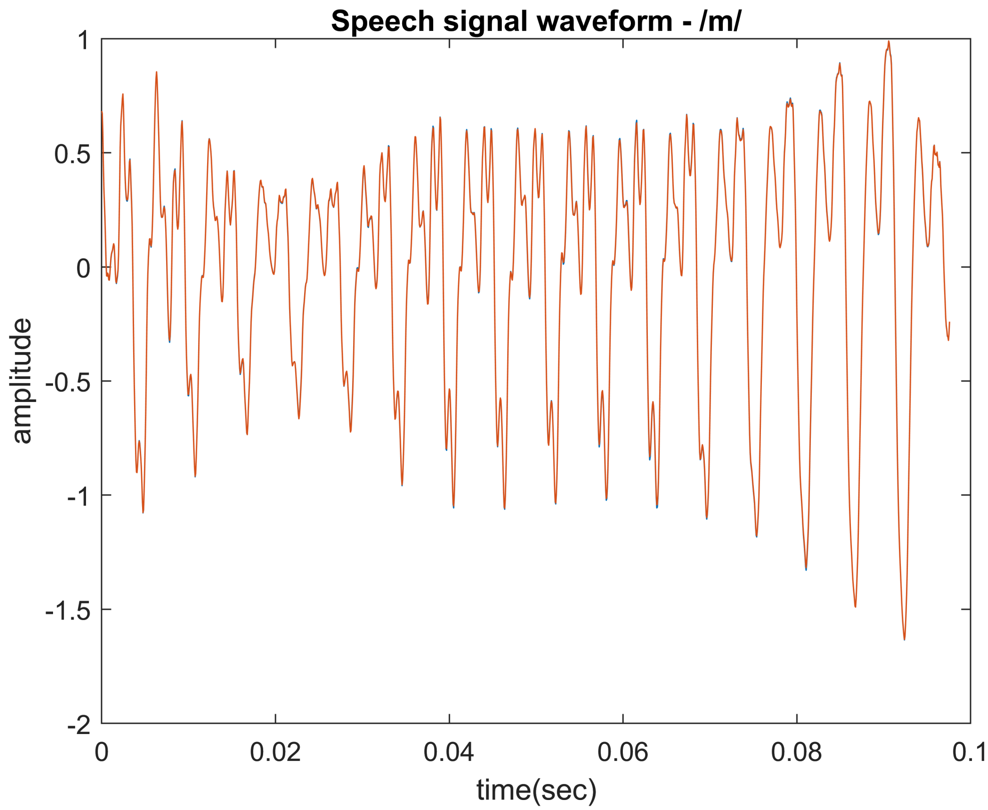
```
%Matlab program to load and plot time domain waveform stored in
% wav file format
%file name is lab5_n.wav and full path is given for /n/ sound (nasal)
[y,fs]=audioread('lab5_n.wav');

%normalising the signal amplitudes to be in -1 to 1
y_n=y./(1.01*abs(max(y)));
%plotting waveform of the speech signal
t = 0 : 1 / fs : (length(y_n) - 1) / fs;
plot(t, y_n);
xlabel('time(sec)');
ylabel('amplitude');
title('Speech signal waveform - /n/');
```




```
%file name is lab5_m.wav and full path is given for /m/ sound (nasal)
[y,fs]=audioread('lab5_m.wav');

%normalising the signal amplitudes to be in -1 to 1
y_m=y./(1.01*abs(max(y)));
%plotting waveform of the speech signal
t = 0 : 1 / fs : (length(y_m) - 1) / fs;
plot(t, y_m);
xlabel('time(sec)');
ylabel('amplitude');
title('Speech signal waveform - /m/');
```



Magnitude spectrum:

We will use the wavesurfer for analysing the plots. We can note down the 25ms duration of the segment at the centre of the sound. The time-stamps for each sound is obtained from wavesurfer.

```
%/n/
y_n = y(ceil(0.007*fs) : floor(0.032*fs));
%/m/
y_m = y(ceil(0.012*fs) : floor(0.039*fs));
```

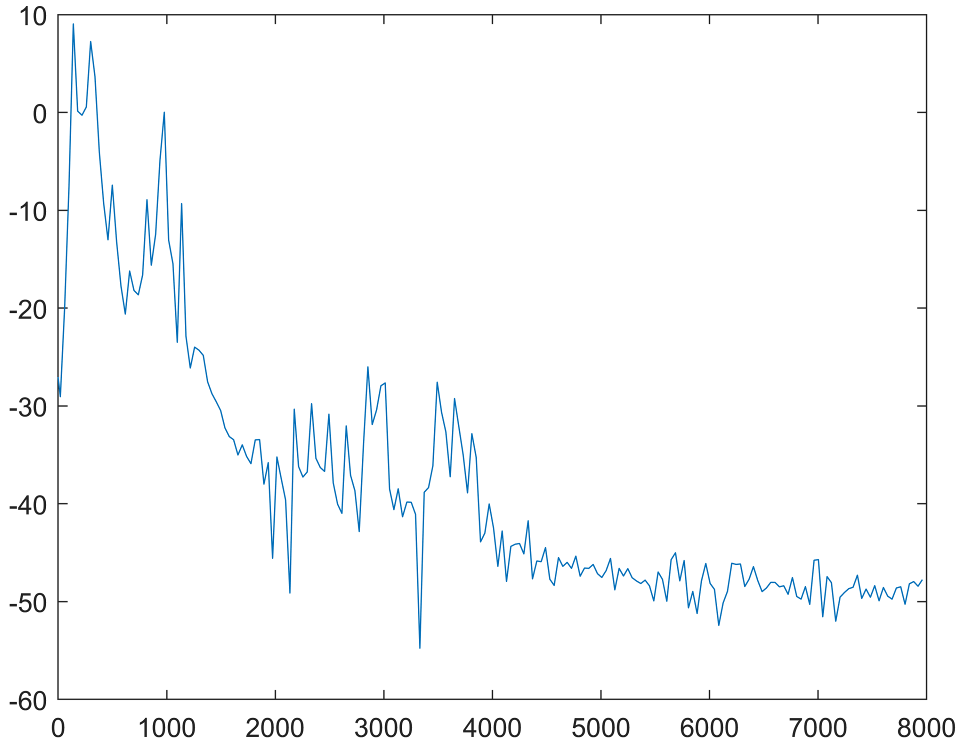
Now we plot the magnitude spectrum plots of the speech signal. We use the same method as specified to compute the N-point DFT.

```
Y_n = fftshift(fft(y_n));
Y_m = fftshift(fft(y_m));
```

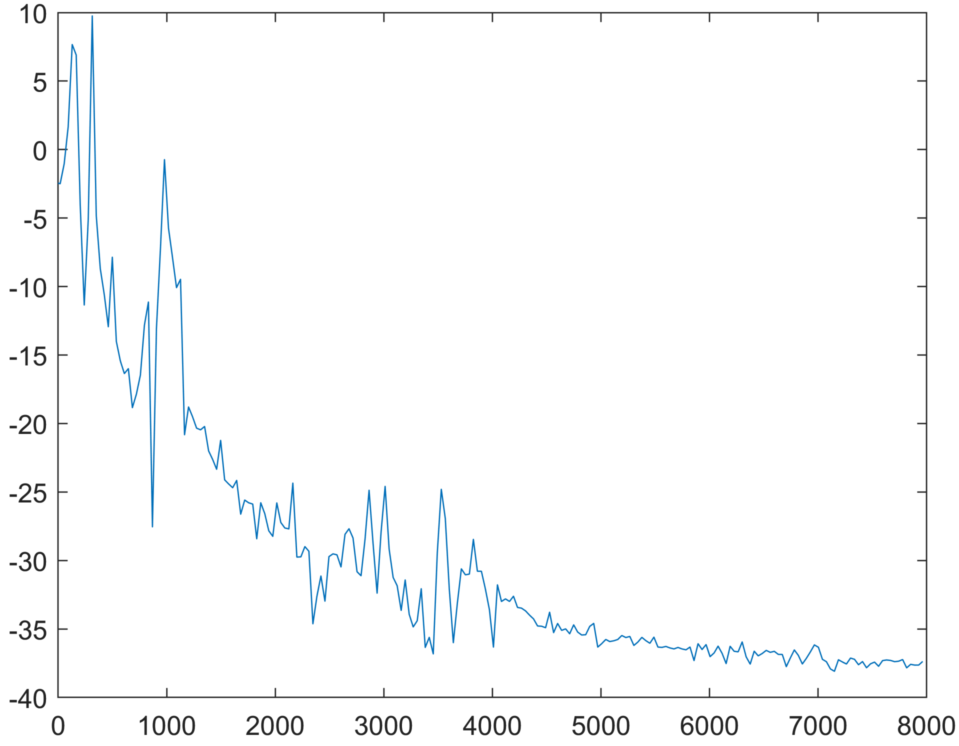
We have obtained the N-point DFTs of all the sounds. Now we plot the frequency spectrum for positive frequencies.

```
F_n = -fs/2 : fs/length(Y_n) : fs/2 - fs/length(Y_n);
F_m = -fs/2 : fs/length(Y_m) : fs/2 - fs/length(Y_m);

% Plots
% /n/
plot(F_n, 20*log10(abs(Y_n)));
xlim([0, fs/2]);
```

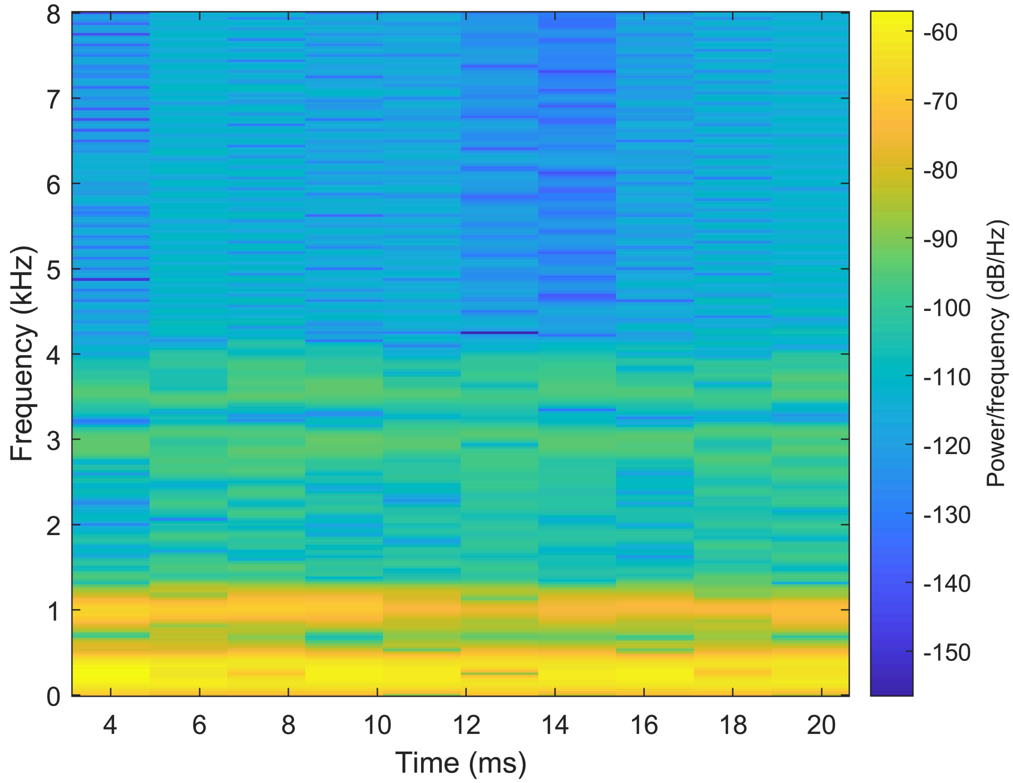



```
% /m/
plot(F_m, 20*log10(abs(Y_m)));
xlim([0, fs/2]);
```

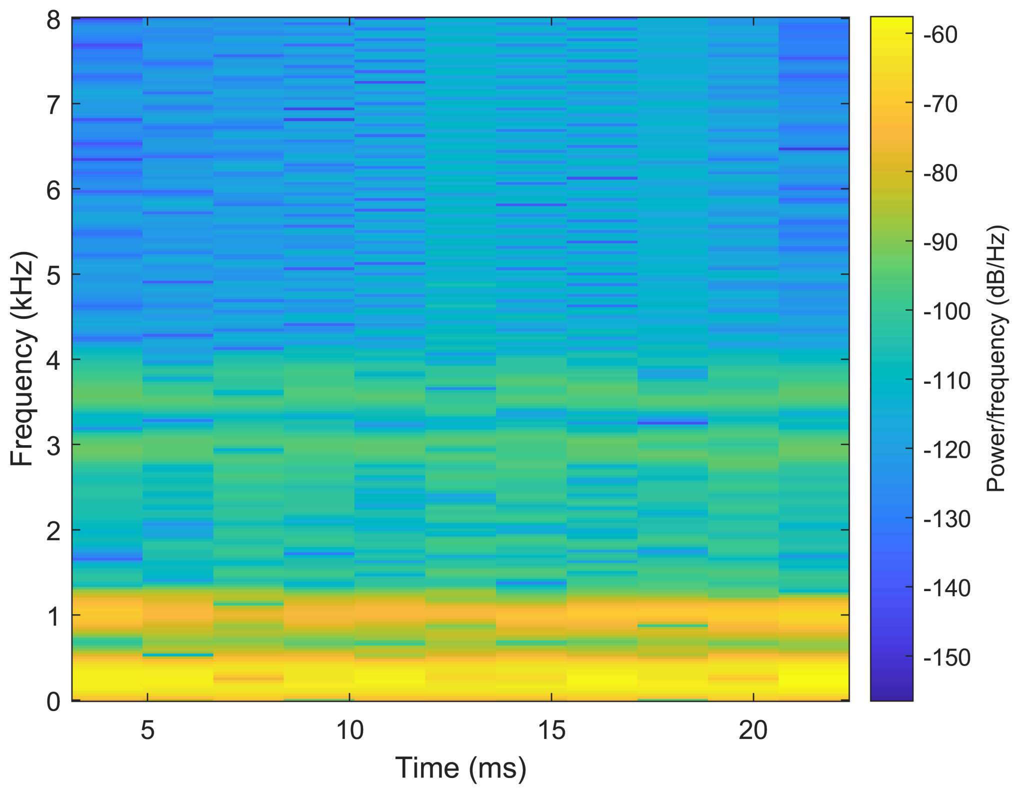


Spectrogram:

```
spectrogram(y_n,  128, (100), 512, fs,'yaxis');
```




```
spectrogram(y_m,  128, (100), 512, fs,'yaxis');
```



Observations:

The nasal sounds are voiced and have some aspiration noise due to airflow through the nose.The sounds /n/ & /m/ are periodic. The formant energy of nasal sounds are less compared to that of the vowels. We can observe anti resonance property in the spectrogram.

D. Semi-Vowels

a. Record the sounds of any two semi-vowels and plot their time domain waveform, the magnitude spectrum and the spectrogram.

b. Inspect the above plots and write your observations. Comment on how these vary from the vowel sounds.

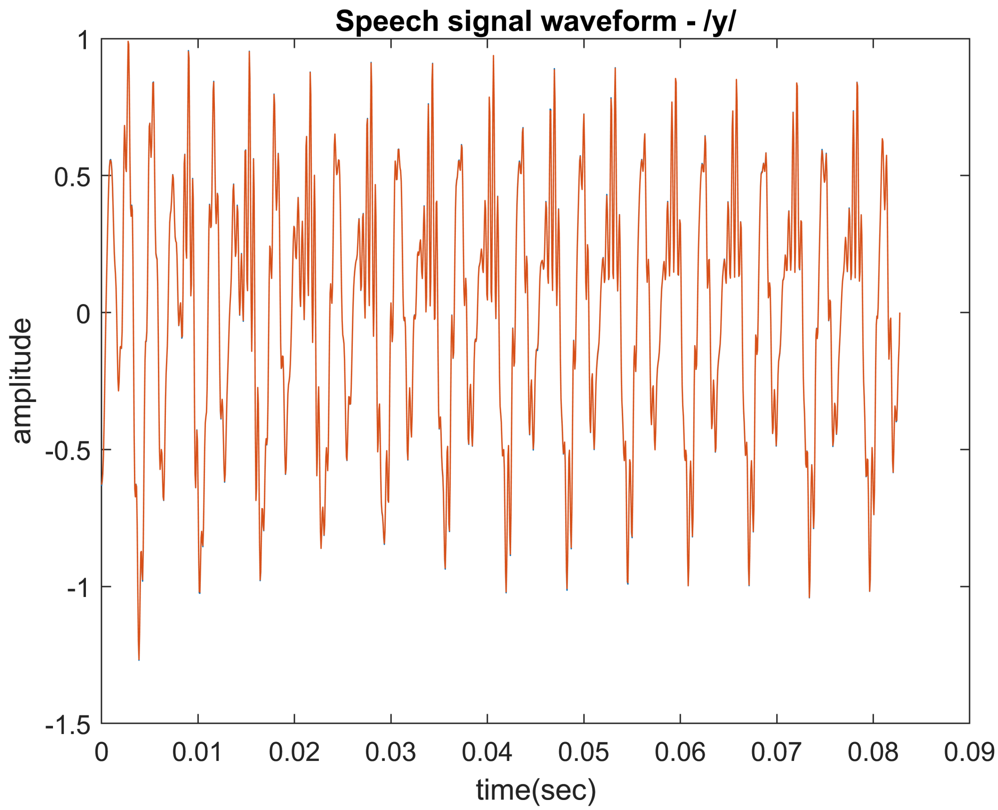
PART : D

We can use wavesurfer to record the sounds and convert into sampling frequency of 16kHz and bit resolution as 16bits/sample. Then here we plot the waveforms.

Time domain waveforms:

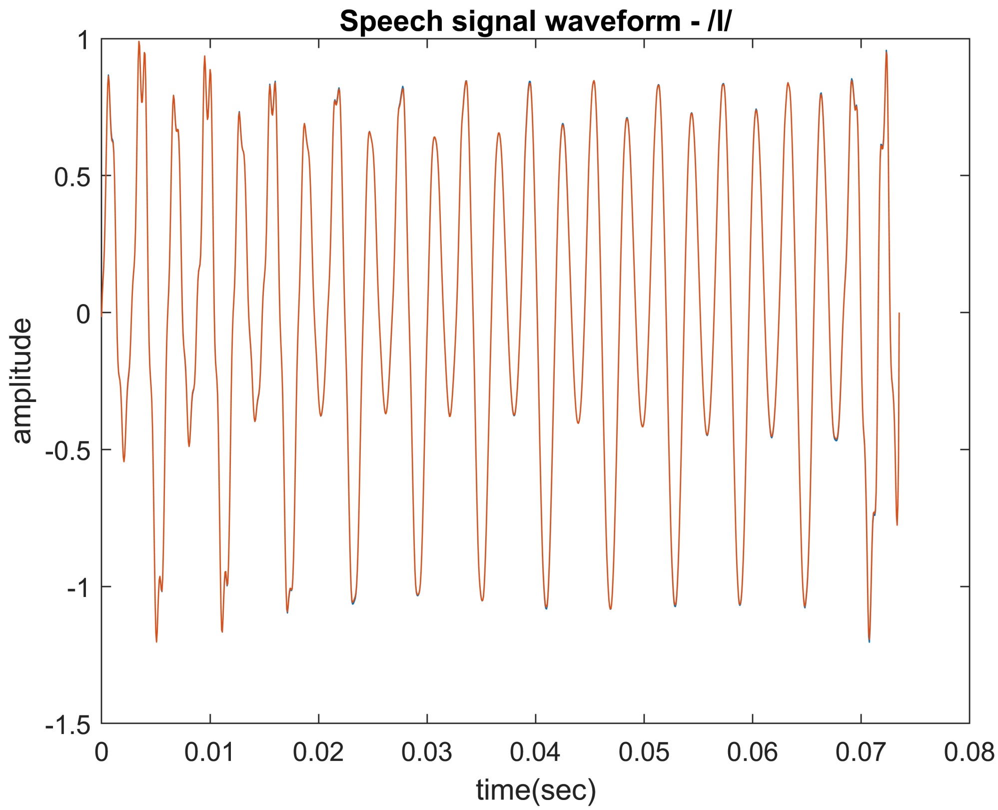
```
%Matlab program to load and plot time domain waveform stored in
% wav file format
%file name is lab5_y.wav and full path is given for /y/ sound (semi vowel)
[y,fs]=audioread('lab5_y.wav');

%normalising the signal amplitudes to be in -1 to 1
y_y=y./(1.01*abs(max(y)));
%plotting waveform of the speech signal
t = 0 : 1 / fs : (length(y_y) - 1) / fs;
plot(t, y_y);
xlabel('time(sec)');
ylabel('amplitude');
title('Speech signal waveform - /y/');
```




```
%file name is lab5_l.wav and full path is given for /l/ sound (semi vowel)
[y,fs]=audioread('lab5_l.wav');

%normalising the signal amplitudes to be in -1 to 1
y_l=y./(1.01*abs(max(y)));
%plotting waveform of the speech signal
t = 0 : 1 / fs : (length(y_l) - 1) / fs;
plot(t, y_l);
xlabel('time(sec)');
ylabel('amplitude');
title('Speech signal waveform - /l/');
```



Magnitude spectrum:

We will use the wavesurfer for analysing the plots. We can note down the 25ms duration of the segment at the centre of the sound. The time-stamps for each sound is obtained from wavesurfer.

```
%/y/
y_y = y(ceil(0.010*fs) : floor(0.036*fs));
%/l/
y_l = y(ceil(0.024*fs) : floor(0.052*fs));
```

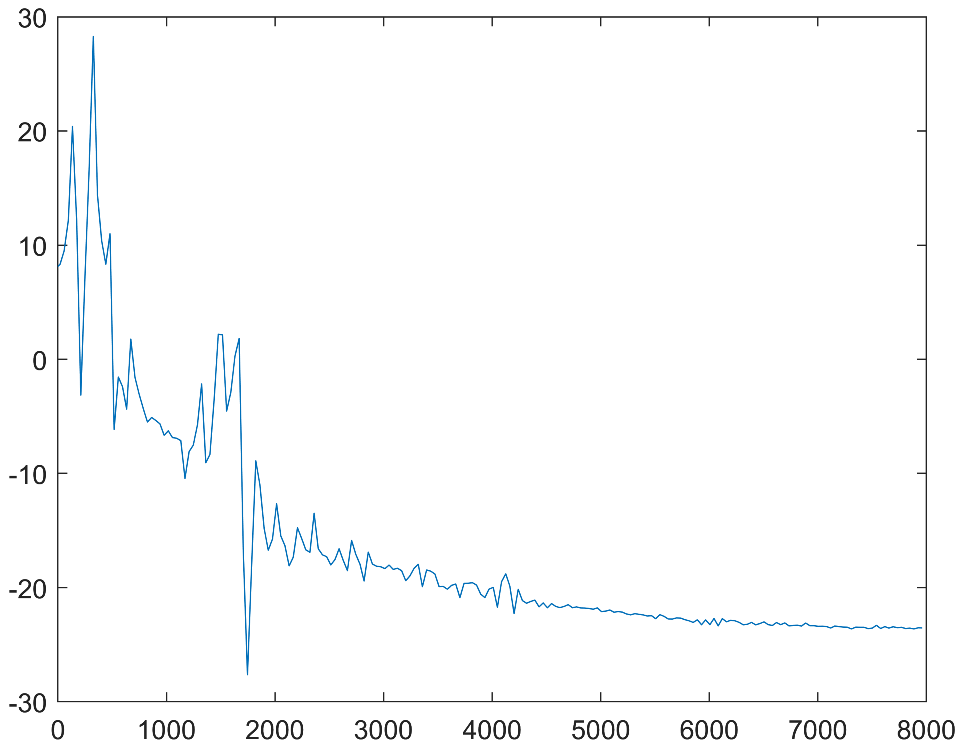
Now we plot the magnitude spectrum plots of the speech signal. We use the same method as specified to compute the N-point DFT.

```
Y_y = fftshift(fft(y_y));
Y_l = fftshift(fft(y_l));
```

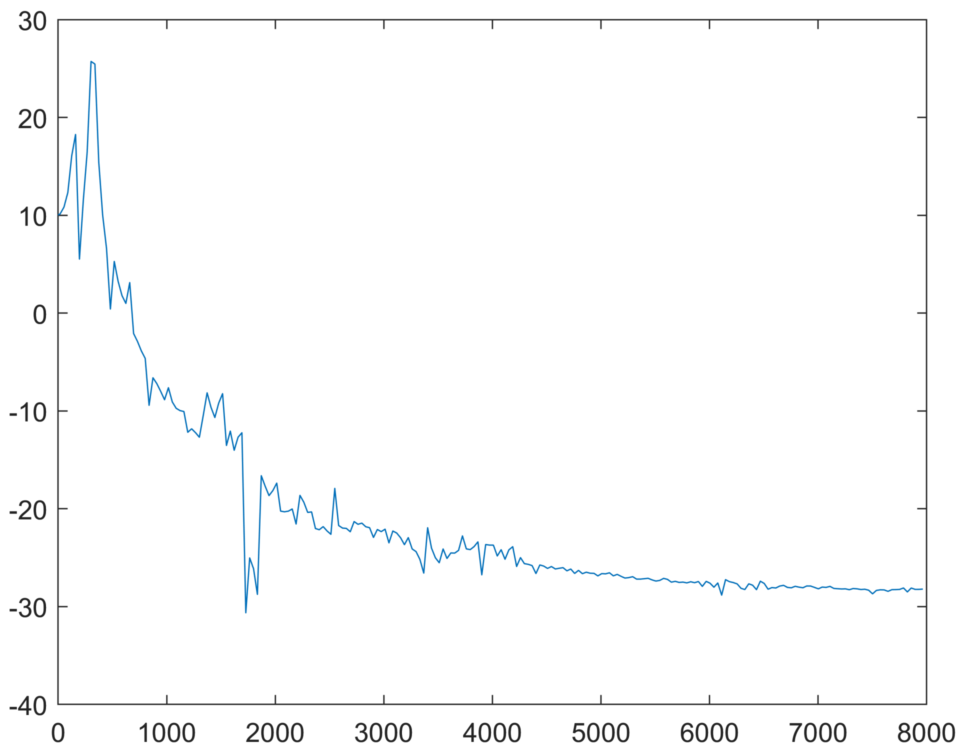
We have obtained the N-point DFTs of all the sounds. Now we plot the frequency spectrum for positive frequencies.

```
F_y = -fs/2 : fs/length(Y_y) : fs/2 - fs/length(Y_y);
F_l = -fs/2 : fs/length(Y_l) : fs/2 - fs/length(Y_l);
```

```
% Plots
% /y/
plot(F_y, 20*log10(abs(Y_y)));
xlim([0, fs/2]);
```

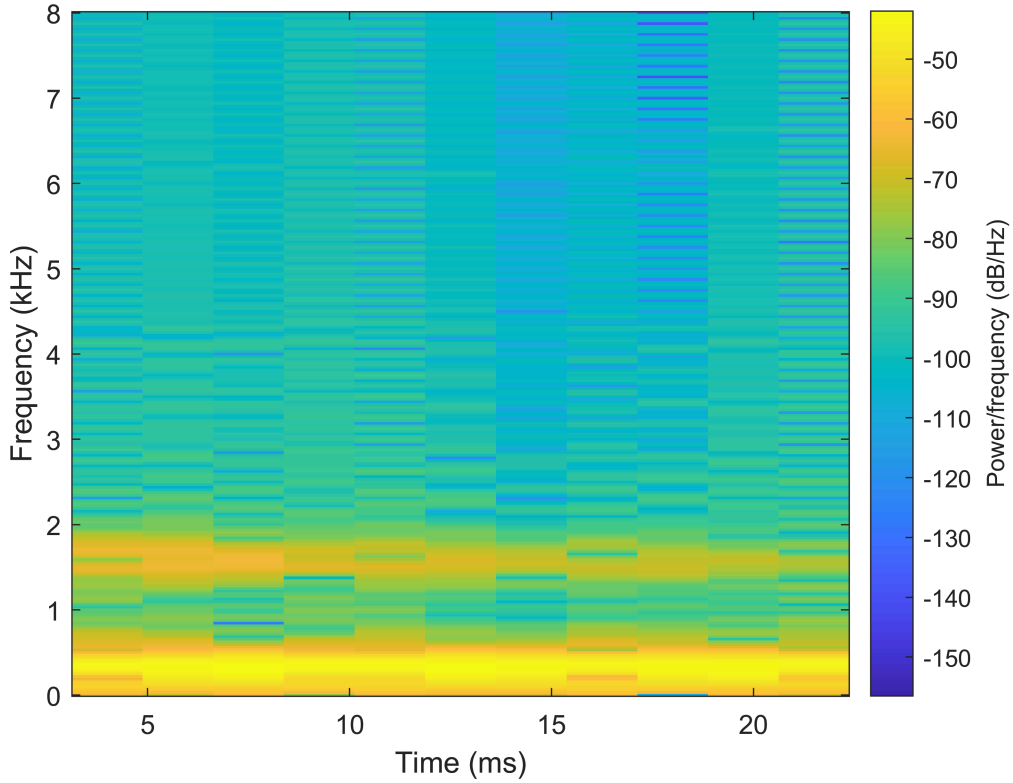



```
% /l/
plot(F_l, 20*log10(abs(Y_l)));
xlim([0, fs/2]);
```

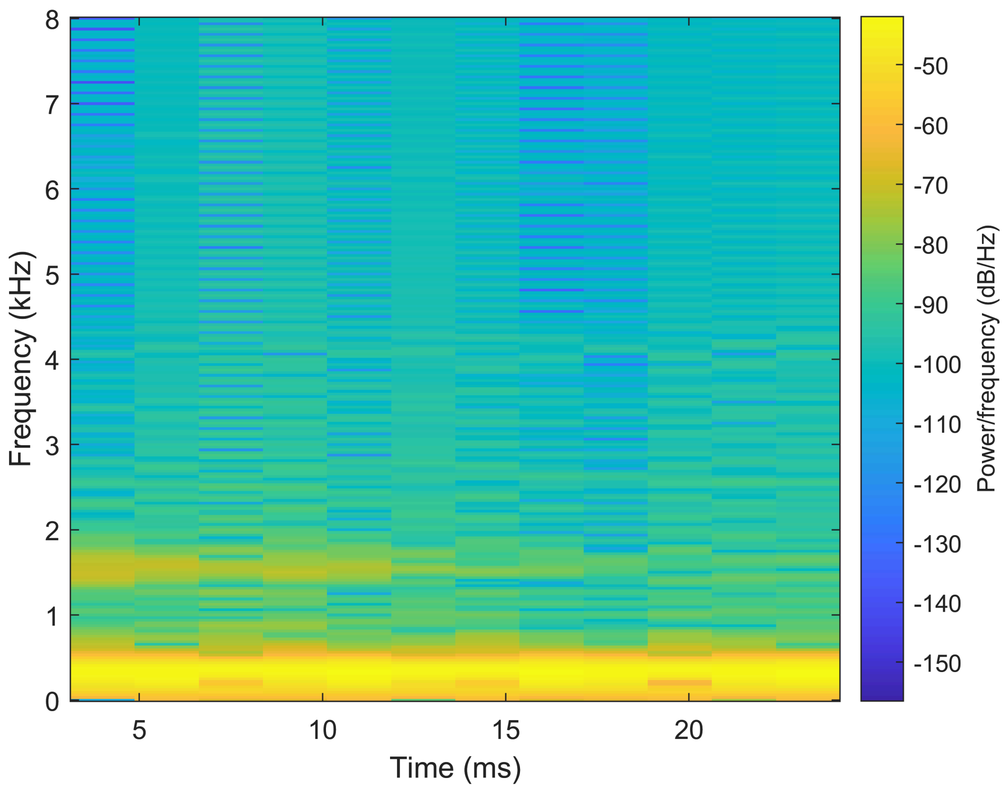


Spectrogram:

```
spectrogram(y_y,  128, (100), 512, fs,'yaxis');
```




```
spectrogram(y_l,  128, (100), 512, fs,'yaxis');
```



Observations:

The semivowels are weakly periodic as compared to the vowels and having lower energy as compared to vowels. /y/ is aspirated and /l/ is unaspirated.

E. Fricatives

a. Pick up any two fricatives having different positions of constrictions. Record these sounds and plot the time-domain waveform, the magnitude spectrum and the spectrogram.

b. Inspect the above plots and write your observations.

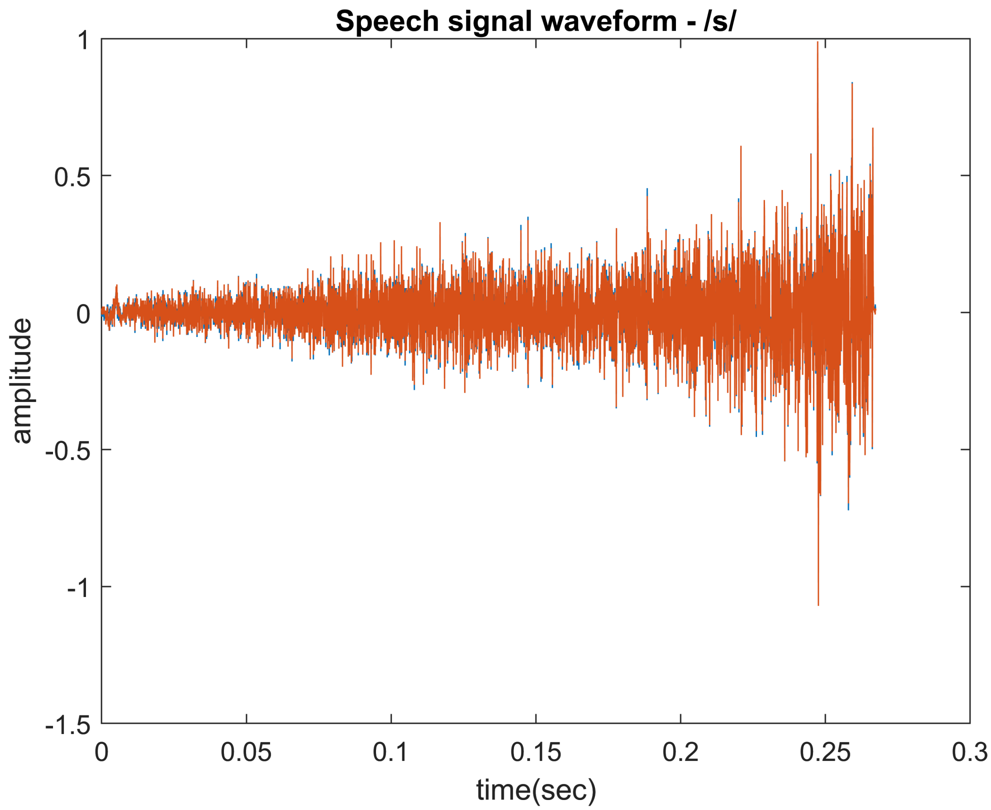
PART : E

We can use wavesurfer to record the sounds and convert into sampling frequency of 16kHz and bit resolution as 16bits/sample. Then here we plot the waveforms.

Time domain waveforms:

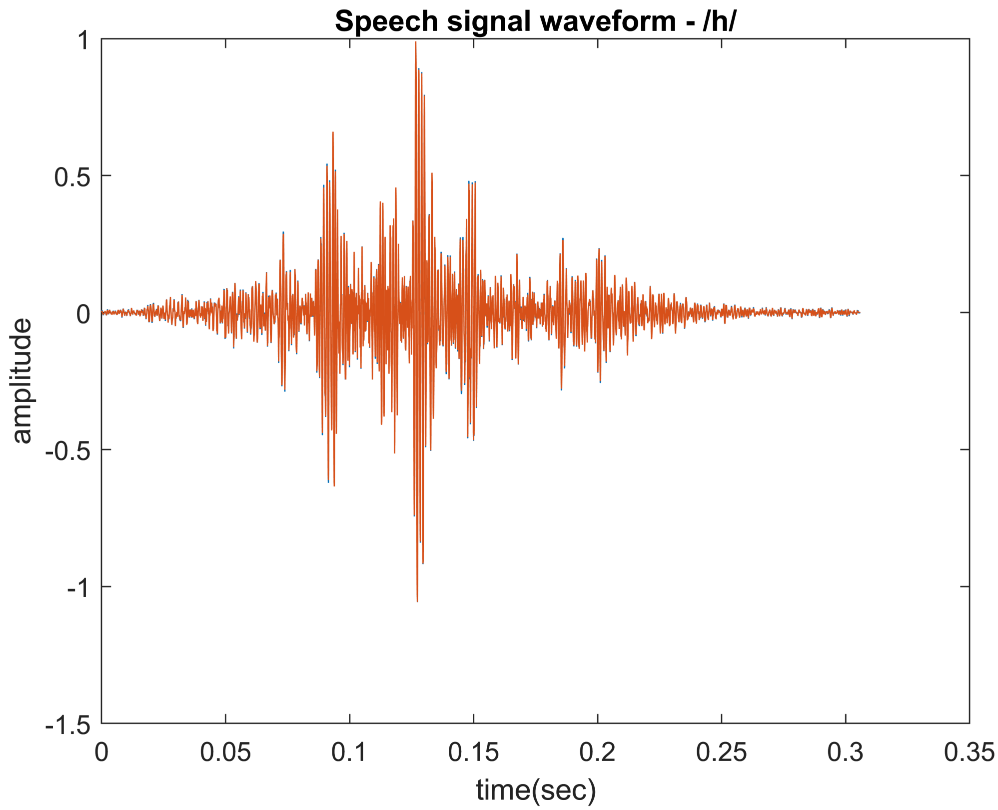
```
%Matlab program to load and plot time domain waveform stored in
% wav file format
%file name is lab5_s.wav and full path is given for /s/ sound (fricative)
[y,fs]=audioread('lab5_s.wav');

%normalising the signal amplitudes to be in -1 to 1
y_s=y./(1.01*abs(max(y)));
%plotting waveform of the speech signal
t = 0 : 1 / fs : (length(y_s) - 1) / fs;
plot(t, y_s);
xlabel('time(sec)');
ylabel('amplitude');
title('Speech signal waveform - /s/');
```




```
%file name is lab5_h.wav and full path is given for /h/ sound (fricative)
[y,fs]=audioread('lab5_h.wav');

%normalising the signal amplitudes to be in -1 to 1
y_h=y./(1.01*abs(max(y)));
%plotting waveform of the speech signal
t = 0 : 1 / fs : (length(y_h) - 1) / fs;
plot(t, y_h);
xlabel('time(sec)');
ylabel('amplitude');
title('Speech signal waveform - /h/');
```



Magnitude spectrum:

We will use the wavesurfer for analysing the plots. We can note down the 25ms duration of the segment at the centre of the sound. The time-stamps for each sound is obtained from wavesurfer.

```
%/s/
y_s = y(ceil(0.097*fs) : floor(0.122*fs));
%/h/
y_h = y(ceil(0.116*fs) : floor(0.142*fs));
```

Now we plot the magnitude spectrum plots of the speech signal. We use the same method as specified to compute the N-point DFT.

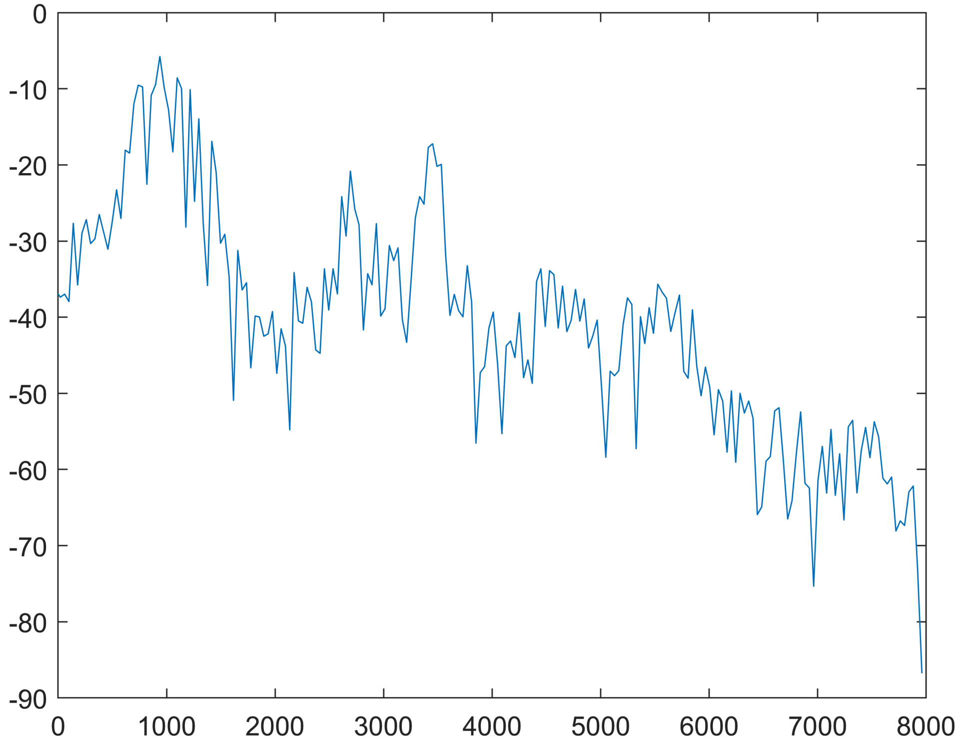
```
Y_s = fftshift(fft(y_s));
Y_h = fftshift(fft(y_h));
```

We have obtained the N-point DFTs of all the sounds. Now we plot the frequency spectrum for positive frequencies.

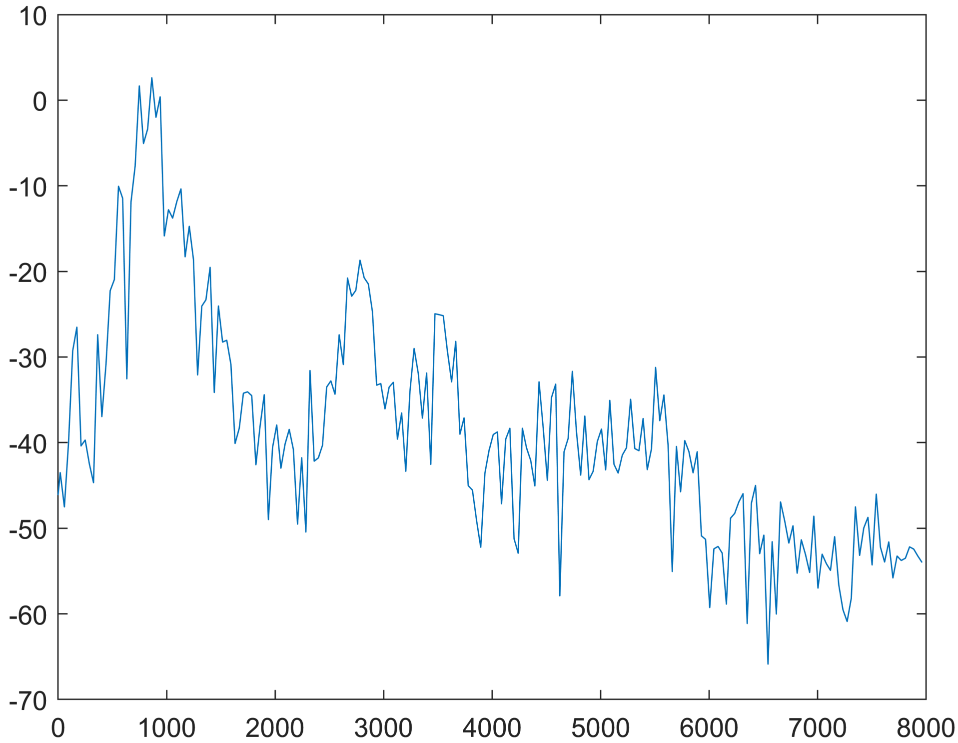
```
F_s = -fs/2 : fs/length(Y_s) : fs/2 - fs/length(Y_s);
F_h = -fs/2 : fs/length(Y_h) : fs/2 - fs/length(Y_h);

% Plots
```

```
% /s/
plot(F_s, 20*log10(abs(Y_s)));
xlim([0, fs/2]);
```

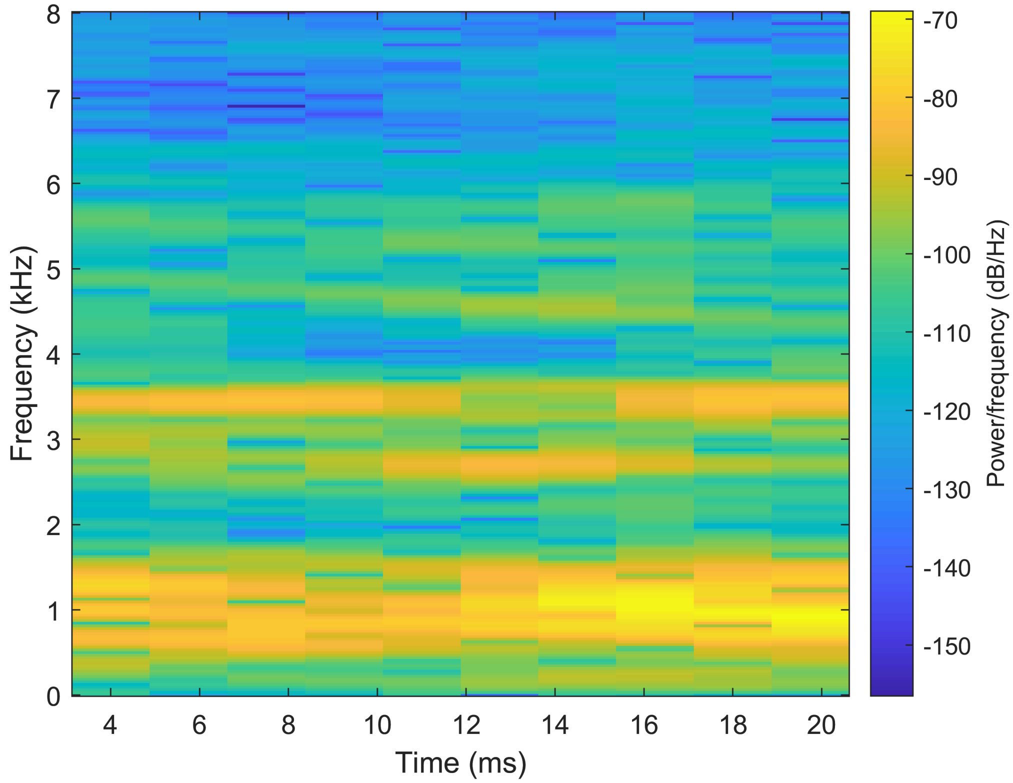



```
% /h/
plot(F_h, 20*log10(abs(Y_h)));
xlim([0, fs/2]);
```

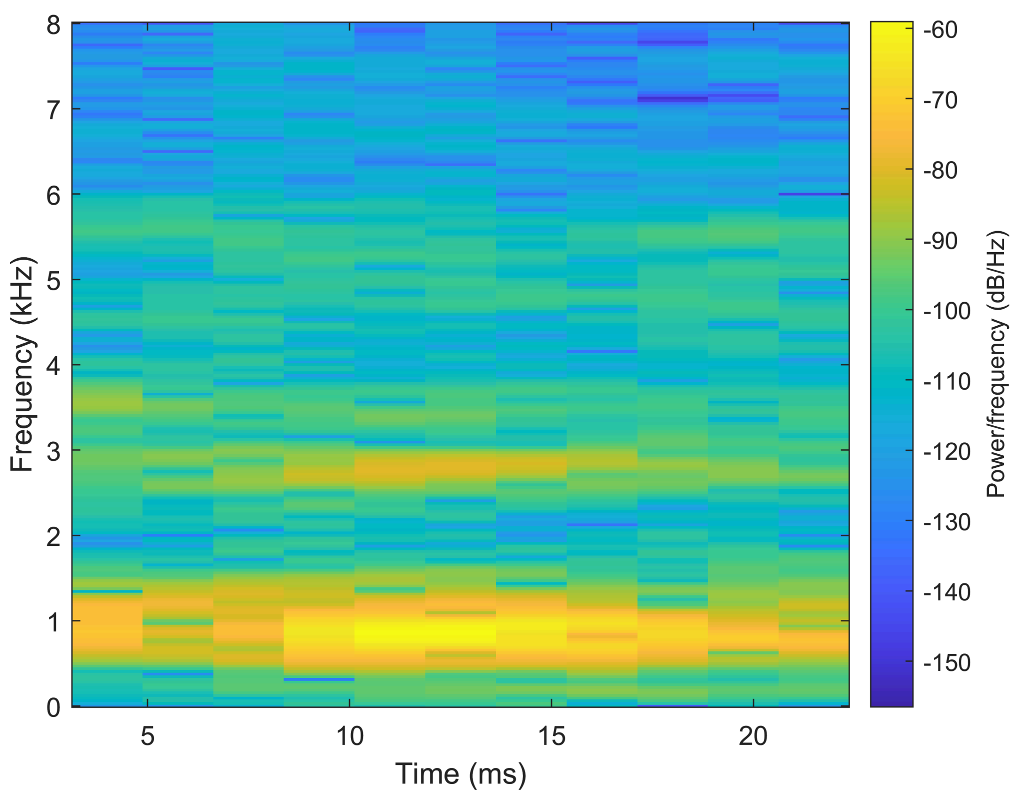


Spectrogram:

```
spectrogram(y_s,  128, (100), 512, fs,'yaxis');
```




```
spectrogram(y_h,  128, (100), 512, fs,'yaxis');
```



Observation:

The fricatives sounds are produced by a narrow constriction somewhere along the length of the vocal tract. The closure will be partial & narrow in case of fricatives. |s| is a dental fricative. |h| is a velar fricative. The waveform looks like a random noise having energy mostly concentrated at higher frequency regions.

F. Affricates

a. Record any one affricate sound and plot the time domain waveform, the magnitude spectrum and the spectrogram.

b. Inspect the plots and write down your observations.

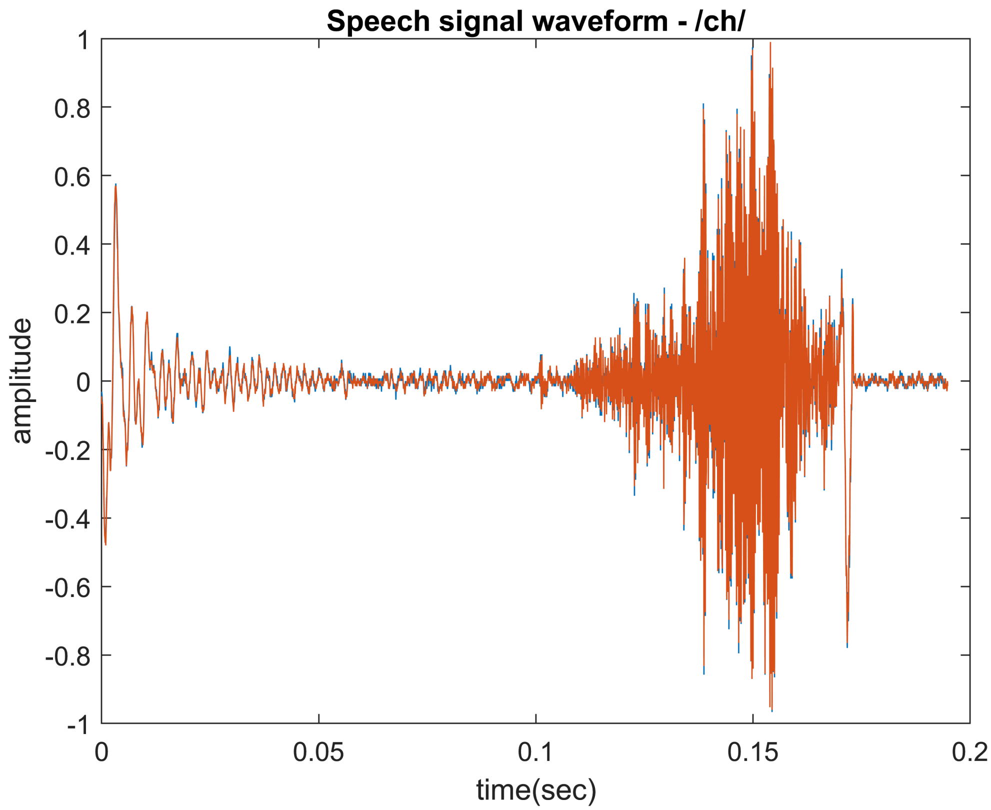
<u>PART : F</u>

We can use wavesurfer to record the sounds and convert into sampling frequency of 16kHz and bit resolution as 16bits/sample. Then here we plot the waveforms.

Time domain waveforms:

```
%Matlab program to load and plot time domain waveform stored in
% wav file format
%file name is lab5_ch.wav and full path is given for /ch/ sound (affricate)
[y,fs]=audioread('lab5_ch.wav');

%normalising the signal amplitudes to be in -1 to 1
y_ch=y./(1.01*abs(max(y)));
%plotting waveform of the speech signal
t = 0 : 1 / fs : (length(y_ch) - 1) / fs;
plot(t, y_ch);
xlabel('time(sec)');
ylabel('amplitude');
title('Speech signal waveform - /ch/');
```



Magnitude spectrum:

We will use the wavesurfer for analysing the plots. We can note down the 25ms duration of the segment at the centre of the sound. The time-stamps for each sound is obtained from wavesurfer.

```
%/ch/
y_ch = y(ceil(0.121*fs) : floor(0.146*fs));
```

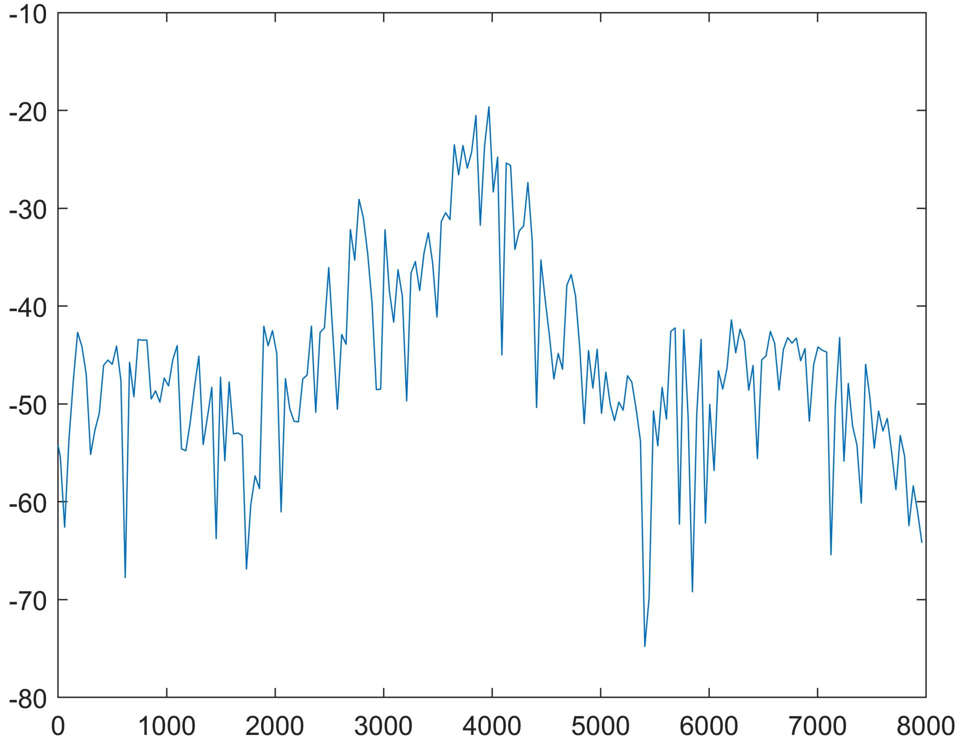
Now we plot the magnitude spectrum plots of the speech signal. We use the same method as specified to compute the N-point DFT.

```
Y_ch = fftshift(fft(y_ch));
```

We have obtained the N-point DFTs of all the sounds. Now we plot the frequency spectrum for positive frequencies.

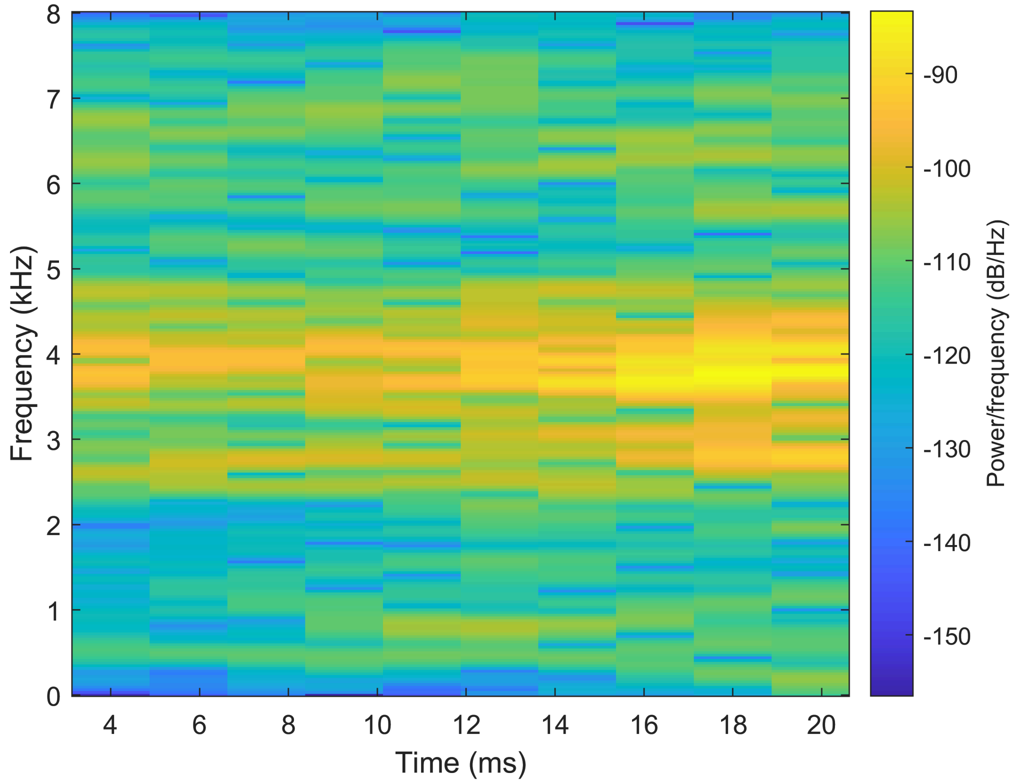
```
F_ch = -fs/2 : fs/length(Y_ch) : fs/2 - fs/length(Y_ch);

% Plots
% /ch/
plot(F_ch, 20*log10(abs(Y_ch)));
xlim([0, fs/2]);
```



Spectrogram:

```
spectrogram(y_ch,  128, (100), 512, fs,'yaxis');
```



Observation:

Unlike other unvoiced and unaspirated stop consonants, /ch/ is aperiodic and turbulent in nature. We can observe closure, burst & frication.