EE 679 : SPEECH PROCESSING
Computing Assignment 1
ANKIT AGRAWAL
10D070027

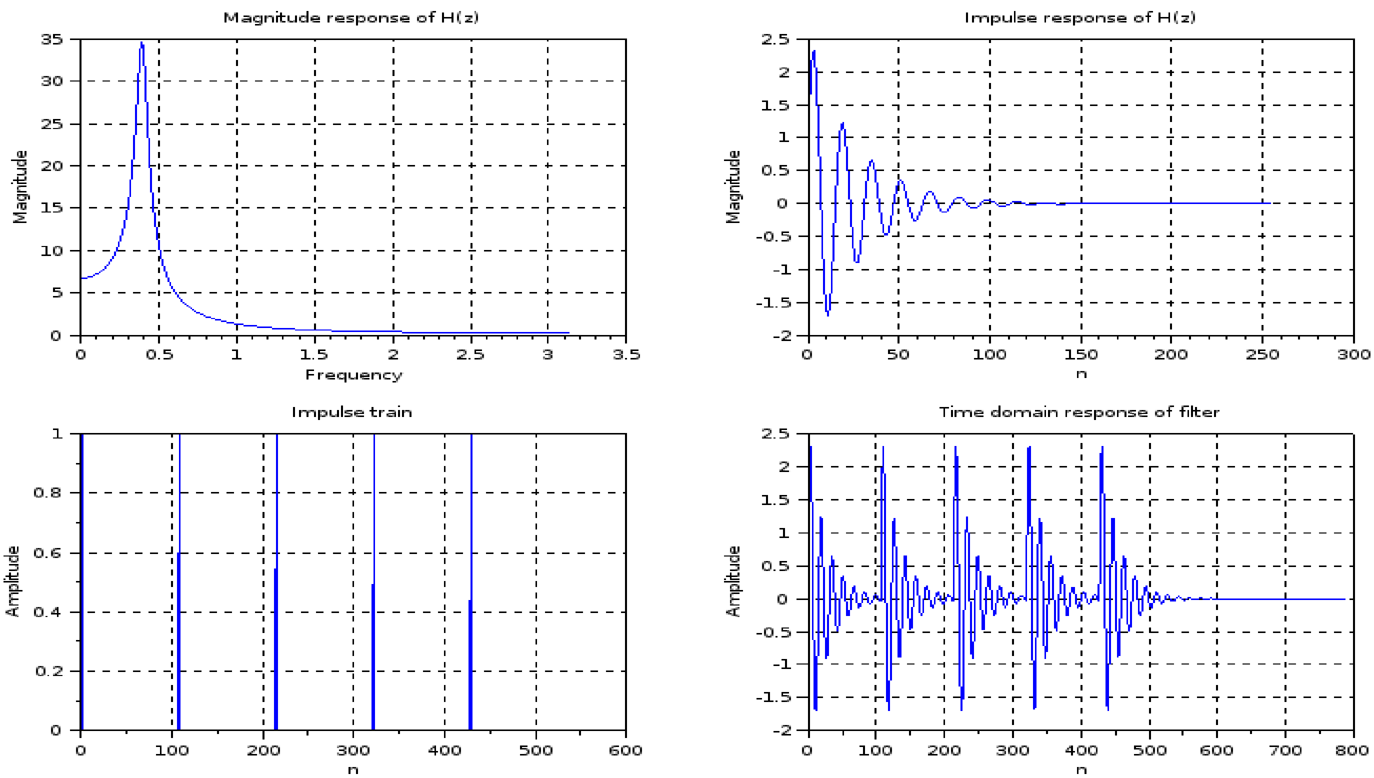
1) 1. Given the following specification for a single-formant resonator, obtain the transfer function of the filter H(z). Plot its magnitude response (dB magnitude versus frequency) and impulse response.
F1 (formant) = 1 kHz
B1(bandwidth) = 200 Hz
Fs (sampling freq) = 16 kHz
Ans : Single formant resonator can be modelled as

H(z) = 1 / ((1 - r*exp(j*theta)/z)*(1 - r*exp(-j*theta)/z))

where r = exp(-πB/Fs) and theta = 2πF1/Fs.

This gives r = 0.96 and theta = π / 8. Impulse response can be calculated after taking the inverse z transform of H(z).

h(n)=(r^n*sin((n+1)theta)/sin(theta))



2. Excite the above resonator (filter) with a source given by an impulse train of F0 = 150 Hz. Compute the output of the source-filter system over the duration of 0.5 second. Plot the time domain waveform. Also play it out and comment on the sound quality.

Ans : Check the plot in Q.1 for the time domain plots. The sound seems to have low frequency components. Sounds like /a/.

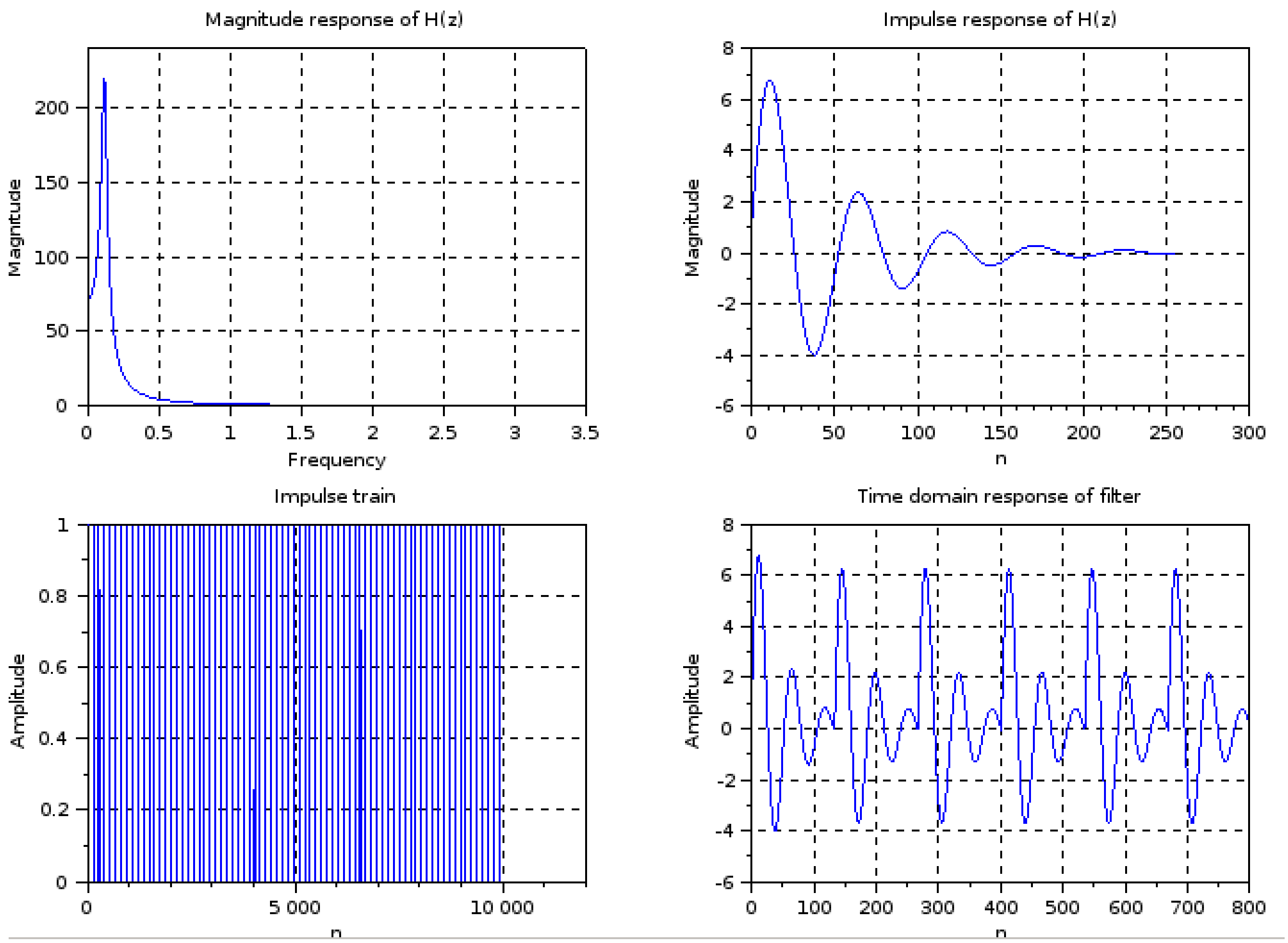
3. Vary the parameters as indicated below and comment on the differences in waveform and sound quality for the different parameter combinations.

(a) F1 = 300 Hz, B1 = 100 Hz, F0 = 120 Hz

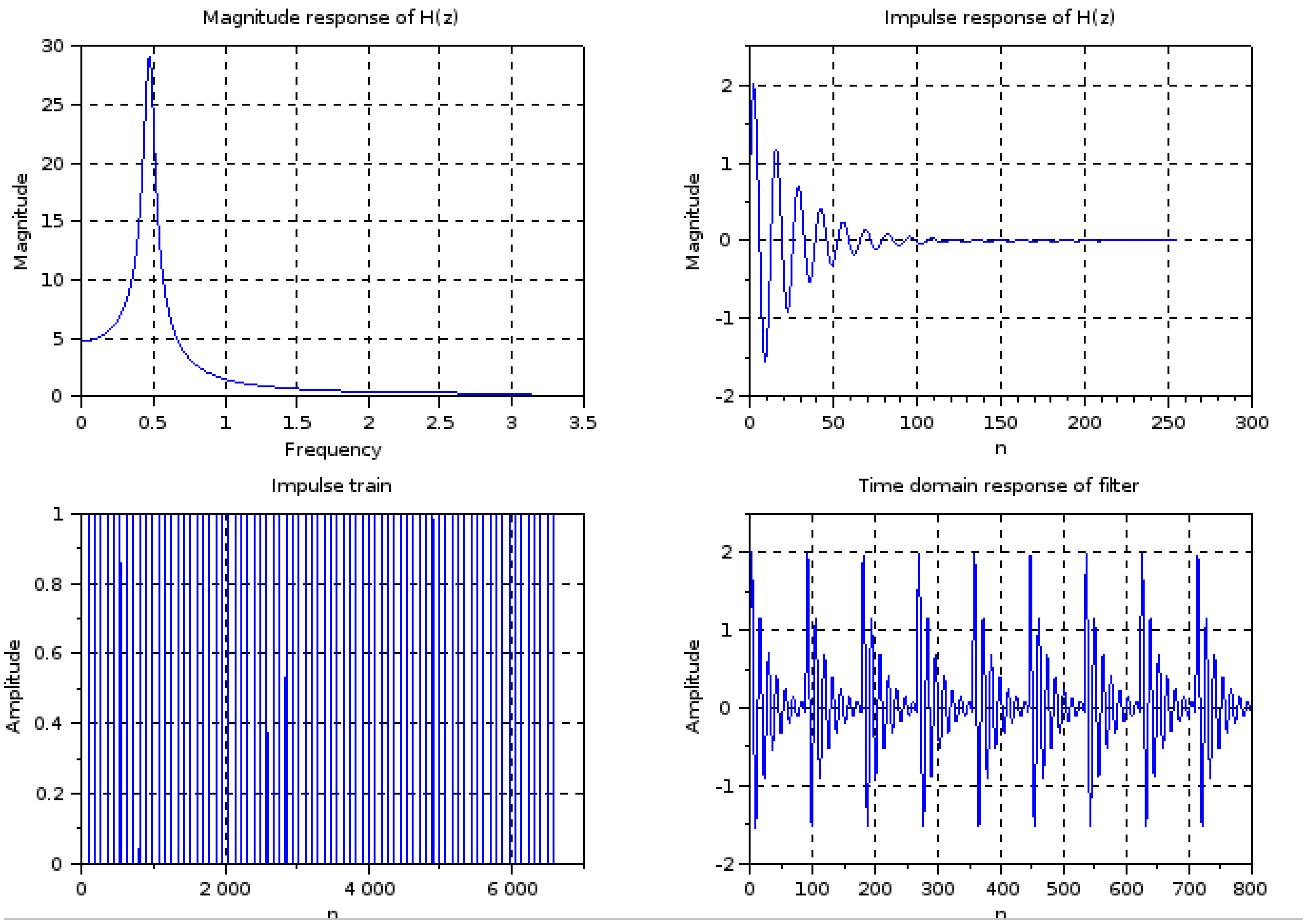
(b) F1=1200 Hz, B1 = 200 Hz, F0 = 180 Hz

Ans :

a)


Magnitude response of H(z)
Magnitude
Frequency
Impulse response of H(z)
Magnitude
n
Impulse train
Amplitude
n
Time domain response of filter
Amplitude
n

b)



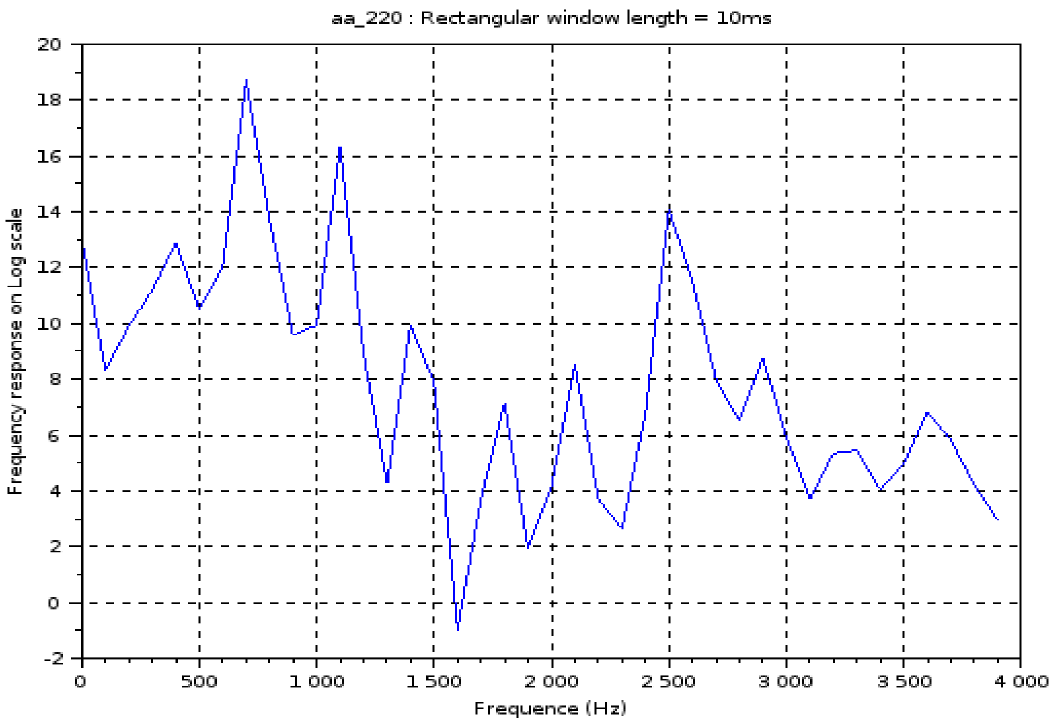
The first difference observed is in the pitch. The first one has a lower pitch than the second one. This is combined by the presence of more higher frequency components from the frequency spectrum.

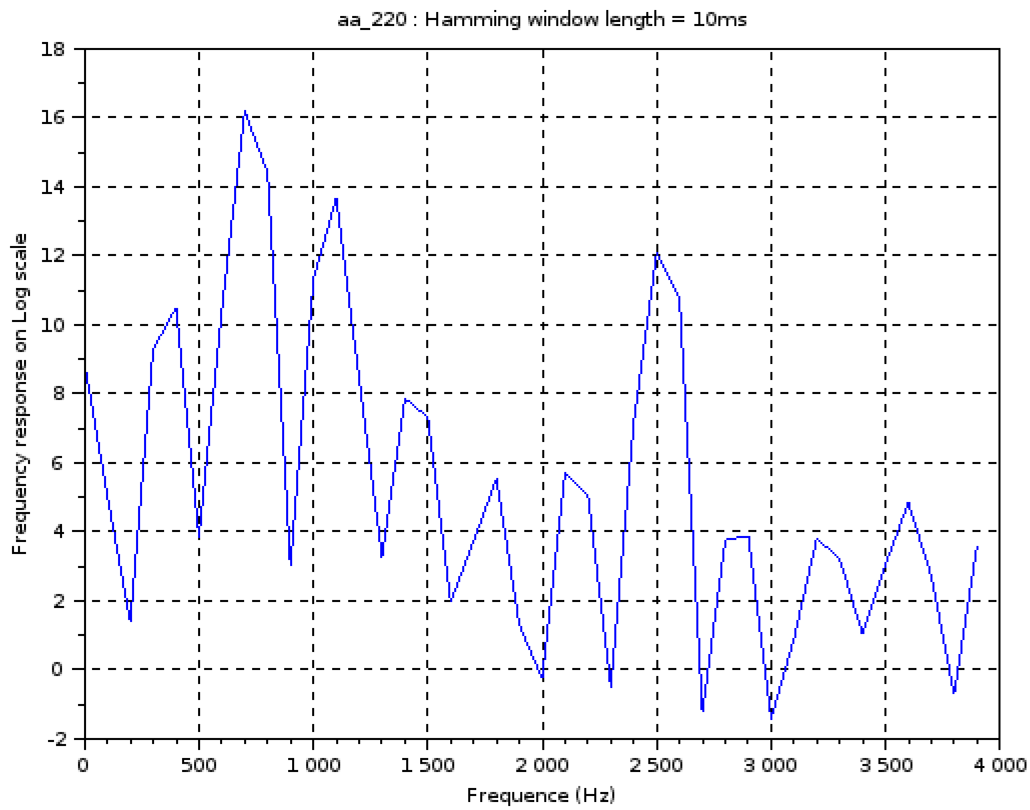
4) In place of the simple single-resonance signal, synthesize the following more realistic vowel sounds at two distinct pitches (F0 = 120 Hz, F0 = 220 Hz). Keep the bandwidths constant at 100 Hz for all formants. Duration of sound: 0.5 sec
Vowel F1, F2, F3
/a/ 730, 1090, 2440
/i/ 270, 2290, 3010
/u/ 300, 870, 2240
Ans : Please execute the code to generate all the plots and the sound files.

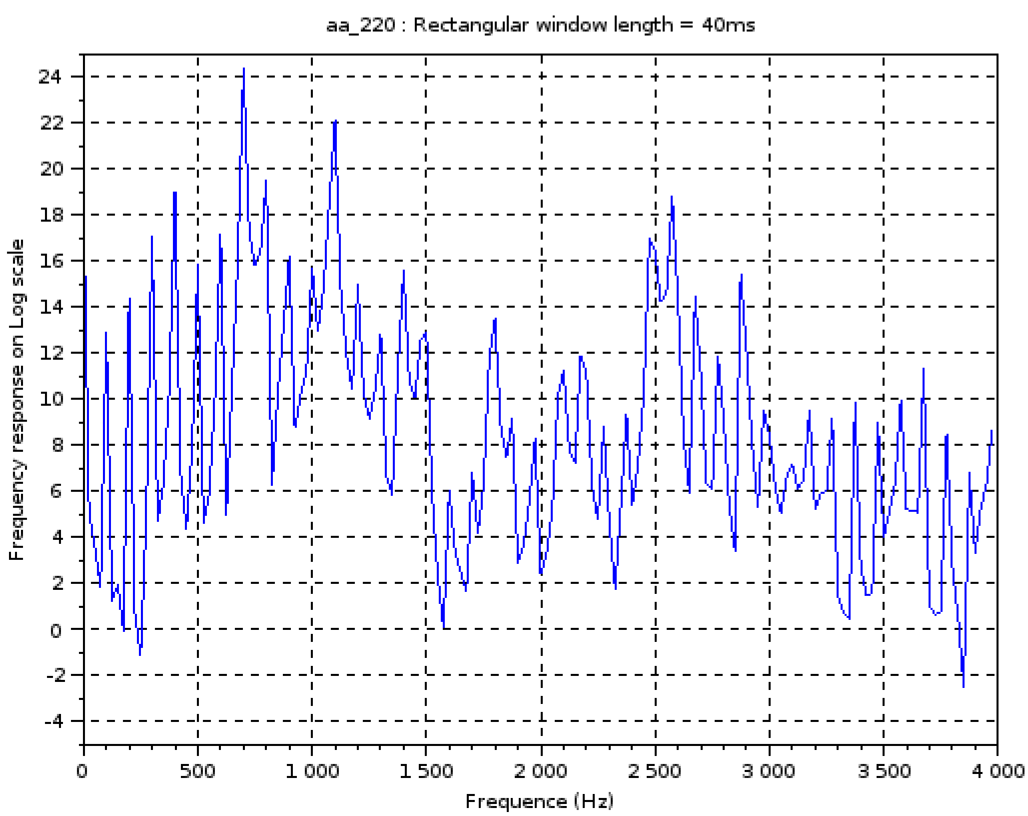
5. Signal Analysis:

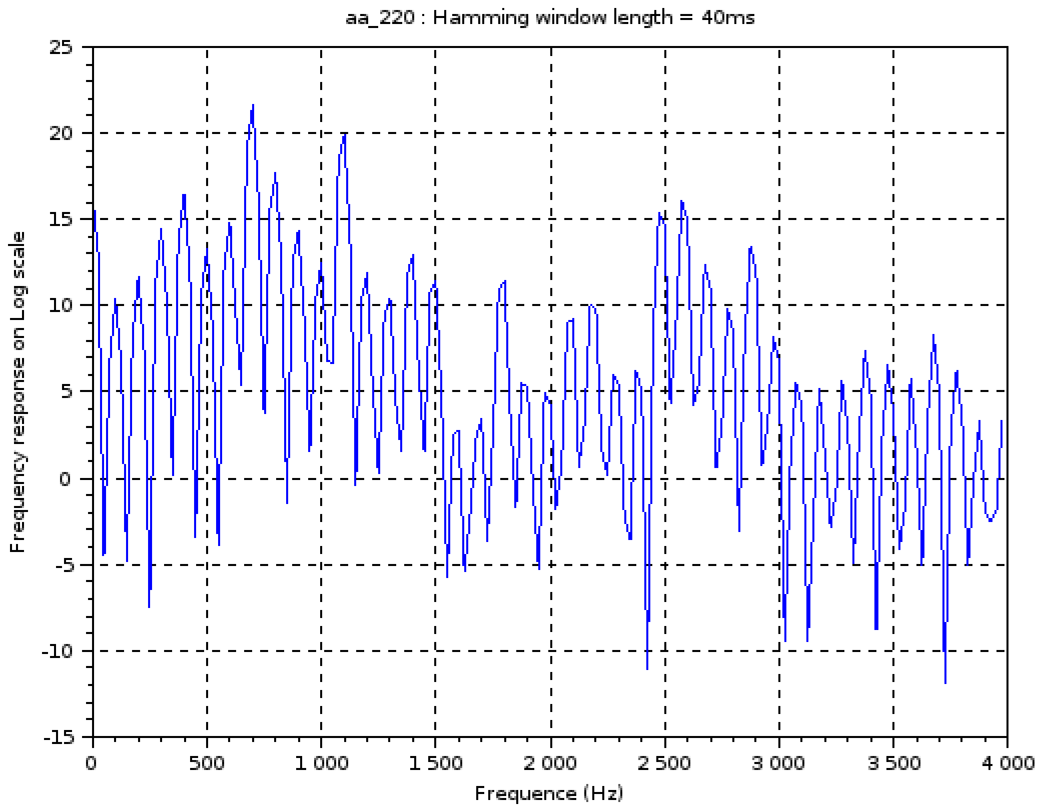
Compute the DTFT magnitude (dB) spectrum of any 2 of the vowel sounds you have synthesized. Use rectangular and Hamming windows of lengths: 10 ms, 40 ms, 100 ms. (i) Comment on the similarities and differences between the different spectra. (ii) Estimate the signal parameters from each of the spectra and compare with the ground-truth.

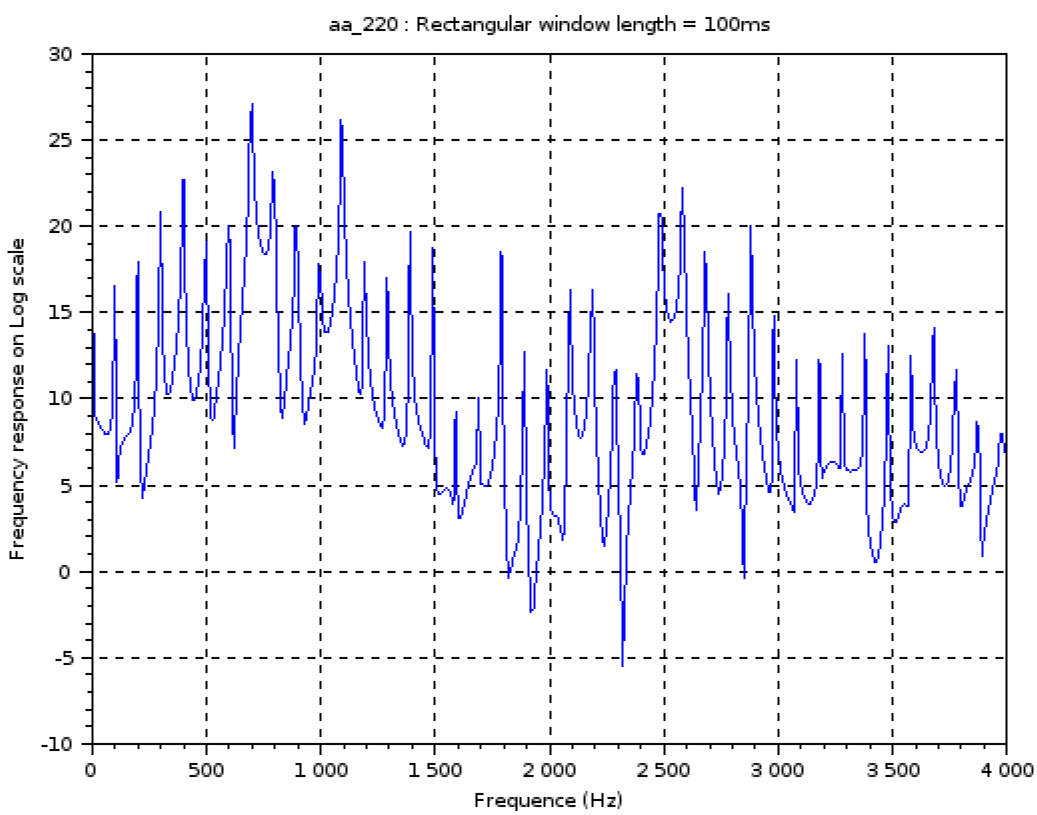
Ans :

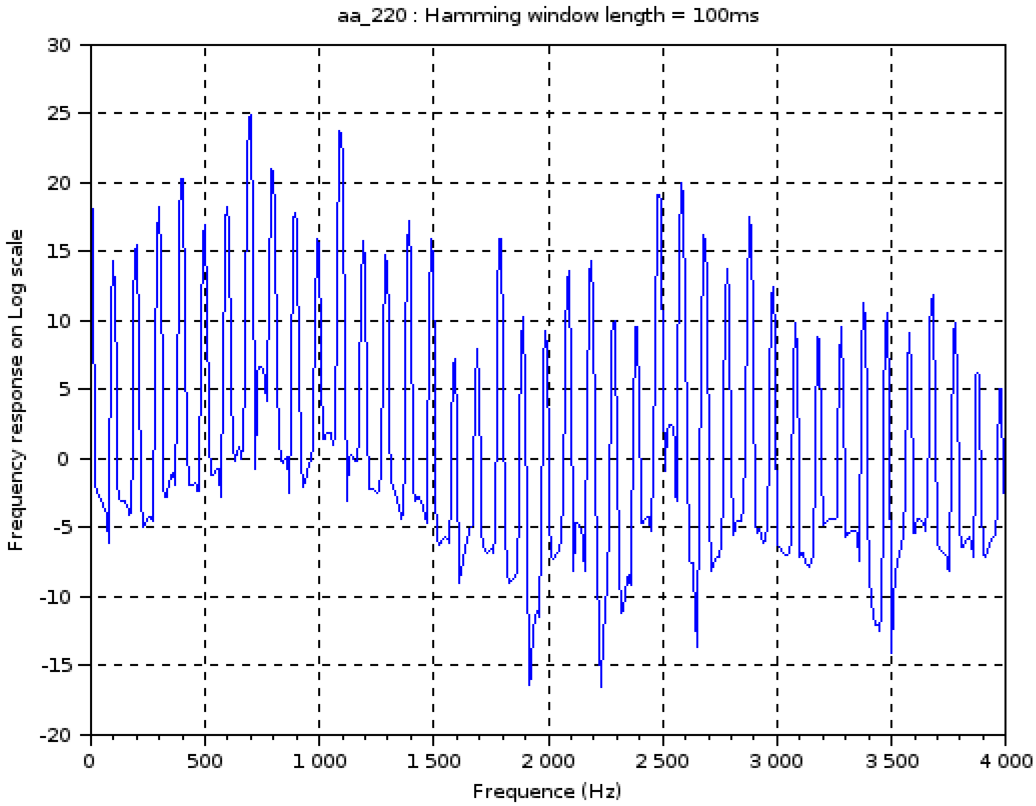


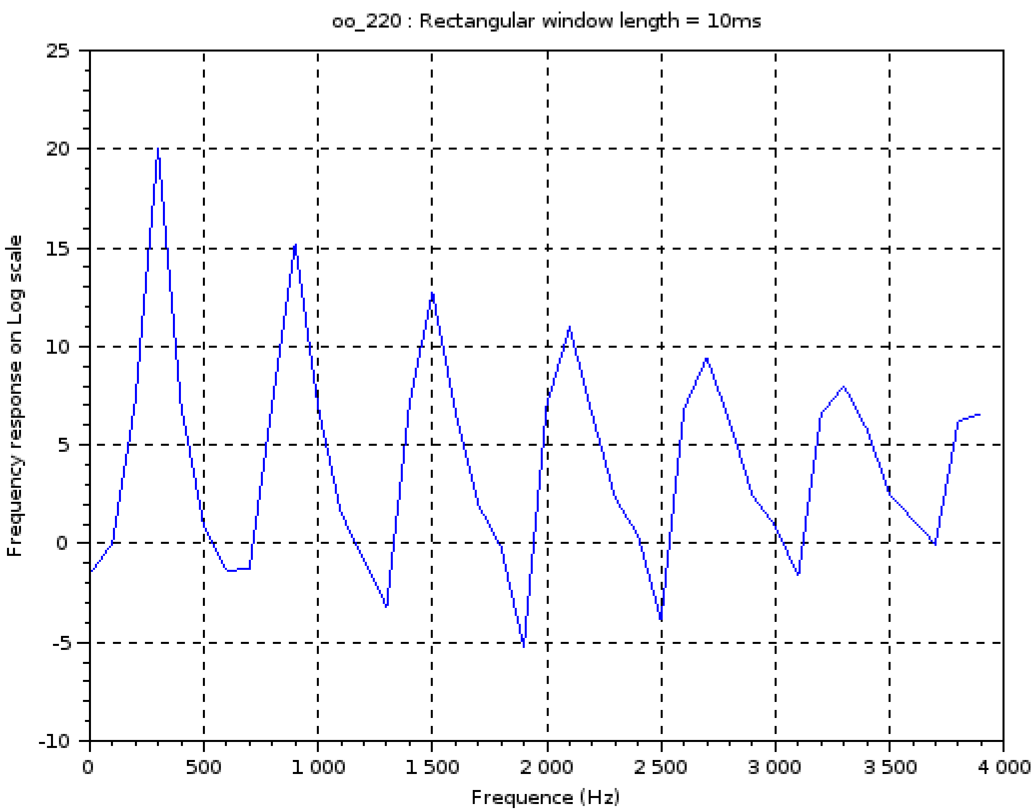

aa_220 : Hamming window length = 10ms
Frequency response on Log scale
Frequence (Hz)
18
16
14
12
10
8
6
4
2
0
-2
0
500
1 000
1 500
2 000
2 500
3 000
3 500
4 000

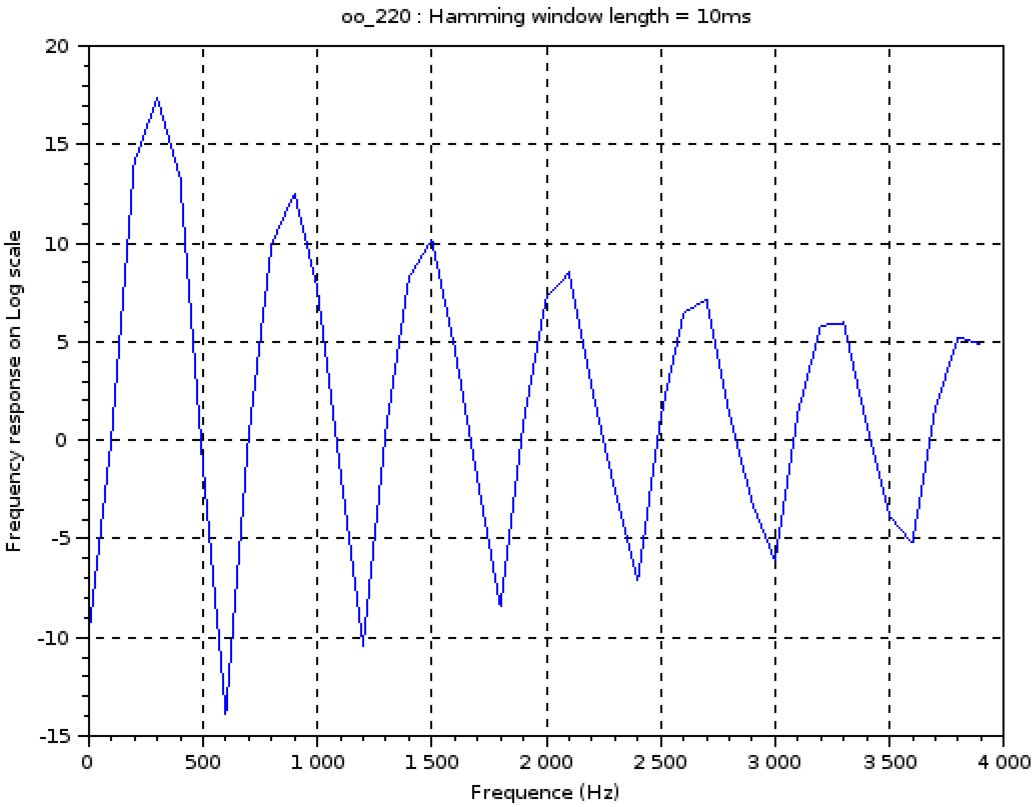

aa_220 : Rectangular window length = 40ms
Frequency response on Log scale
Frequence (Hz)
24
22
20
18
16
14
12
10
8
6
4
2
0
-2
-4
0
500
1 000
1 500
2 000
2 500
3 000
3 500
4 000

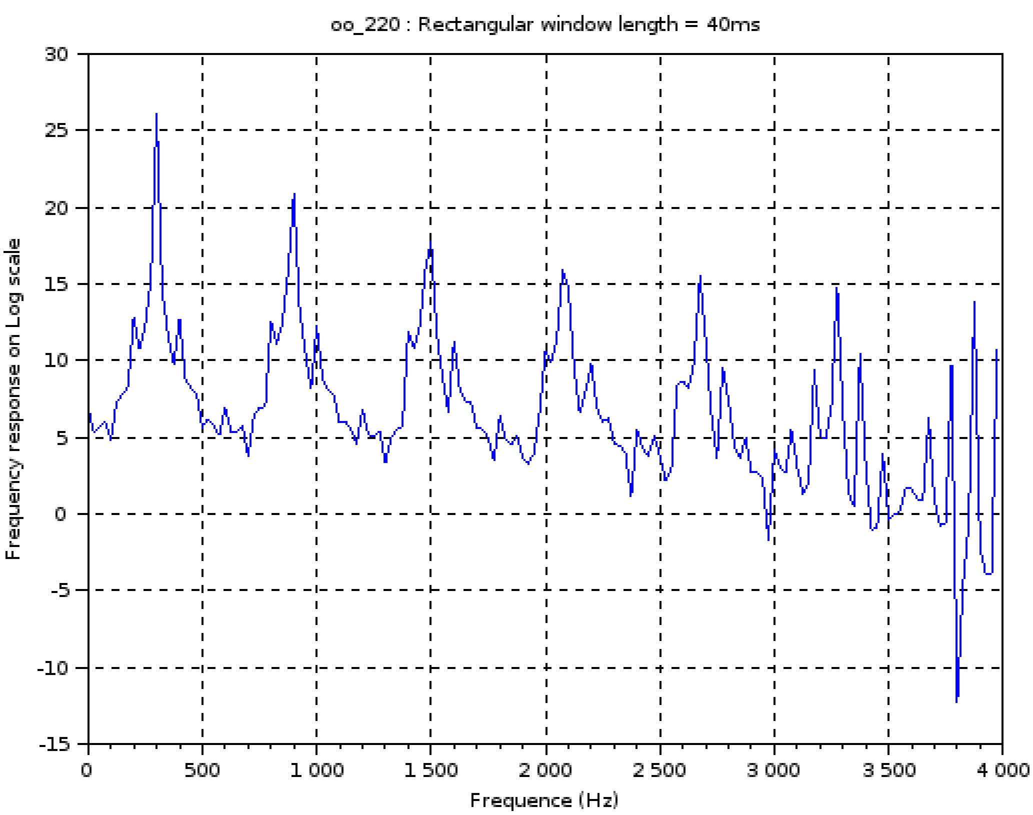

aa_220 : Hamming window length = 40ms
Frequency response on Log scale
25
20
15
10
5
0
-5
-10
-15
0
500
1 000
1 500
2 000
2 500
3 000
3 500
4 000
Frequence (Hz)

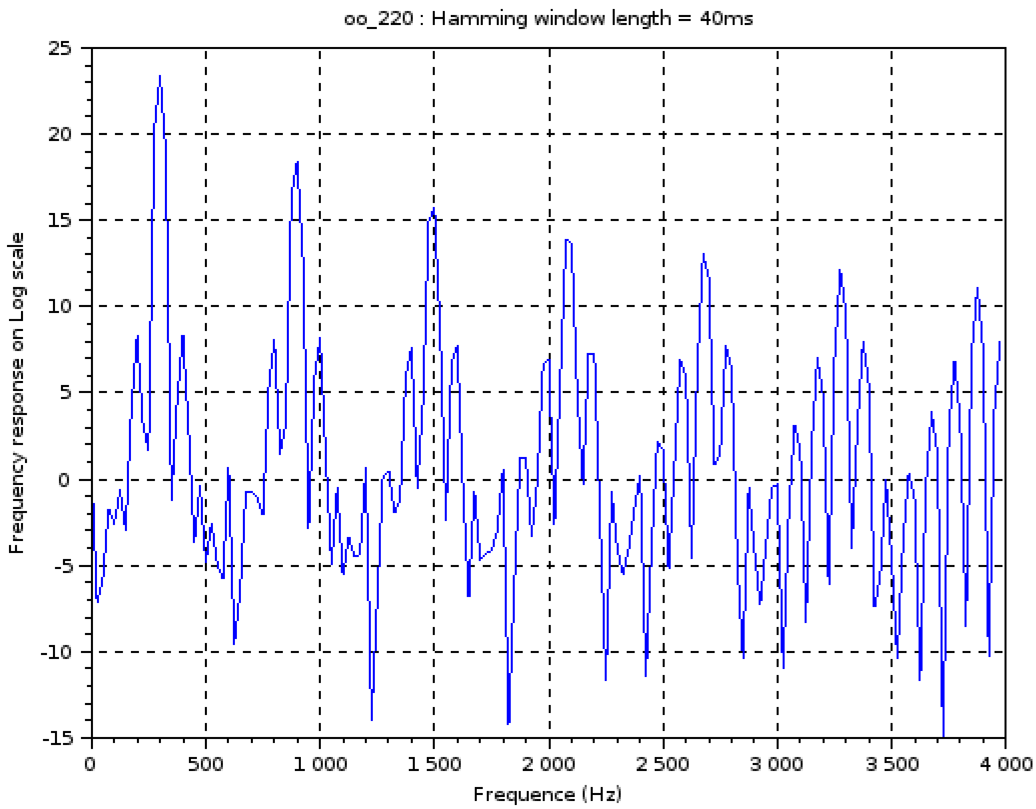

aa_220 : Rectangular window length = 100ms
Frequency response on Log scale
30
25
20
15
10
5
0
-5
-10
0
500
1 000
1 500
2 000
2 500
3 000
3 500
4 000
Frequence (Hz)


aa_220 : Hamming window length = 100ms
Frequency response on Log scale
Frequence (Hz)
30
25
20
15
10
5
0
-5
-10
-15
-20
0
500
1 000
1 500
2 000
2 500
3 000
3 500
4 000

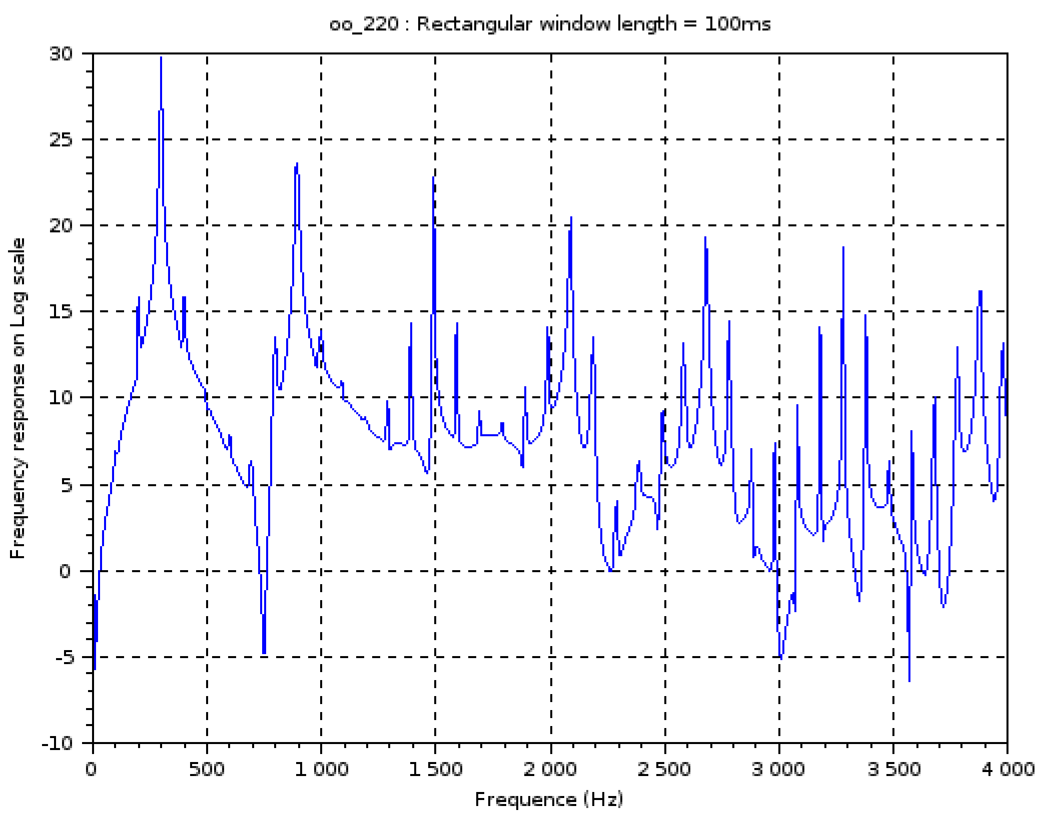

oo_220 : Rectangular window length = 10ms
Frequency response on Log scale
Frequence (Hz)
25
20
15
10
5
0
-5
-10
0
500
1 000
1 500
2 000
2 500
3 000
3 500
4 000


oo_220 : Hamming window length = 10ms
Frequency response on Log scale
Frequence (Hz)
20
15
10
5
0
-5
-10
-15
0
500
1 000
1 500
2 000
2 500
3 000
3 500
4 000

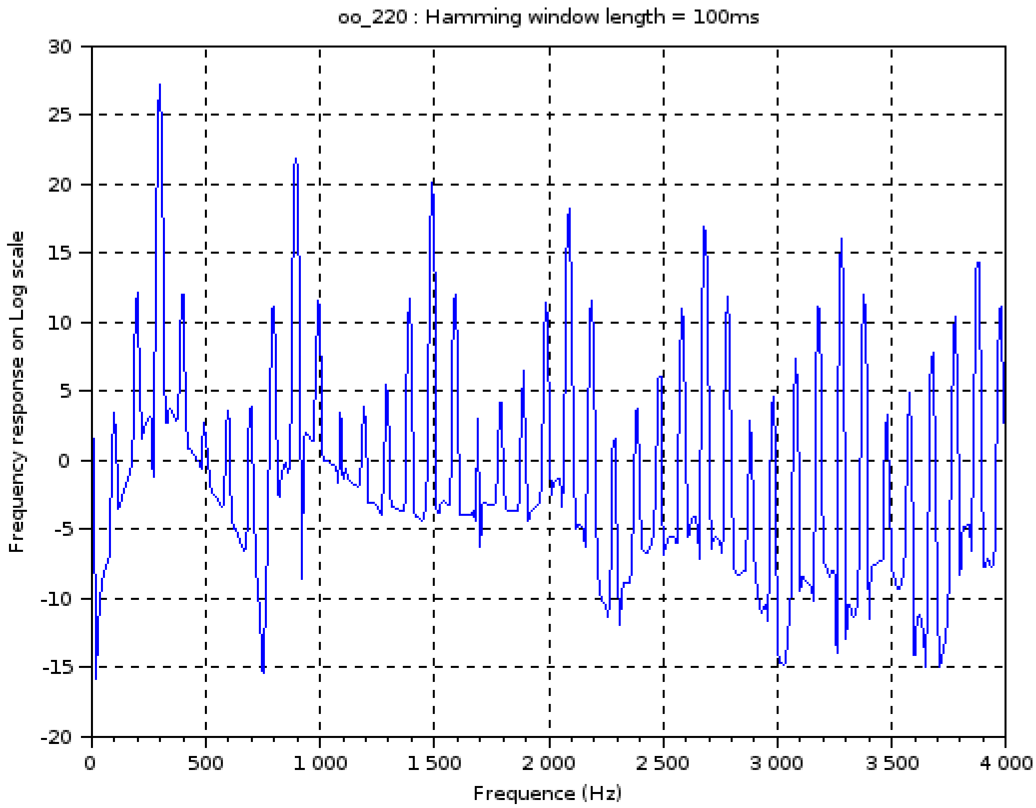

oo_220 : Rectangular window length = 40ms
Frequency response on Log scale
Frequence (Hz)
30
25
20
15
10
5
0
-5
-10
-15
0
500
1 000
1 500
2 000
2 500
3 000
3 500
4 000



oo_220 : Hamming window length = 40ms
Frequency response on Log scale
Frequence (Hz)



oo_220 : Rectangular window length = 100ms
Frequency response on Log scale
Frequence (Hz)



The sounds considered here for analysis are /aa/ at 220 Hz and /oo/ at 220 Hz.
The formants in the spectrum can be conveniently analyzed in all the above spectrums, but can be relatively better analyzed in the

i) Smaller time windows as compared to larger time windows because of less undulations(because of low no. of samples)
ii) Hamming window as compared to Rectangular because of less spectral leakage.

The first two formants in /oo/ seem to be around 300 Hz and 850 Hz while that of /aa/ seem to be near 700 Hz and 1100 Hz and hence are in accordance with the first two formant frequencies used for synthesis. The observed values reinforce the Vowel Quadrilateral model where the first formant is more discriminative for the above two vowels. The third formant for /oo/ is observed at around 1500 Hz which is quite far to the third formant frequency used for synthesis(2240 Hz). The third formant for is observed at around 2550 Hz which is closed to the one used for synthesis(2440 Hz).