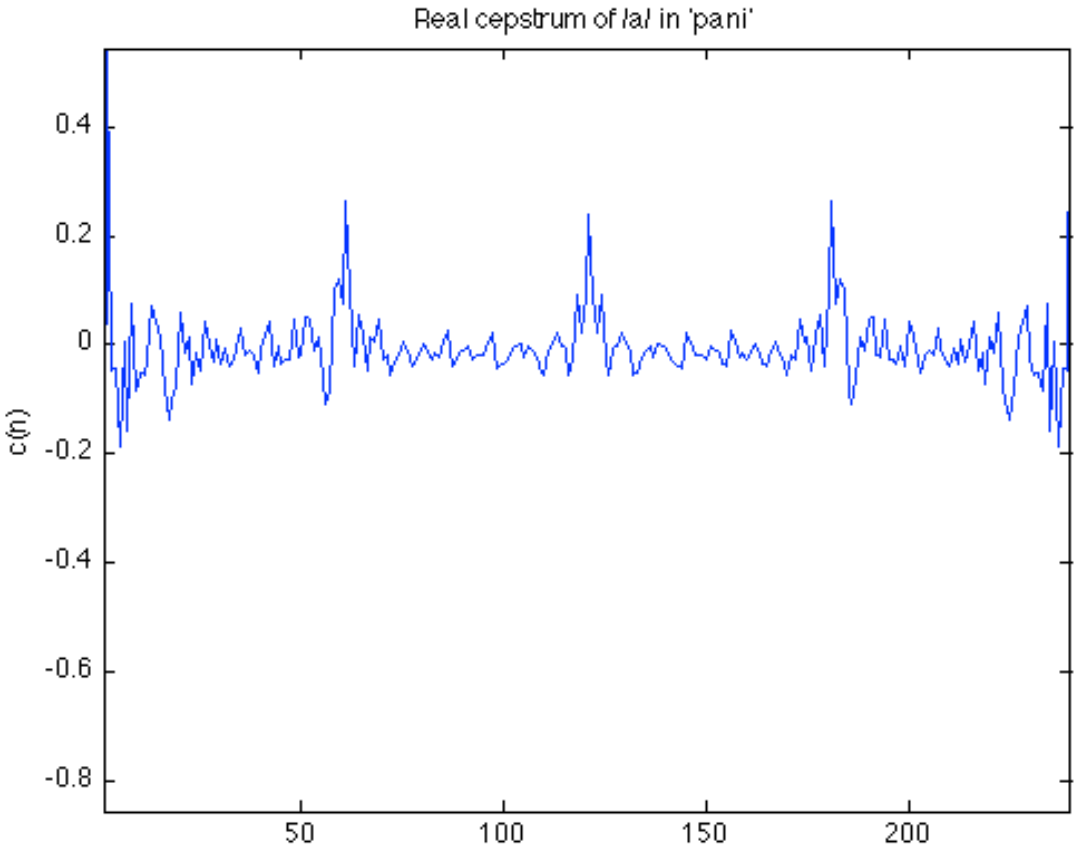
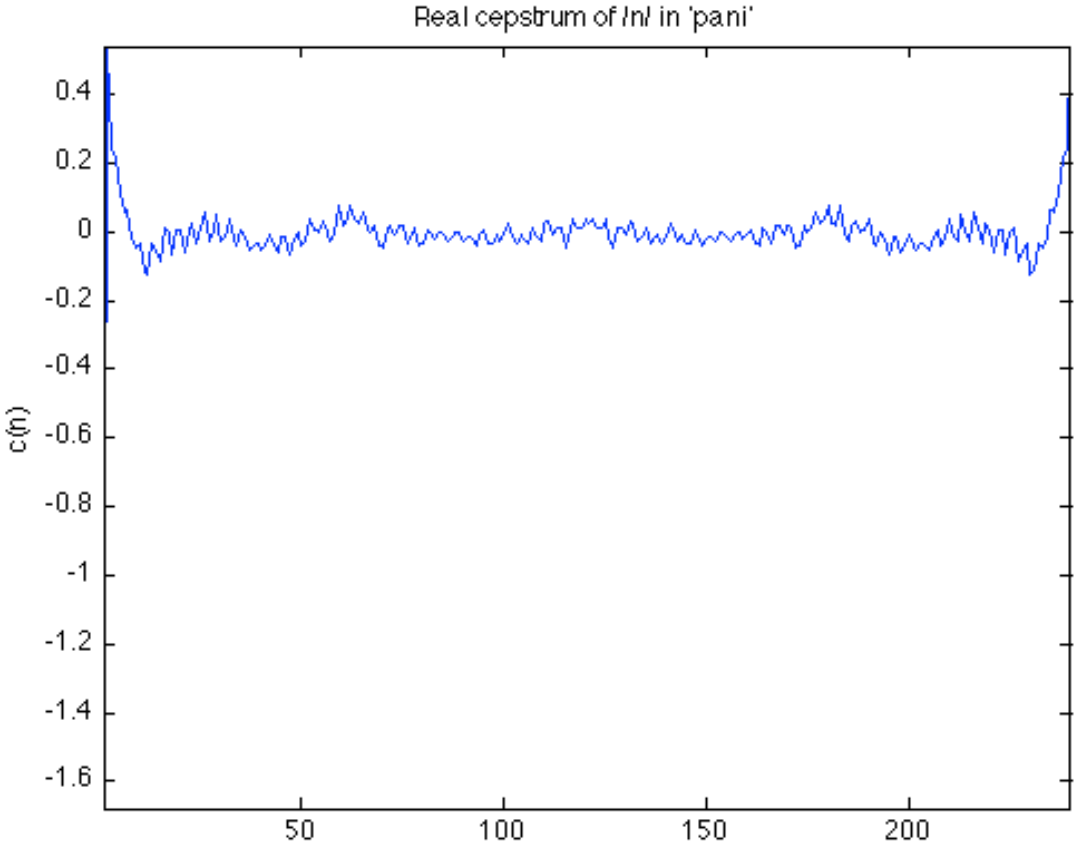
EE 679 Computing Assignment 4

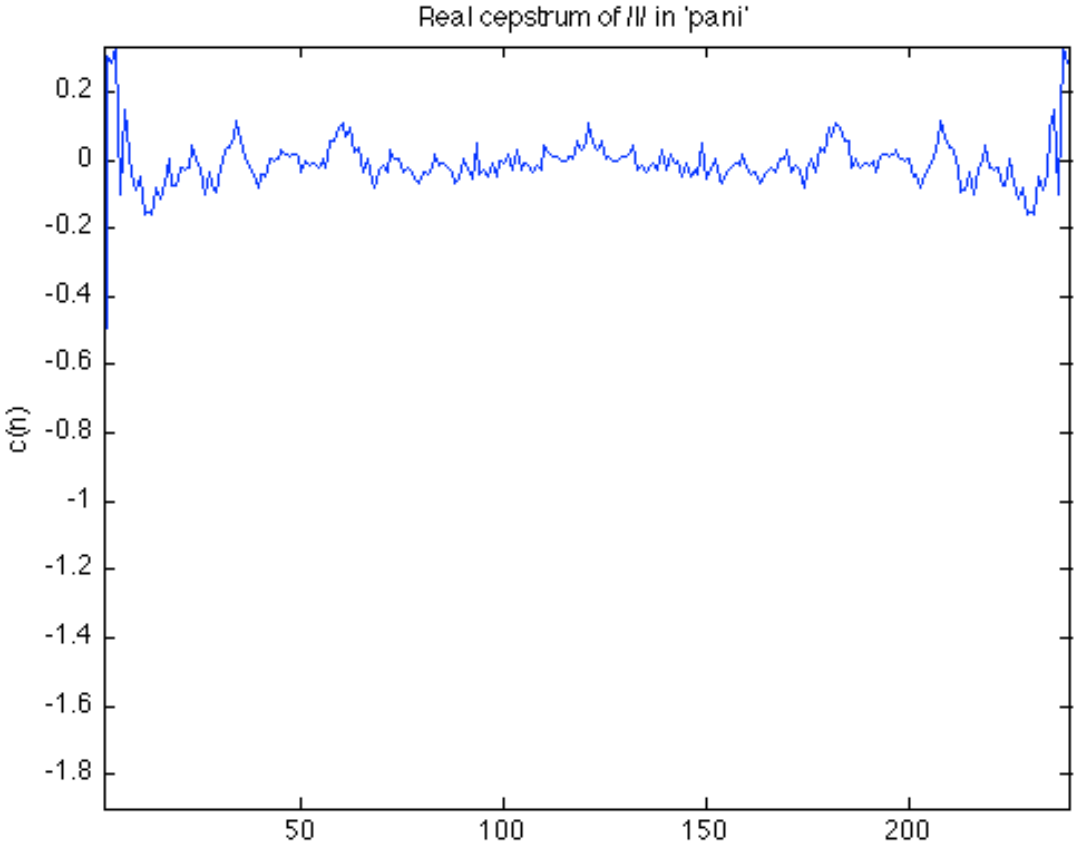
Name: Swrangsar Basumatary Roll: 09d07040

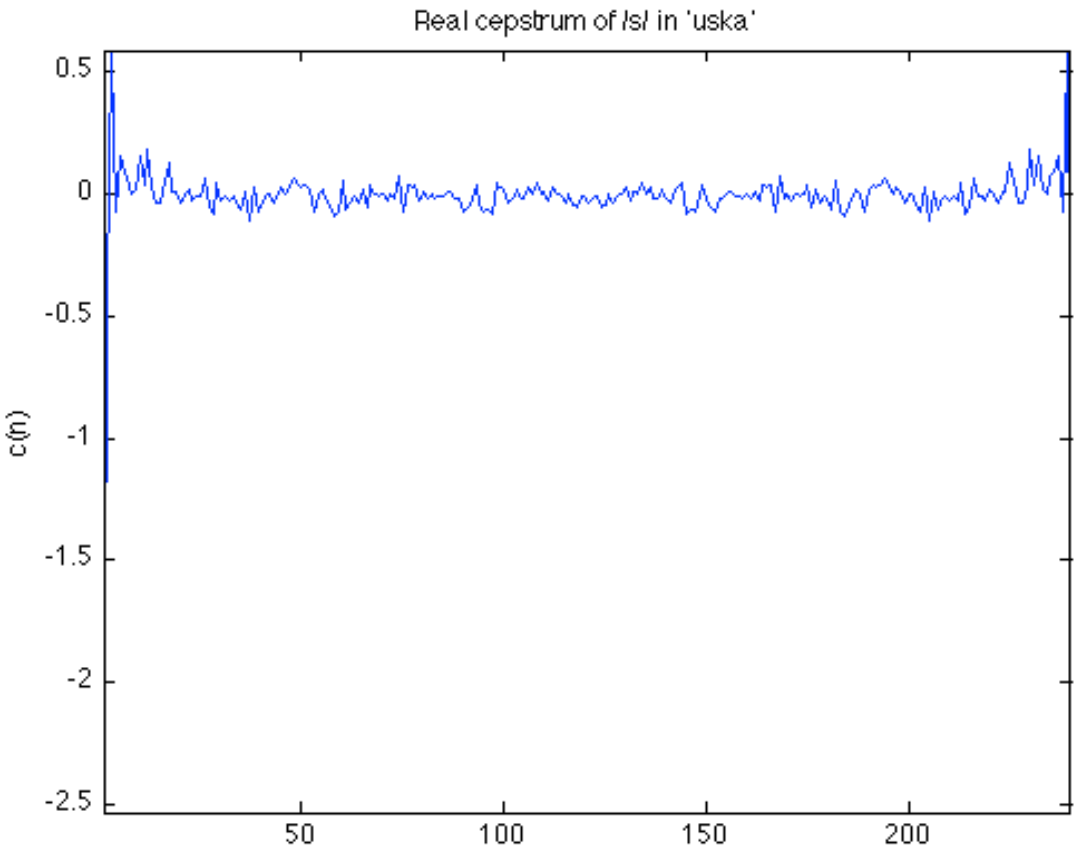
Question 1

The plots:




Real cepstrum of /n/ in 'pani'
c(n)
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
50
100
150
200


Real cepstrum of /ɳ/ in 'pani'
c(n)
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
-1.8
50
100
150
200



The functions:

```
function [realCepstrum, fs] = getRealCepstrum(speechSegment)

[windowedSignal, fs] = getWindowedSignal(speechSegment);
logMagnitudeSpectrum = getLogMagnitudeSpectrum(windowedSignal);
realCepstrum = ifft(logMagnitudeSpectrum);

end


%% get a 30 ms window of speech signal

function [windowedSignal, fs] = getWindowedSignal(inputFile)

windowDuration = 0.030; % in ms

[y, fs] = wavread(inputFile);
length = size(y, 1);
centralIndex = round(length/2);
M = round(windowDuration * fs);
```

```
startIndex = round(centralIndex - M/2);
windowedSignal = y(startIndex:startIndex + M-1);

end

%% get logarithm of the magnitude spectrum

function logMagnitudeSpectrum =
getLogMagnitudeSpectrum(windowedSignal)

mag = fft(windowedSignal);
logMagnitudeSpectrum = log(abs(mag));

end
```

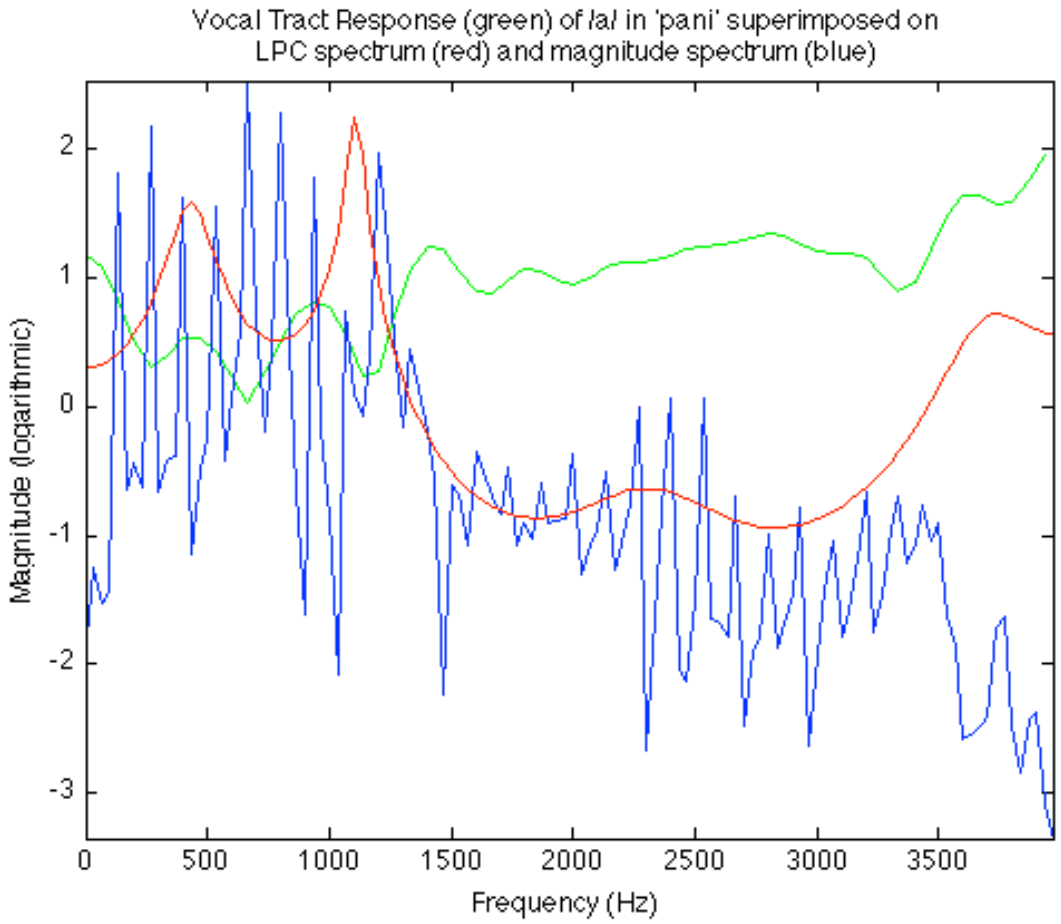
The script:

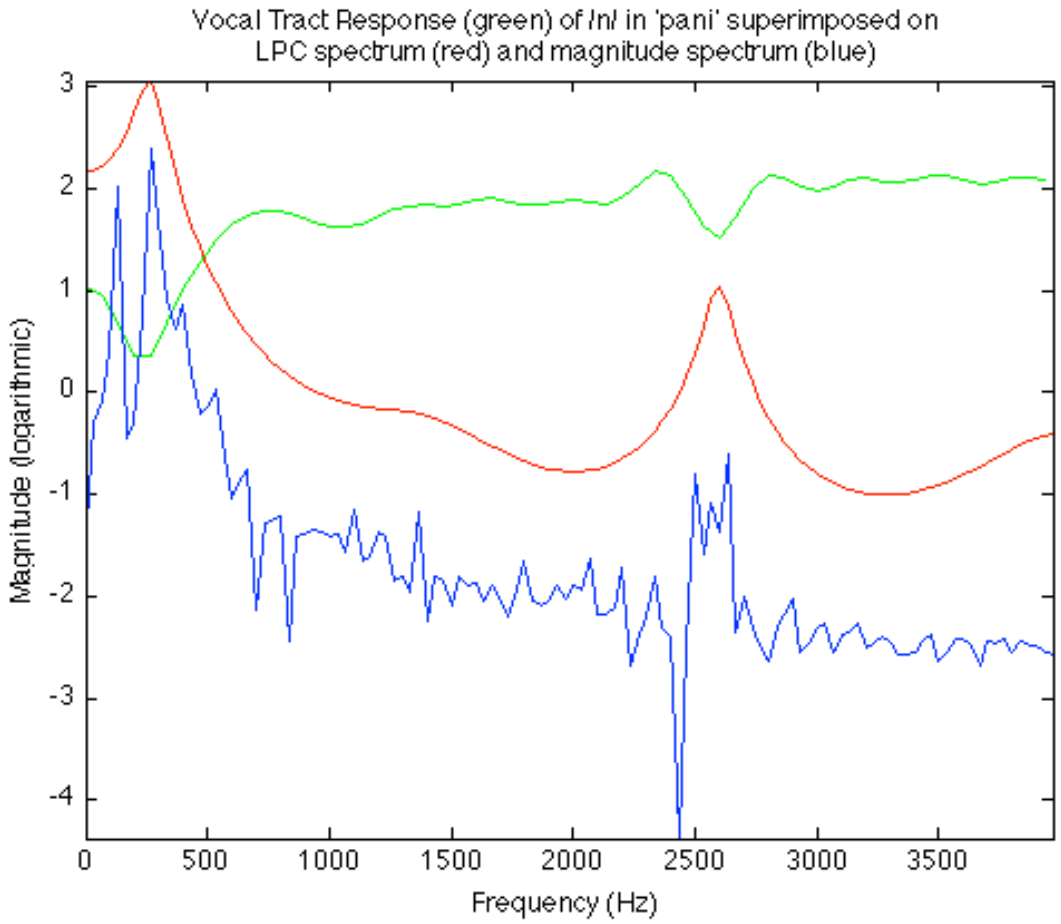
```
close all; clear all;

addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
dataFiles/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q1/

%% plot real cepstrum of /a/ in 'pani'

rc1 = getRealCepstrum('a_pani.wav');
figure(100); clf;
plot(rc1); axis tight;
title('Real cepstrum of /a/ in ''pani''');
ylabel('c(n)');

%% plot real cepstrum of /n/ in 'pani'

rc2 = getRealCepstrum('n_pani.wav');
figure(200); clf;
plot(rc2); axis tight;
title('Real cepstrum of /n/ in ''pani''');
ylabel('c(n)');

%% plot real cepstrum of /I/ in 'pani'

rc3 = getRealCepstrum('i_pani.wav');
figure(300); clf;
plot(rc3); axis tight;
```

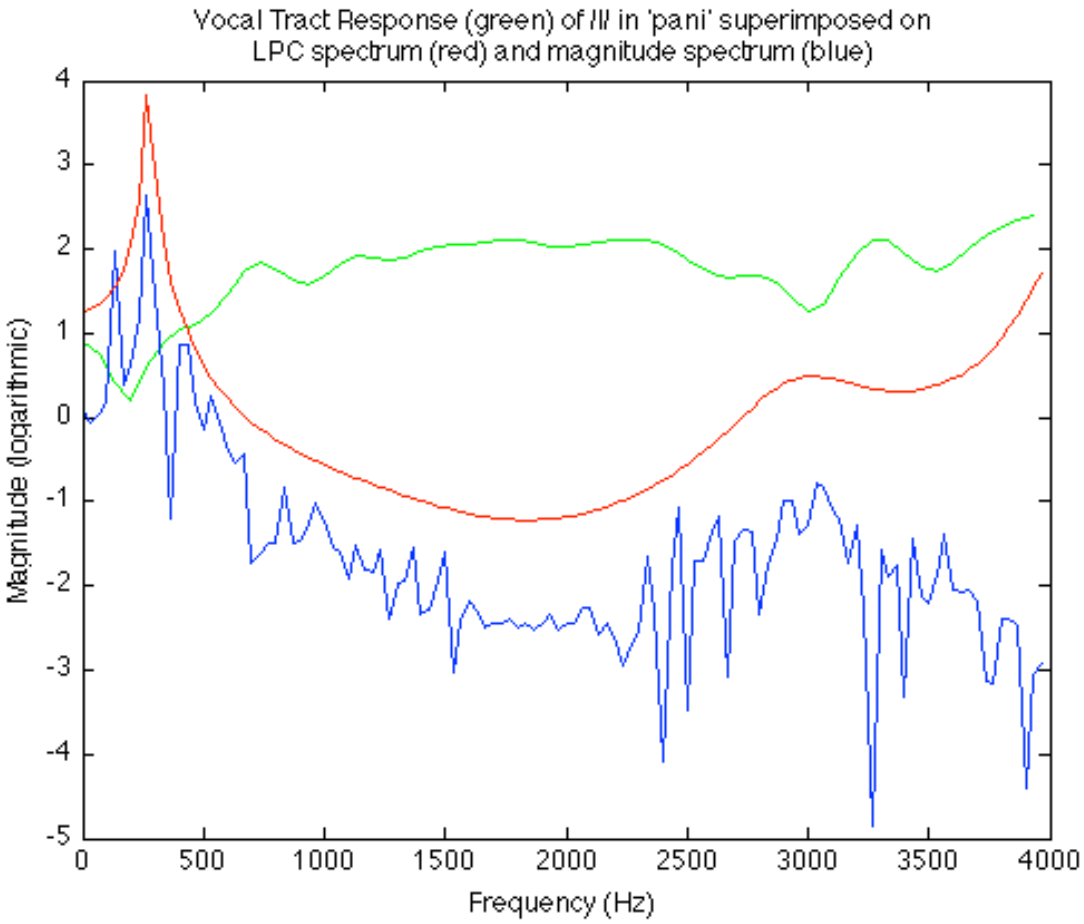
```
title('Real cepstrum of /I/ in ''pani''');
ylabel('c(n)');

%% plot real cepstrum of /s/ in 'uska'

rc4 = getRealCepstrum('s_uska.wav');
figure(400); clf;
plot(rc4); axis tight;
title('Real cepstrum of /s/ in ''uska''');
ylabel('c(n)');


rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q1/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
dataFiles/
```

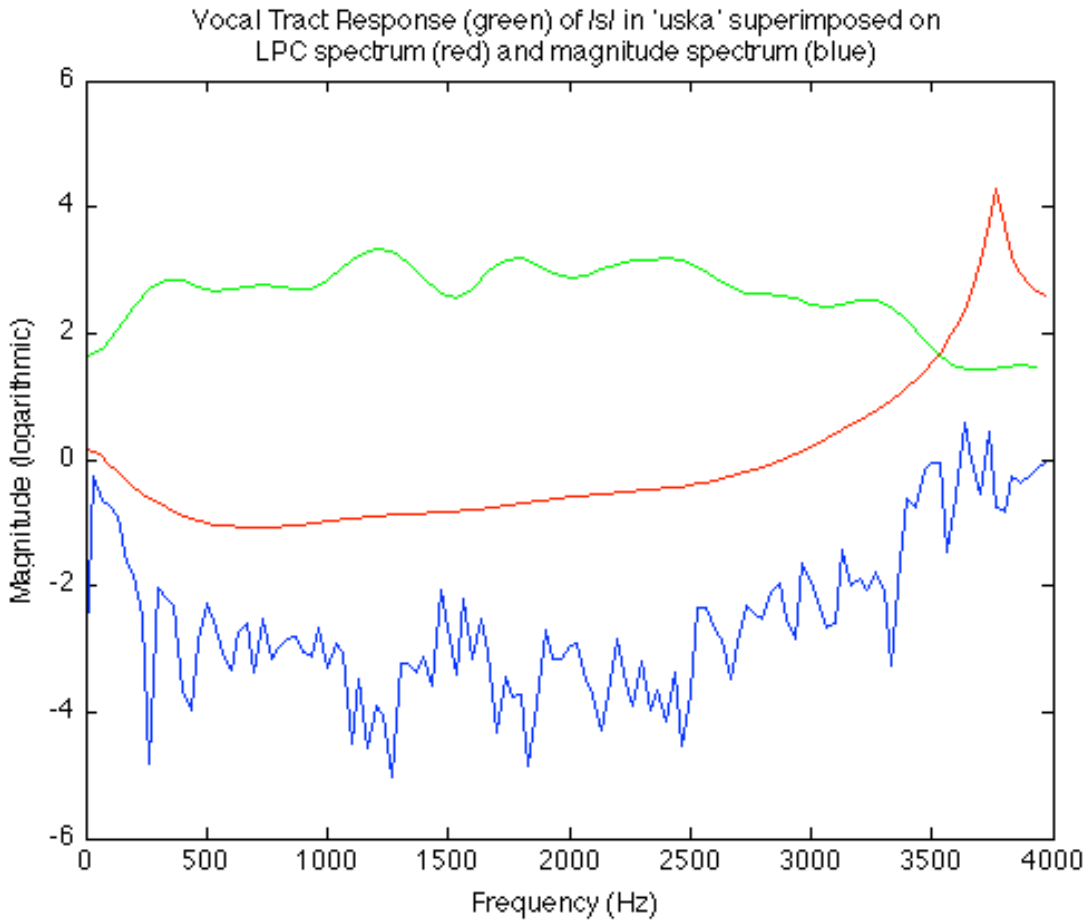
Question 2

The plots:


Vocal Tract Response (green) of /a/ in 'pani' superimposed on
LPC spectrum (red) and magnitude spectrum (blue)
Magnitude (logarithmic)
Frequency (Hz)
2
1
0
-1
-2
-3
0
500
1000
1500
2000
2500
3000
3500


Vocal Tract Response (green) of /n/ in 'pani' superimposed on
LPC spectrum (red) and magnitude spectrum (blue)
Magnitude (logarithmic)
Frequency (Hz)
3
2
1
0
-1
-2
-3
-4
0
500
1000
1500
2000
2500
3000
3500


Vocal Tract Response (green) of /i/ in 'pani' superimposed on
LPC spectrum (red) and magnitude spectrum (blue)
Magnitude (logarithmic)
Frequency (Hz)
4
3
2
1
0
-1
-2
-3
-4
-5
0
500
1000
1500
2000
2500
3000
3500
4000



The functions:

```
function [vocalTractResponse, w] =
getVocalTractResponse(speechSegment)

[speechCepstrum, fs] = getRealCepstrum(speechSegment);
lowTimeCepstrum = getLowTimeCepstrum(speechCepstrum, fs);
vocalTractResponse = abs(fft(lowTimeCepstrum));
vocalTractResponse = vocalTractResponse(1:
(length(vocalTractResponse)/2));

M = length(vocalTractResponse);
frequencies = (fs/M) * (0:M-1);
w = frequencies(:)/2;

end

%% do low-time liftering for vocal tract estimation
```

```
function lowTimeCepstrum = getLowTimeCepstrum(speechCepstrum,
fs)

cutoffLength = fs * .003; % take 20 or 15
speechCepstrum = speechCepstrum(1:round(length(speechCepstrum)/
2));
lowTimeCepstrum = speechCepstrum;
lowTimeCepstrum((cutoffLength+1):end) = 0;

end
```

The LPC function (adapted from previous assignment):

```
function [LPCSpectrum, narrowbandSpectrum, w] =
getLPCSpectrum(speechSegment, poleOrder)

if nargin < 2, poleOrder = 10; end

[windowedSignal, fs] = getWindowedSignal(speechSegment);
M = length(windowedSignal);
% the narrowband spectrum first
narrowbandSpectrum = log(abs(fft(windowedSignal)));
narrowbandSpectrum = narrowbandSpectrum(1:round(M/2));

preEmpdSignal = preEmphasize(windowedSignal);
LPCoeffs = getLPCoefficients(preEmpdSignal, poleOrder);
frequencies = (fs/M) * (0:M-1);

denominator = 0;
numerator = 1;

for k = 1:poleOrder
    denominator = denominator + (LPCoeffs(k) * (exp(-1i * 2 * pi
* frequencies ./ fs) .^ (k-1)));
end

H = numerator ./ denominator;
LPCSpectrum = log(abs(H));
LPCSpectrum = LPCSpectrum(1:round(M/2));
w = frequencies(1:round(M/2));
LPCSpectrum = LPCSpectrum(:);
narrowbandSpectrum=narrowbandSpectrum(:);
w=w(:);

end
```

```
%% get a 30 ms window of speech signal

function [windowedSignal, fs] = getWindowedSignal(speechSegment)

windowDuration = 0.030; % in ms
[y, fs] = wavread(speechSegment);
length = size(y, 1);
centralIndex = round(length/2);
M = round(windowDuration * fs);
startIndex = round(centralIndex - M/2);
windowedSignal = y(startIndex:startIndex + M-1);

end

%% pre emphasize the speech segment

function preEmpdSignal = preEmphasize(windowedSignal)

y = windowedSignal;
length = size(y, 1);
for k = 1:length
    if k > 1
        y(k) = y(k) - (0.97*y(k-1));
    end
end
preEmpdSignal = y;

end


%% get autocorrelation coefficients

function autocorrVector = getAutoCorrelation(preEmpdSignal,
poleOrder)

ACCoeff = zeros(poleOrder+1, 1);
for p = 0:poleOrder
    for k = 0:(length(preEmpdSignal)-1)
        valueToBeAdded = 0;
        if k-p >= 0
            valueToBeAdded = preEmpdSignal(k+1) .*
preEmpdSignal(k+1-p);
        end
        ACCoeff(p+1) = ACCoeff(p+1) + valueToBeAdded;
    end
end
```

```
autocorrVector = ACCoeff;

end


%% get LPC Coefficients

function LPCoeffs = getLPCoefficients(preEmpdSignal, poleOrder)

autocorrVector = getAutoCorrelation(preEmpdSignal, poleOrder);
LPCoeffs = levinson(autocorrVector);

end
```

The script for the plots:

```
close all; clear all;


addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
dataFiles/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q1/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q2/

%% plot vocal tract response of /a/ in 'pani'

[vtr1, ww1] = getVocalTractResponse('a_pani.wav');
[LPCSpectrum1, narrowbandSpectrum1, w1] =
getLPCSpectrum('a_pani.wav');

figure(100); clf;
vtrPlot1 = plot(ww1, vtr1);
hold on;
narrowbandPlot1 = plot(w1, narrowbandSpectrum1);
LPCPlot1 = plot(w1, LPCSpectrum1);
set(vtrPlot1, 'Color', 'green');
set(LPCPlot1,'Color','red');
set(narrowbandPlot1, 'Color', 'blue');
hold off;
axis tight;
title({'Vocal Tract Response (green) of /a/ in ''pani''
superimposed on'; ...
    'LPC spectrum (red) and magnitude spectrum (blue)'});
xlabel('Frequency (Hz)');
ylabel('Magnitude (logarithmic)');
```

```
%% plot vocal tract response of /n/ in 'pani'

[vtr2, ww2] = getVocalTractResponse('n_pani.wav');
[LPCSpectrum2, narrowbandSpectrum2, w2] =
getLPCSpectrum('n_pani.wav');

figure(200); clf;
vtrPlot2 = plot(ww2, vtr2);
hold on;
narrowbandPlot2 = plot(w2, narrowbandSpectrum2);
LPCPlot2 = plot(w2, LPCSpectrum2);
set(vtrPlot2, 'Color', 'green');
set(LPCPlot2,'Color','red');
set(narrowbandPlot2, 'Color', 'blue');
hold off;
axis tight;
title({'Vocal Tract Response (green) of /n/ in ''pani''
superimposed on'; ...
    'LPC spectrum (red) and magnitude spectrum (blue)'});
xlabel('Frequency (Hz)');
ylabel('Magnitude (logarithmic)');


%% plot vocal tract response of /I/ in 'pani'

[vtr3, ww3] = getVocalTractResponse('i_pani.wav');
[LPCSpectrum3, narrowbandSpectrum3, w3] =
getLPCSpectrum('i_pani.wav');

figure(300); clf;
vtrPlot3 = plot(ww3, vtr3);
hold on;
narrowbandPlot3 = plot(w3, narrowbandSpectrum3);
LPCPlot3 = plot(w3, LPCSpectrum3);
set(vtrPlot3, 'Color', 'green');
set(LPCPlot3,'Color','red');
set(narrowbandPlot3, 'Color', 'blue');
hold off;
title({'Vocal Tract Response (green) of /I/ in ''pani''
superimposed on'; ...
    'LPC spectrum (red) and magnitude spectrum (blue)'});
xlabel('Frequency (Hz)');
ylabel('Magnitude (logarithmic)');
```

```
%% plot vocal tract response of /s/ in 'uska'

[vtr4, ww4] = getVocalTractResponse('s_uska.wav');
[LPCSpectrum4, narrowbandSpectrum4, w4] =
getLPCSpectrum('s_uska.wav');

figure(400); clf;
vtrPlot4 = plot(ww4, vtr4);
hold on;
narrowbandPlot4 = plot(w4, narrowbandSpectrum4);
LPCPlot4 = plot(w4, LPCSpectrum4);
set(vtrPlot4, 'Color', 'green');
set(LPCPlot4,'Color','red');
set(narrowbandPlot4, 'Color', 'blue');
hold off;
title({'Vocal Tract Response (green) of /s/ in ''uska''
superimposed on'; ...
    'LPC spectrum (red) and magnitude spectrum (blue)'});
xlabel('Frequency (Hz)');
ylabel('Magnitude (logarithmic)');


rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q2/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q1/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
dataFiles/
```

The vocal tract response is the DFT of the first 3-4ms window of the cepstrum of the signal. We did a low-time liftering of the cepstrum using a 3 ms window and then took the DFT which is supposed to be equal to the log of abs(V(z)). Here z = e^(jω). The excitation component is supposed to be in the high quefrencies of the cepstrum.

s(n) = e(n) * v(n) and S(z) = E(z)V(z)

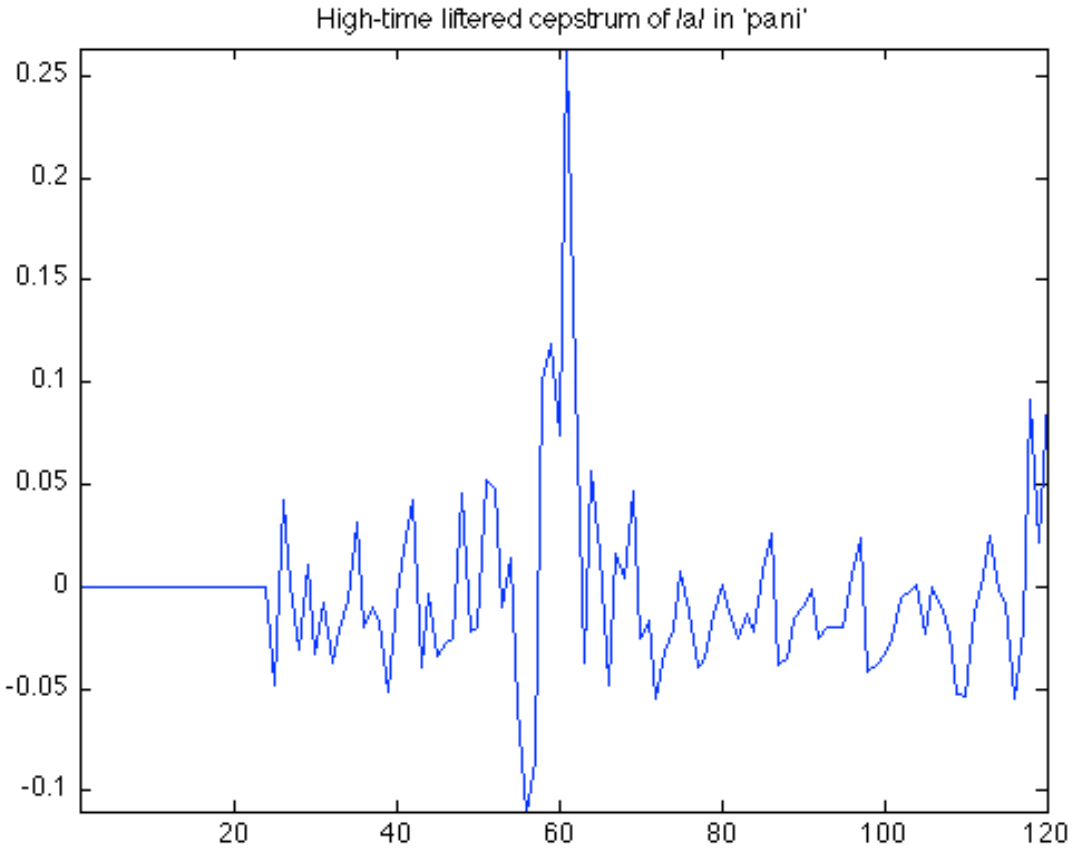
log(|S(z)|) = log(|E(z)|) + log(|V(z)|)
e'(n) and v'(n) are the IDFT of log(|E(z)|) and log(|V(z)|) respectively. v'(n) corresponds to the first 3-4ms window of the cepstrum. e'(n) is impulsive and periodic for voiced sounds.

Question 3

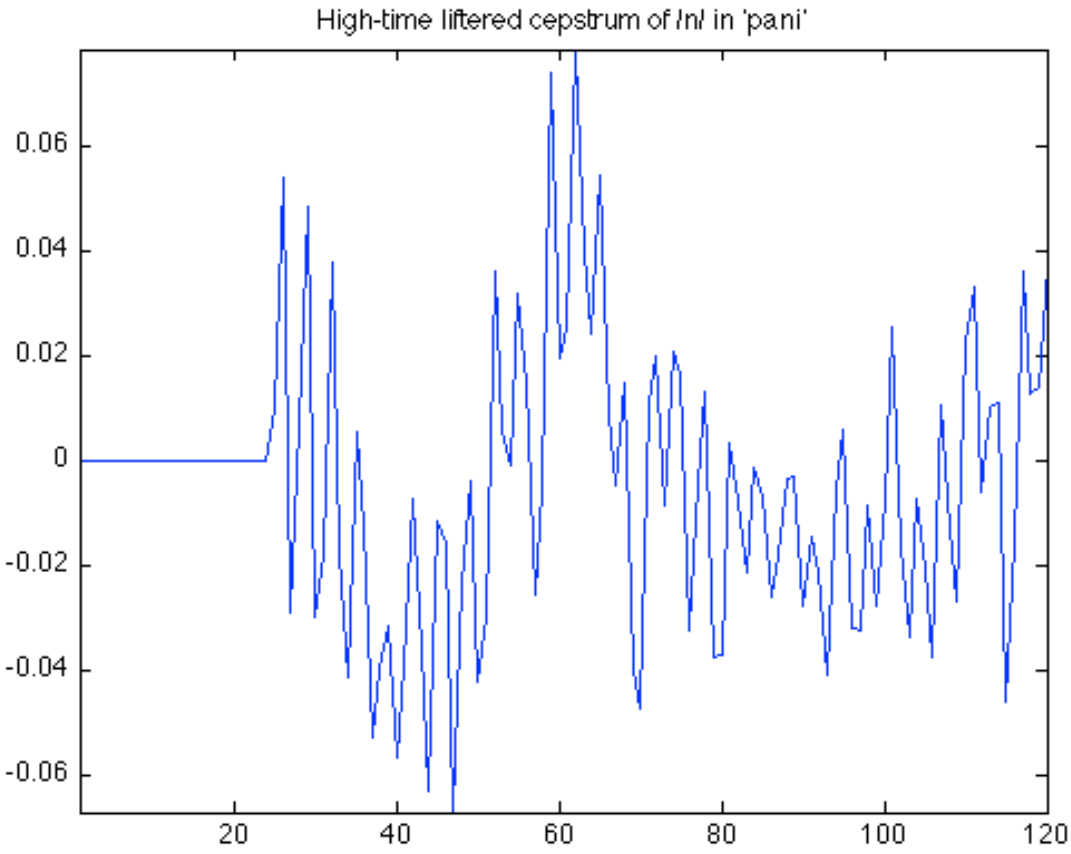
The pitch estimation is done by finding the maximum of the peaks in the high-time liftered part of the cepstrum. The pitch period is the time corresponding to the maximum peak from time '0'. If there are more than one maximum peaks then we consider the first one.

Pitch Estimation

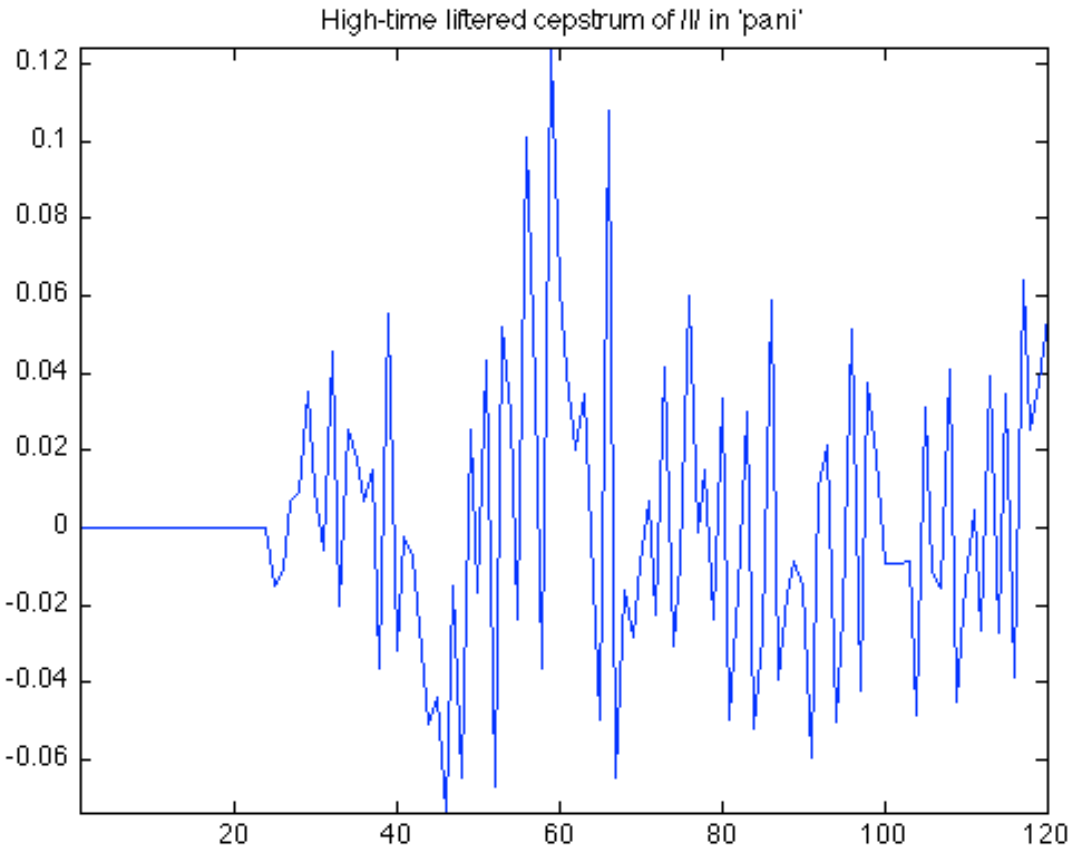
Pitch for /a/ in 'pani': 133.3333Hz



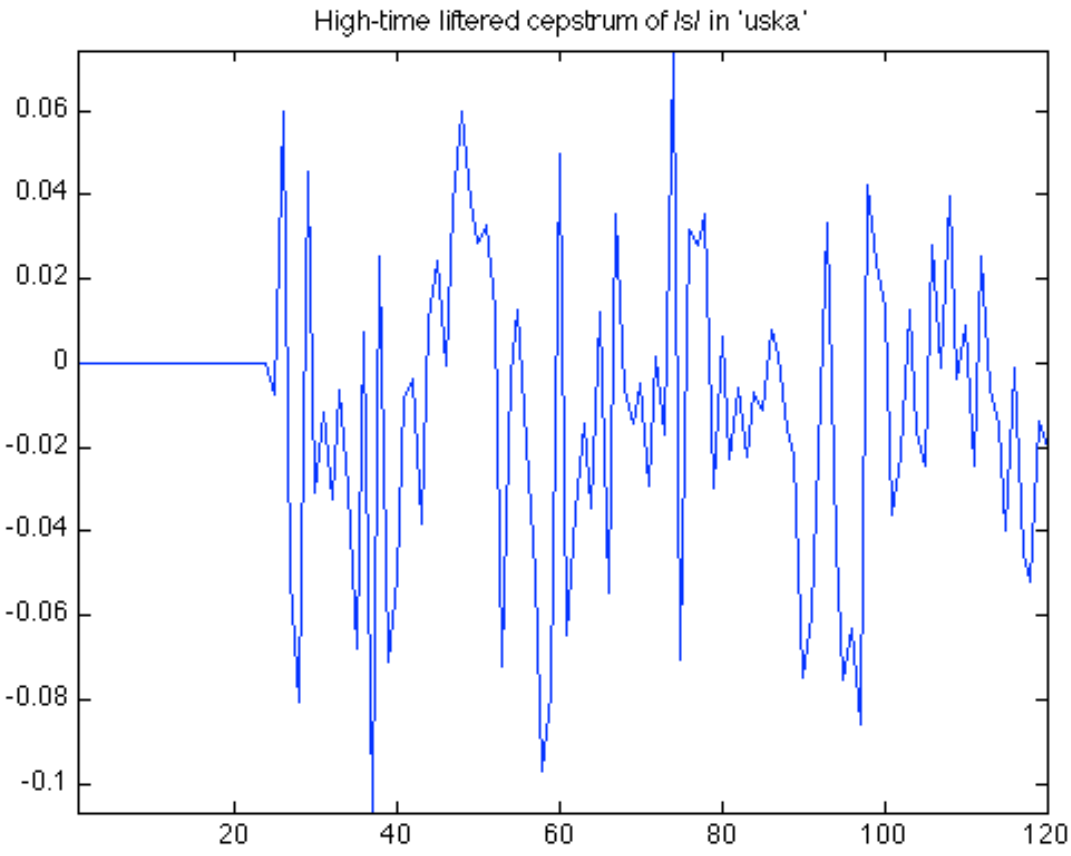
Pitch for /n/ in ‘pani’: 131.1475Hz



Pitch for /l/ in ‘pani’: 137.9310Hz

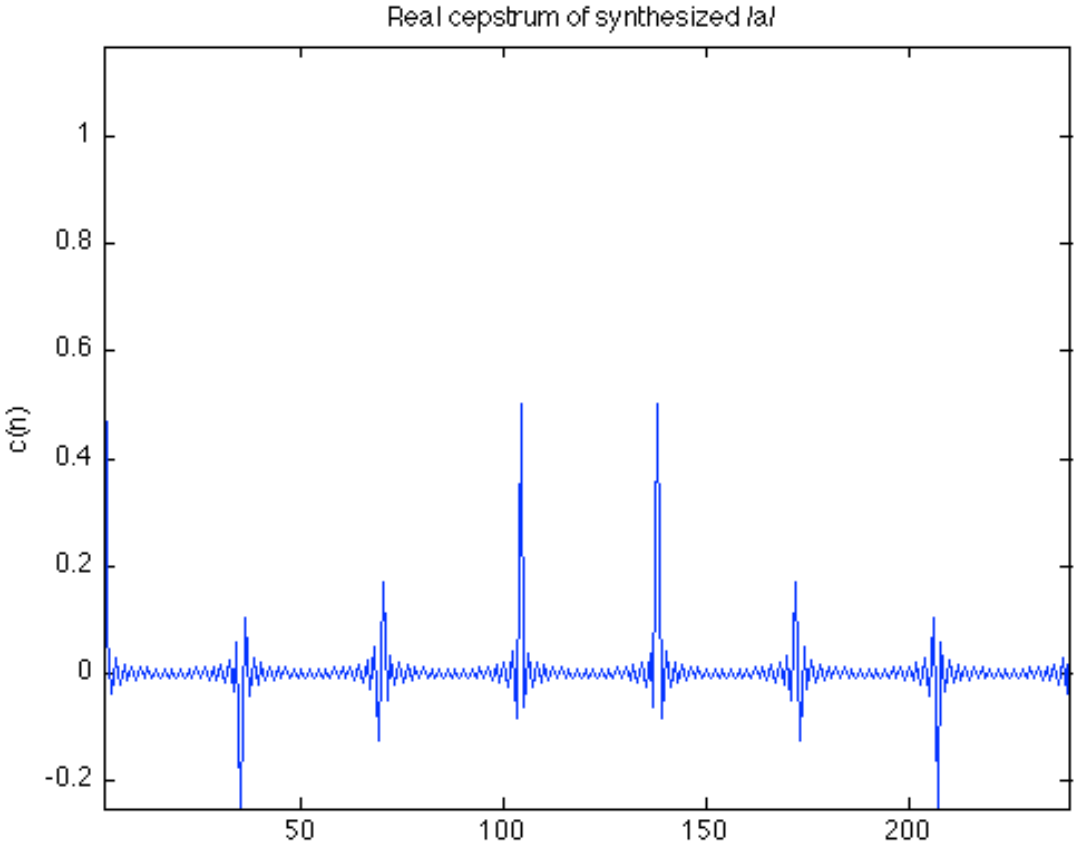


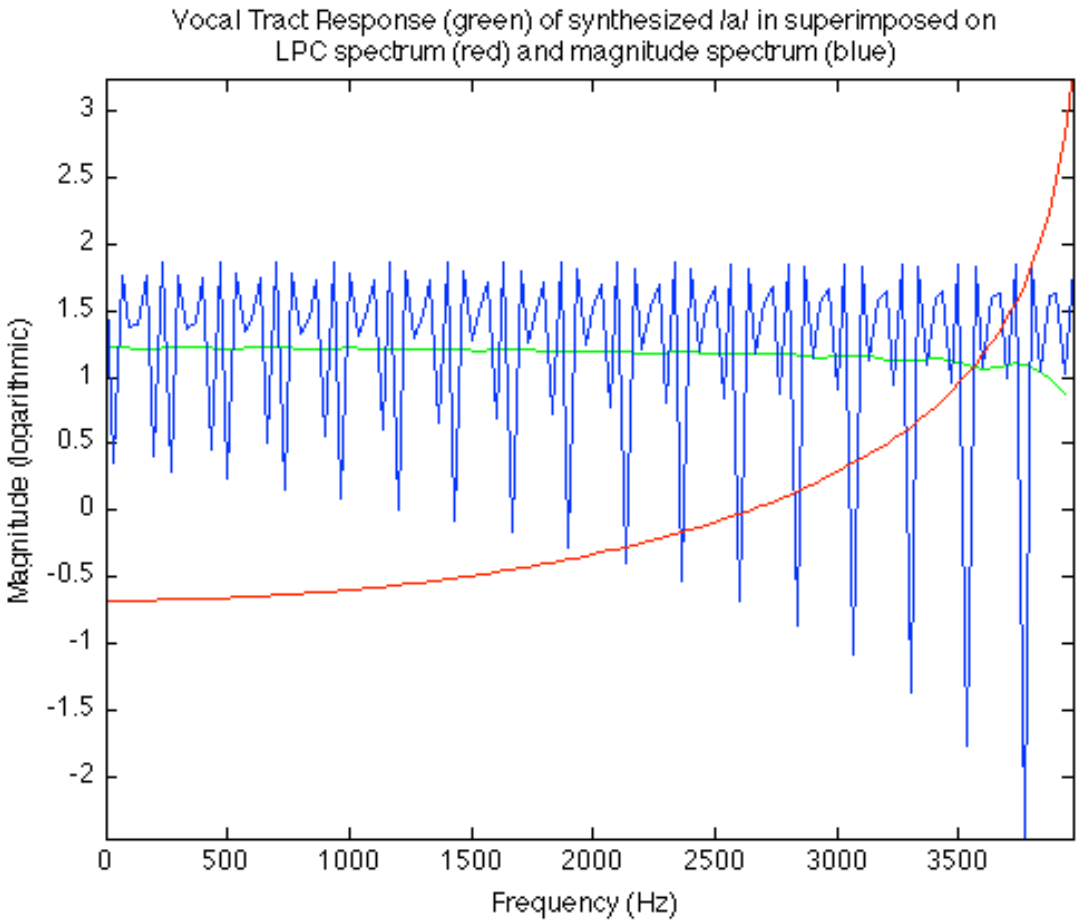
Pitch for /s/ in ‘uska’: 109.5890Hz

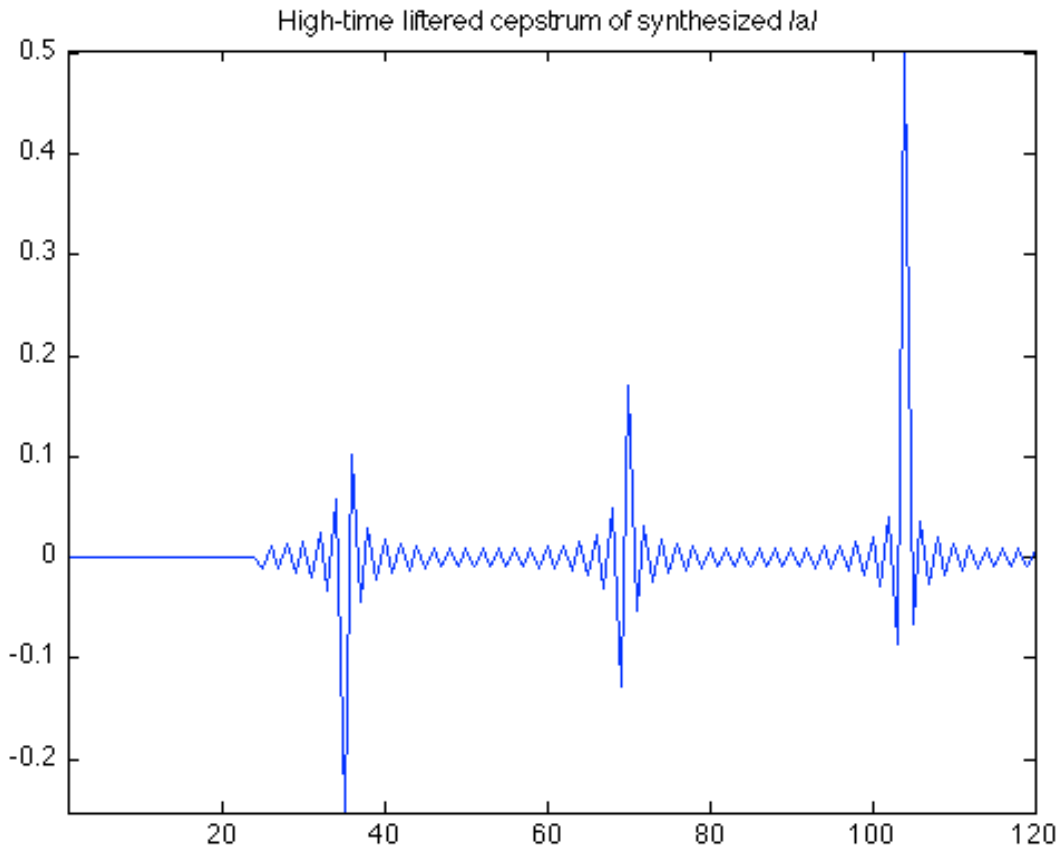


Question 4

The plots:


Real cepstrum of synthesized /a/
c(n)
1
0.8
0.6
0.4
0.2
0
-0.2
50
100
150
200


Vocal Tract Response (green) of synthesized /a/ in superimposed on
LPC spectrum (red) and magnitude spectrum (blue)
Magnitude (logarithmic)
Frequency (Hz)
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
500
1000
1500
2000
2500
3000
3500



The script:

```
close all; clear all;

addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
dataFiles/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q1/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q2/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q3/

%% plot real cepstrum of synthesized /a/

rc1 = getRealCepstrum('a_pani_synth.wav');
figure(100); clf;
plot(rc1); axis tight;
title('Real cepstrum of synthesized /a/');
ylabel('c(n)');

%% plot vocal tract response of synthesized /a/

[vtr1, ww1] = getVocalTractResponse('a_pani_synth.wav');
```

```
[LPCSpectrum1, narrowbandSpectrum1, w1] =
getLPCSpectrum('a_pani_synth.wav');

figure(200); clf;
vtrPlot1 = plot(ww1, vtr1);
hold on;
narrowbandPlot1 = plot(w1, narrowbandSpectrum1);
LPCPlot1 = plot(w1, LPCSpectrum1);
set(vtrPlot1, 'Color', 'green');
set(LPCPlot1,'Color','red');
set(narrowbandPlot1, 'Color', 'blue');
hold off;
axis tight;
title({'Vocal Tract Response (green) of synthesized /a/ in
superimposed on'; ...
    'LPC spectrum (red) and magnitude spectrum (blue)'});
xlabel('Frequency (Hz)');
ylabel('Magnitude (logarithmic)');

%% get the pitch of synthesized /a/

[pitchOfSynthesizedA, htl1] =
getPitchUsingCepstrum('a_pani_synth.wav');
pitchOfSynthesizedA
figure(300); clf;
plot(htl1); axis tight;
title('High-time liftered cepstrum of synthesized /a/');



rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q3/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q2/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q1/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
dataFiles/
```

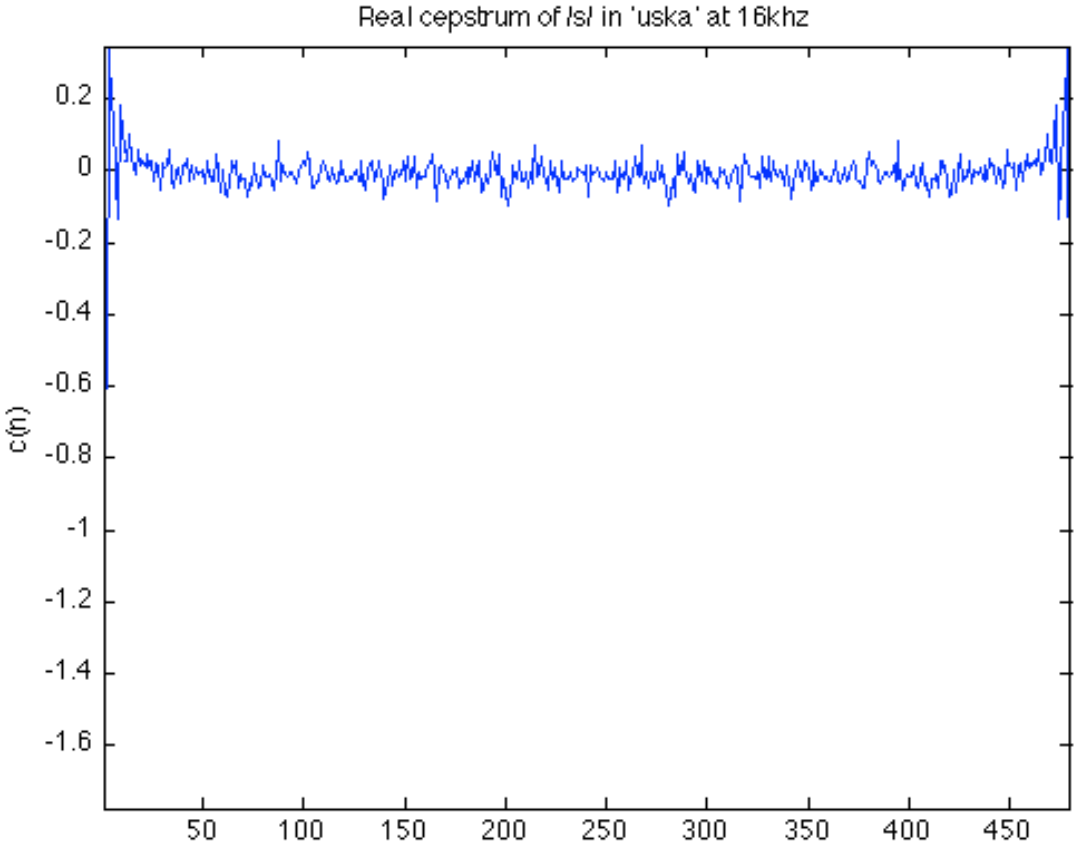
Results and discussion:

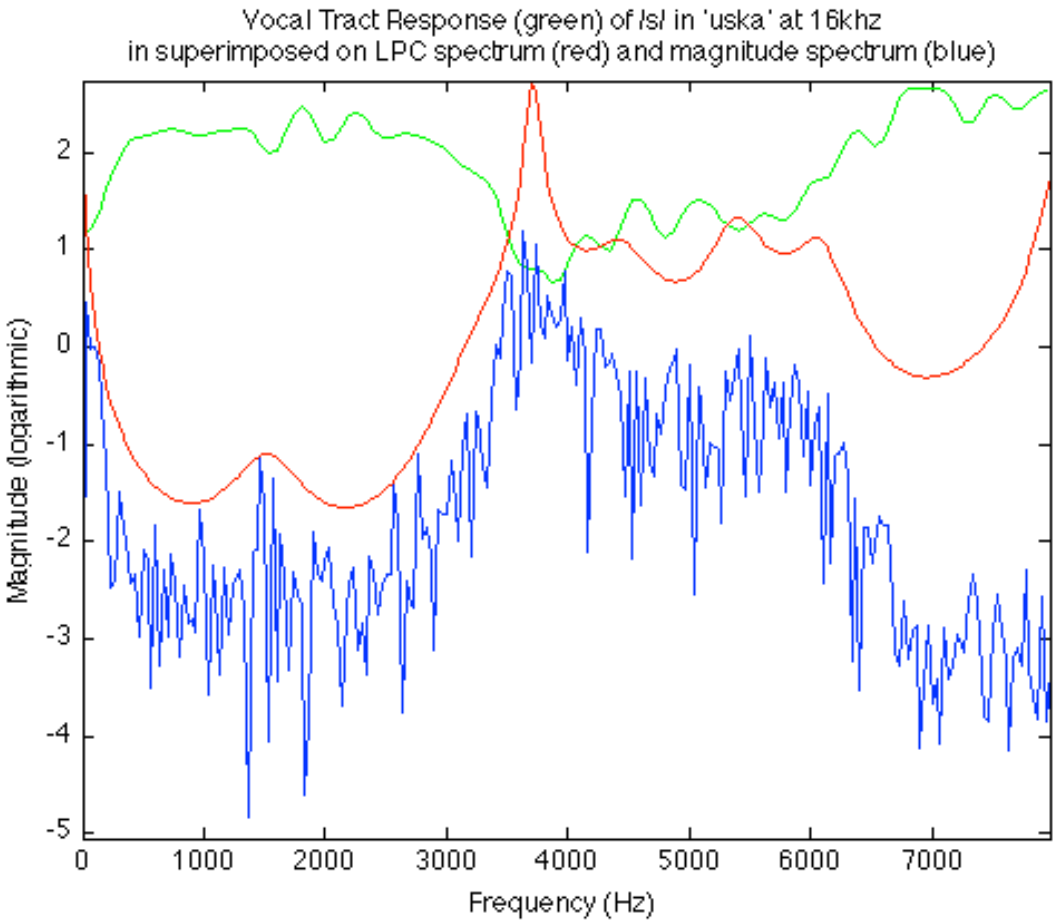
pitch of synthesized /a/ = 77.6699Hz

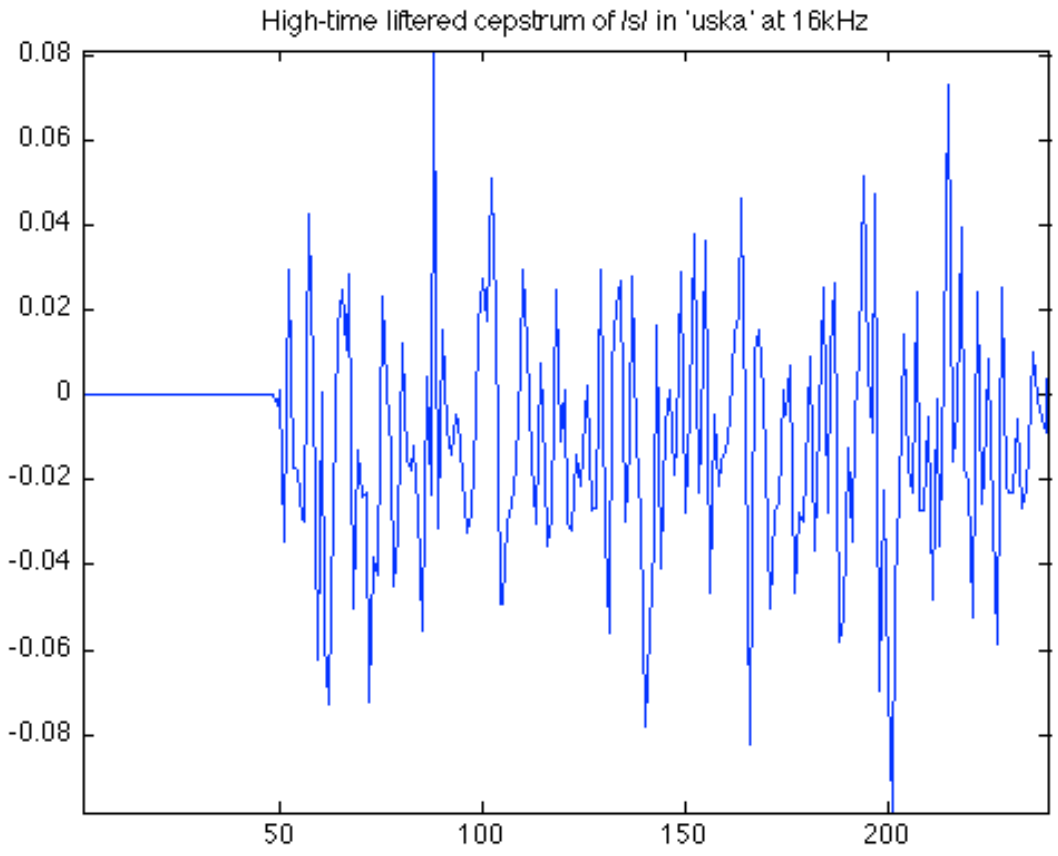
This result is quite close to the pitch value the sound was generated from in the previous assignment. The estimated parameters for ‘a_pani.wav’ (**in the previous assignment**):

```
pitch1 = 78.1250Hz
gain1 =1.6950
LPCCoeffs1 =
  Columns 1 through 7
    0.4349    0.6849   -0.6386   -0.2396    0.0940    0.1864    0.2259
  Columns 8 through 10
   -0.3754   -0.1771    0.4051
```

Question 5


Real cepstrum of /s/ in 'uska' at 16khz
c(n)
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
50
100
150
200
250
300
350
400
450


Vocal Tract Response (green) of /s/ in 'uska' at 16khz
in superimposed on LPC spectrum (red) and magnitude spectrum (blue)
Magnitude (logarithmic)
Frequency (Hz)
2
1
0
-1
-2
-3
-4
-5
0
1000
2000
3000
4000
5000
6000
7000



The script:

```
close all; clear all;

addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
dataFiles/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q1/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q2/
addpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q3/

%% plot real cepstrum of /s/ in 'uska' at 16khz

rc1 = getRealCepstrum('s_uska_16k.wav');
figure(100); clf;
plot(rc1); axis tight;
title('Real cepstrum of /s/ in ''uska'' at 16khz');
ylabel('c(n)');

%% plot vocal tract response of /s/ in 'uska' at 16khz

[vtr1, ww1] = getVocalTractResponse('s_uska_16k.wav');
```

```
[LPCSpectrum1, narrowbandSpectrum1, w1] =
getLPCSpectrum('s_uska_16k.wav', 18);

figure(200); clf;
vtrPlot1 = plot(ww1, vtr1);
hold on;
narrowbandPlot1 = plot(w1, narrowbandSpectrum1);
LPCPlot1 = plot(w1, LPCSpectrum1);
set(vtrPlot1, 'Color', 'green');
set(LPCPlot1,'Color','red');
set(narrowbandPlot1, 'Color', 'blue');
hold off;
axis tight;
title({'Vocal Tract Response (green) of /s/ in ''uska'' at
16khz';...
    'in superimposed on LPC spectrum (red) and magnitude
spectrum (blue)'});
xlabel('Frequency (Hz)');
ylabel('Magnitude (logarithmic)');

%% get the pitch of /s/ in 'uska' at 16khz

[pitchS16, htl1] = getPitchUsingCepstrum('s_uska_16k.wav');
pitchS16
figure(300); clf;
plot(htl1); axis tight;
title('High-time liftered cepstrum of /s/ in ''uska'' at
16kHz');



rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q3/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q2/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/q1/
rmpath /Users/swrangsarbasumatary/Desktop/speechAssignment4/
dataFiles/
```

Pitch for /s/ in ‘uska’ at 16kHz: 183.9080Hz