ECE 251A End-Term Project: Excitation Signal Characterization

Wenyu Zhang

TOTAL POINTS

95 / 100

QUESTION 1

1 Topic II-B **95 / 100**

- **- 0 pts** Correct
- **- 5 pts** Figure for Part (a) incorrect/ missing - Time series of frame
- **- 5 pts** Figure for Part(a) incorrect/ missing - Conventional spectral estimates
- **- 5 pts** Figure for Part (a) incorrect/ missing - LPC spectral estimate
- **- 5 pts** Figure for Part (b) incorrect/ missing - Time series of excitation signal estimate
- **- 5 pts** Figure for Part (b) incorrect/ missing - Conventional spectral estimates
- **- 5 pts** Figure for Part (b) incorrect/ missing - autocorrelation functions of excitation signal estimates
- **- 5 pts** Figure for Part (c) incorrect/ missing - time series of entire data
- ✓ **- 5 pts** **Figure for Part (c) incorrect/ missing - time series of speech phrase**
- **- 5 pts** Figure for Part (c) incorrect/ missing - spectrogram of the speech phrase
- **- 5 pts** Figure for Part (c) incorrect/ missing - time-evolving LPC analysis plot
- **- 5 pts** Figure for Part (c) incorrect/ missing - pitch period vs time
- **- 5 pts** Figure for Part (c) incorrect/ missing - excitation signal power vs time
- **- 5 pts** Part(b) pitch period estimates of speech vowel frame missing/ incorrect
- **- 2 pts** pitch period unit missing
- **- 15 pts** didn't analyze at least two frame for each vowel
- **- 5 pts** Not enough analysis in results part

End-Term Project: Excitation Signal Characterization

Wenyu Zhang
ECE 251A

March 22, 2018

1 Text

1.1 Objective

In this project, we generate excitation signals of speech series and plot the pitch period and power with evolving time to extract the voiced frames and analyze their characteristics.

1.2 Background

Speech signal can be classified into voiced, unvoiced and silence regions. The near periodic vibration of vocal folds is excitation for the production of voiced speech. The random like excitation is present for unvoiced speech. There is no excitation during silence region.
Majority of speech regions are voiced in nature that include vowels, semivowels and other voiced components. The voiced regions looks like a near periodic signal in the time domain representation. The periodicity associated with such segments is defined as 'pitch period T_0' in the time domain and 'Pitch frequency or Fundamental Frequency F_0'. Pitch is an important attribute of voiced speech. It contains speaker-specific information. It is also needed for speech coding task. Thus estimation of pitch is one of the important issue in speech processing.
There are a large set of methods that have been developed in the speech processing area for the estimation of pitch. The most typical among them is single inverse filtering technique (SIFT) pitch estimation, a method derived from autocorrelation method but introducing excitation signal [1].

1.3 Approach

In this project, we follow steps as below:
1. For the vowel /i/ and /u/, plot the original time series, its power spectrum (conventional analysis (dB) using a single $|\text{FFT}|^2$), and LPC estimate of its power spectrum.
2. Pass the original speech waveform through the inverse filter, thus extracting an estimate of the excitation signal. Plot the estimate of the excitation signal and compute both its power spectrum (conventional analysis (dB) using a single $|\text{FFT}|^2$) and autocorrelation function. Determine the pitch period of the excitation signal.
3. Extend this analysis to the phrase "we were away" (speaker JCN), generating a plot of pitch period vs. time (use 50% overlapped, 256-point frames). Also, plot the excitation signal power (dB) vs. time. Include a spectrogram (with color bar (dB)) of the phrase providing a time-evolving conventional spectral analysis look at the data.

1.4 Results and Analysis

1.4.1 Part A. Conventional and LPC Spectrum Estimate of Vowels

1. 256-point time series
The plot of 256-point series of vowel /i/ and /u/ are shown in **Figure 1**.
We pick two frames(n=6000:6255 and n=256:511) of each vowel to verify pitch period as a good property of vowel regardless of time shift.
Comment:
From **Figure 1**, we can see that the speech series is periodic, which indicates existence of frequency component on spectrum.

2. Conventional Power Spectrum of 256-point FFT

The plot of magnitude of 256-point series' conventional power spectrum is shown in **Figure 2**.
Comment:
With theoretical frequency components we know, we can check the peaks at certain frequency. From the plot, we can see the peaks exist near the theoretical value. However, as personal influence and noise to the sample, there are some ripples that vibrate through the whole frequency interval.

3. LPC Estimate of Power Spectrum
The plot of LPC Power Spectrum is shown in **Figure 3**.
Comment:
Compared it with **Figure 2**, we can see the inverse filtering extract the frequency components we want near the theoretical frequency values without ripples, which is similar to that in Homework 8.

1.4.2 Part B. Excitation Signal Analysis of Vowels

1. Time Series of Excitation Signal
We pass the speech series through its inverse filter to generate the 256-point excitation signal, which is shown in **Figure 4**.
Comment:
Compare **Figure 4** with **Figure 1**, we can see the excitation signal is also periodical with the same period as its original time series. In fact, excitation signal is the output of inverse filter(a LTI system) with input of original time series.

2. Conventional Power Spectrum of Excitation Signal
The plot of conventional power spectrum is shown in **Figure 5**.
Comment:
The conventional power spectrum is like white noise with no certain frequency components. The effect of the inverse filter is to transform the input signal into the best estimate of white noise (in the least squares sense) [2].
Theoretically, it is seen that the result is roughly a white noise spectrum with a periodic component superimposed upon it as Fig.1(E) in [2]. However, our plot is not that perfect to observe the period.

3. Autocorrelation of Excitation Signal and Pitch Period Detection
The plot of autocorrelation of excitation signal is shown in **Figure 6**. To find the pitch period, we specifically look into its first circle(i.e. 50-150Lags or 5-15ms) and try to find the peak, which is shown in **Figure 7**.
Comment:
Figure 6 is similar to fig.1(F) in [2]. Blow are the pitch periods from finding peaks.
Vowel /i/ has peak at 88Lags, i.e. pitch period 8.8ms in frame n = 6000:6255.
Vowel /i/ has peak at 86Lags, i.e. pitch period 8.6ms in frame n = 256:511.
Vowel /u/ has peak at 96Lags, i.e. pitch period 9.6ms in frame n = 6000:6255.
Vowel /u/ has peak at 92Lags, i.e. pitch period 9.2ms in frame n = 256:511.

1.4.3 Part C. Excitation Analysis of Phrase "we were away"

1. Time Series, Power Spectrogram and LPC Analysis of Phrase
The 15360-point time series, Spectrogram and LPC Analysis of Phrase are separately shown in **Figure 8**, **Figure 9** and **Figure 10**.
Comment:
From **Figure 8**, we can see that the beginning of the phrase is unvoiced with little amplitude. And

three components after it represent three words of the phrase.
From **Figure 9**, we can see that the power spectrum density of certain time frame, the yellow parts indicate high power density, i.e. the power spectral property of certain word.
According to **Figure 8** and **Figure 9**, we can cut off the beginning part of the phrase series to extract the voiced speech. As a result, we keep 4201-14440 points of the phrase series, i.e. 0.42s-1.4439s, and get 79 frames with 256-point length and 50% overlap.
Then we do LPC analysis on it and draw color-bar graph. From **Figure 10**, we can see the LPC power estimate of each frame. The overall shape yields spectral characteristics in **Figure 9**.

2. Pitch Period vs. Evolving Time
The plot of pitch period vs. evolving time is shown in **Figure 11**. We draw both scatter graph and line graph.
Comment:
Overall, the pitch period increases with time evolution from 6.5ms to 9.5 ms.
To detect pitch period of different frames, we should set threshold first. And considering different frames may have different autocorrelation amplitude, we set the threshold as dynamic, which is 80% of the peak, and if there are more than 2 locations over the threshold, we say this frame is unvoiced and set pitch period to 0. With this method, we get the scatter graph.
From the scatter graph, we see 0 period at beginning and ending of evolving interval, this is obvious since those frames are close to unvoiced part that we have cut, so it is the residual of unvoiced part. Also note that the 27th and 28th frame are also 0 period. To illustrate that, see that those two frames are word transition from ”we” to ”were”. So the power of voice is small and not enough to exceeds noise a lot. So there are more locations over 80% peak.
However, to ensure the continuity of line graph, we just take the highest value as peak ignoring that there are others with near amplitude to the peak. With this change, we get the line graph.

3. Power(dB) vs. Evolving Time
The plot of power(dB) vs. evolving time is shown in **Figure 12**.
Comment:
From **Figure 12**, we can see that there are three power components, mapping the three words. The power of word ”were” are more averaged while ”we” and ”away” have significant peak and rapid power change.

1.5 Summary

In this project, we extend the speech process to analysis of excitation signal. With excitation signal, we can better separate the voiced and unvoiced part compared with autocorrelation just from original time series. We also explain the phrase and words from time, power density and averaged power perspectives. Finally, We developed a method to detect peak and return pitch period dynamically.

2 Plots

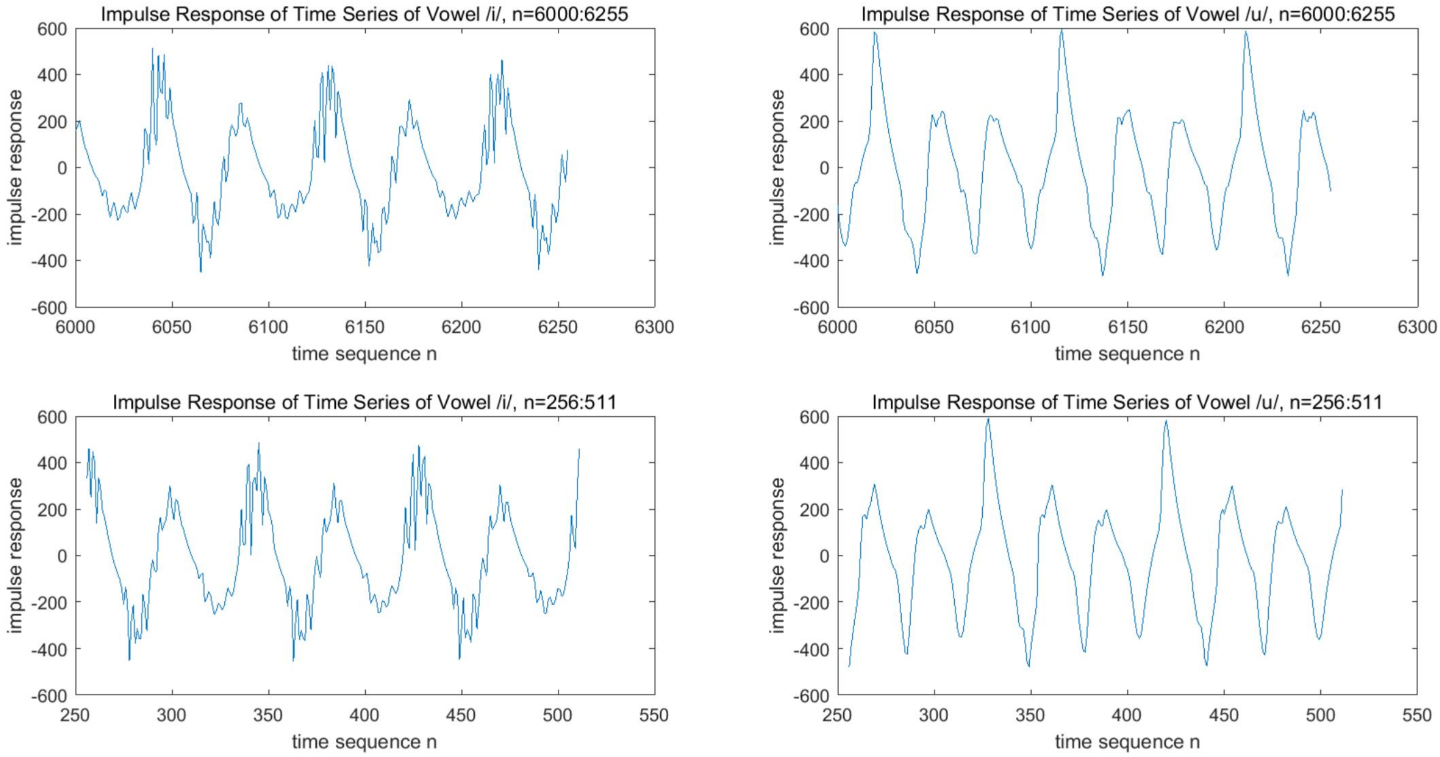

Figure 1: Time Series of Vowels /i/ /u/

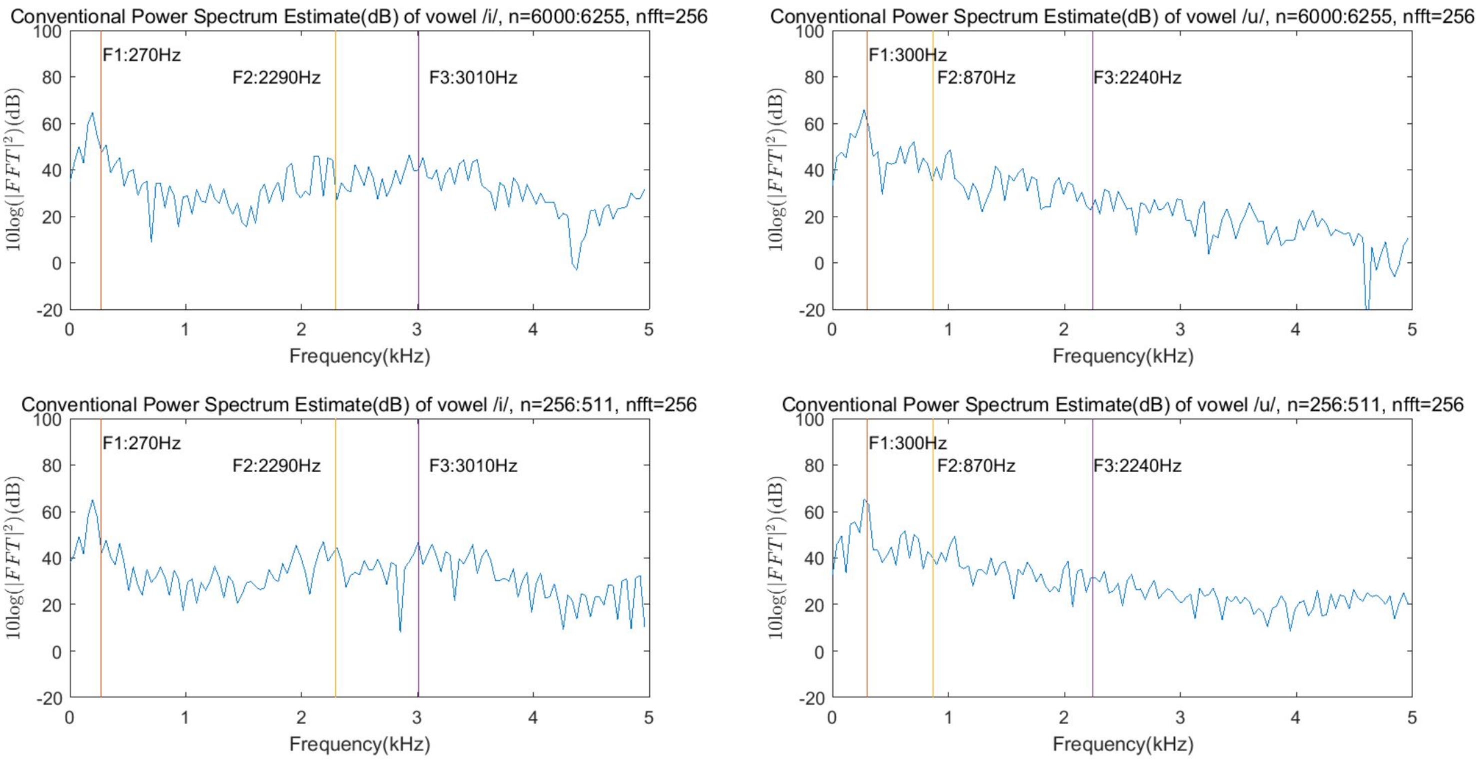

Figure 2: Conventional Power Spectrum of Vowels /i/ /u/

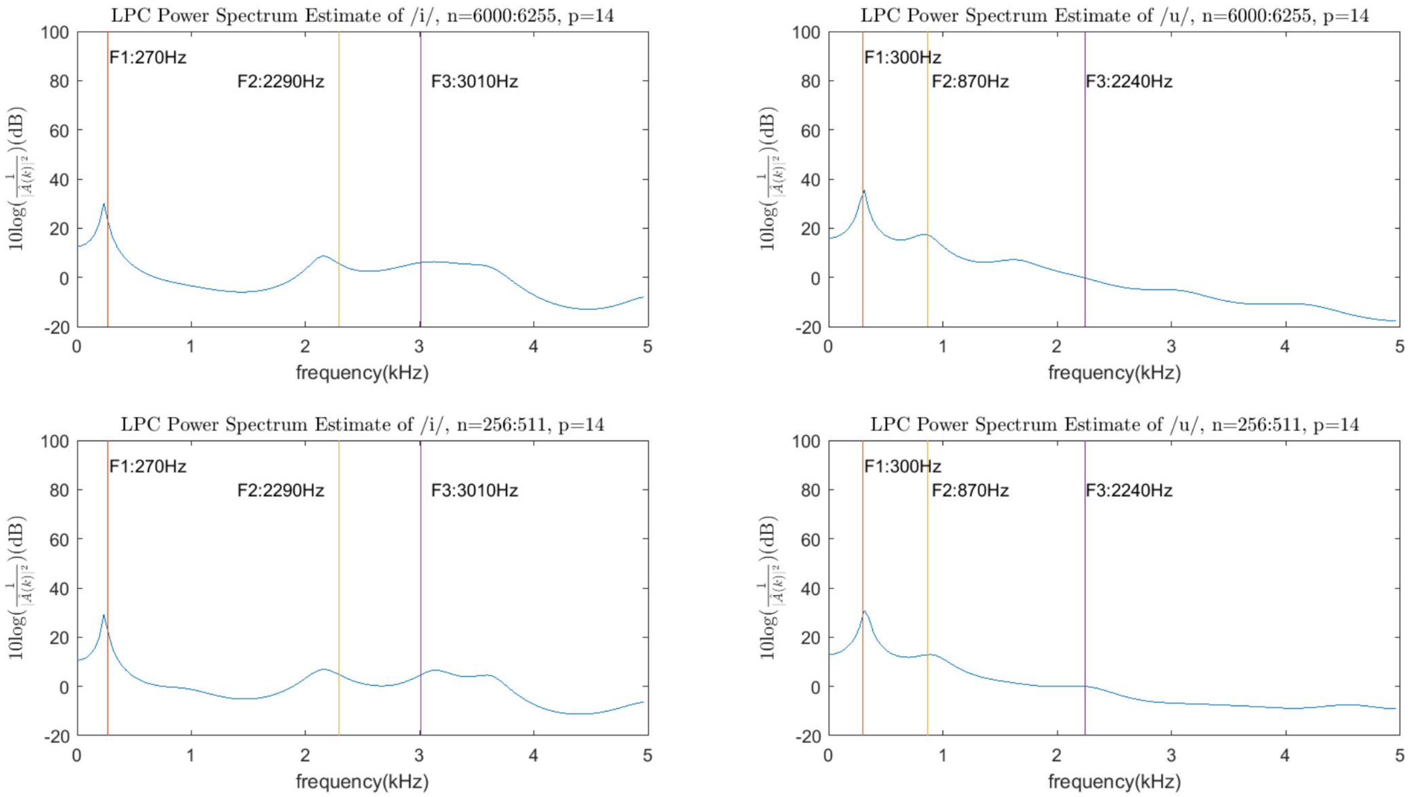

Figure 3: LPC Power Spectrum of Vowels /*i*/ /*u*/

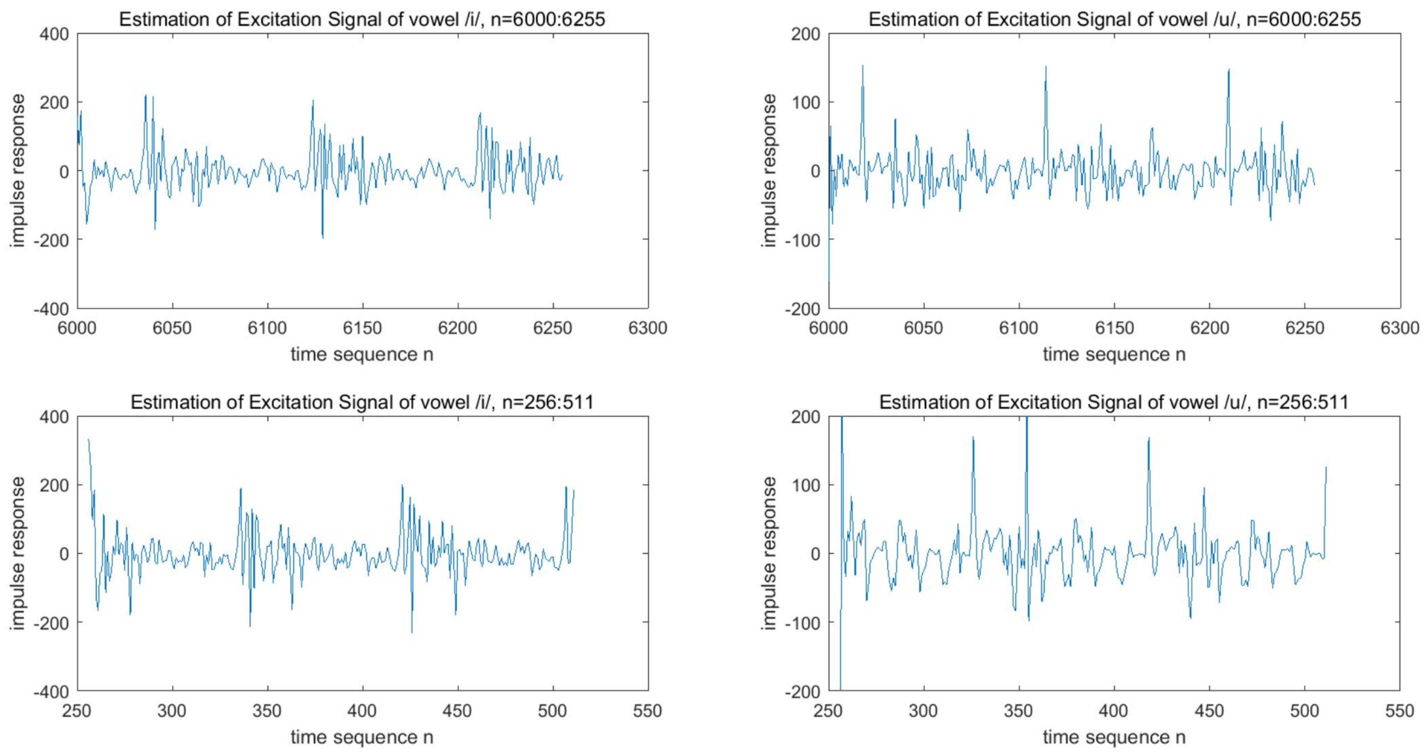

Figure 4: Excitation Signal of Vowels /*i*/ /*u*/

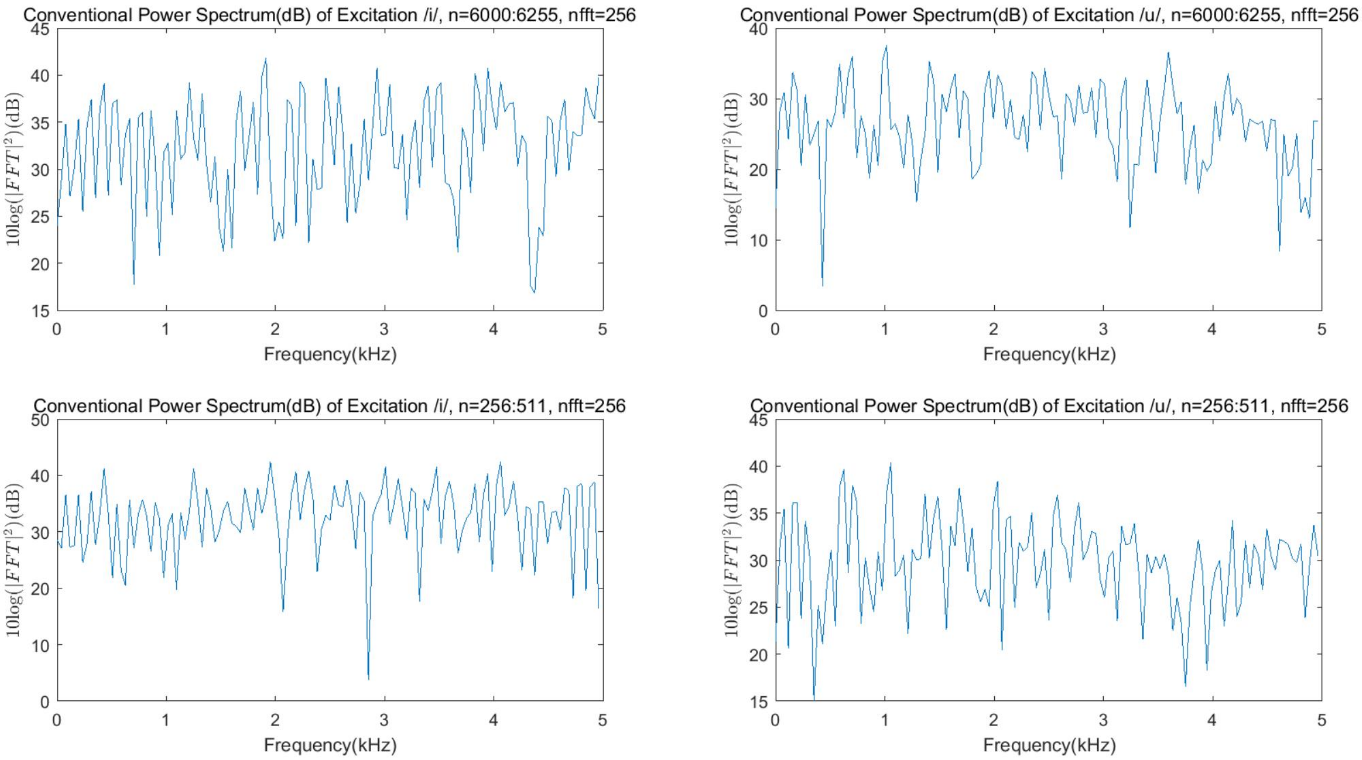

Figure 5: Conventional Power Spectrum of Excitation of Vowels /i/ /u/

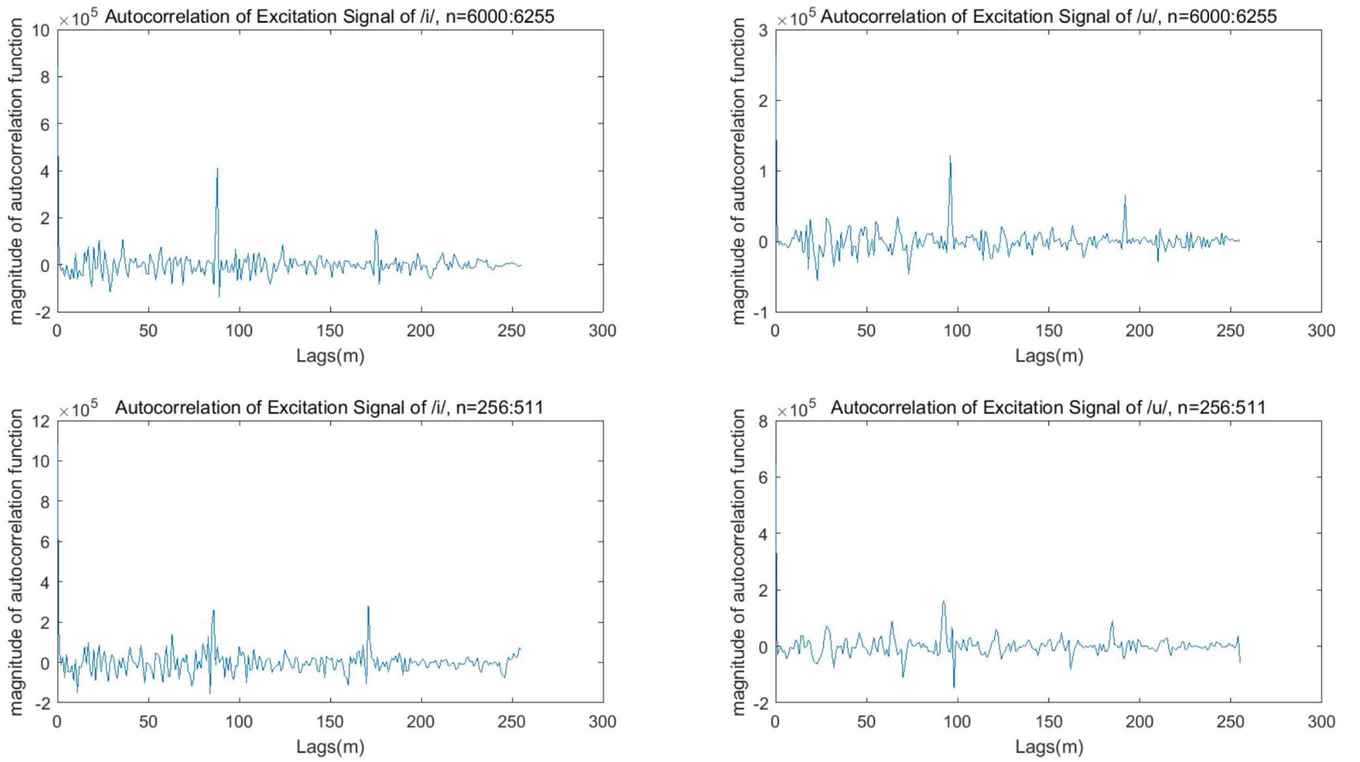

Figure 6: Autocorrelation of Excitation of /i/ /u/7

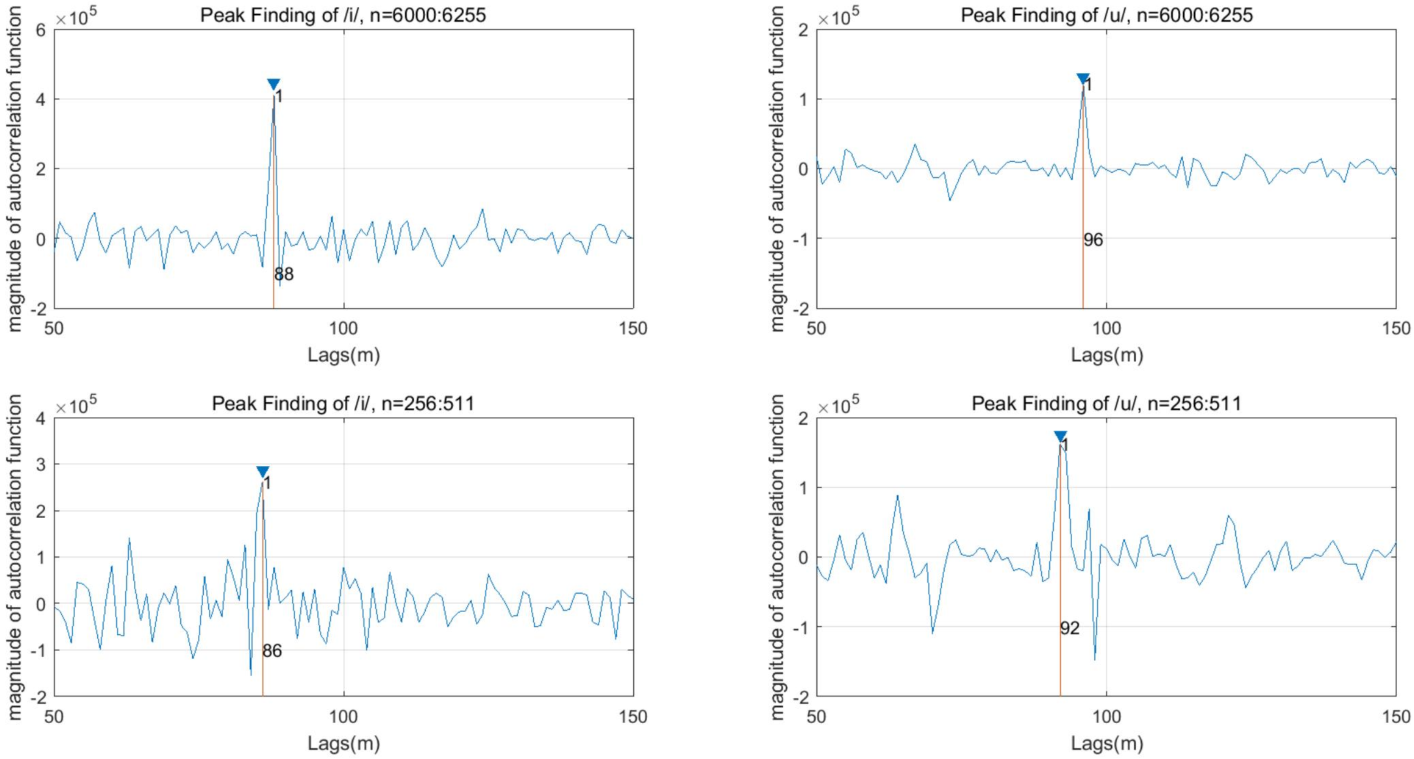

Figure 7: Peak Finding of Autocorrelation of Vowels /i/ /u/

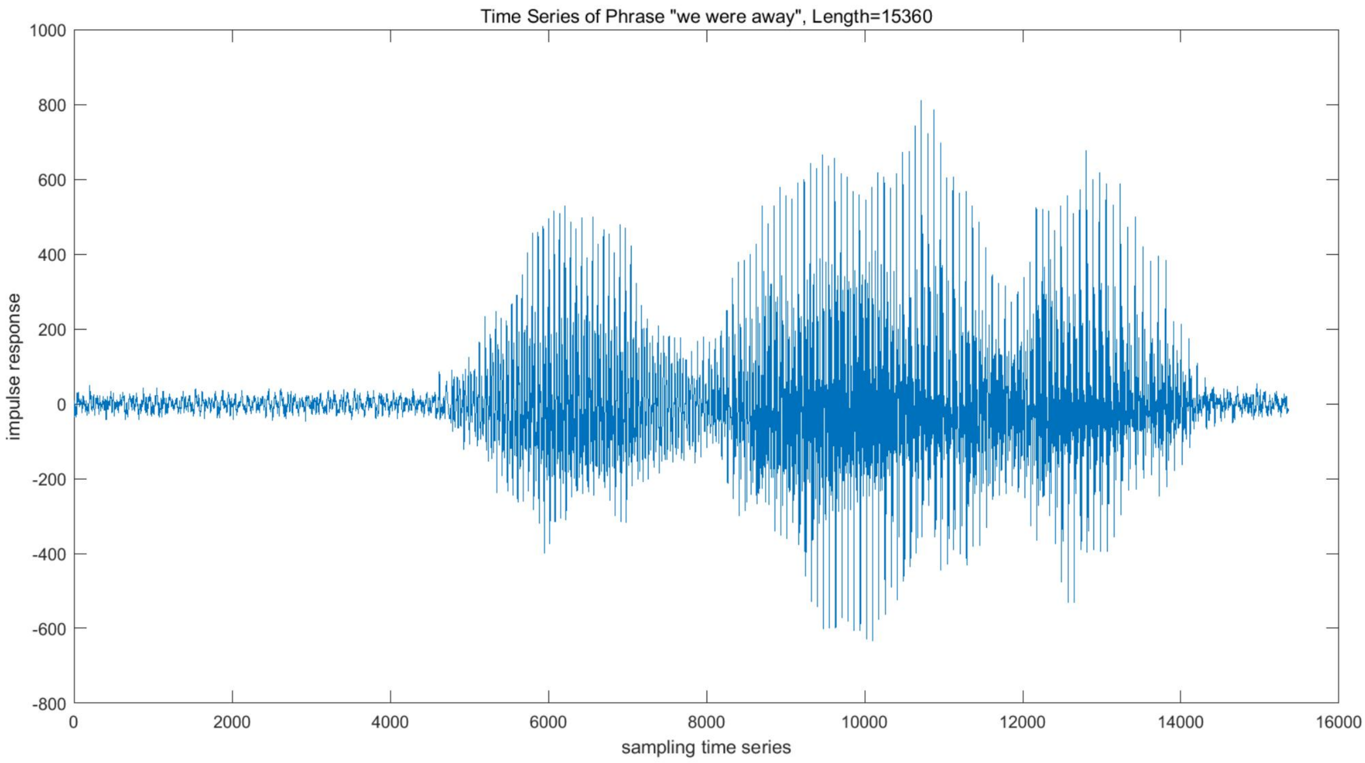

Figure 8: Time Series of Whole Phrase

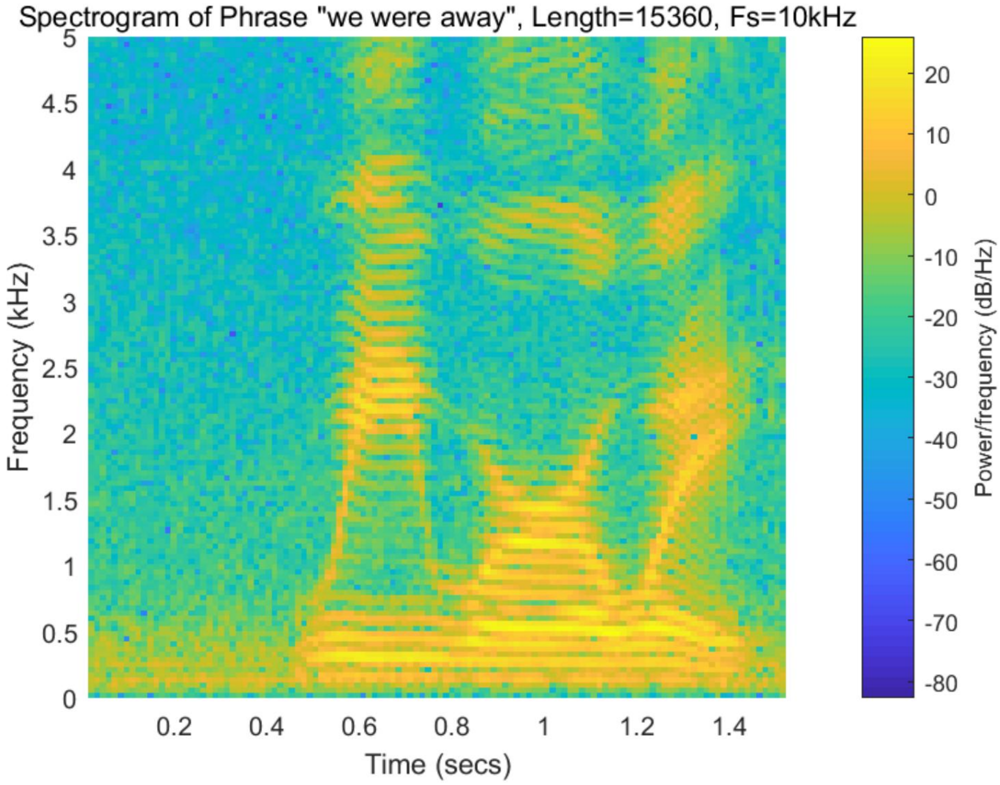

Figure 9: Power Spectrogram of Whole Phrase

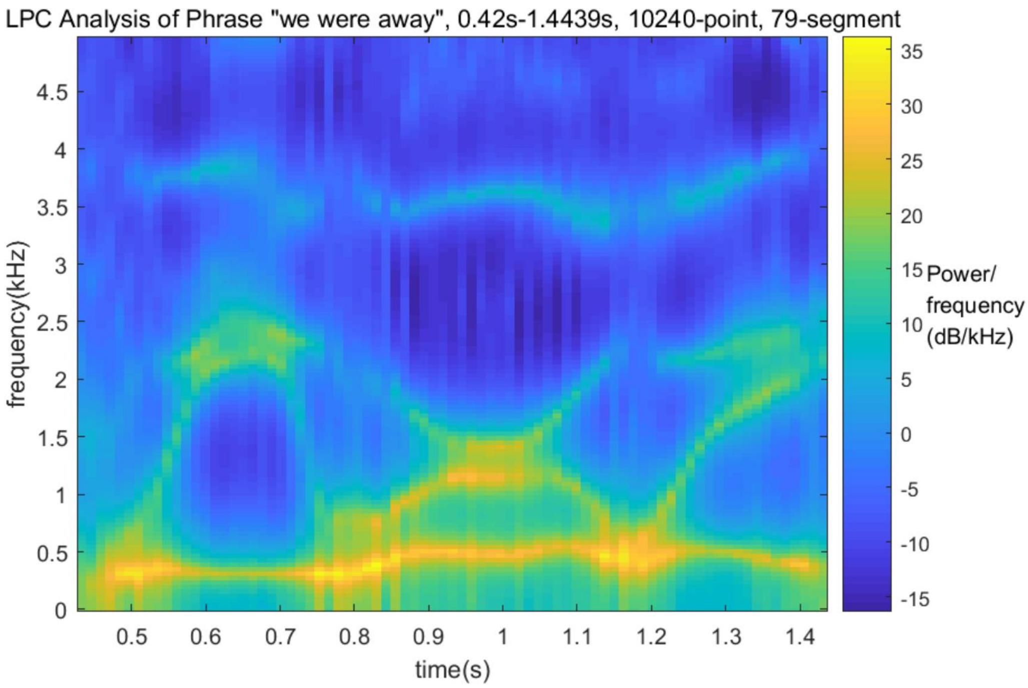

Figure 10: LPC Power Spectrum of Voiced Frames with Color Bar

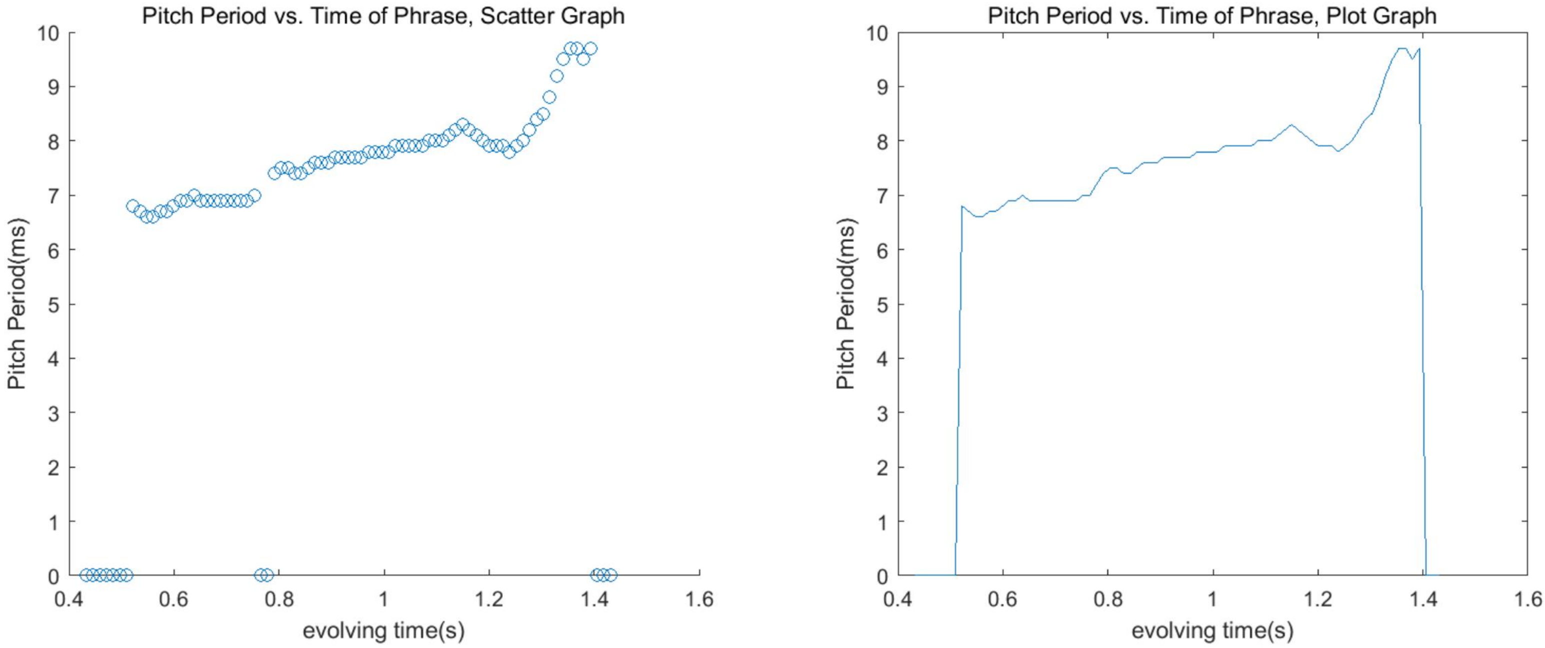

Figure 11: Pitch Period vs. Evolving Time

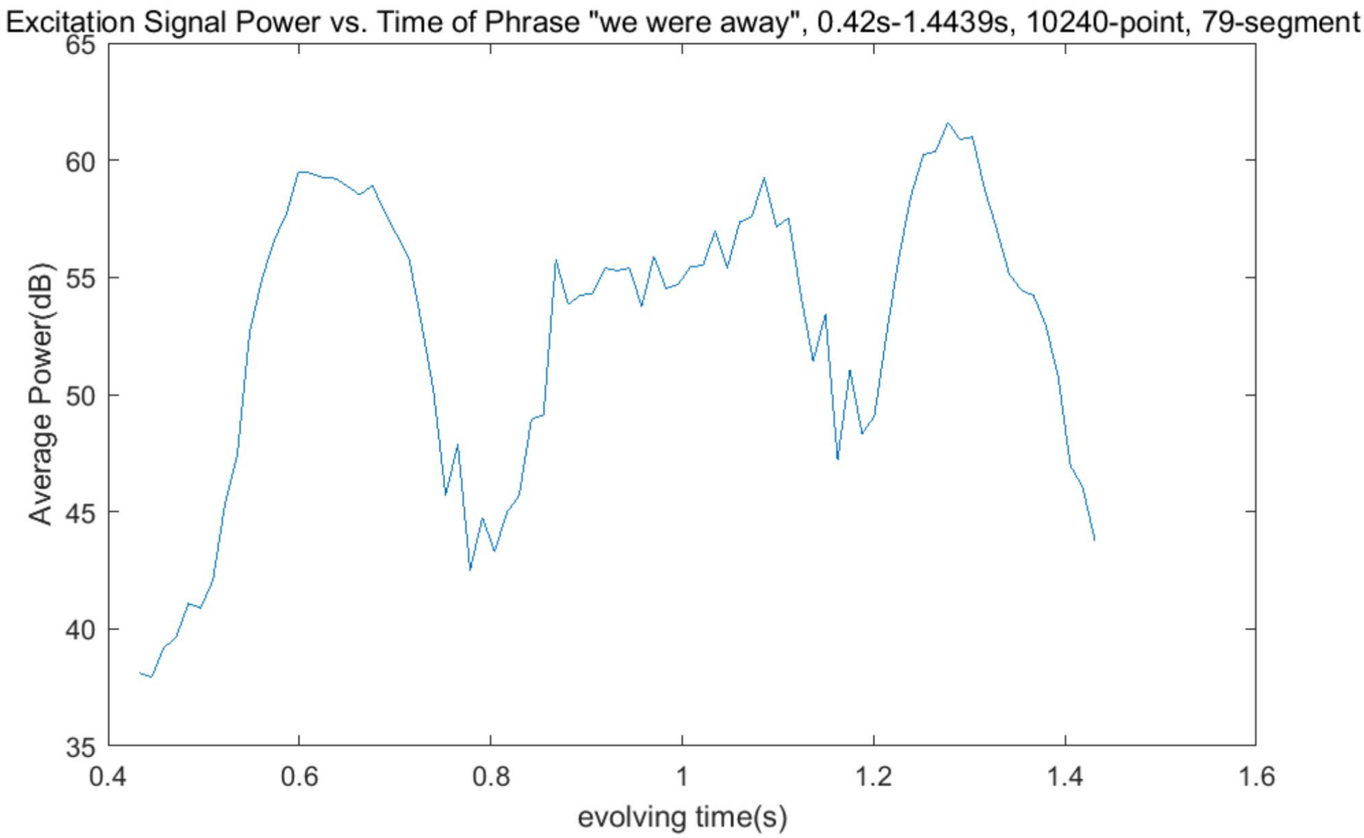

Figure 12: Power(dB) vs. Evolving Time

3 Appendix

```
% End-Term Project: Excitation Signal Characterization

% Excitation Signal Generation
I1 = xlsread('phonemes.xlsx','A6001:A6256');
U1 = xlsread('phonemes.xlsx','B6001:B6256');
I2 = xlsread('phonemes.xlsx','A257:A512');
U2 = xlsread('phonemes.xlsx','B257:B512');
% plot(a)
% Time Series
n1 = 6000:6255;
n2 = 256:511;
figure(1);
subplot(2,2,1);
plot(n1,I1);
xlabel('time sequence n');
ylabel('impulse response');
title('Impulse Response of Time Series of Vowel /i/, n=6000:6255');
subplot(2,2,2);
plot(n1,U1);
xlabel('time sequence n');
ylabel('impulse response');
title('Impulse Response of Time Series of Vowel /u/, n=6000:6255');
subplot(2,2,3);
plot(n2,I2);
xlabel('time sequence n');
ylabel('impulse response');
title('Impulse Response of Time Series of Vowel /i/, n=256:511');
subplot(2,2,4);
plot(n2,U2);
xlabel('time sequence n');
ylabel('impulse response');
title('Impulse Response of Time Series of Vowel /u/, n=256:511');
% Conventional Power Spectrum
nfft = 256;
N = 0:127;
Nf = N*10/256;
p_i1 = 0.5*mag2db(pwelch(I1,nfft,0,nfft,1,'centered'));
p_u1 = 0.5*mag2db(pwelch(U1,nfft,0,nfft,1,'centered'));
p_i2 = 0.5*mag2db(pwelch(I2,nfft,0,nfft,1,'centered'));
p_u2 = 0.5*mag2db(pwelch(U2,nfft,0,nfft,1,'centered'));
figure(2)
subplot(2,2,1);
plot(Nf,p_i1(129:256),[0.27,0.27],[-20,100],[2.29,2.29],[-
20,100],[3.01,3.01],[-20,100]);
text(0.28,90,'F1:270Hz');
text(1.4,80,'F2:2290Hz');
text(3.1,80,'F3:3010Hz');
xlabel('Frequency(kHz)');
ylabel('10log($|FFT|^2$)(dB)','Interpreter','latex','FontSize',11);
```

```
title('Conventional Power Spectrum Estimate(dB) of vowel /i/, n=6000:6255,
nfft=256');
subplot(2,2,2);
plot(Nf,p_u1(129:256),[0.3,0.3],[-20,100],[0.87,0.87],[-
20,100],[2.24,2.24],[-20,100]);
text(0.31,90,'F1:300Hz');
text(0.9,80,'F2:870Hz');
text(2.25,80,'F3:2240Hz');
xlabel('Frequency(kHz)');
ylabel('10log($|FFT|^2$)(dB)','Interpreter','latex','FontSize',11);
title('Conventional Power Spectrum Estimate(dB) of vowel /u/, n=6000:6255,
nfft=256');
subplot(2,2,3);
plot(Nf,p_i2(129:256),[0.27,0.27],[-20,100],[2.29,2.29],[-
20,100],[3.01,3.01],[-20,100]);
text(0.28,90,'F1:270Hz');
text(1.4,80,'F2:2290Hz');
text(3.1,80,'F3:3010Hz');
xlabel('Frequency(kHz)');
ylabel('10log($|FFT|^2$)(dB)','Interpreter','latex','FontSize',11);
title('Conventional Power Spectrum Estimate(dB) of vowel /i/, n=256:511,
nfft=256');
subplot(2,2,4);
plot(Nf,p_u2(129:256),[0.3,0.3],[-20,100],[0.87,0.87],[-
20,100],[2.24,2.24],[-20,100]);
text(0.31,90,'F1:300Hz');
text(0.9,80,'F2:870Hz');
text(2.25,80,'F3:2240Hz');
xlabel('Frequency(kHz)');
ylabel('10log($|FFT|^2$)(dB)','Interpreter','latex','FontSize',11);
title('Conventional Power Spectrum Estimate(dB) of vowel /u/, n=256:511,
nfft=256');
% LPC Estimate of Power Spectrum
% p=14
I1_win = I1.*kaiser(256,1.5);
U1_win = U1.*kaiser(256,1.5);
I2_win = I2.*kaiser(256,1.5);
U2_win = U2.*kaiser(256,1.5);
a_i1 = zeros(1,256);
[a_i1(1:15),~] = lpc(I1_win,14);
N = 0:127;
Nf = N*10/256;
A_i1 = fft(a_i1,256);
L_i1 = zeros(1,128);
for i = 1:128
    L_i1(i) = 0.5*mag2db(1/((abs(A_i1(i)))^2));
end
a_u1 = zeros(1,256);
[a_u1(1:15),~] = lpc(U1_win,14);
N = 0:127;
Nf = N*10/256;
A_u1 = fft(a_u1,256);
L_u1 = zeros(1,128);
for i = 1:128
```

```
    L_u1(i) = 0.5*mag2db(1/((abs(A_u1(i)))^2));
end
a_i2 = zeros(1,256);
[a_i2(1:15),~] = lpc(I2_win,14);
N = 0:127;
Nf = N*10/256;
A_i2 = fft(a_i2,256);
L_i2 = zeros(1,128);
for i = 1:128
    L_i2(i) = 0.5*mag2db(1/((abs(A_i2(i)))^2));
end
a_u2 = zeros(1,256);
[a_u2(1:15),~] = lpc(U2_win,14);
N = 0:127;
Nf = N*10/256;
A_u2 = fft(a_u2,256);
L_u2 = zeros(1,128);
for i = 1:128
    L_u2(i) = 0.5*mag2db(1/((abs(A_u2(i)))^2));
end
figure(3)
subplot(2,2,1)
plot(Nf,L_i1,[0.27,0.27],[-20,100],[2.29,2.29],[-20,100],[3.01,3.01],[-
20,100]);
text(0.28,90,'F1:270Hz');
text(1.4,80,'F2:2290Hz');
text(3.1,80,'F3:3010Hz');
xlabel('frequency(kHz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)(dB)','Interpreter','latex','Font
Size',11);
title('LPC Power Spectrum Estimate of /i/, n=6000:6255,
p=14','Interpreter','latex');
subplot(2,2,2)
plot(Nf,L_u1,[0.3,0.3],[-20,100],[0.87,0.87],[-20,100],[2.24,2.24],[-
20,100]);
text(0.31,90,'F1:300Hz');
text(0.9,80,'F2:870Hz');
text(2.25,80,'F3:2240Hz');
xlabel('frequency(kHz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)(dB)','Interpreter','latex','Font
Size',11);
title('LPC Power Spectrum Estimate of /u/, n=6000:6255,
p=14','Interpreter','latex');
subplot(2,2,3)
plot(Nf,L_i2,[0.27,0.27],[-20,100],[2.29,2.29],[-20,100],[3.01,3.01],[-
20,100]);
text(0.28,90,'F1:270Hz');
text(1.4,80,'F2:2290Hz');
text(3.1,80,'F3:3010Hz');
xlabel('frequency(kHz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)(dB)','Interpreter','latex','Font
Size',11);
title('LPC Power Spectrum Estimate of /i/, n=256:511,
p=14','Interpreter','latex');
```

```
subplot(2,2,4)
plot(Nf,L_u2,[0.3,0.3],[-20,100],[0.87,0.87],[-20,100],[2.24,2.24],[-
20,100]);
text(0.31,90,'F1:300Hz');
text(0.9,80,'F2:870Hz');
text(2.25,80,'F3:2240Hz');
xlabel('frequency(kHz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)(dB)','Interpreter','latex','Font
Size',11);
title('LPC Power Spectrum Estimate of /u/, n=256:511,
p=14','Interpreter','latex');

% plot(b)
% Excitation Signal
I1_F = filter(a_i1(1:15),1,I1);
U1_F = filter(a_u1(1:15),1,U1);
I2_F = filter(a_i2(1:15),1,I2);
U2_F = filter(a_u2(1:15),1,U2);
figure(4);
subplot(2,2,1);
plot(n1,I1_F);
xlabel('time sequence n');
ylabel('impulse response');
title('Estimation of Excitation Signal of vowel /i/, n=6000:6255');
subplot(2,2,2);
plot(n1,U1_F);
xlabel('time sequence n');
ylabel('impulse response');
title('Estimation of Excitation Signal of vowel /u/, n=6000:6255');
subplot(2,2,3);
plot(n2,I2_F);
xlabel('time sequence n');
ylabel('impulse response');
title('Estimation of Excitation Signal of vowel /i/, n=256:511');
subplot(2,2,4);
plot(n2,U2_F);
xlabel('time sequence n');
ylabel('impulse response');
title('Estimation of Excitation Signal of vowel /u/, n=256:511');
% Conventional Power Spectrum
nfft = 256;
N = 0:127;
Nf = N*10/256;
p_i1_F = 0.5*mag2db(pwelch(I1_F,nfft,0,nfft,1,'centered'));
p_u1_F = 0.5*mag2db(pwelch(U1_F,nfft,0,nfft,1,'centered'));
p_i2_F = 0.5*mag2db(pwelch(I2_F,nfft,0,nfft,1,'centered'));
p_u2_F = 0.5*mag2db(pwelch(U2_F,nfft,0,nfft,1,'centered'));
figure(5)
subplot(2,2,1);
plot(Nf,p_i1_F(129:256),[0.27,0.27],[-20,100],[2.29,2.29],[-
20,100],[3.01,3.01],[-20,100]);
text(0.28,90,'F1:270Hz');
text(1.4,80,'F2:2290Hz');
text(3.1,80,'F3:3010Hz');
```

```
xlabel('Frequency(kHz)');
ylabel('10log($|FFT|^2$)(dB)','Interpreter','latex','FontSize',11);
title('Conventional Power Spectrum(dB) of Excitation /i/, n=6000:6255,
nfft=256');
subplot(2,2,2);
plot(Nf,p_u1_F(129:256),[0.3,0.3],[-20,100],[0.87,0.87],[-
20,100],[2.24,2.24],[-20,100]);
text(0.31,90,'F1:300Hz');
text(0.9,80,'F2:870Hz');
text(2.25,80,'F3:2240Hz');
xlabel('Frequency(kHz)');
ylabel('10log($|FFT|^2$)(dB)','Interpreter','latex','FontSize',11);
title('Conventional Power Spectrum(dB) of Excitation /u/, n=6000:6255,
nfft=256');
subplot(2,2,3);
plot(Nf,p_i2_F(129:256),[0.27,0.27],[-20,100],[2.29,2.29],[-
20,100],[3.01,3.01],[-20,100]);
text(0.28,90,'F1:270Hz');
text(1.4,80,'F2:2290Hz');
text(3.1,80,'F3:3010Hz');
xlabel('Frequency(kHz)');
ylabel('10log($|FFT|^2$)(dB)','Interpreter','latex','FontSize',11);
title('Conventional Power Spectrum(dB) of Excitation /i/, n=256:511,
nfft=256');
subplot(2,2,4);
plot(Nf,p_u2_F(129:256),[0.3,0.3],[-20,100],[0.87,0.87],[-
20,100],[2.24,2.24],[-20,100]);
text(0.31,90,'F1:300Hz');
text(0.9,80,'F2:870Hz');
text(2.25,80,'F3:2240Hz');
xlabel('Frequency(kHz)');
ylabel('10log($|FFT|^2$)(dB)','Interpreter','latex','FontSize',11);
title('Conventional Power Spectrum(dB) of Excitation /u/, n=256:511,
nfft=256');
% Autocorrelation Function
C_I1 = xcorr(I1_F);
C_U1 = xcorr(U1_F);
C_I2 = xcorr(I2_F);
C_U2 = xcorr(U2_F);
nc = 0:255;
figure(6)
subplot(2,2,1);
plot(nc,C_I1(256:511));
xlabel('Lags(m)');
ylabel('magnitude of autocorrelation function');
title('Autocorrelation of Excitation Signal of /i/, n=6000:6255');
subplot(2,2,2);
plot(nc,C_U1(256:511));
xlabel('Lags(m)');
ylabel('magnitude of autocorrelation function');
title('Autocorrelation of Excitation Signal of /u/, n=6000:6255');
subplot(2,2,3);
plot(nc,C_I2(256:511));
xlabel('Lags(m)');
```

```
ylabel('magnitude of autocorrelation function');
title('Autocorrelation of Excitation Signal of /i/, n=256:511');
subplot(2,2,4);
plot(nc,C_U2(256:511));
xlabel('Lags(m)');
ylabel('magnitude of autocorrelation function');
title('Autocorrelation of Excitation Signal of /u/, n=256:511');
% Pitch Period of Excitation Signal
figure(7)
subplot(2,2,1);
[psor_I1,lsor_I1] =
findpeaks(C_I1(306:406),50:150,'MinPeakHeight',2e5,'SortStr','descend');
findpeaks(C_I1(306:406),50:150,'MinPeakHeight',2e5,'SortStr','descend');
hold on;
plot([lsor_I1(1),lsor_I1(1)],[-500000,psor_I1(1)]);
text(lsor_I1+.02,psor_I1(1),num2str(1));
text(lsor_I1+.02,[-100000],num2str(lsor_I1(1)'));
hold off;
xlabel('Lags(m)');
ylabel('magnitude of autocorrelation function');
title('Peak Finding of /i/, n=6000:6255');
subplot(2,2,2);
[psor_U1,lsor_U1] =
findpeaks(C_U1(306:406),50:150,'MinPeakHeight',1e5,'SortStr','descend');
findpeaks(C_U1(306:406),50:150,'MinPeakHeight',1e5,'SortStr','descend');
hold on;
plot([lsor_U1(1),lsor_U1(1)],[-500000,psor_U1(1)]);
text(lsor_U1+.02,psor_U1(1),num2str(1));
text(lsor_U1+.02,[-100000],num2str(lsor_U1(1)'));
hold off;
xlabel('Lags(m)');
ylabel('magnitude of autocorrelation function');
title('Peak Finding of /u/, n=6000:6255');
subplot(2,2,3);
[psor_I2,lsor_I2] =
findpeaks(C_I2(306:406),50:150,'MinPeakHeight',2e5,'SortStr','descend');
findpeaks(C_I2(306:406),50:150,'MinPeakHeight',2e5,'SortStr','descend');
hold on;
plot([lsor_I2(1),lsor_I2(1)],[-500000,psor_I2(1)]);
text(lsor_I2+.02,psor_I2(1),num2str(1));
text(lsor_I2+.02,[-100000],num2str(lsor_I2(1)'));
hold off;
xlabel('Lags(m)');
ylabel('magnitude of autocorrelation function');
title('Peak Finding of /i/, n=256:511');
subplot(2,2,4);
[psor_U2,lsor_U2] =
findpeaks(C_U2(306:406),50:150,'MinPeakHeight',1e5,'SortStr','descend');
findpeaks(C_U2(306:406),50:150,'MinPeakHeight',1e5,'SortStr','descend');
hold on;
plot([lsor_U2(1),lsor_U2(1)],[-500000,psor_U2(1)]);
text(lsor_U2+.02,psor_U2(1),num2str(1));
text(lsor_U2+.02,[-100000],num2str(lsor_U2(1)'));
hold off;
```

```
xlabel('Lags(m)');
ylabel('magnitude of autocorrelation function');
title('Peak Finding of /u/, n=256:511');

% plot(c)
% Voiced Extraction
load -ascii jcnwwa_single_col.txt
phrase = jcnwwa_single_col;
n_p = 0:(length(phrase)-1);
figure(8)
plot(n_p,phrase);
xlabel('sampling time series');
ylabel('impulse response');
title('Time Series of Phrase "we were away", Length=15360');
figure(9)
spectrogram(phrase,256,128,256,1e4);
title('Spectrogram of Phrase "we were away", Length=15360, Fs=10kHz');
% LPC Analysis
% Voiced Phrase: 4201-14440(10240-point, 0.42s-1.4439s)
phrase_v = phrase(4201:14440);
frame_num = 79;
A_P = zeros(79,256);
L_P = zeros(79,128);
for i = 1:79
    a_P = zeros(1,256);
    [a_P(1:15),~] = lpc(phrase_v(128*(i-
1)+1:128*(i+1)).*kaiser(256,1.5),14);
    A_P(i,1:256) = fft(a_P,256);
    for j = 1:128
        L_P(i,j) = 0.5*mag2db(1/((abs(A_P(i,j)))^2));
    end
end
N = 0:127;
Nf = N*10/256;
C = L_P';
y = Nf;
x = zeros(1,79);
for i = 1:79
    x(i) = i*0.0128+0.42;
end
figure (10)
imagesc(x,y,C), axis xy
colorbar
xlabel('time(s)');
ylabel('frequency(kHz)');
title('LPC Analysis of Phrase "we were away", 0.42s-1.4439s, 10240-point,
79-segment');
% Pitch Period v.s. Time

% Excitation Signal Power(dB) v.s. Time
Period_Phrase = zeros(1,79);
Power_Phrase = zeros(1,79)
C_Phrase = zeros(79,511);
for i = 1:79
```

```
    a_P = zeros(1,256);
    [a_P(1:15),~] = lpc(phrase_v(128*(i-
1)+1:128*(i+1)).*kaiser(256,1.5),14);
    Phrase_F = filter(a_P(1:15),1,phrase_v(128*(i-1)+1:128*(i+1)));
    C_Phrase(i,1:511) = xcorr(Phrase_F);
end
for i =1:79
    % [psor,lsor] =
findpeaks(C_Phrase(i,306:406),50:150,'MinPeakHeight',1e5,'SortStr','descen
d');
    % Period_Phrase(i) = lsor;
    Power_Phrase(i) = 0.5*mag2db(C_Phrase(i,256));

    [maxi_P,lsor] =
findpeaks(C_Phrase(i,306:376),50:120,'SortStr','descend','NPeaks',1);
    [maxi_PP,lsor] =
findpeaks(C_Phrase(i,306:376),50:120,'MinPeakHeight',0.5*maxi_P,'SortStr',
'descend');
    % figure(10+i)
    %
findpeaks(C_Phrase(i,306:406),50:150,'MinPeakHeight',0.5*maxi_P,'SortStr',
'descend');
    if length(maxi_PP) > 2
        Period_Phrase(i) = 0;
    else
        Period_Phrase(i) = lsor(1);
    end
    % plot([lsor,lsor],[-500000,psor]);
    % text(lsor+.02,psor,num2str((1:numel(psor))'));
    % text(lsor+.02,-100000,num2str(lsor));
    % hold off;
    % xlabel('Lags(m)');
    % ylabel('magnitude of autocorrelation function');
    % title(['Pitch Period of Segment',num2str(i),'is',num2str(lsor)]);
end
figure(11)
subplot(1,2,1);
scatter(x,Period_Phrase/10);
xlabel('evolving time(s)');
ylabel('Pitch Period(ms)');
title('Pitch Period vs. Time of Phrase, Scatter Graph');
figure(12)
plot(x,Power_Phrase);
xlabel('evolving time(s)');
ylabel('Average Power(dB)');
title('Excitation Signal Power vs. Time of Phrase "we were away", 0.42s-
1.4439s, 10240-point, 79-segment');
for i = 27:28
    [maxi_P,Period_Phrase(i)] =
findpeaks(C_Phrase(i,306:376),50:120,'SortStr','descend','NPeaks',1);
    figure(i)
    findpeaks(C_Phrase(i,306:376),50:120,'SortStr','descend','NPeaks',1)
end
figure(11)
```

```
subplot(1,2,2);
plot(x,Period_Phrase/10);
xlabel('evolving time(s)');
ylabel('Pitch Period(ms)');
title('Pitch Period vs. Time of Phrase, Plot Graph');
```

Reference

[1] Estimation Of Pitch From Speech Signals (Reference) : Speech Signal Processing Laboratory : Electronics & Communications : IIT GUWAHATI Virtual Lab, iitg.vlab.co.in/?sub=59&brch=164&sim=1012&cnt=6.

[2] J. Markel, "Digital Inverse Filtering - A New Tool for Formant Trajectory Estimation", IEEE Trans. on Audio and Electroacoustics, AU-20: 129-127 (1972)

1 Topic II-B **95 / 100**

- **- 0 pts** Correct
- **- 5 pts** Figure for Part (a) incorrect/ missing - Time series of frame
- **- 5 pts** Figure for Part(a) incorrect/ missing - Conventional spectral estimates
- **- 5 pts** Figure for Part (a) incorrect/ missing - LPC spectral estimate
- **- 5 pts** Figure for Part (b) incorrect/ missing - Time series of excitation signal estimate
- **- 5 pts** Figure for Part (b) incorrect/ missing - Conventional spectral estimates
- **- 5 pts** Figure for Part (b) incorrect/ missing - autocorrelation functions of excitation signal estimates
- **- 5 pts** Figure for Part (c) incorrect/ missing - time series of entire data
- ✓ **- 5 pts Figure for Part (c) incorrect/ missing - time series of speech phrase**
- **- 5 pts** Figure for Part (c) incorrect/ missing - spectrogram of the speech phrase
- **- 5 pts** Figure for Part (c) incorrect/ missing - time-evolving LPC analysis plot
- **- 5 pts** Figure for Part (c) incorrect/ missing - pitch period vs time
- **- 5 pts** Figure for Part (c) incorrect/ missing - excitation signal power vs time
- **- 5 pts** Part(b) pitch period estimates of speech vowel frame missing/ incorrect
- **- 2 pts** pitch period unit missing
- **- 15 pts** didn't analyze at least two frame for each vowel
- **- 5 pts** Not enough analysis in results part