

# EE-414 Speech Processing Lab

## Lab-5

### 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.

### Introduction:

Speech signals are composed of different sound units which differ from language to language. It is important to get a feel of how these sound units are produced, by analysing its time, spectral properties and spectrogram. There are different categories of speech sounds such as vowels(short vowels, long vowels, diphthongs) and consonants(short consonants, nasals, semi-vowels, fricatives, affricates). Speech signals are formed due to the combinations of different sound units of a particular language, hence it is important to study these sound units individually. The sounds produced are either voiced(when there is no interruption of air flow from glottis to lips by voiced excitation) or unvoiced (where there is narrow constriction/obstruction in the air flow by unvoiced/voiced excitation).

### A.

#### Short vowels, Long vowels and Diphthongs

1. Record the sounds of any one short vowel sound, long vowel sound and a diphthong (Also, record the two sounds present in the diphthong).
2. Plot the time domain waveform, magnitude spectrum and the spectrogram for each of the above sounds.
3. Inspect each of the above plots and write your observations comparing them.

### Theory:

Short vowels, long vowels and diphthongs are the categories of classification for vowels. Short vowels and long vowels are produced in a similar manner however the duration of long vowels is much longer (about twice as long) as compared to short vowels. Diphthongs are combinations of different vowels, where there is a change from one vowel to another vowel during the production of sound (i.e. the vocal tract changes as the two sounds are produced)

### Procedure:

We are analysing "u" (short vowel), "uu" (long vowel) and "ou" (diphthong). We record these sounds via the software *Audacity*, with 16 bits/sample and 16 kHz sampling frequency. We then plot time domain signal using user-defined function *plot\_time\_signal()* and frequency spectrum using user-defined function *plot\_freq\_spectrum()*. We use inbuilt matlab function *spectrogram()* to plot the spectrogram. We then analyse the plots and make some observations.

### Code and Plots:

```
close all; clc; clear all;  
  
[u_sound,fs]=audioread('u.wav');
```

```

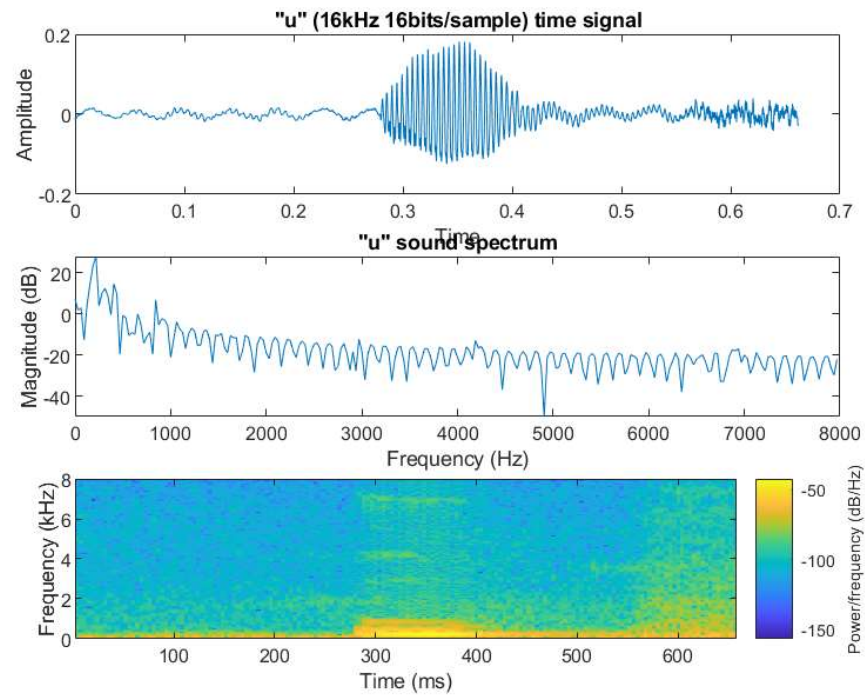
[uu_sound,fs]=audioread('uu.wav');

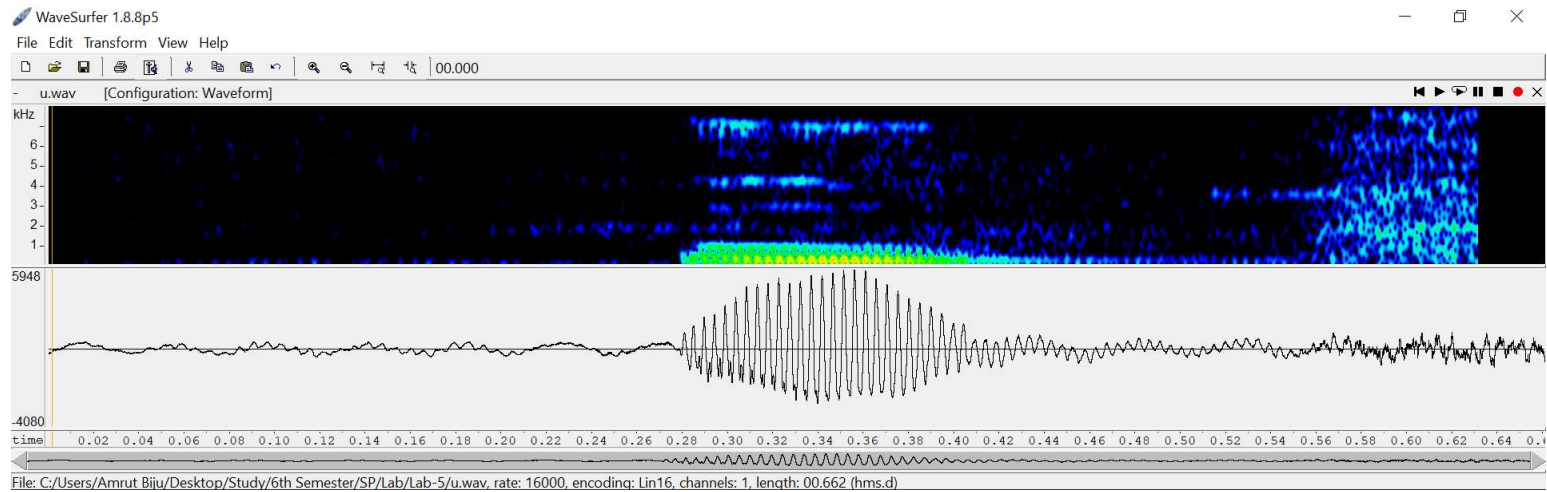
[ou_sound,fs]=audioread('ou.wav');

[a_sound,fs]=audioread('a.wav');

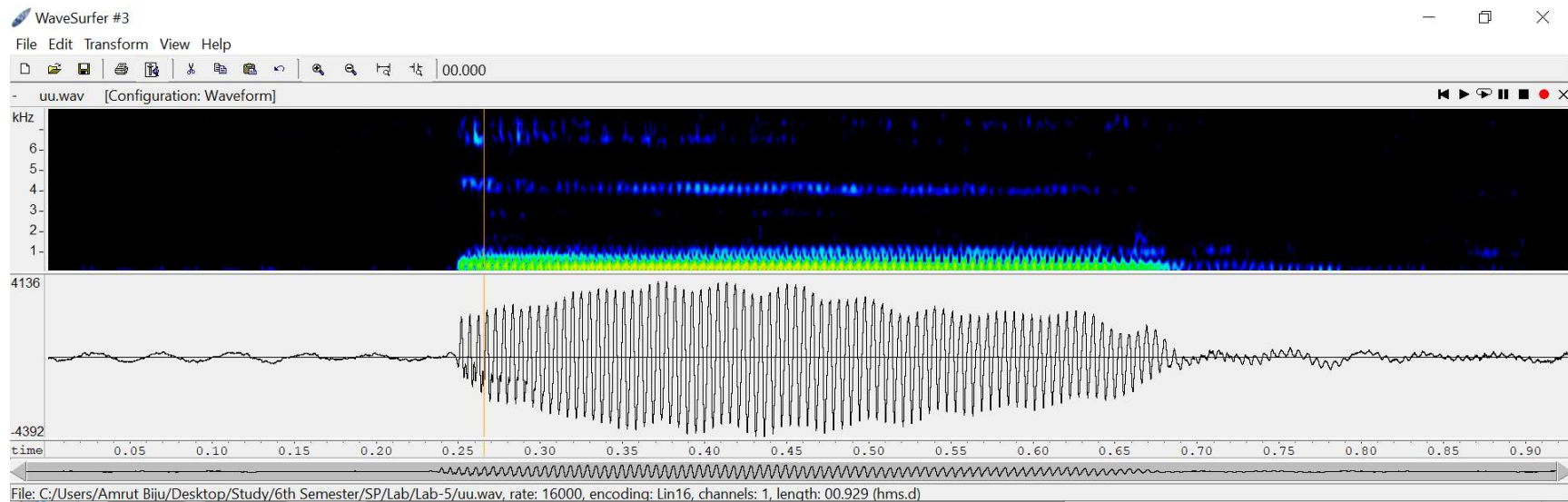
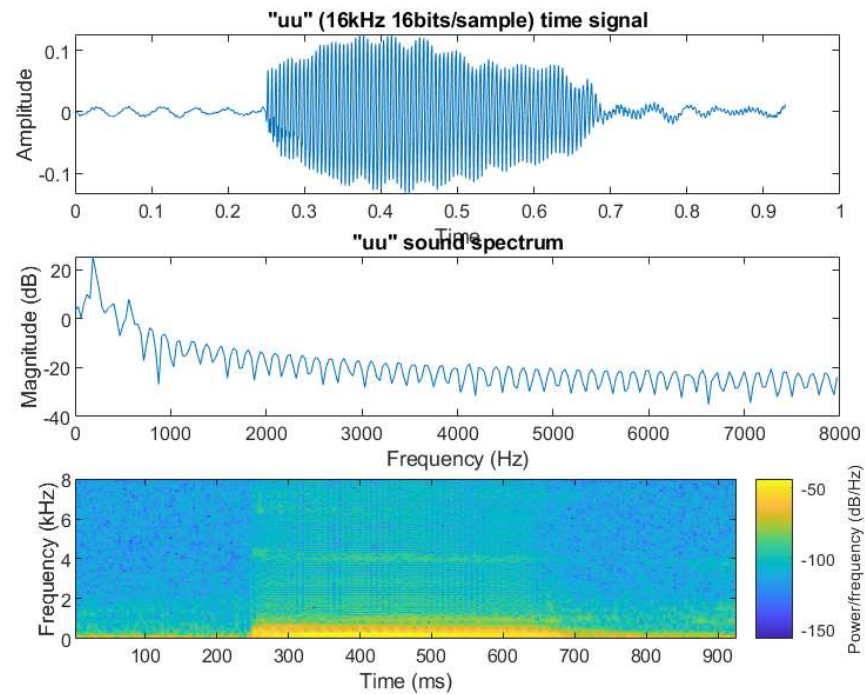
% "u"
subplot(311);
plot_time_signal(u_sound, fs, '"u" (16kHz 16bits/sample) time signal');
subplot(312);
plot_freq_spectrum(u_sound((0.32*fs):(int64((0.32+0.025)*fs))), fs, '"u" sound spectrum');
subplot(313);
spectrogram(u_sound, 128, [], 512, fs, 'yaxis');

```





```
% "uu"
close hidden
subplot(311);
plot_time_signal(uu_sound, fs, '"uu" (16kHz 16bits/sample) time signal');
subplot(312);
plot_freq_spectrum(uu_sound((0.30*fs):(int64((0.30+0.025)*fs))), fs, '"uu" sound spectrum');
subplot(313);
spectrogram(uu_sound, 128, [], 512, fs, 'yaxis');
```

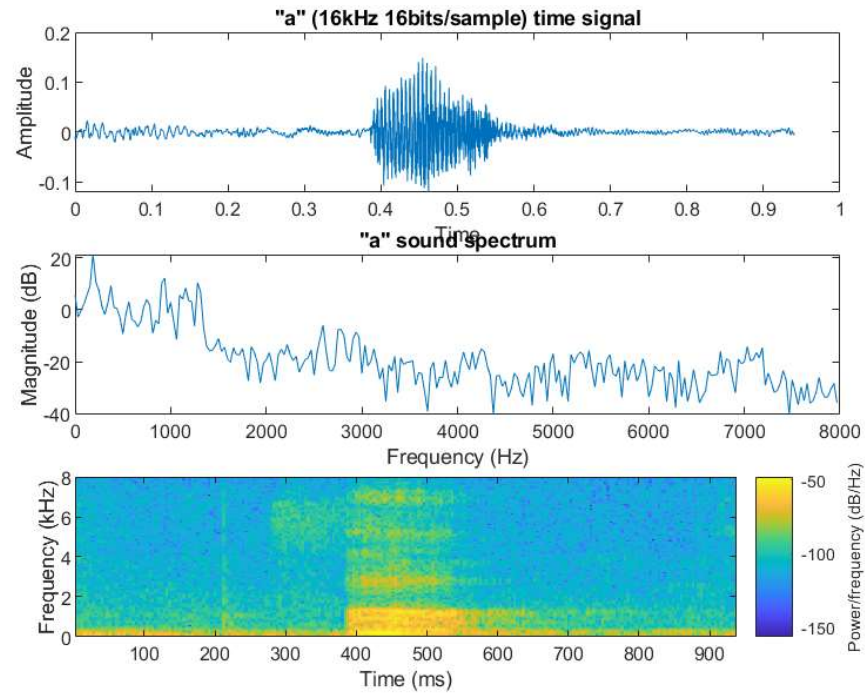


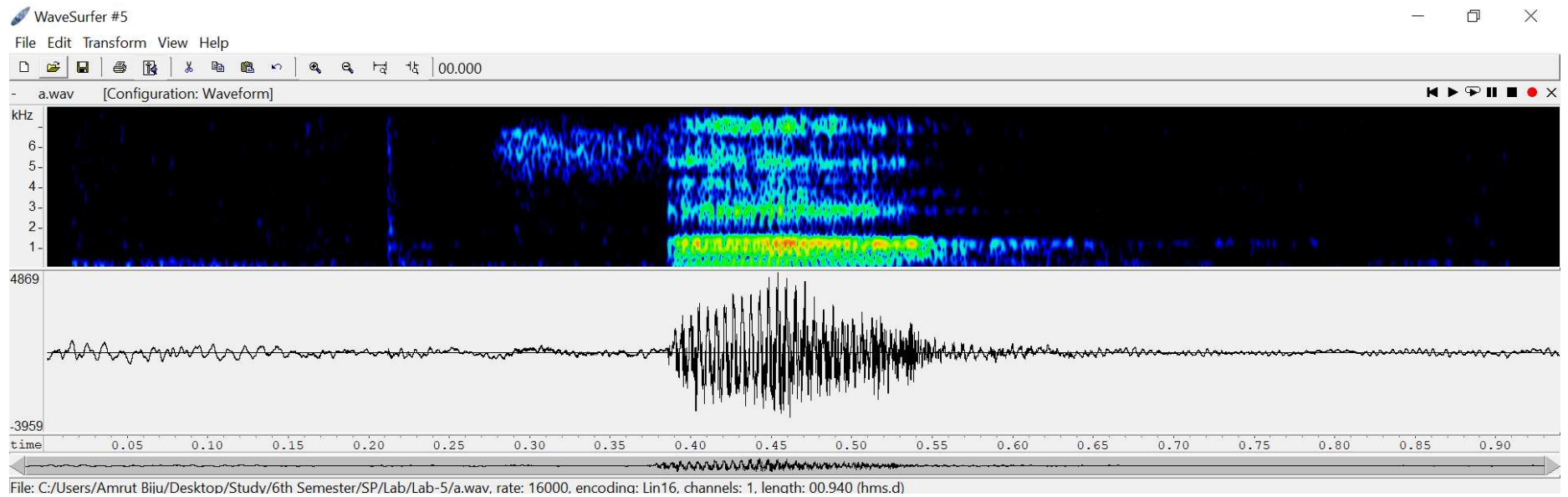
```
% "a" (for comparison, for diphtong part)
close hidden
subplot(311);
```

```

plot_time_signal(a_sound, fs, '"a" (16kHz 16bits/sample) time signal');
subplot(312);
plot_freq_spectrum(a_sound((0.40*fs):(int64((0.40+0.025)*fs))), fs, '"a" sound spectrum');
subplot(313);
spectrogram(a_sound, 128, [], 512, fs, 'yaxis');

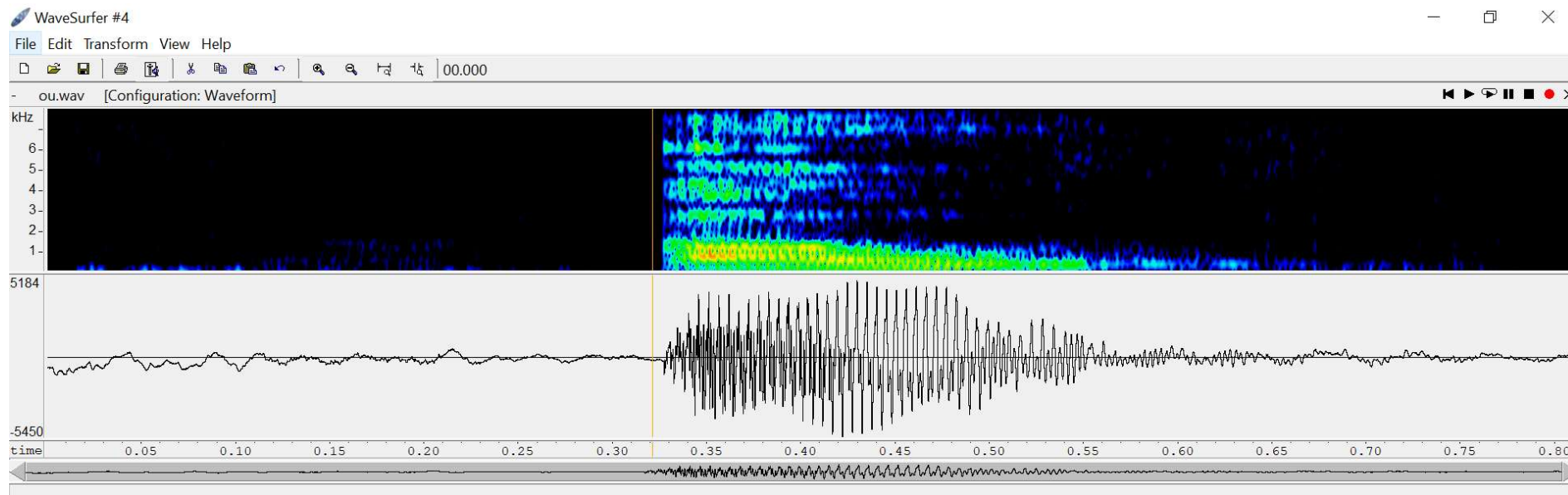
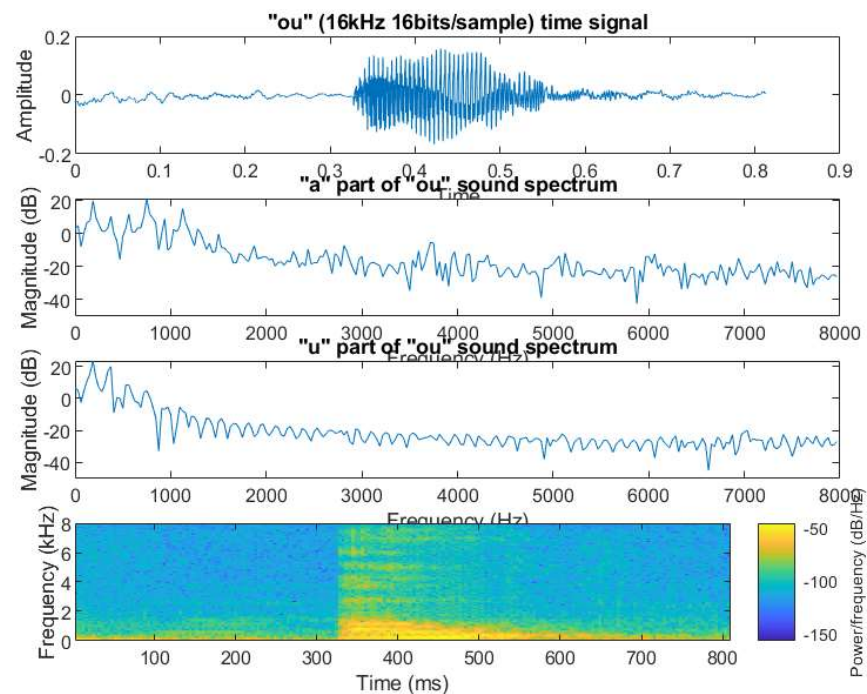
```





```
% "ou"
close hidden
subplot(411);
plot_time_signal(ou_sound, fs, '"ou" (16kHz 16bits/sample) time signal');
subplot(412);
plot_freq_spectrum(ou_sound((0.35*fs):(int64((0.35+0.025)*fs))), fs, '"a" part of "ou" sound spectrum');
subplot(413);
plot_freq_spectrum(ou_sound((0.47*fs):(int64((0.47+0.025)*fs))), fs, '"u" part of "ou" sound spectrum');
subplot(414);
spectrogram(ou_sound, 128, [], 512, fs, 'yaxis');
```





**Observations:**

We see that "u" and "uu" have same time domain structure except that for "uu" the duration is longer. We can also see that the spectral properties are also same through frequency spectrum and spectrogram. For the diphthong "ou", we see that it has "a" and "u" sounds by seeing the spectrograms and the frequency spectrum. We see that first half has the same frequency spectrum as "a" sound and 2nd half has same frequency spectrum as "u" sound.

## **B.**

### **Stop Consonants**

1. Pick up any one of the POA(Position of Articulation) types and record the sounds present in the respective row for all the MOA(Manner of Articulation) types.
2. Plot the time domain waveform, the magnitude spectrum and the spectrogram for each of the above sounds.
3. Inspect the above plots and describe the various sub phonetic events that take place, their relative duration and how they vary across different kinds of MOA.

### **Theory:**

There is stoppage of air flow in the vocal tract when stop consonant sounds are produced (hence the name stop consonant). The stop consonant are categorized by Place of Articulation (POA) and Manner of Articulation (MOA)

POA is categorized as Velar, Palatal, Alveolar, Dental and Bilabial. These categories show us where the stoppage of air flow occurs.

MOA is categorized as unvoiced unaspirated (UU), unvoiced aspirated (UA), voiced unaspirated (VU) and voiced aspirated (VA). These categories show how the excitation happens in the vocal tract (either voicing or aspiration takes place or not).

### **Procedure:**

We are analysing the velar sounds i.e. Ka, Kha, Ga, Gha. We record these sounds via the software *Audacity*, with 16 bits/sample and 16 kHz sampling frequency. We then plot time domain signal using user-defined function *plot\_time\_signal()* and frequency spectrum using user-defined function *plot\_freq\_spectrum()*. We use inbuilt matlab function *spectrogram()* to plot the spectrogram. We then analyse the plots and make some observations.

### **Code and Plots:**

```
close all; clc; clear all;  
close hidden
```

```
[ka_sound,fs]=audioread('ka.wav');
```

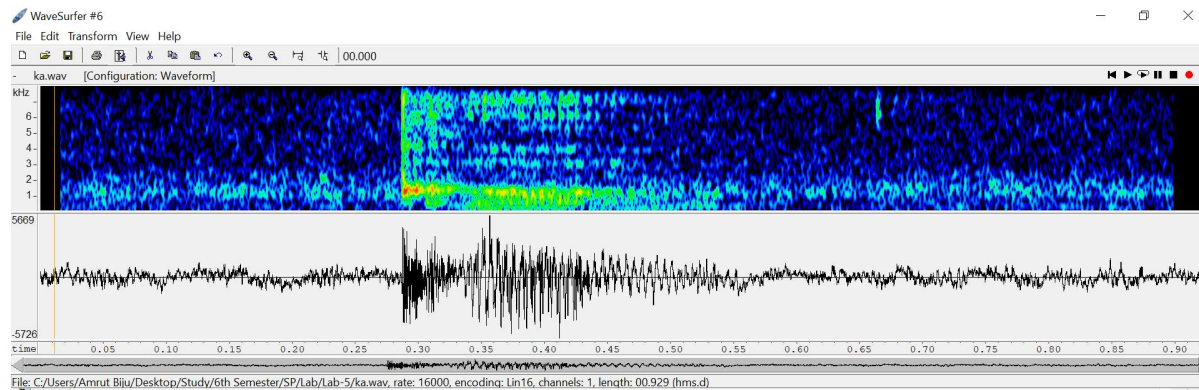
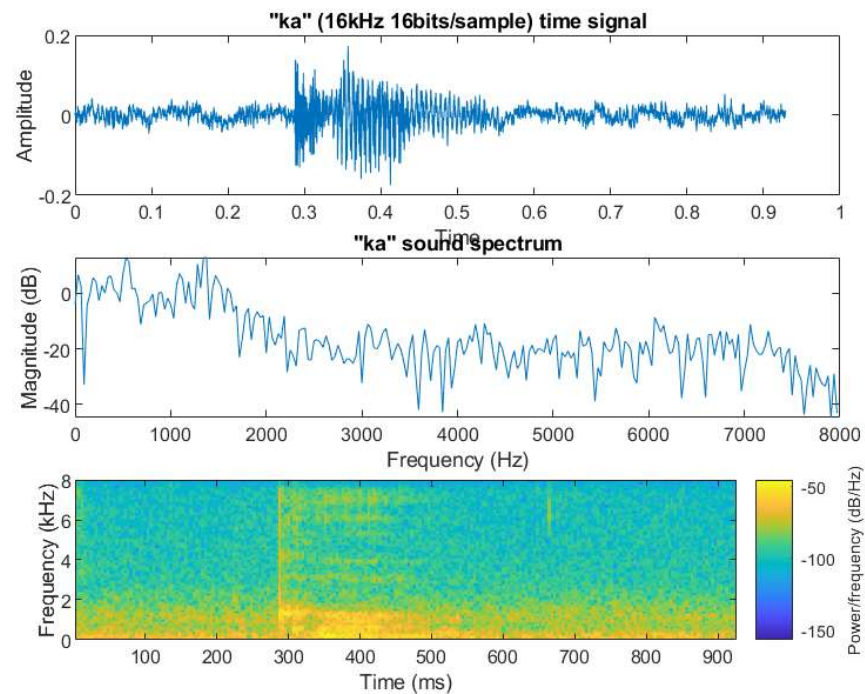
```
[kha_sound,fs]=audioread('kha.wav');
```

```
[ga_sound,fs]=audioread('ga.wav');
```

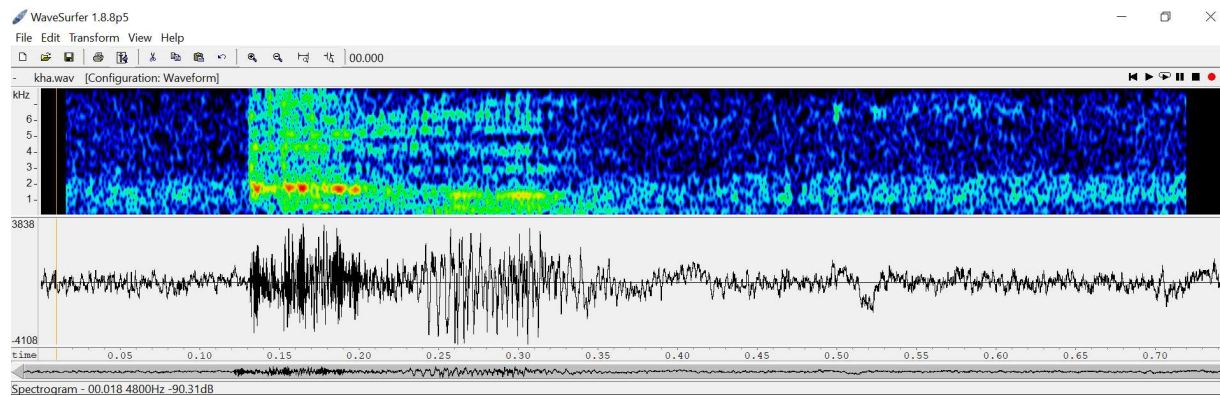
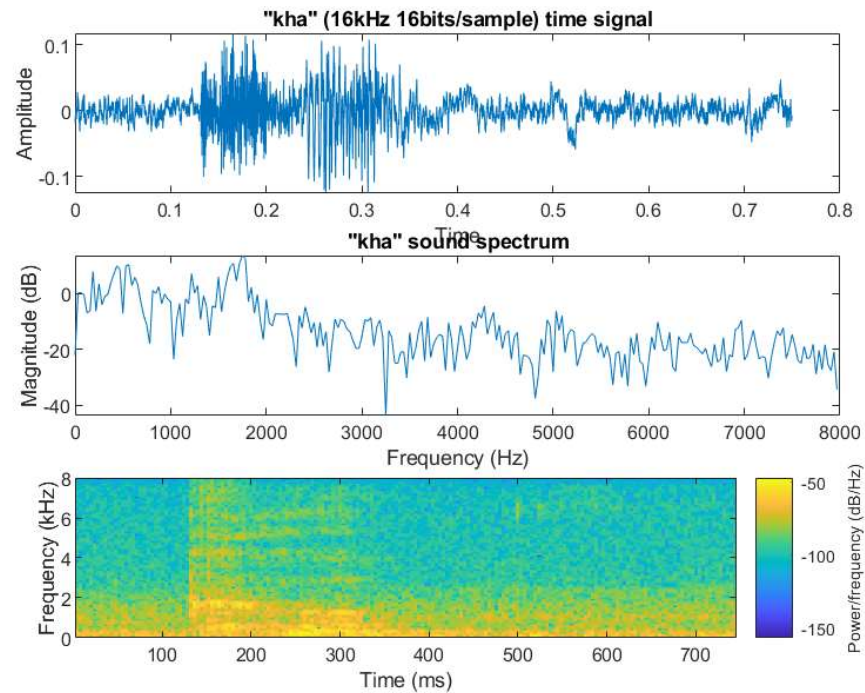
```
[gha_sound,fs]=audioread('gha.wav');
```

```
subplot(311);  
plot_time_signal(ka_sound, fs, '"ka" (16kHz 16bits/sample) time signal');  
subplot(312);  
plot_freq_spectrum(ka_sound((0.3*fs):(int64((0.3+0.025)*fs))), fs, '"ka" sound spectrum');  
subplot(313);  
spectrogram(ka_sound, 128, [], 512, fs, 'yaxis');
```

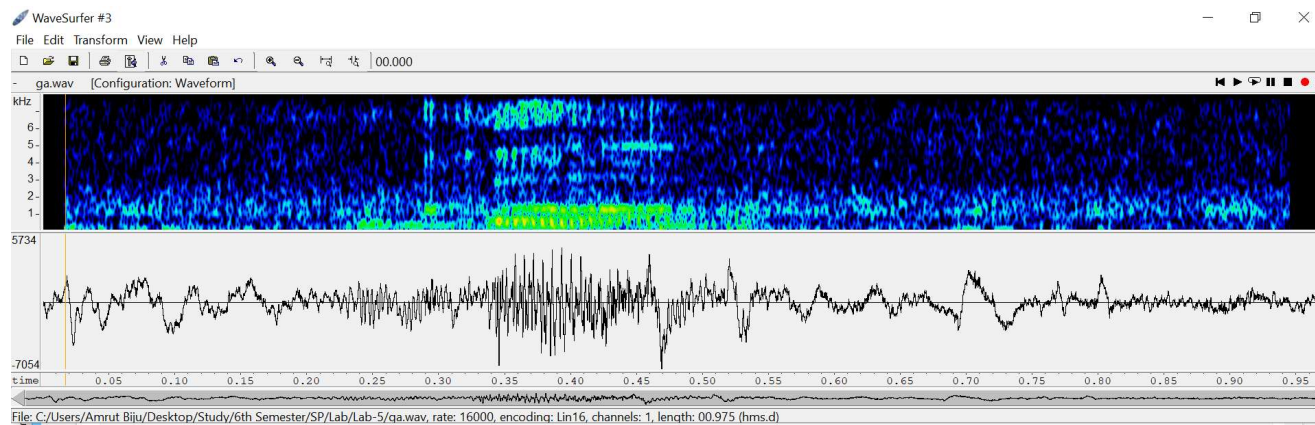
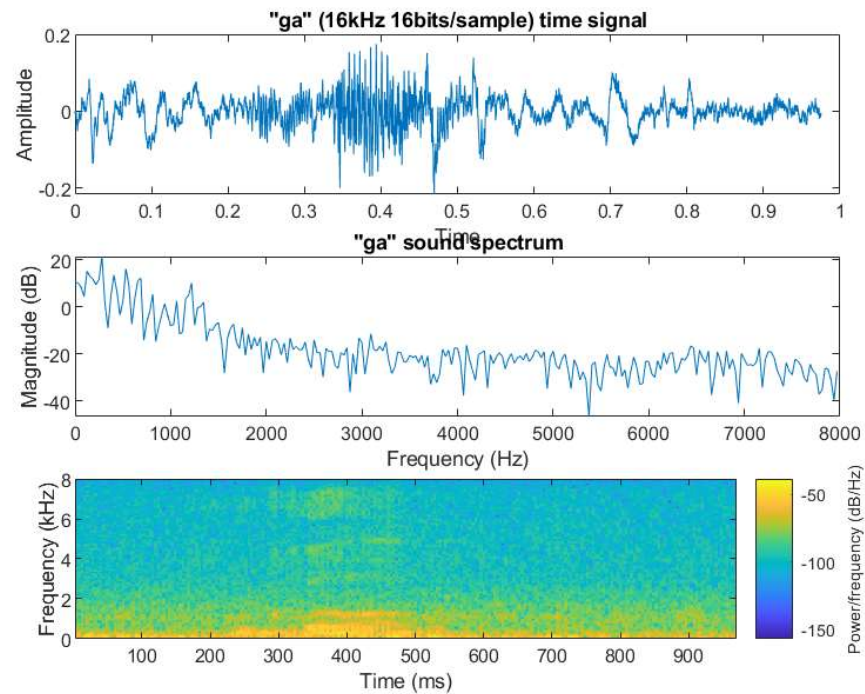




```
close hidden
subplot(311);
plot_time_signal(kha_sound, fs, '"kha" (16kHz 16bits/sample) time signal');
subplot(312);
plot_freq_spectrum(kha_sound((0.15*fs):(int64((0.15+0.025)*fs))), fs, '"kha" sound spectrum');
subplot(313);
spectrogram(kha_sound, 128, [], 512, fs, 'yaxis');
```

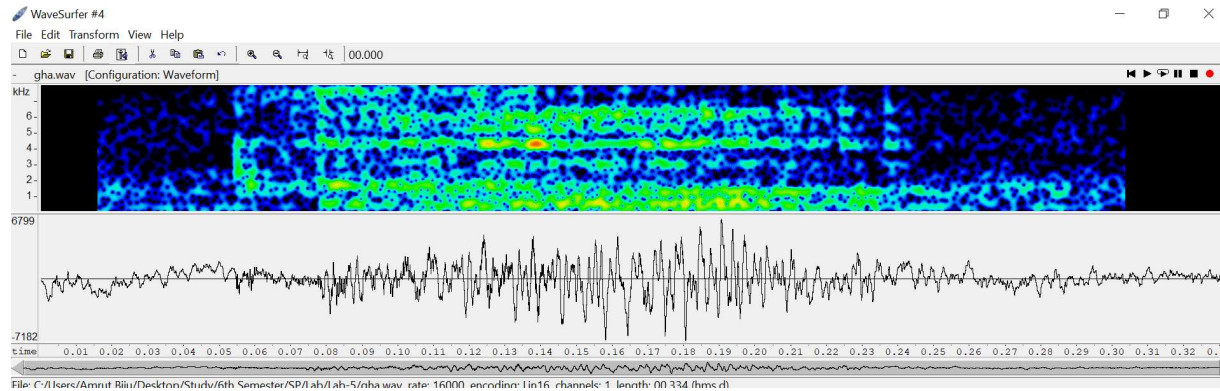
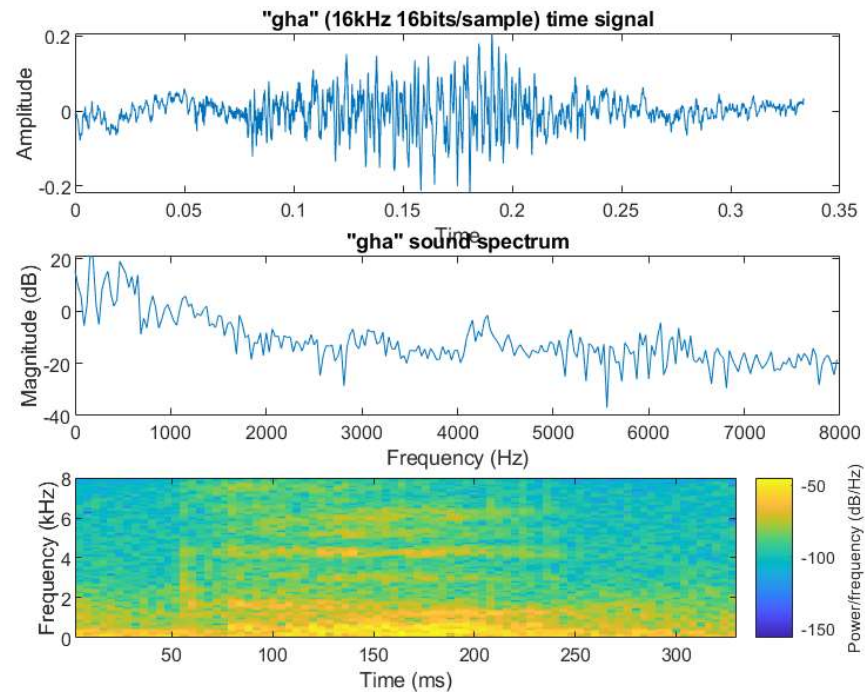


```
close hidden
subplot(311);
plot_time_signal(ga_sound, fs, "ga" (16kHz 16bits/sample) time signal');
subplot(312);
plot_freq_spectrum(ga_sound((0.4*fs):(int64((0.4+0.025)*fs))), fs, "ga" sound spectrum');
subplot(313);
spectrogram(ga_sound, 128, [], 512, fs, 'yaxis');
```



```
close hidden
subplot(311);
plot_time_signal(ga_sound, fs, '"ga" (16kHz 16bits/sample) time signal');
subplot(312);
plot_freq_spectrum(ga_sound((0.15*fs):(int64((0.15+0.025)*fs))), fs, '"ga" sound spectrum');
subplot(313);
spectrogram(ga_sound, 128, [], 512, fs, 'yaxis');
```





### Observations:

Ka: (UU) There is no periodicity (from time domain plot). We can see that there is a closure 1st then a burst (wide range of frequencies with considerable energy) is followed.

Kha: (UA) There is no periodicity (from time domain plot). We can see an additional event along with closure and burst, which is aspiration (a lot of frequency components with considerable present).

Ga: (VU) There is periodicity (from time domain plot). We see a lesser range of frequency components covered as compared to UU sounds (as here it is voiced). Here also we see a closure and burst but with lesser frequencies as compared to UU sounds.

Gha: (VA) There is periodicity (from time domain plot). We see a lesser range of frequency components covered as compared to UA sounds (as here it is voiced). Here also we see a closure, burst and aspiration but with lesser frequencies as compared to UA sounds.

## **Function Definitions**

```
function plot_time_signal(y, fs, title_name) % PLOT TIME DOMAIN PLOT FOR ANY SIGNAL
    plot((0:length(y)-1)/fs,y);
    xlabel("Time");
    ylabel("Amplitude");
    title(title_name);
end

% PLOT FREQUENCY SPECTRUM FOR ANY SIGNAL
function plot_freq_spectrum(x, fs, title_name)
    N = length(x);
    n = pow2(nextpow2(N)); % Transforming the length so that the number of samples is a power of 2.
    f = fs*[0: n-1]/n;
    X_mags = 20*log10(abs(fft(x, n)));
    N_2 = ceil(n/2);
    plot(f(1:N_2), (X_mags(1:N_2)));
    xlabel('Frequency (Hz)')
    ylabel('Magnititude (dB)');
    title(title_name);
end
```