

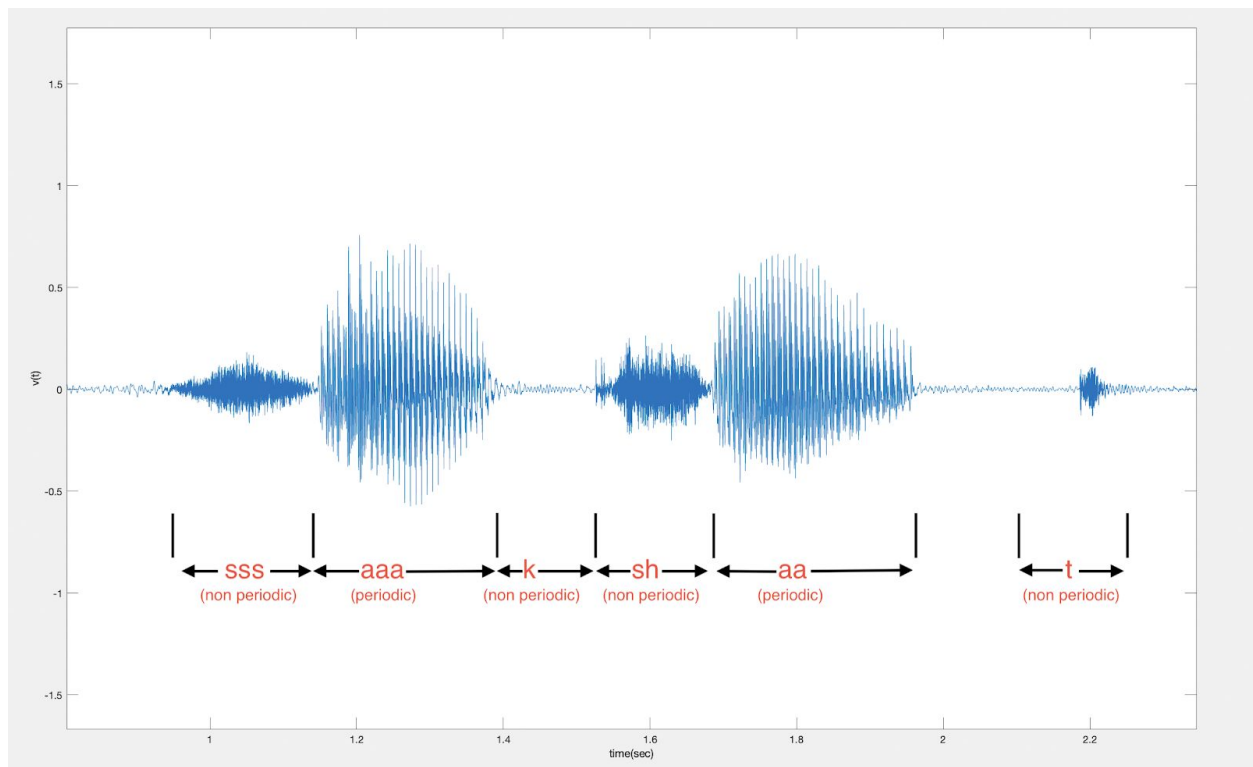
Speech Processing

Lab-1 Report

Rajat Tyagi 180020029

Task A

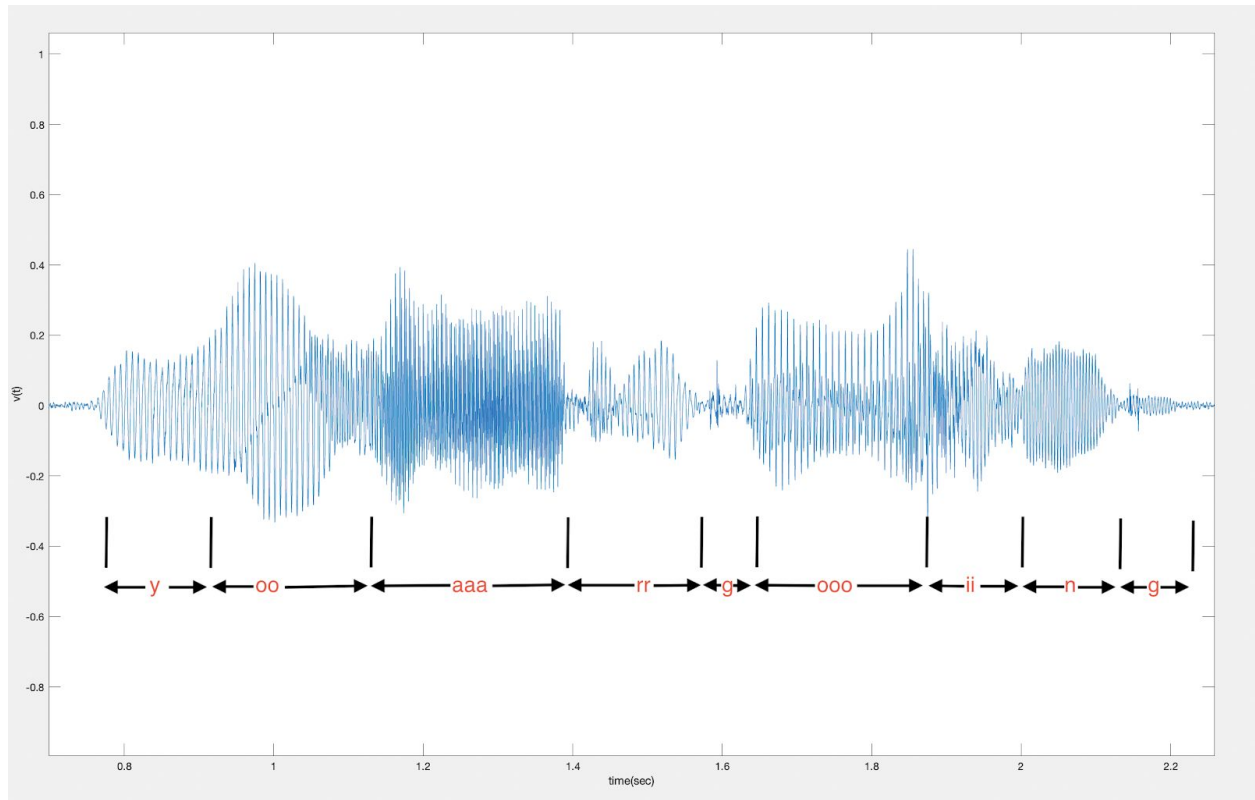
Recording of the phrase “**Sakshat**” with appropriate speech labels.



- 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.

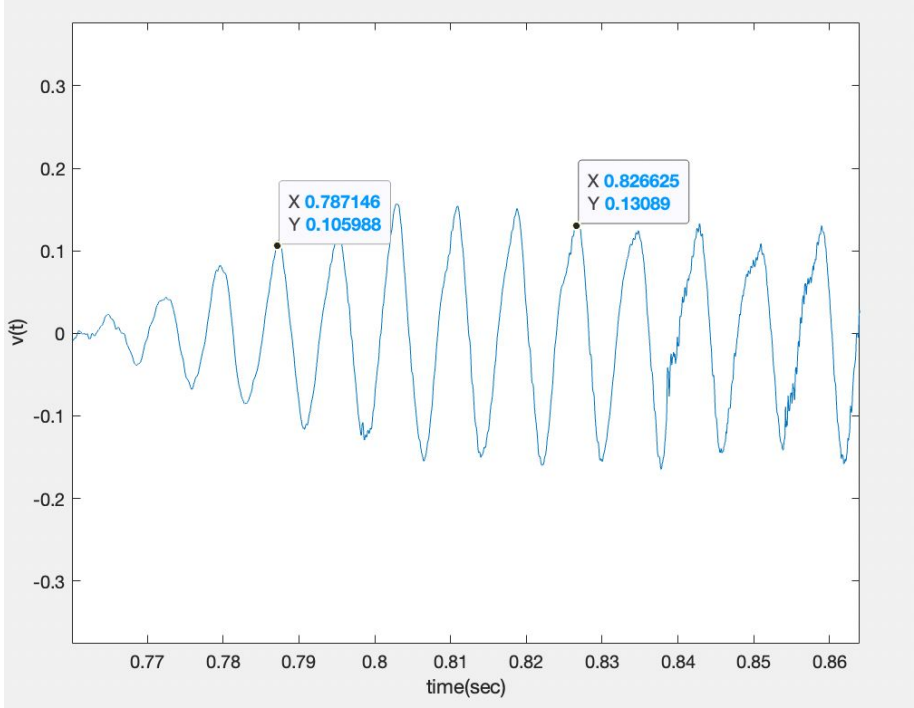
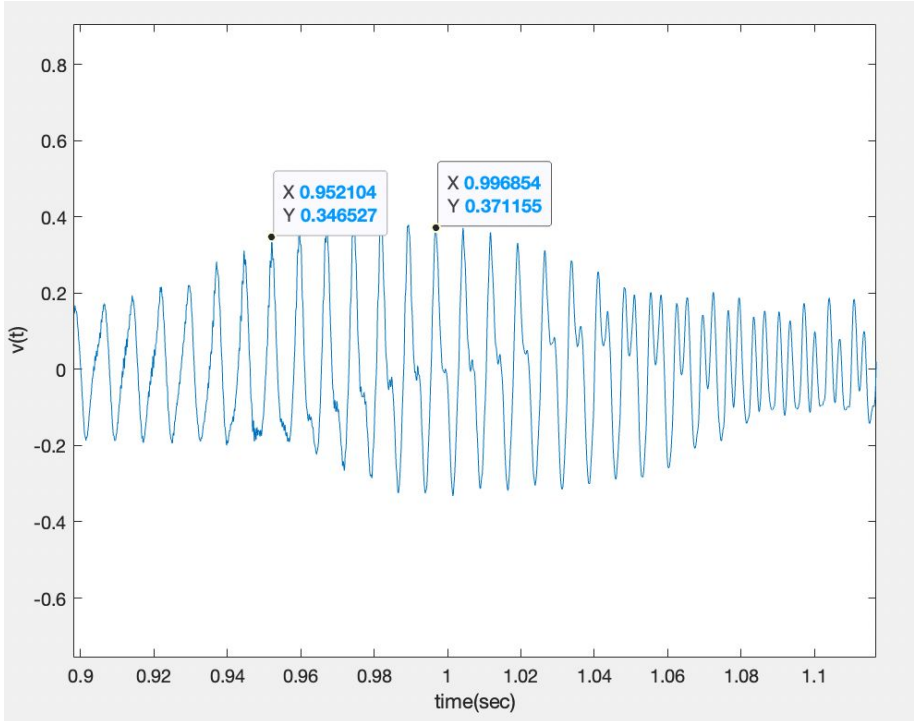
Task B

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



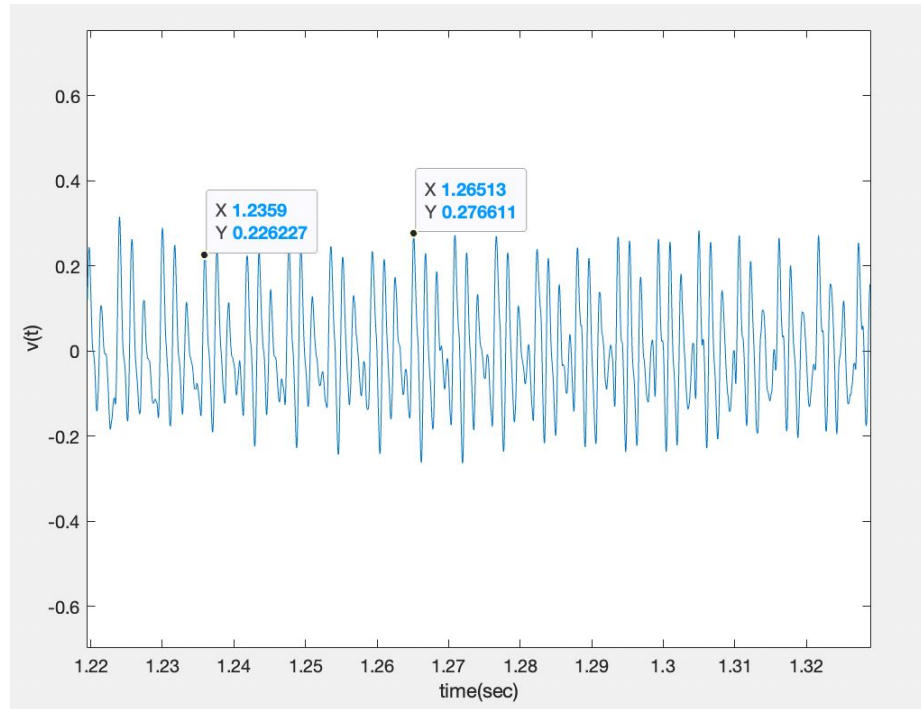
- We can see the phrase “Where are you going?” can be divided into 9 parts.

Part by part zoomed in waveform:

	Sound	Zoomed in Waveform									
1	y	 <p>The waveform for sound 'y' is a blue line plot of voltage $v(t)$ versus time in seconds. The x-axis ranges from 0.77 to 0.86 with major ticks every 0.01 seconds. The y-axis ranges from -0.3 to 0.3 with major ticks every 0.1 units. Two specific points on the waveform are highlighted with data labels:</p> <table border="1"><thead><tr><th>Point</th><th>X (time(sec))</th><th>Y ($v(t)$)</th></tr></thead><tbody><tr><td>1</td><td>0.787146</td><td>0.105988</td></tr><tr><td>2</td><td>0.826625</td><td>0.13089</td></tr></tbody></table>	Point	X (time(sec))	Y ($v(t)$)	1	0.787146	0.105988	2	0.826625	0.13089
Point	X (time(sec))	Y ($v(t)$)									
1	0.787146	0.105988									
2	0.826625	0.13089									
2	u	 <p>The waveform for sound 'u' is a blue line plot of voltage $v(t)$ versus time in seconds. The x-axis ranges from 0.9 to 1.1 with major ticks every 0.02 seconds. The y-axis ranges from -0.6 to 0.8 with major ticks every 0.2 units. Two specific points on the waveform are highlighted with data labels:</p> <table border="1"><thead><tr><th>Point</th><th>X (time(sec))</th><th>Y ($v(t)$)</th></tr></thead><tbody><tr><td>1</td><td>0.952104</td><td>0.346527</td></tr><tr><td>2</td><td>0.996854</td><td>0.371155</td></tr></tbody></table>	Point	X (time(sec))	Y ($v(t)$)	1	0.952104	0.346527	2	0.996854	0.371155
Point	X (time(sec))	Y ($v(t)$)									
1	0.952104	0.346527									
2	0.996854	0.371155									

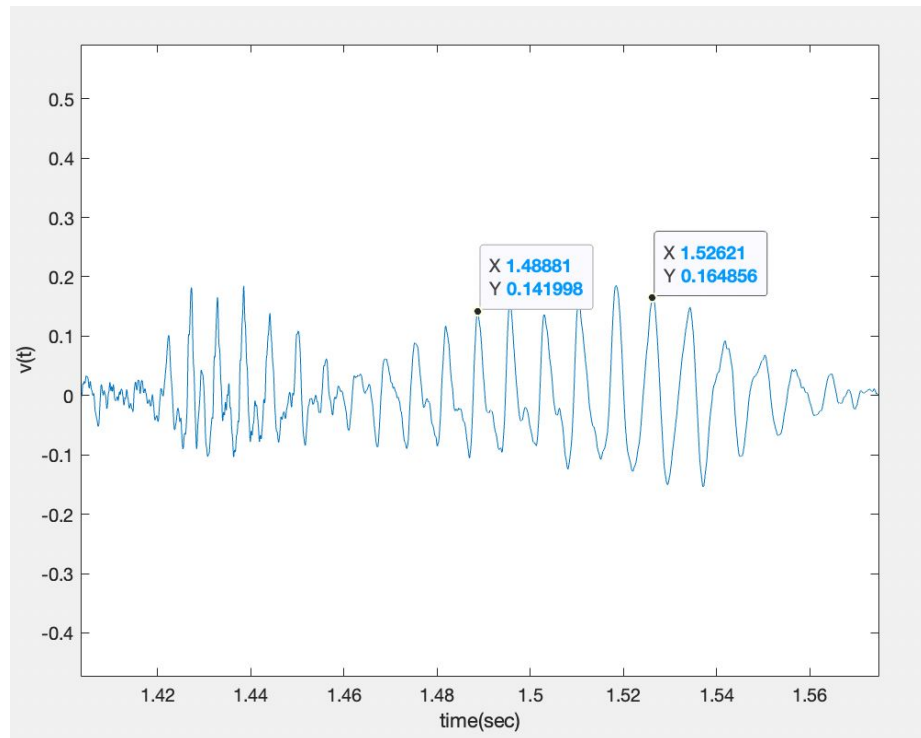
3

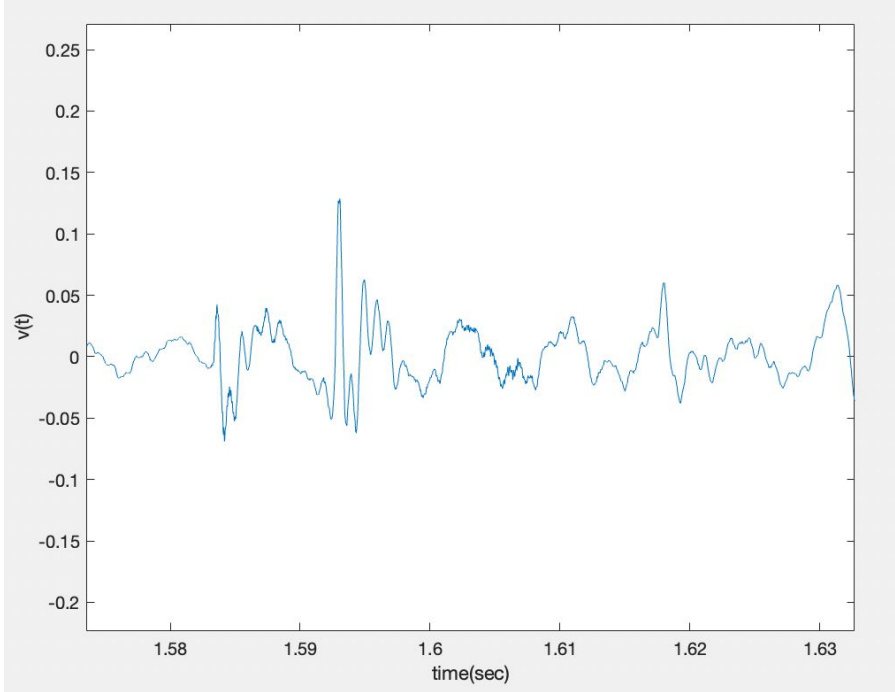
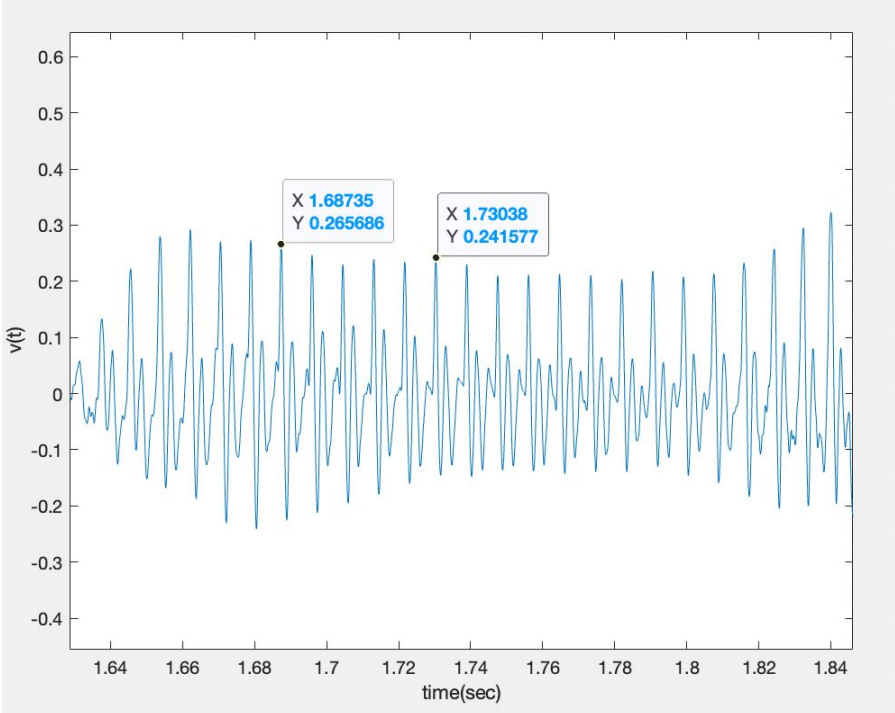
aa

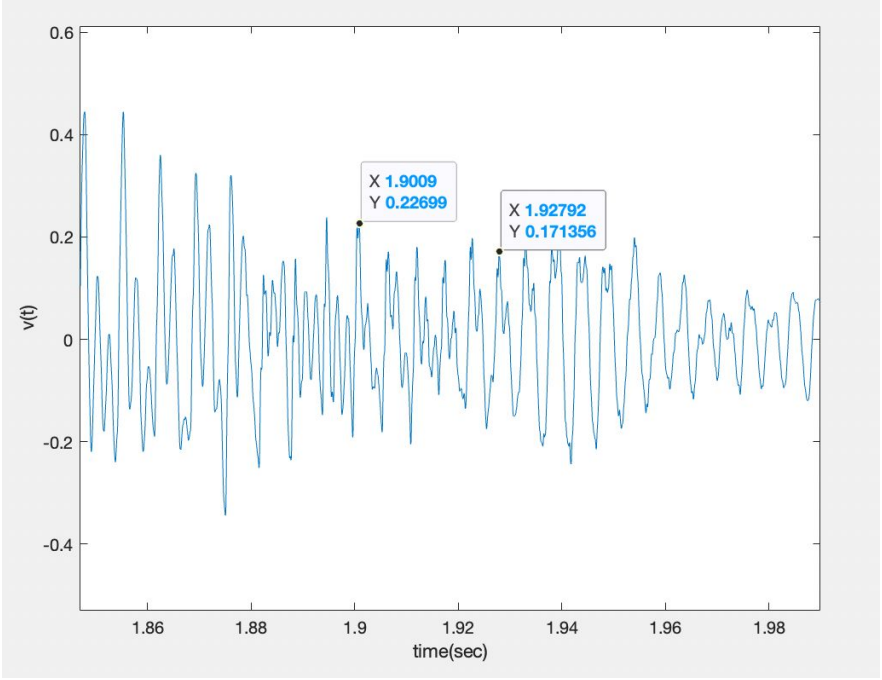
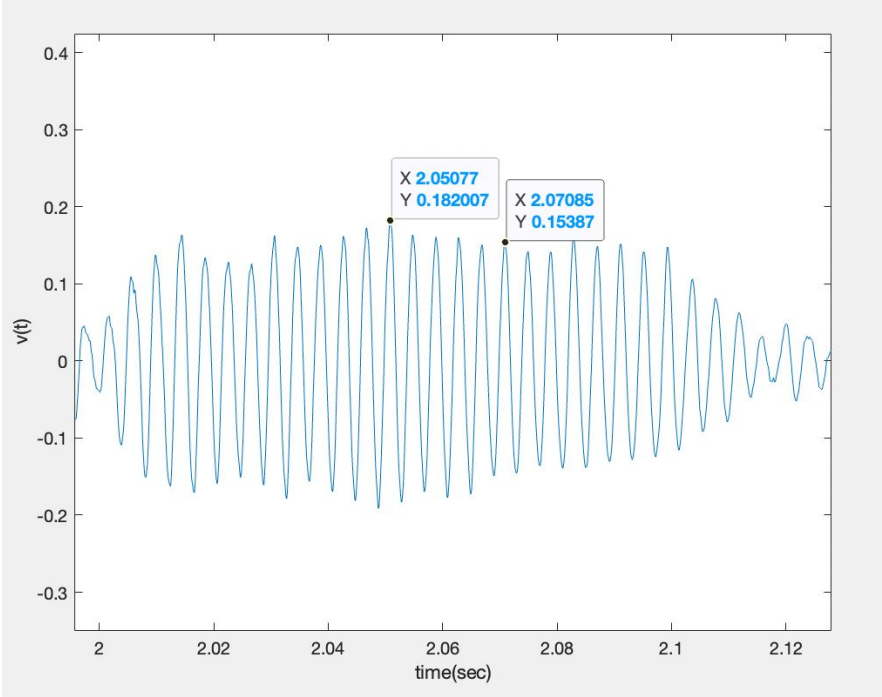


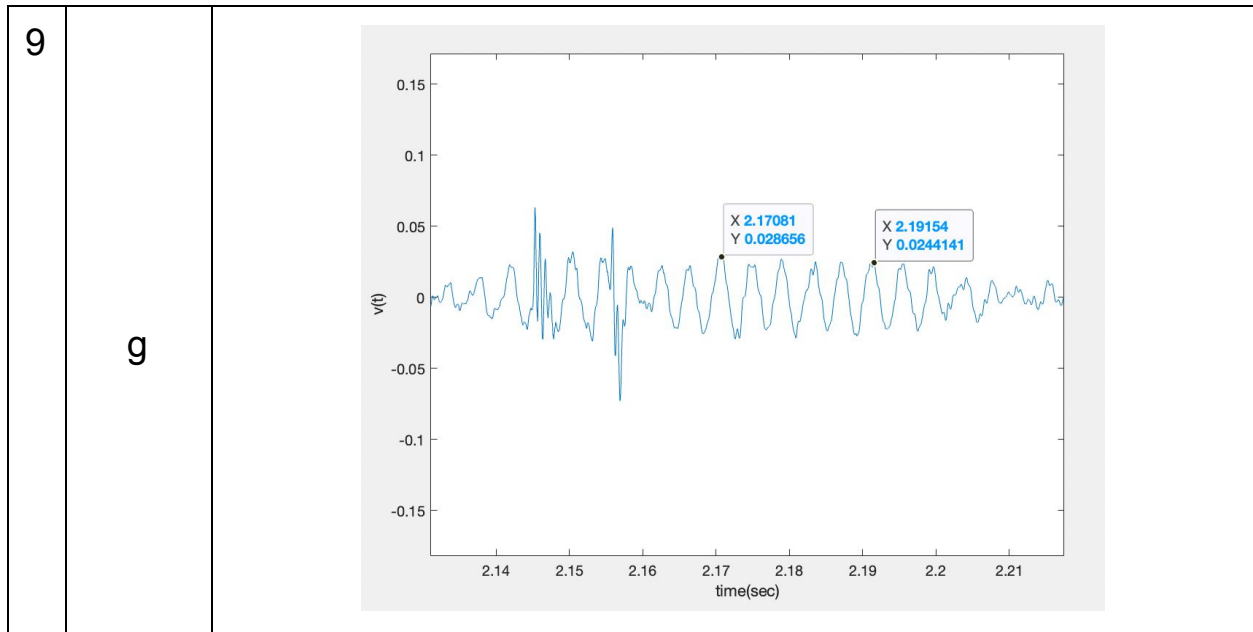
4

rr



5	g	 <p>Plot of $v(t)$ vs time(sec) for g. The y-axis ranges from -0.2 to 0.25, and the x-axis ranges from 1.58 to 1.63. The plot shows a noisy signal with a prominent peak around 1.595 seconds.</p>
6	oo	 <p>Plot of $v(t)$ vs time(sec) for oo. The y-axis ranges from -0.4 to 0.6, and the x-axis ranges from 1.64 to 1.84. The plot shows a highly oscillatory signal. Two data points are highlighted:</p> <ul style="list-style-type: none"> X 1.68735, Y 0.265686 X 1.73038, Y 0.241577

7	ii	 <p>Plot of $v(t)$ vs time(sec) for row 7. The plot shows a decaying oscillatory signal. Two points are highlighted:</p> <ul style="list-style-type: none"> Point 1: X 1.9009, Y 0.22699 Point 2: X 1.92792, Y 0.171356
8	n (nasal)	 <p>Plot of $v(t)$ vs time(sec) for row 8. The plot shows a decaying oscillatory signal. Two points are highlighted:</p> <ul style="list-style-type: none"> Point 1: X 2.05077, Y 0.182007 Point 2: X 2.07085, Y 0.15387



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:

