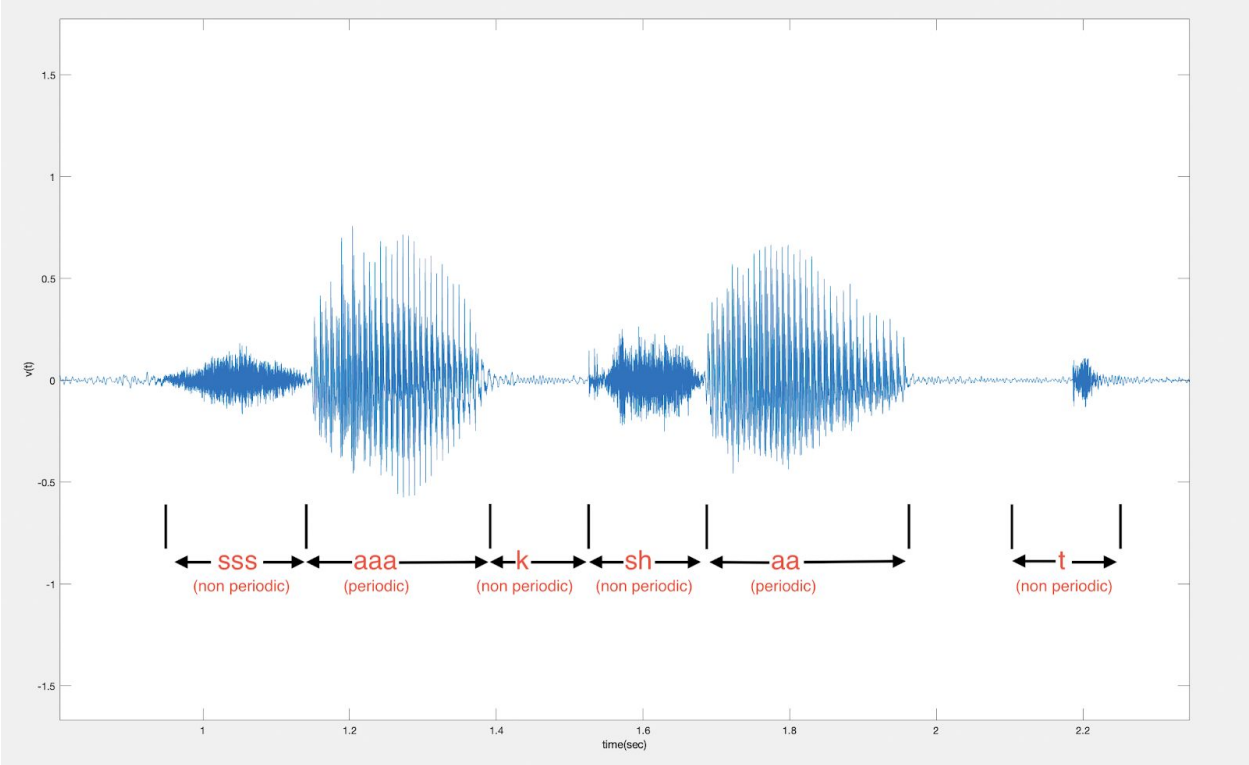
Speech Processing
Lab-1 Report

Rajat Tyagi 180020029

Task A

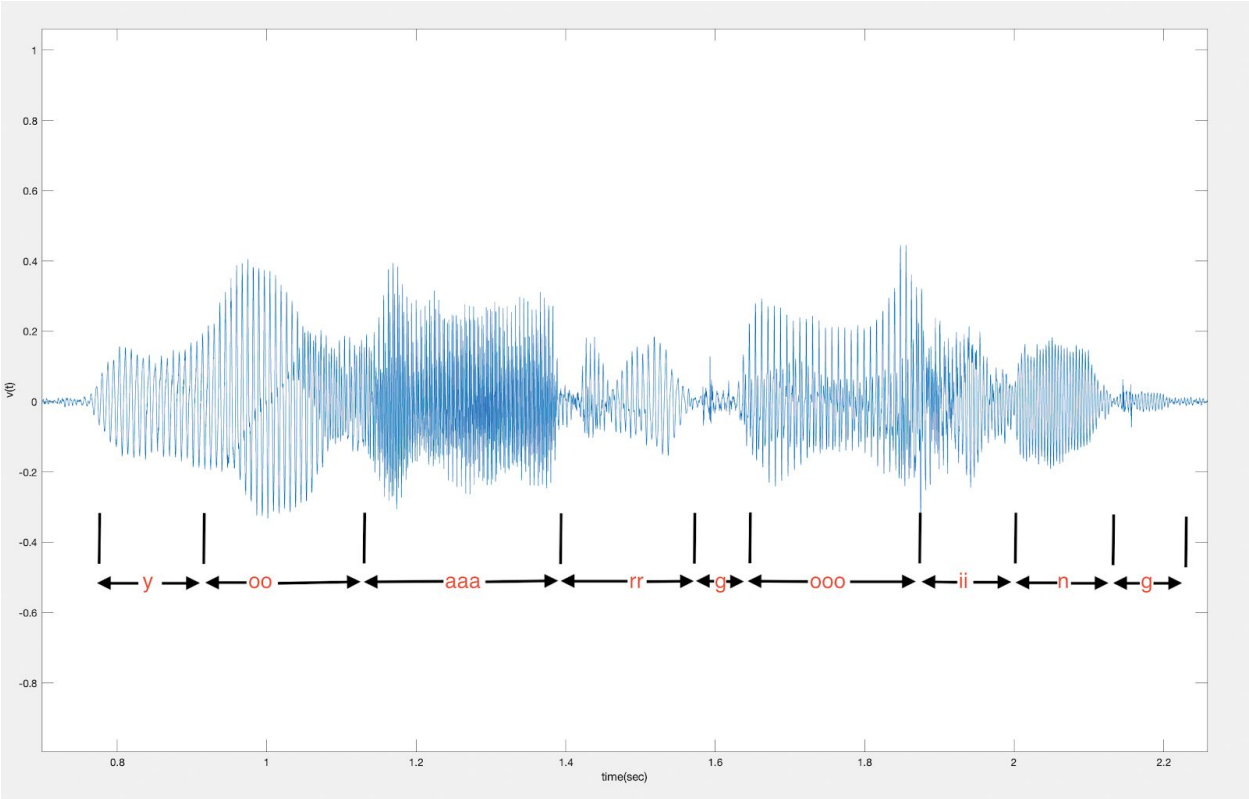
Recording of the phrase **“Sakshat”** with appropriate speech labels**.**



- We can see that the phrase sakshat can be divided into 6 parts as shown in the figure.
- Out of these 6 parts we have 2 parts which are vowels and hence are periodic in nature.

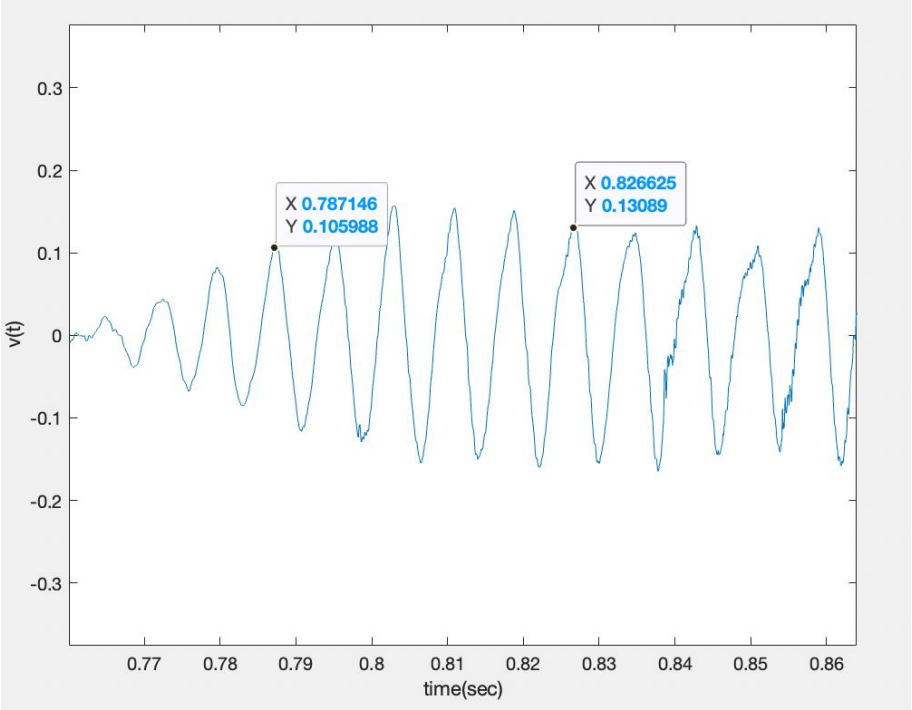
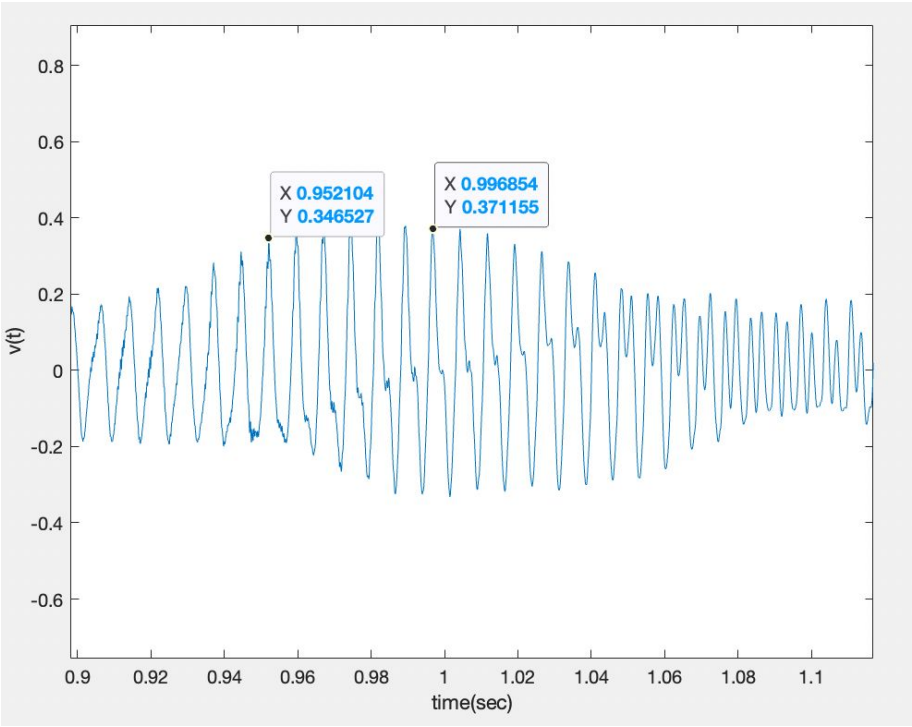
Task B

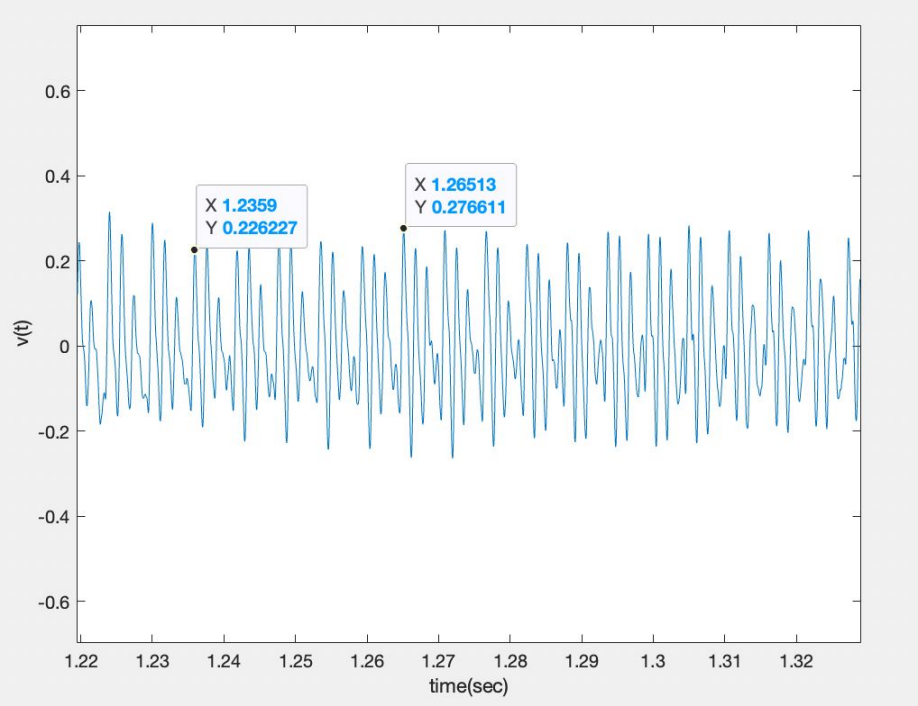
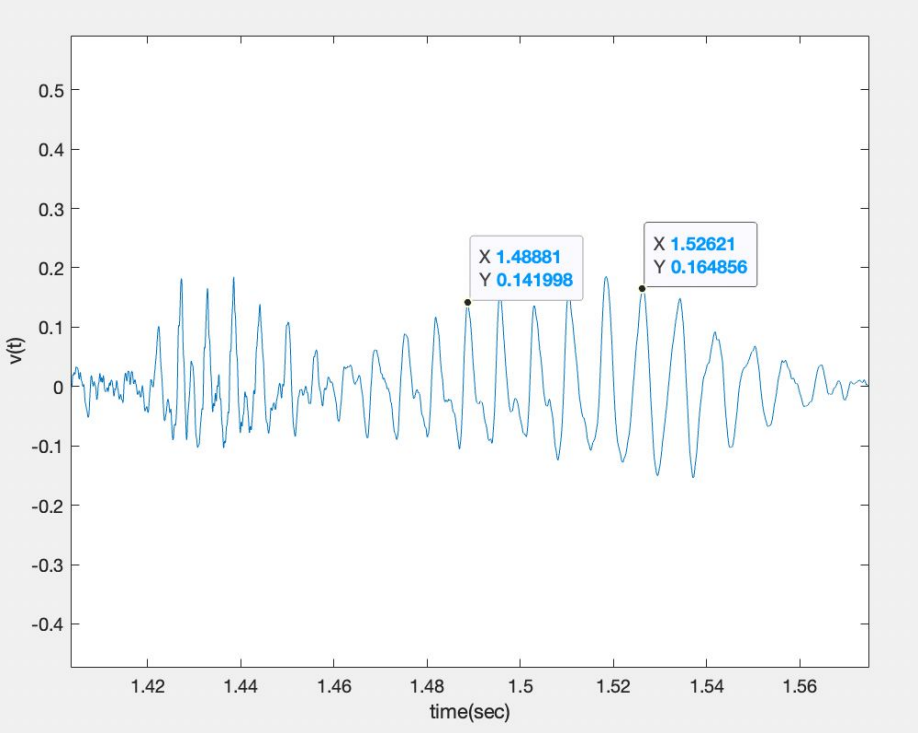
Recording of the interrogative phrase **“Where are you going?”** with appropriate speech labels.

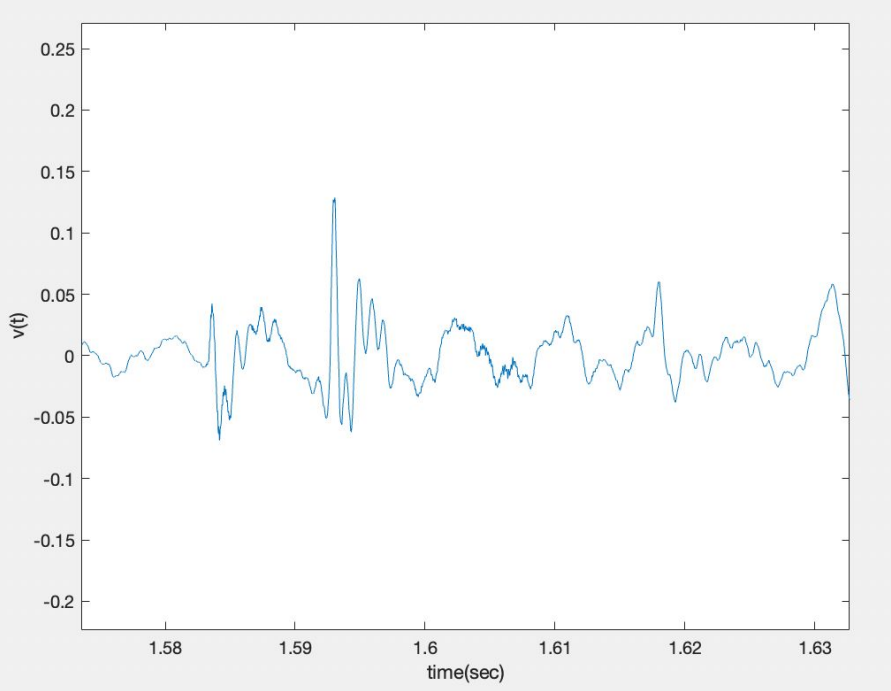
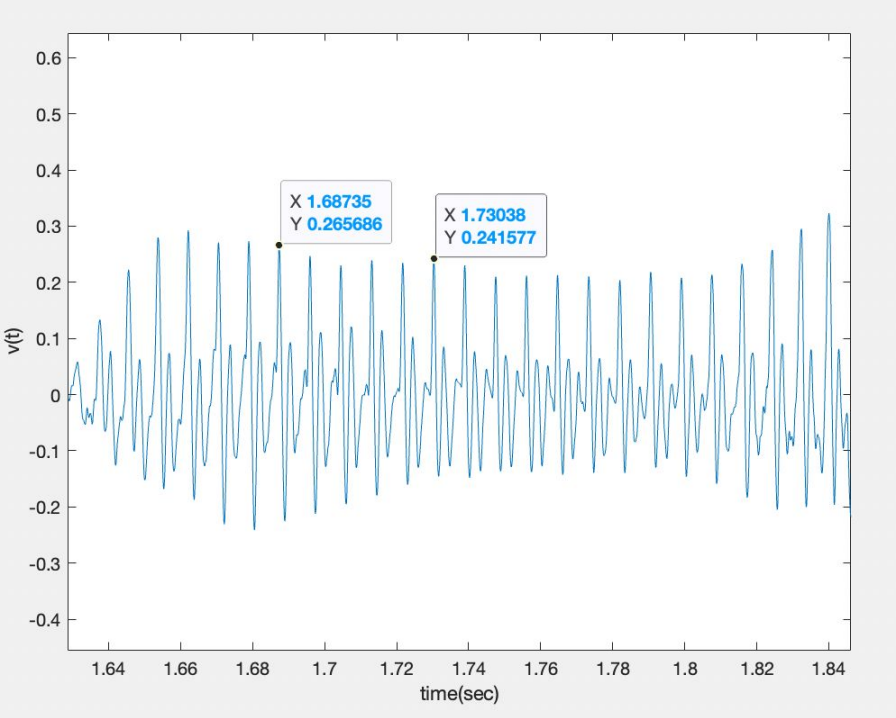


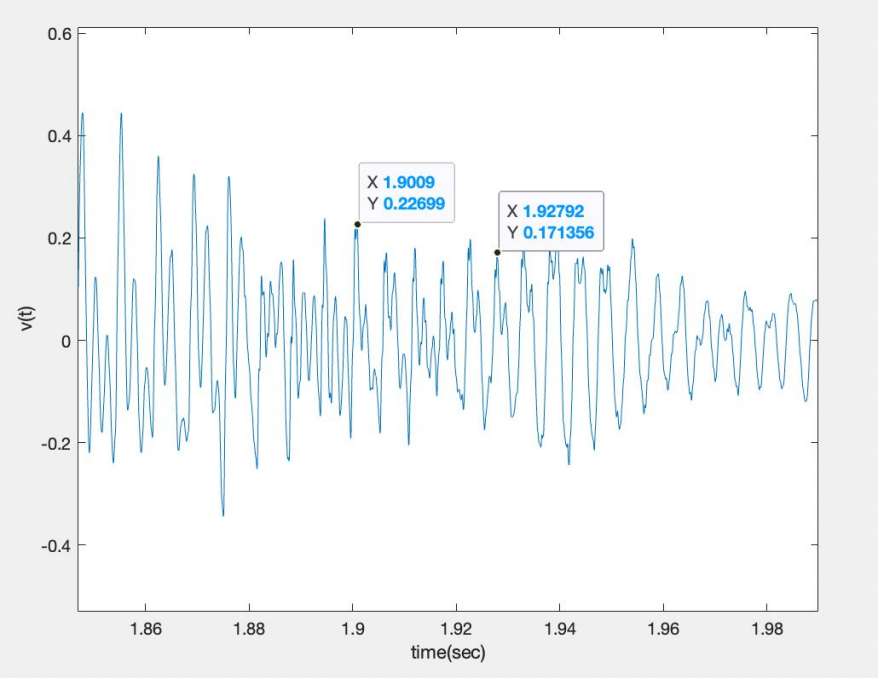
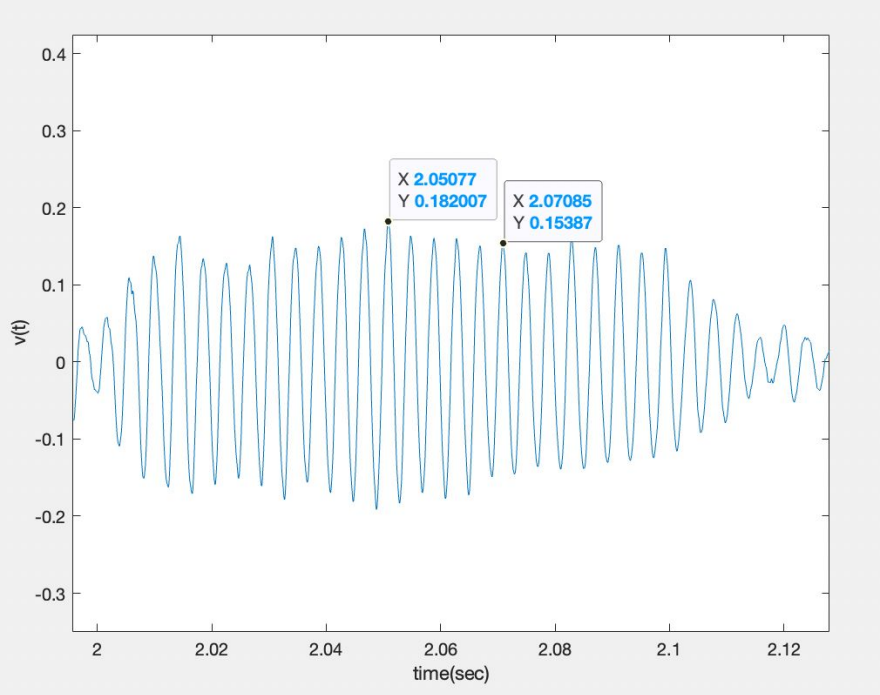
- We can see the phrase “Where are you going?” can be divided into 9 parts.

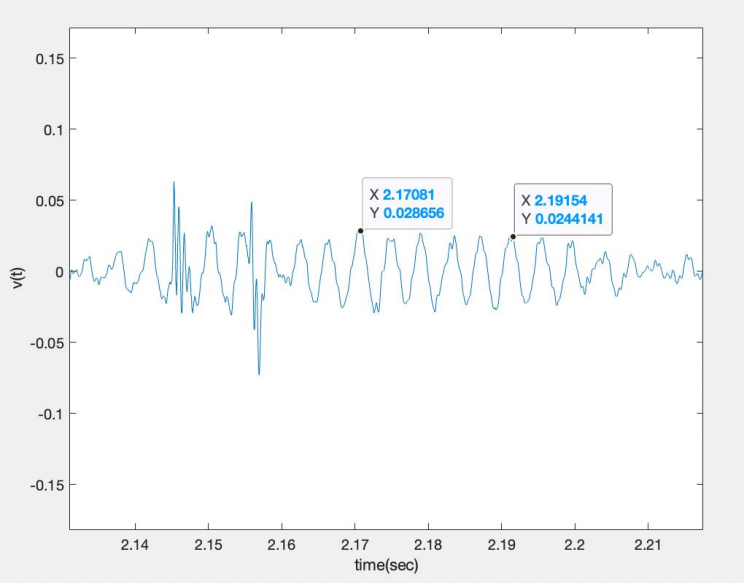
Part by part zoomed in waveform:

	Sound	Zoomed in Waveform
1	y	
2	u	

3	aa	 X 1.2359 Y 0.226227 X 1.26513 Y 0.276611 v(t) time(sec)
4	rr	 X 1.48881 Y 0.141998 X 1.52621 Y 0.164856 v(t) time(sec)

5	g	 0.25 0.2 0.15 0.1 0.05 0 -0.05 -0.1 -0.15 -0.2 v(t) 1.58 1.59 1.6 1.61 1.62 1.63 time(sec)
6	oo	 X 1.68735 Y 0.265686 X 1.73038 Y 0.241577 0.6 0.5 0.4 0.3 0.2 0.1 0 -0.1 -0.2 -0.3 -0.4 v(t) 1.64 1.66 1.68 1.7 1.72 1.74 1.76 1.78 1.8 1.82 1.84 time(sec)

7	ii	
8	n (nasal)	

9	g	

Part by part frequency calculation:

Method:

To calculate the time interval between 2 peaks we note time at a peak A and then again note the time at peak B which is 5th from peak A.
We do this to minimize error.

	Sound	Point A	Point B	Period $\tau = \frac{(B-A)}{5}$	Frequency $f = \frac{1}{\tau}$
1	y	0.787146 s	0.826625 s	7.89 ms	126.64 Hz
2	u	0.952104 s	0.996854 s	8.95 ms	111.73 Hz
3	aa	1.23590 s	1.26513 s	5.84 ms	171.05 Hz
4	rr	1.48881 s	1.52621 s	7.48 ms	133.68 Hz
5	g	-	-	-	-
6	oo	1.68735 s	1.73038 s	8.60 ms	116.19 Hz
7	ii	1.90090 s	1.92792 s	5.40 ms	185.04 Hz
8	n(nasal)	2.05077 s	2.07085 s	4.16 ms	249.00 Hz
9	g	-	-	-	-

Plot of frequency and sound: