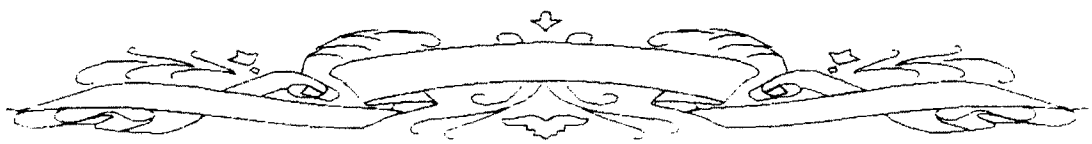


CHAPTER-IV

SPEECH WAVEFORM ANALYSIS

4.1 INTRODUCTION:

R.K. Balani has described some discoveries found while analysing human vioce through amptitude versus time wave forms. The waveforms is split into two parts, the engine part and the wagon part. Engine part represents about 20 % of voice time and it almost characteristic of the consonant. This part corresponds to the duration taken by the movements of the lips and tongue While uttering a consonant and carries the maximum intelligence. In this transition period one cannot visualise, define or measure the pitch period.

Wagon's part represents the remaining 80 % of the time and it is noticed that this part consists of number of identical pitch waveforms (wagon). Which are characteristics of the particular vowel but independent of the consonant preceding it. It has been observed that 80 % of voice pattern of day, kay, may, say is similar to that of vowel A and the voice pattern for C,D, G,T is similar to that of vowel E.

For synthesis it is therefore not necessary to store latter 80 % of every phoneme in full and waste memory space. Accurate pitch estimation in real time is a crucial aspect for synthesis of good quality speech in vocoders. There are a variety of pitch detection algorithms (2-6).

Implementation and comparison of different pitch algorithms on DSP TMS 32010 have been reported by M.R. Kesheorey et al. (7). They have used the following algorithms.

i. Simplified inverse filtering technique.

ii. Autocorrelation with centre clipping and

ii. Autocorrelation method.

We have carried on analysis of various sounds based on single cycle analysis.

4.2 SPEECH WAVEFORM ANALYSIS METHODS:

Speech waveform can be analyzed by various methods. In speech synthesis, it is carried out by various ways, e.g., time domain analysis/synthesis, frequency domain analysis/synthesis, formant analysis/synthesis.

In time domain speech synthesis, the waveform is analyzed by digitizing it periodically. That is the waveform it sampled periodically and encoded. In frequency domain analysis/synthesis, frequency spectrum of waveform is analyzed. As well as in formant analysis/synthesis, the formant frequency characteristics of a spoken waveform are digitally coded and used for synthesis.

There is one method of waveform analysis via curve fitting. It is a software that enables waveform analysis via curve fitting. The technique described herein utilize ordered pairs of data values (X_k, Y_k), X_k generally is a specific time and Y_k is the signal amplitude at that time. Then there is a restriction for the calculation of amplitude for missing time. Here, by utilizing two distinct sequences and by fitting a curve to the data provided in (X_k) and (Y_k) an estimate of the signal amplitude at missing time can be obtained.

Curve fitting routines are also useful when the measured data is noisy.

4.2.1 Single Cycle Analysis Method:

In single cycle analysis technique the cycle by cycle analysis leads to evaluation of period and hence formant frequency. The power and the energy content in a waveform can also be estimated by the known relations. The spikes superimposed on the formant frequency and their variation in time form very important information while synthesizing.

We have attempted analysis of following sounds 'one', 'on', 'e', 'o', 'u', etc. by this technique.

4.3 RECORDING PROCEDURE OF SPEECH SOUNDS WITH THE HELP OF DSO:

Digital storage oscilloscope: Features:

1. L & T Gould storage
2. Frequency - 20 MHz
3. Model 1425
4. Time - base selection
5. Accuracy - upto 500 n
6. Sampled data interpolation through sample hold - linear straight line method.
7. RS 432 interface available on back-panel

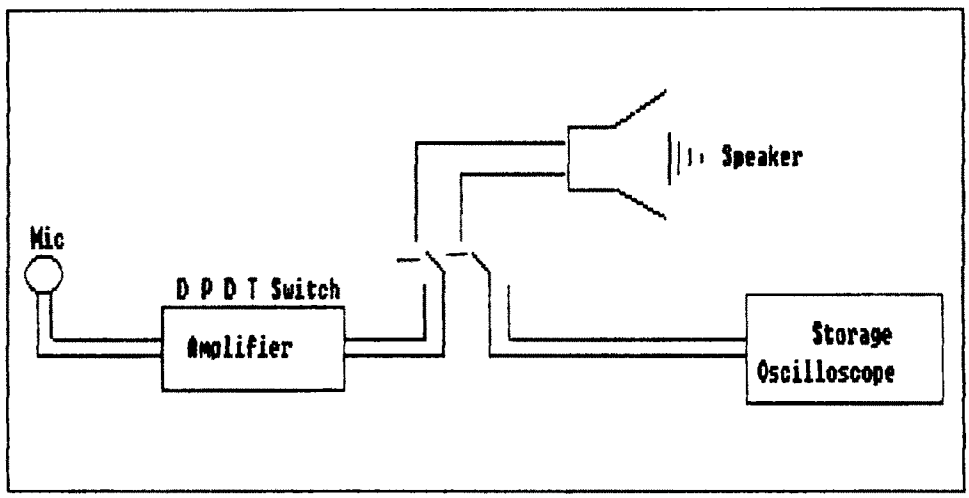


Fig. 4.1 Hardware for storage of speech data.

Steps in getting the data.

1. Put the CRO in store and roll mode.
2. Verify the word on the speaker (only for synthesizer storage) through DPDT switch.
3. See the slow entry of the speech waveform on the screen of CRO.

4. The waveform of the complete word (selected for storage) should fit in the frame of CRO screen.

5. Select proper time base to fulfil point no. (4).

6. Get visual display system (on the screen) "ON' and select the proper channel, i.e. of data.

7. Set the "Datum" line at the start speech wave form with the help of cursor movement.

8. Now with further movement of cursor from the 'Datum' line will visualize the time from "Datum" and voltage w.r.t.'Datum' points voltage.

9. Note down time and voltage and get the next data by again cursor push.

10. Repeating points (8) and (9) cover the complete speech waveform.

Note: By the proper interface with computer it is possible to list the time and data of the waveform. It requires a small software which should match with the storage system monitor software which is already with the oscilloscope.

4.4 PROCEDURE OF ANALYSIS:

We have carried out the analysis by single cycle analysis method. The procedure sequence is as follows:

a. Graphical representation of speech sounds with the help of amplitude Vs time graph, e.g. 'one' 'on' 'e' 'o' 'u'.

b. Spike calculations using fourier series.

c. Graphical representation of spikes.

d. Estimation of amplitude and formant frequencies from the graph.

e. Estimation of power and energy using single cycle analysis.

4.4.1 Spike Calculations Using Fourier Series:

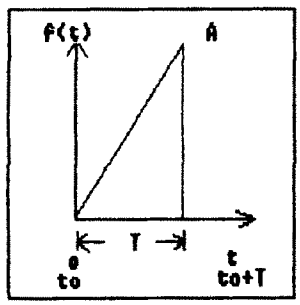
Taking amplitude, occurring time and period, fourier series is calculated.

. .

. . fourier series of a ramp signal is given by,

$$f(t) = a_0 + a_1 \cos 2\pi t + a_2 \cos 4\pi t + \ldots\ldots\ldots + a_n \cos 2\pi nt$$
$$+ b_1 \sin 2\pi t + b_2 \sin 4\pi t + \ldots\ldots\ldots + b_n \sin 2\pi nt$$

$$f(t) = At \quad (0 < t < t)$$



$$\therefore W_0 = 2\pi/T$$

$$\therefore a_0 = \frac{1}{T}\int_0^T f(t)\,dt = AT/2$$

$$\therefore a_n = \frac{2}{T}\int_{to}^{to+T} f(t) \cos nw_0 t\,dt$$

$$\therefore a_n = \frac{AT}{2\pi^2 n^2}\left[\frac{2\pi nt}{T}\sin\frac{2\pi nt}{T} + \cos\frac{2\pi nt}{T}\right]_0^T$$

$$\therefore b_n = \frac{AT}{2\pi^2 n^2}\left[\sin\frac{2\pi nt}{T} - \frac{2\pi nt}{T} \times \cos\frac{2\pi nt}{T}\right]_0^T$$

Program to evaluate fourier series:

PROGRAM:

```
10    REM PROGRAM TO FIND fs
20    For I = 1 to 38
30    PRINT " Fourier Series - Spike"; I
40    INPUT A, P, T IN
50    PRINT "Input Values of amp and period"; A,P, TIN
60    FS1 = (A * P)/2
70    For N = 1 to 30 STEP 1
80    T = P/30 * N + TIN
90    Z = (2 * 3.142 * N * T)/ P
100   X = SIN (z)
110   Y = cos (z)
120   D = (A * P)/(2 * 3.142 * 3.142 * N * N)
```

```
130   E = D * [(Z * X) + Y]
140   F = D * [X - (Z * X)]
150   G = cos (2 * 3.142 * N * T)
160   N = SIN (2 * 3.142 * N * T)
170   FS2 = E * G
180   FS3 = F * H
190   FS4 = FS2 + FS3
200   FS = FS1 + FS4
210   PRINT FS, T
220   NEXT N
230   NEXT I
240   END
```

4.4.2 Cycle analysis for 'one' and representation

For peaks				For spikes			
Time	Period	Amplitude		Time	Period	Amplitude	Side
ms	ms	+ve	-ve	ms	ms	mv	
0-5.0	5.0	140	-	-	-	-	-
5-6.0	1.0	70	-	-	-	-	-
6-7.5	1.5	70	-	-	-	-	-
7.5	4.5	200	200	-	-	-	-
13.0	5.0	330	480	-	-	-	-
18.0	5.0	700	980	-	-	-	-
23.0	6.0	1000	1400	30.0	2.0	380	-ve
30.0	6.0	1680	1600	36.0	1.5	280	-ve
37.5	4.0	1760	1330	-	-	-	-
41.5	6.0	2000	1400	47.0	1.0	280	-ve
48.0	6.5	2800	1330	53.0	1.5	700	-ve
54.5	5.5	2500	1000	56.5	1.5	360	+ve
-	-	-	-	58.0	2.0	1120	-ve

(Continued.....)

For peaks				For spikes			
Time	Period	Amplitude		Time	Period	Ampli-tude	Side
ms	ms	+ve	-ve	ms	ms	mv	
60.5	4.5	2260	1000	65.0	1.0	1340	-ve
66.0	6.5	2400	1000	69.0	1.0	700	+ve
-	-	-	-	71.0	1.5	1320	-ve
72.5	7.5	2380	940	75.0	1.0	920	+ve
-	-	-	-	77.0	2.0	1600	-ve
-	-	-	-	79.0	1.0	580	-ve
80.0	5.0	216	1800	81.0	1.0	1000	+ve
85.0	5.0	216	940	88.0	1.5	1000	+ve

Upto this the phase was starting from +ve direction, i.e. upto 90 ms of speech waveform. After this phase inversion takes place.

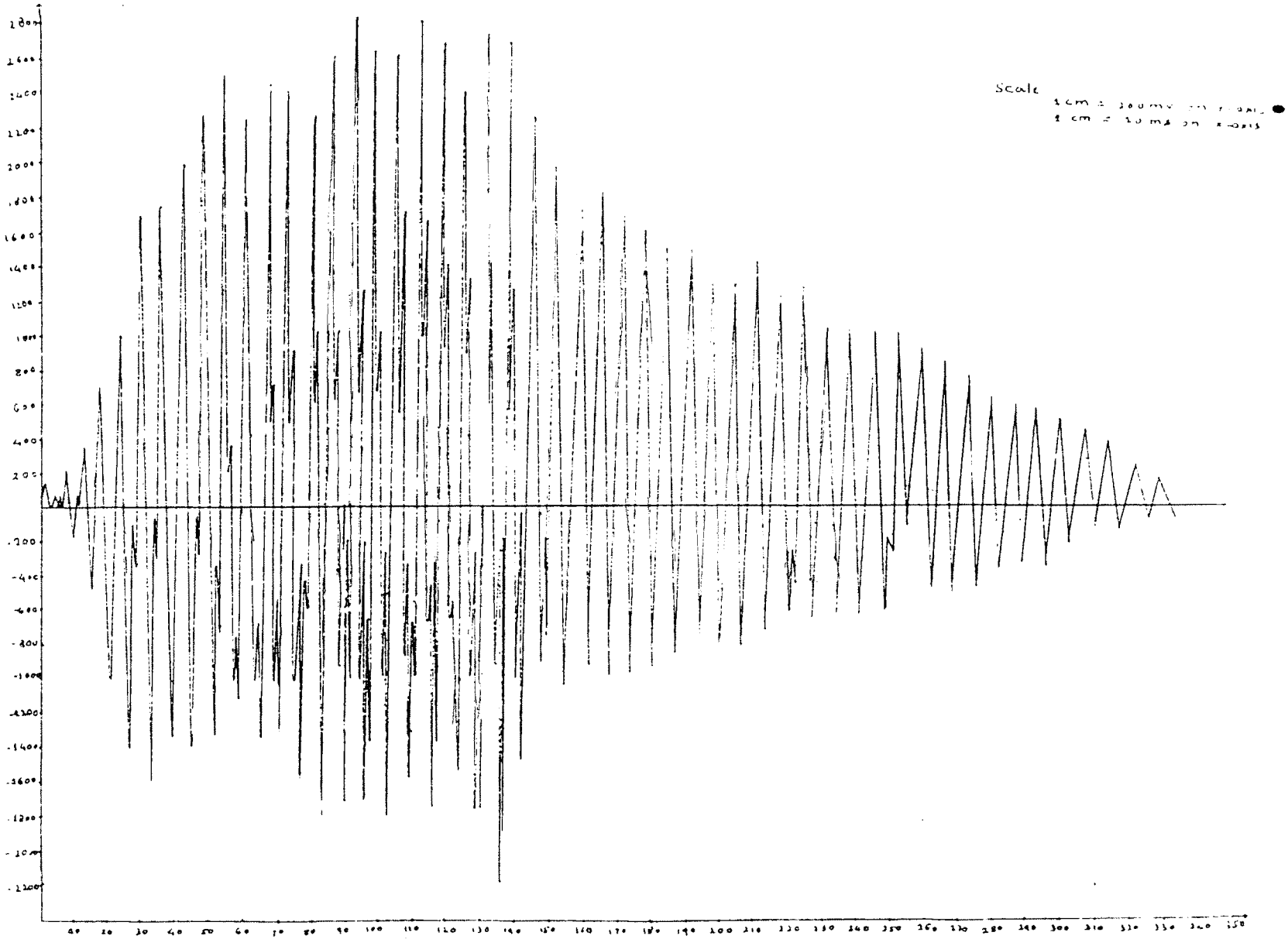
For peaks				For spikes			
Time	Period	Amplitude		Time	Period	Ampli-tude	Side
ms	ms	+ve	-ve	ms	ms	mv	
90.0	5.0	2820	1720	91.0	1.0	1000	-ve
-	-	-	-	94.0	1.0	1400	+ve
95.5	5.5	216	1700	95.5	1.0	1000	-ve
-	-	-	-	97.5	1.0	1300	-ve
-	-	-	-	100.0	1.0	1000	+ve

(Contined.....)

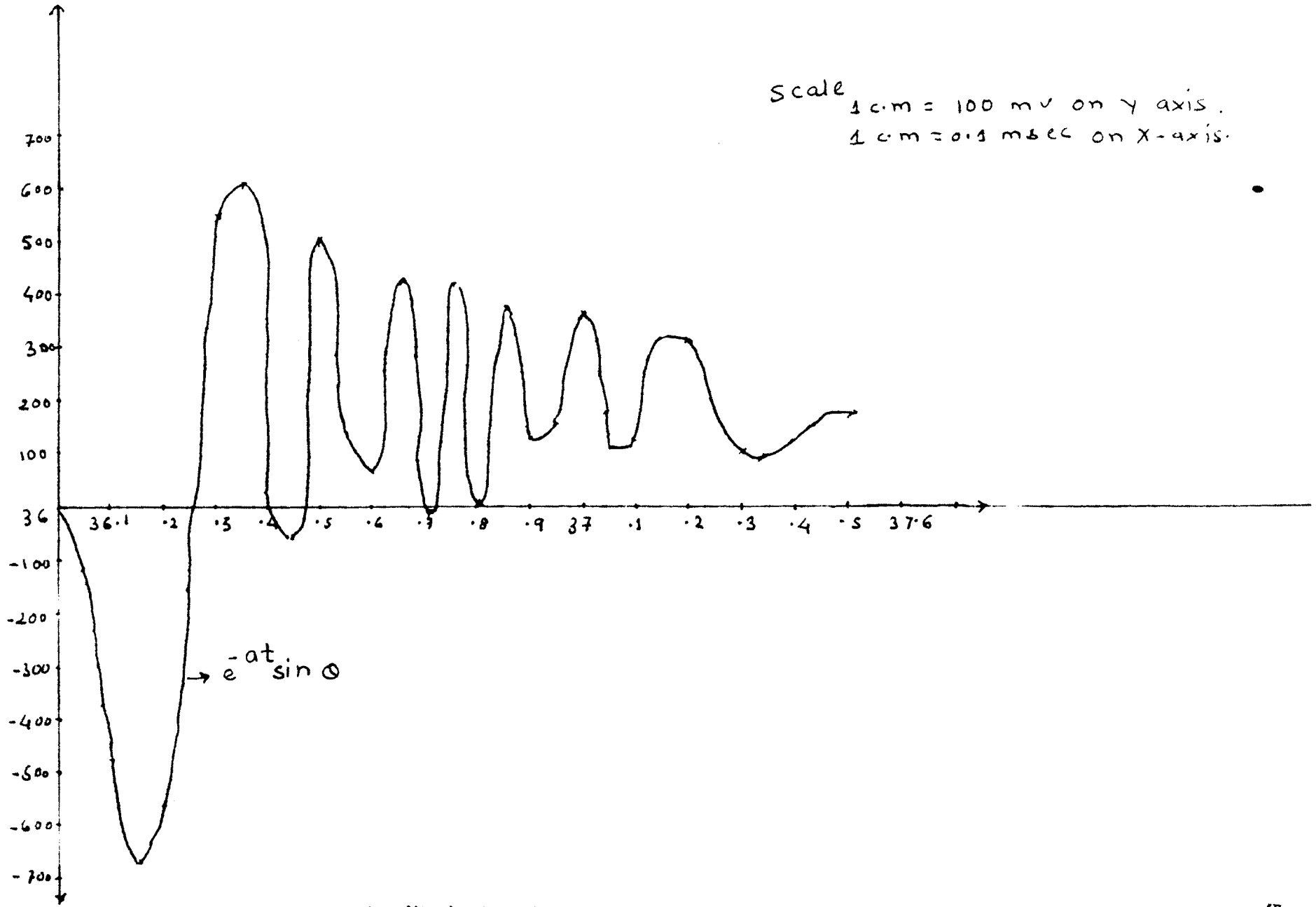
For peaks				For spikes			
Time ms	Period ms	Amplitude +ve	Amplitude -ve	Time ms	Period ms	Amplitude mv	Side
101.5	7	216	1800	101.5	1.0	1000	-ve
-	-	-	-	106.0	2.0	1700	+ve
108.0	5.5	2800	1600	108.0	1.0	880	-ve
-	-	-	-	110.0	1.0	1000	-ve
-	-	-	-	112.0	1.5	1660	+ve
113.5	3.5	2480	1740	113.5	1.0	760	-ve
-	-	-	-	116.5	1.0	1380	-ve
-	-	-	-	119.0	1.0	1400	+ve
120.0	6.5	2400	1320	120.5	1.0	660	-ve
-	-	-	-	123.0	1.0	1540	-ve
-	-	-	-	126.0	1.0	1320	+ve
-	-	-	-	127.0	1.5	1000	-ve
-	-	-	-	127.0	1.5	1760	-ve
130.0	3.0	2740	1800	132.0	1.0	1400	+ve
133.0	6.5	2700	2200	133.5	1.0	920	-ve
-	-	-	-	136.0	1.0	1900	-ve
-	-	-	-	137.0	1.0	1260	+ve
139.5	6.5	2400	1000	141.5	1.5	1440	-ve
146.5	6.5	1960	920	148.0	1.0	780	-ve
153.0	7.0	1760	1200	-	-	-	-
160.0	6.0	1700	940	-	-	-	-

(Continued.....)

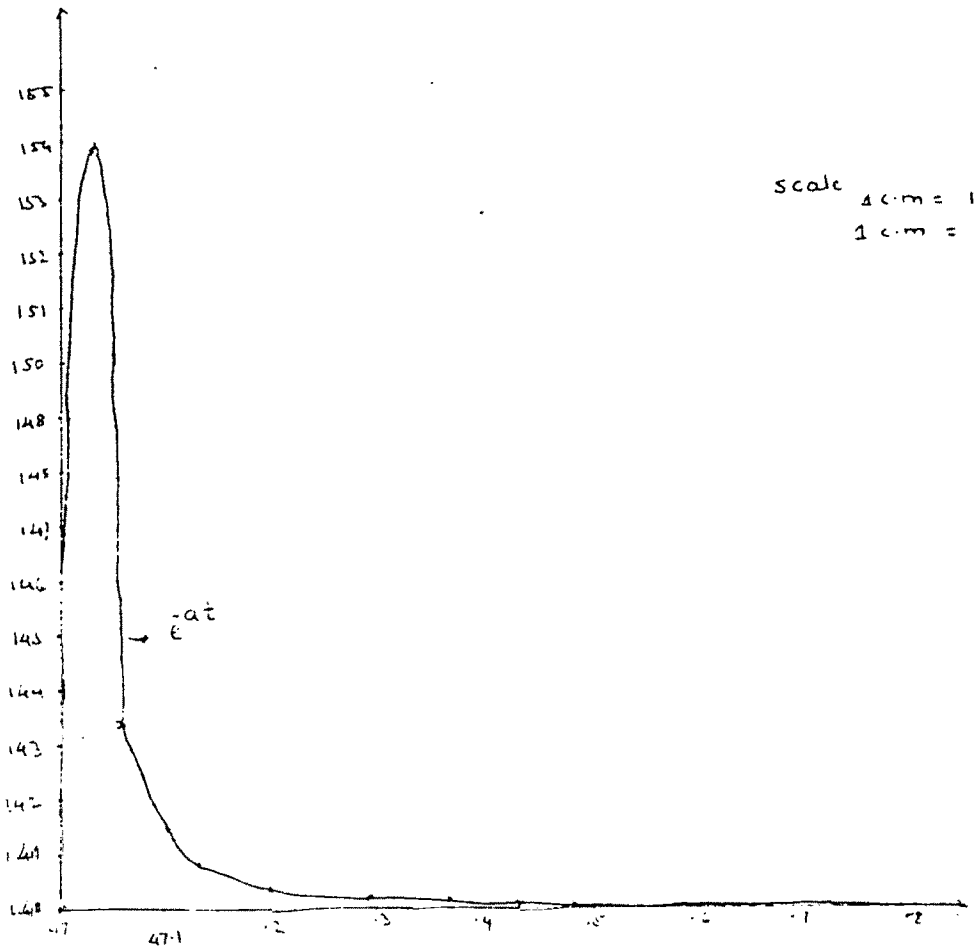
		For peaks			For spikes		
Time	Period	Amplitude		Time	Period	Ampli-tude	Side
ms	ms	+ve	-ve	ms	ms	mv	
166.0	6.5	1760	1000	-	-	-	-
172.5	6.5	1700	1000	-	-	-	-
179.0	7.0	1480	940	-	-	-	-
186.0	7.0	1500	840	-	-	-	-
193.0	6.0	1480	840	-	-	-	-
199.0	7.0	1280	820	-	-	-	-
206.0	7.0	1400	800	-	-	-	-
213.0	7.0	1200	700	-	-	-	-
220.0	6.0	1250	640	222.0	1.0	460	-ve
226.0	7.0	1000	640	-	-	-	-
233.0	7.0	1000	640	-	-	-	-
240.0	7.0	1000	640	-	-	-	-
247.0	7.5	1000	600	248.0	3.0	300	-ve
254.5	6.5	920	100	-	-	-	-
261.0	7.0	840	500	-	-	-	-
267.0	7.0	780	500	-	-	-	-
274.0	7.0	600	500	-	-	-	-
281.0	7.0	580	400	-	-	-	-
288.0	7.0	560	360	-	-	-	-
295.0	7.0	480	360	-	-	-	-
302.0	7.0	400	240	-	-	-	-
324.0	8.0	140	70.0	-	-	-	



Amplitude Vs time graph for speech sound "one"



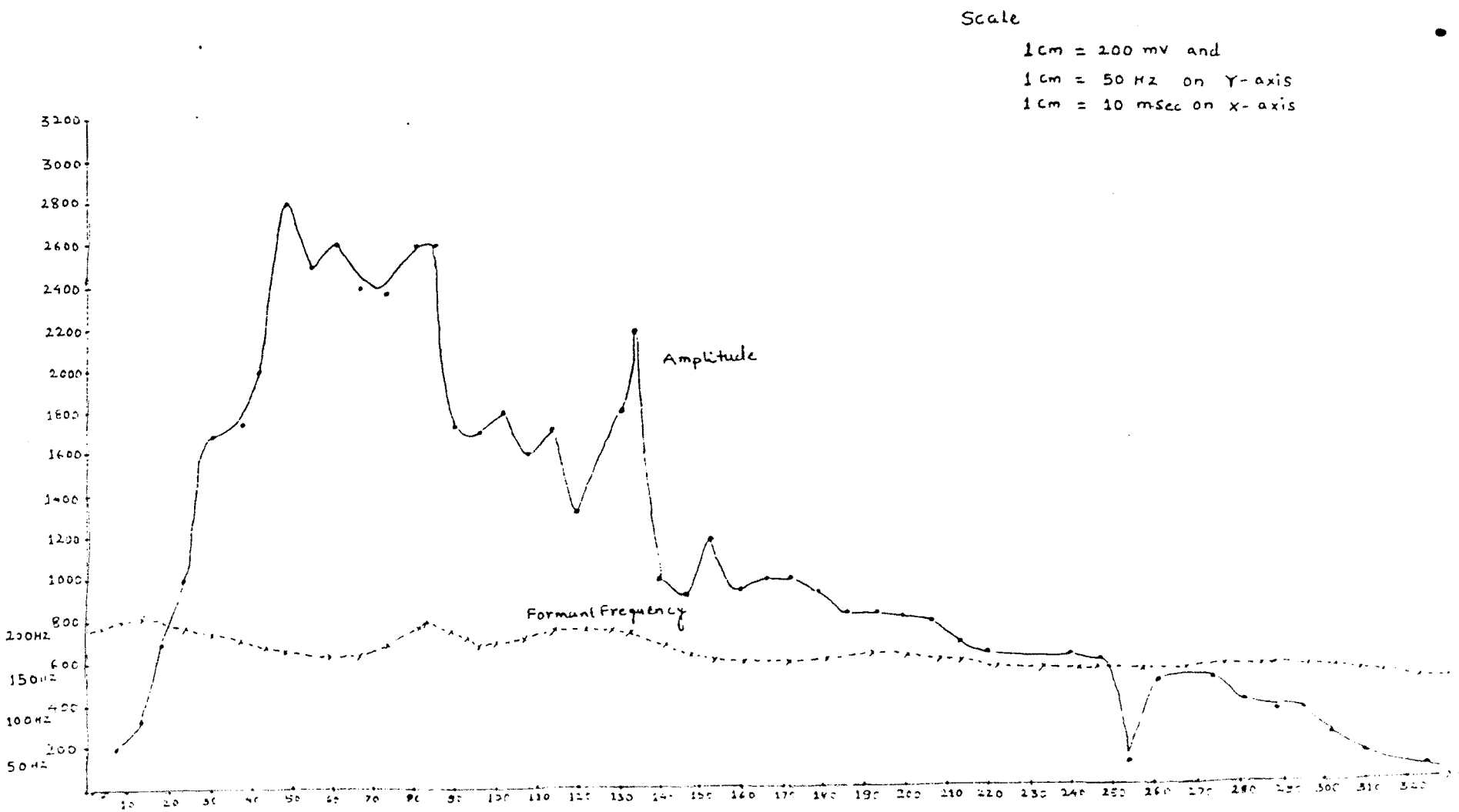
Amplitude Vs time graph for spike having period 1.5 m sec and occuring time as 36 m sec for "one"



Amplitude Vs time graph for spike having period 1.0 m sec and occuring time as 47 m sec for "one"

4.4.3 Formant frequency, energy and power for 'one'

Cycle No.	period	formant frequency	energy	power
1	5.0	200	0.0196	0.0039
2	4.5	222	0.0400	0.0088
3	5.0	200	0.1584	0.0317
4	5.0	200	0.6860	0.1372
5	5.0	200	1.4000	0.2800
6	6.0	166	2.6880	0.4480
7	4.0	250	2.3408	0.5852
8	6.0	166	2.8000	0.4666
9	6.5	154	3.7240	0.5729
10	5.5	181	2.5000	0.4545
11	4.5	222	2.2600	0.5022
12	6.5	154	2.4000	0.3692
13	7.5	133	2.2372	0.2983
14	5.0	200	4.6800	0.9360
15	5.0	200	2.4440	0.4888
16	5.0	200	4.8504	0.9700
17	5.5	181	4.4200	0.8036
18	7.0	142	4.6800	0.6686
19	5.5	181	4.4800	0.8145
20	5.5	181	4.3152	0.7846
21	6.5	154	3.1680	0.4873
22	6.5	154	5.9400	0.9138
23	6.5	154	5.7600	0.8861
24	6.5	154	1.8032	0.2774
25	7.0	142	2.1120	0.3017
26	6.0	166	1.7860	0.2976
27	6.5	154	3.0976	0.4765
28	6.5	154	2.8900	0.4446
29	7.0	142	1.6920	0.2417
30	7.0	142	1.2600	0.1800
31	6.0	166	1.2432	0.2072
32	7.0	142	1.0496	0.1499
33	7.0	142	1.1200	0.1600
34	7.0	142	0.8400	0.1200
35	6.0	166	0.8000	0.1333
36	7.0	142	0.6400	0.0914
37	7.0	142	0.6400	0.9142
38	7.0	142	0.6400	0.0914
39	7.5	133	0.6000	0.0800
40	6.5	154	0.9200	0.1415
41	7.0	142	0.4200	0.0600
42	7.0	142	0.3900	0.0557
43	7.0	142	0.3000	0.0428
44	7.0	142	0.2320	0.0331
45	7.0	142	0.2016	0.0280
46	7.0	142	0.1728	0.0246
47	7.0	142	0.0960	0.0137
48	7.0	142	0.0504	0.0072
49	8.0	125	0.0280	0.0035



Amplitude and Formant frequency for "one"

4.4.4 Cycle analysis for 'on'and representation

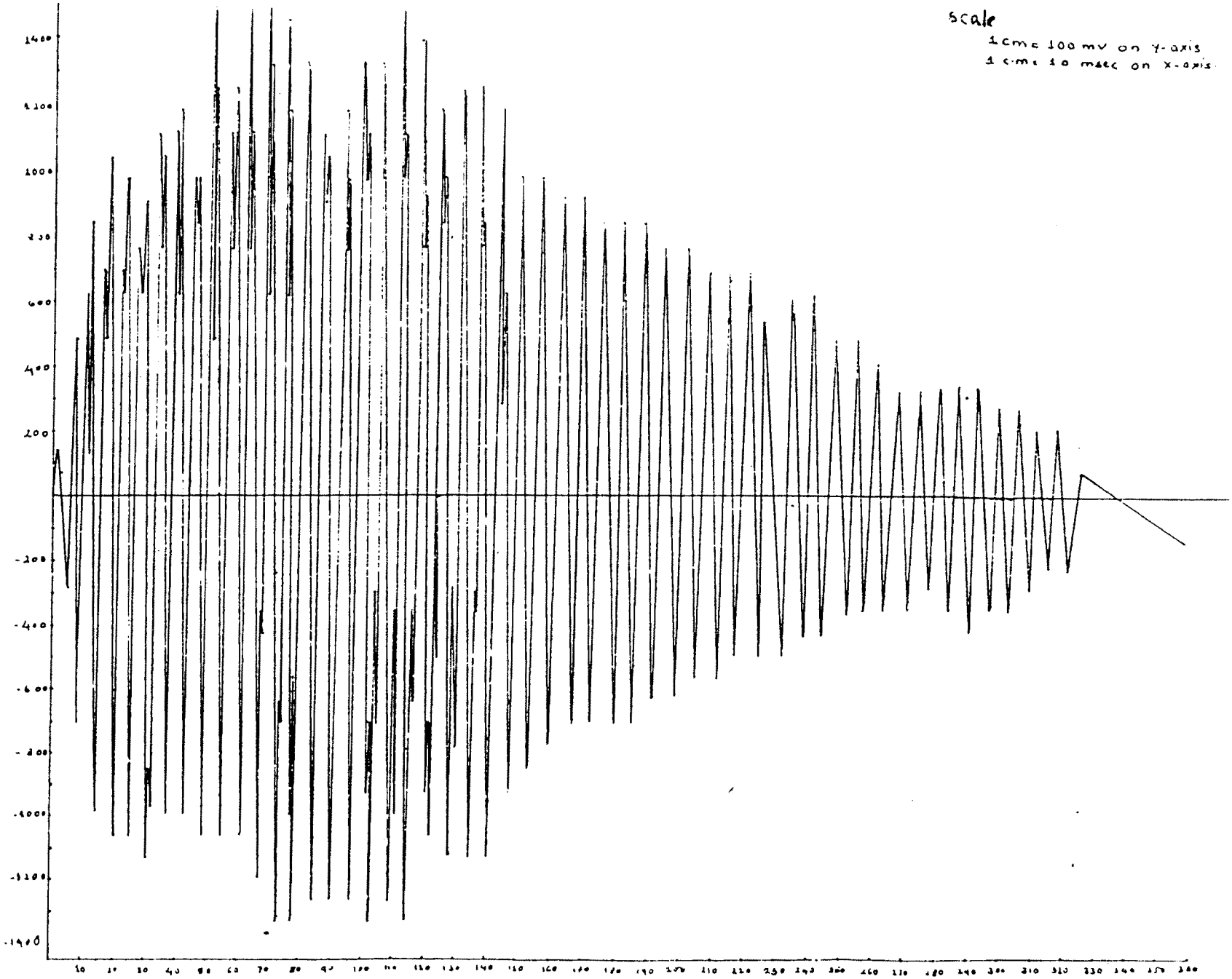
For peaks				For spikes			
Time	Period	Amplitude		Time	Period	Ampli-tude	Side
ms	ms	+ve	-ve	ms	ms	mv	
0.5	5.0	140	280	-	-	-	-
5.0	4.0	490	630	10.5	1.0	840	+ve
9.0	6.0	840	980	-	-	-	-
15	5.0	700	1050	17.0	1.5	1050	+ve
20	5.0	700	1050	22.5	1.5	980	+ve
25	6.0	770	1120	27.5	1.0	910	+ve
-	-	-	-	31.0	1.0	980	-ve
31	5.0	1120	980	33.5	1.0	1050	+ve
37	5.0	1190	980	39.0	1.0	1190	+ve
42	5.0	980	1050	44.0	1.0	980	+ve
47	7.0	1540	1050	50.5	1.0	1260	+ve
-	-	-	-	54.5	1.0	770	-ve
54	6.0	1120	1050	56.5	1.0	1260	+ve
60	6.0	1540	1190	62.0	1.0	1120	+ve
-	-	-	-	66.5	1.0	420	-ve
66	6.0	1750	1330	68.0	1.0	1330	+ve
-	-	-	-	72.0	1.0	700	-ve
72	5.0	1470	1330	74.0	1.0	1190	+ve
-	-	-	-	77.0	1.0	1330	-ve
77	7.0	1330	1260	84.0	1.0	980	-ve
83	6.0	1120	1260	86.0	1.0	1050	+ve
90	6.0	1190	1260	92.5	1.0	980	+ve

(Continued.....)

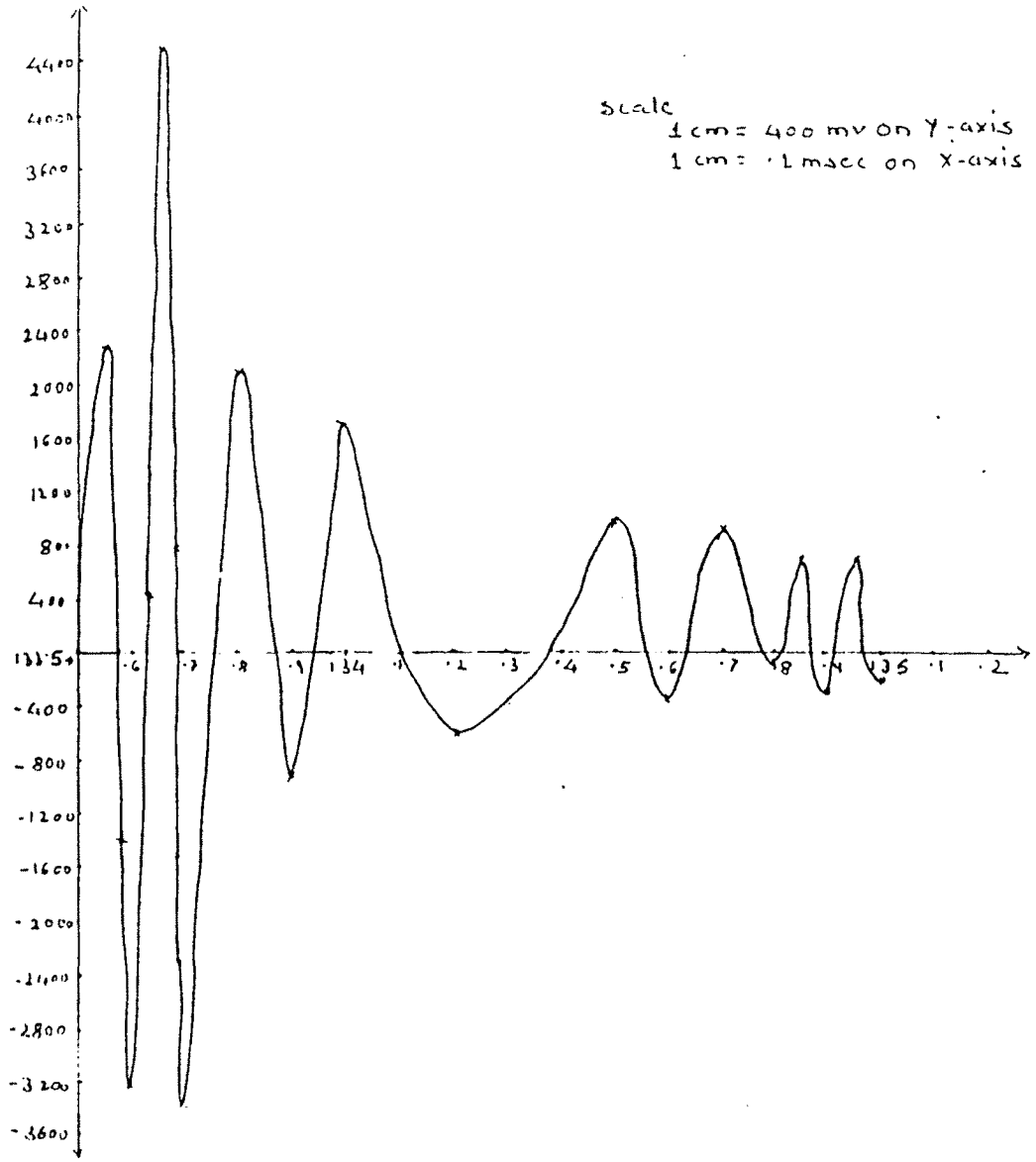
For peaks				For spikes			
Time	Period	Amplitude		Time	Period	Ampli-tude	Side
ms	ms	+ve	-ve	ms	ms	mv	
-	-	-	-	98.5	1.0	1120	+ve
96	6.0	1190	1260	101.5	1.5	1330	-ve
-	-	-	-	102.5	1.0	490	-ve
102	6.0	1330	1260	108.5	1.0	980	-ve
108	6.0	1540	1330	110.0	1.0	1120	+ve
-	-	-	-	114.5	1.0	630	-ve
114	6.0	1400	910	117.0	1.5	910	+ve
-	-	-	-	119.5	1.0	1050	-ve
-	-	-	-	120.5	1.0	490	-ve
120	7.0	1190	1120	123.0	1.0	980	+ve
-	-	-	-	127.0	1.0	770	-ve
127	6.0	1260	1120	133.5	1.5	350	-ve
133	7.0	1260	1120	135.5	1.0	840	+ve
140	6.0	1190	910	142.5	1.5	630	+ve
146	6.0	980	840	-	-	-	-
152	6.0	980	770	-	-	-	-
158	7.0	910	700	-	-	-	-
165	6.0	910	700	-	-	-	-
171	7.0	840	700	-	-	-	-
178	6.0	840	700	-	-	-	-
184	6.0	840	630	-	-	-	-

(Continued.....)

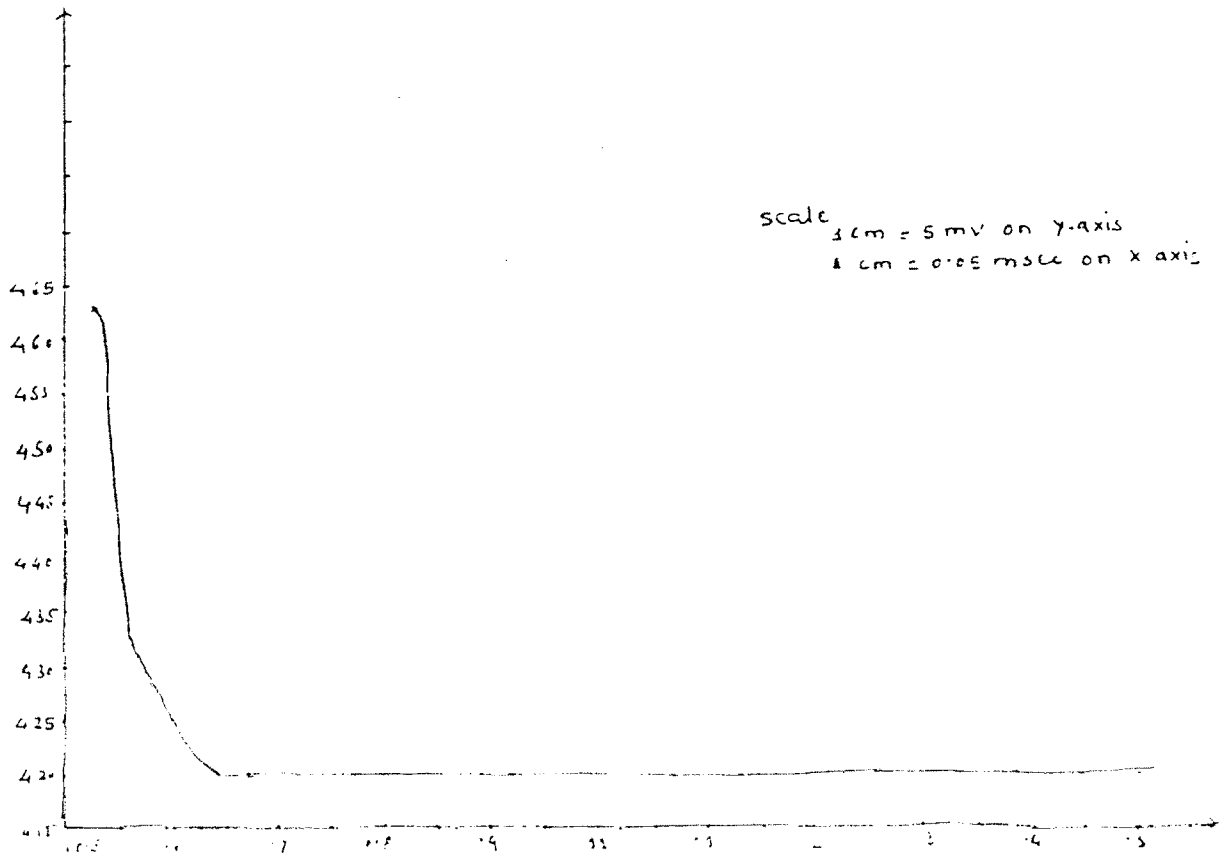
For peaks				For spikes			
Time	Period	Amplitude		Time	Period	Ampli-tude	Side
ms	ms	+ve	-ve	ms	ms	mv	
190	7.0	770	630	-	-	-	-
197	7.0	770	560	-	-	-	-
204	6.0	700	560	-	-	-	-
210	6.0	700	490	-	-	-	-
216	6.0	700	490	-	-	-	-
222	8.0	560	490	-	-	-	-
230	7.0	630	420	-	-	-	-
237	6.0	630	420	-	-	-	-
243	7.0	490	350	-	-	-	-
250	6.0	440	350	-	-	-	-
256	7.0	420	350	-	-	-	-
263	7.0	350	350	-	-	-	-
270	6.0	350	280	-	-	-	-
276	7.0	350	350	-	-	-	-
283	7.0	350	420	-	-	-	-
290	7.0	350	350	-	-	-	-
297	7.0	280	350	-	-	-	-
304	6.0	280	280	-	-	-	-
310	6.0	210	210	-	-	-	-
316	18.0	210	210	-	-	-	-
334	-	70	140	-	-	-	-



Amplitude Vs time graph for speech sound "on"



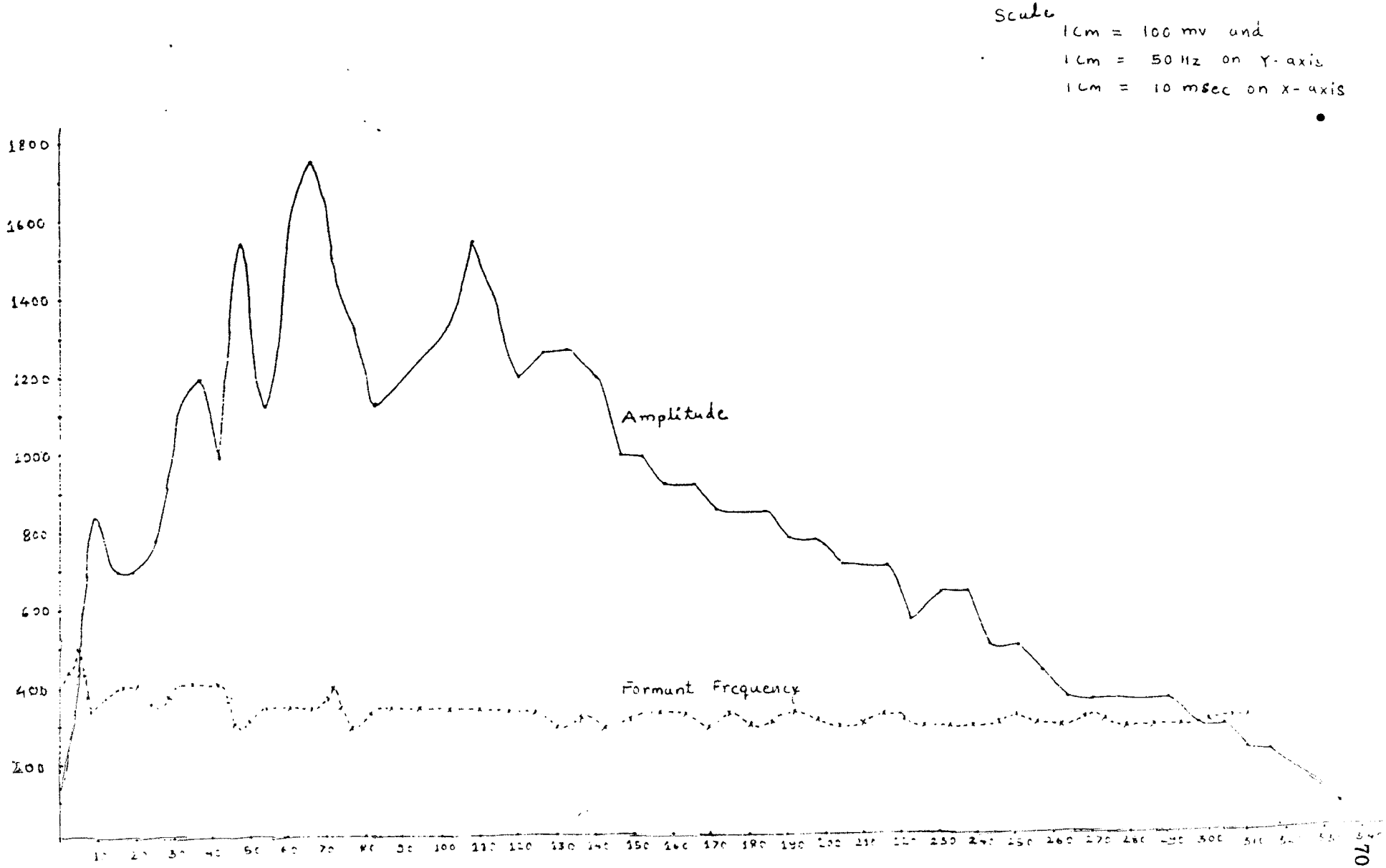
Amplitude Vs time graph for spike having period 1.5 m sec and occuring time as 133.5 m sec for "on"



Amplitude Vs time graph for spike having period 1.0 m sec and occuring time as 10.5 m sec for "on"

4.4.5 Formant frequency, energy and power for 'on'

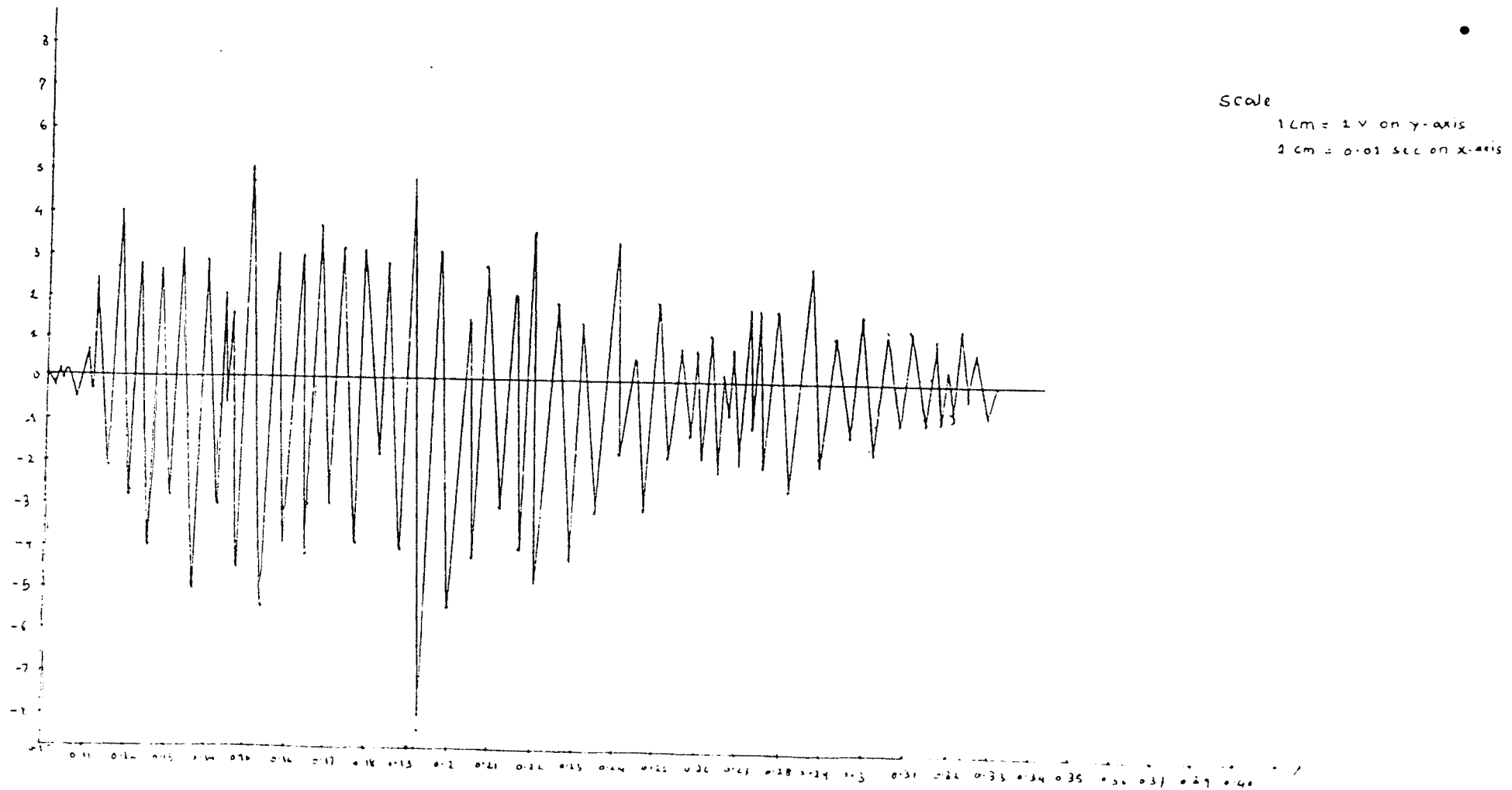
Cycle No.	period	formant frequency	energy	power
1	5	200	0.0392	0.0078
2	4	250	0.3087	0.0772
3	6	166	0.8232	0.1372
4	5	200	0.7350	0.1470
5	5	200	0.7350	0.1470
6	6	166	0.8624	0.1437
7	5	200	1.0976	0.2195
8	5	200	1.1662	0.2332
9	5	200	1.0290	0.2058
10	7	142	1.6170	0.2310
11	6	166	1.1760	0.1960
12	6	166	1.8326	0.3050
13	6	166	2.3275	0.3879
14	5	200	1.9551	0.3910
15	7	142	1.6758	0.2394
16	6	166	1.4112	0.2352
17	6	166	1.4994	0.2499
18	6	166	1.6758	0.2793
19	6	166	2.0480	0.3410
20	6	166	1.2740	0.2123
21	7	142	1.3328	0.1904
22	6	166	1.4112	0.2352
23	7	142	1.4120	0.2016
24	6	166	1.0829	0.1804
25	6	166	0.8232	0.1372
26	6	166	0.7546	0.1257
27	7	142	0.6370	0.0910
28	6	166	0.6370	0.1060
29	7	142	0.5880	0.0840
30	6	166	0.5880	0.0980
31	6	166	0.5290	0.0882
32	7	142	0.4851	0.0693
33	7	142	0.4312	0.0616
34	6	166	0.3920	0.0653
35	6	166	0.3430	0.0571
36	6	166	0.3430	0.0571
37	8	125	0.2744	0.0343
38	7	142	0.2646	0.0376
39	6	166	0.2646	0.0441
40	7	142	0.1715	0.0245
41	6	166	0.1715	0.0285
42	7	142	0.1470	0.0210
43	7	142	0.1225	0.0175
44	6	166	0.0980	0.0163
45	7	142	0.1225	0.0175
46	7	142	0.1470	0.0210
47	7	142	0.1225	0.0175



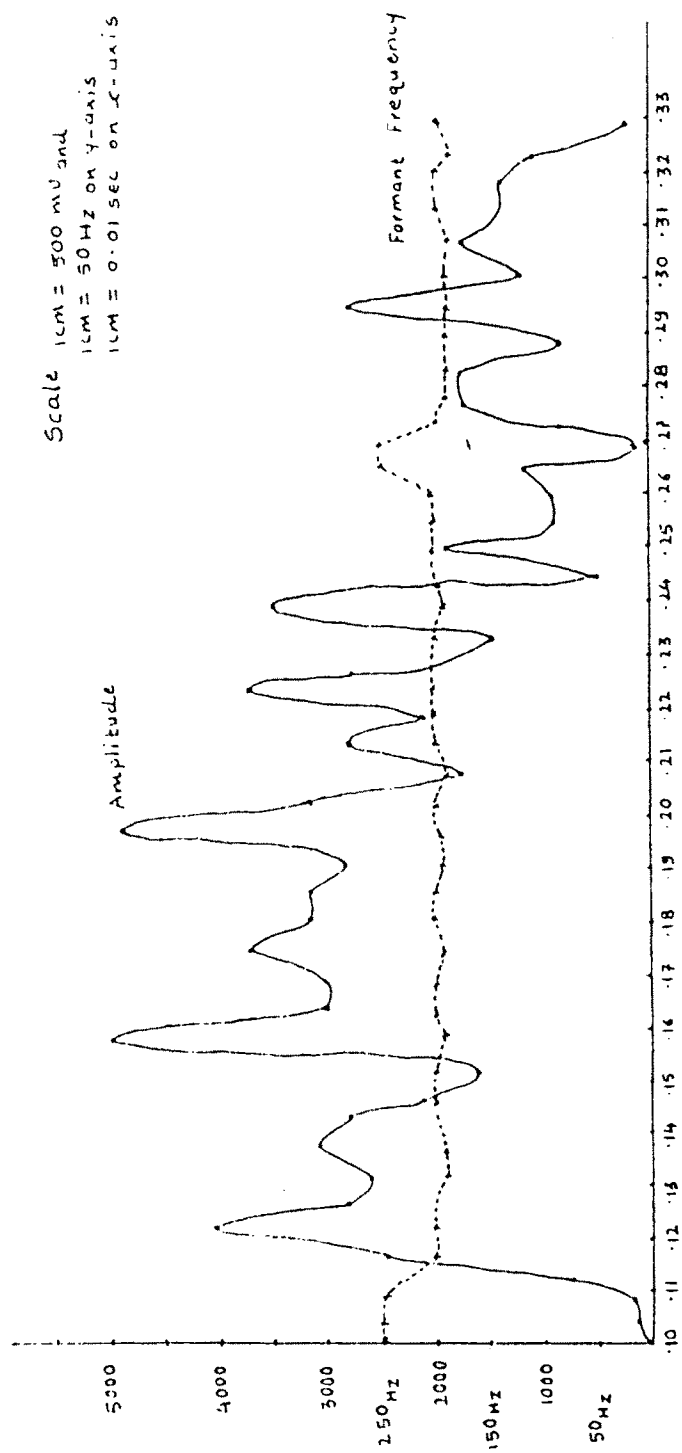
Amplitude and Formant frequency for "on"

4.4.6 Cycle analysis for 'u' and representation

cycle No.	Time	Period	Amplitude +ve	Amplitude -ve	Formant frequency	Energy	power
1	0.1	4	0.0	0.175	250	0.0306	0.0076
2	0.104	4	0.175	0.0	250	0.0306	0.0076
3	0.108	4	0.175	0.525	250	0.0919	0.0229
4	0.112	4	0.700	0.350	250	0.2450	0.0613
5	0.116	5	2.45	2.1	200	5.1450	1.0290
6	0.121	5	4.02	2.8	200	11.2560	2.2512
7	0.126	5	2.8	4.02	200	11.2560	2.2512
8	0.131	6	2.62	2.8	166	7.3360	1.2226
9	0.137	6	3.15	5.07	166	15.9700	2.6617
10	0.143	3	2.8	2.97	333	8.3160	2.7720
11	0.146	5	2.1	0.700	200	1.4700	0.2940
12	0.151	6	1.57	4.5	166	7.0650	1.1775
13	0.157	6	5.7	5.42	166	27.4794	4.5799
14	0.163	5	2.97	4.02	200	11.9394	2.3878
15	0.168	5	2.97	4.20	200	12.4740	2.4948
16	0.174	6	3.07	2.97	166	10.8999	1.8166
17	0.180	5	3.15	3.85	200	12.1275	2.4255
18	0.185	5	3.15	1.57	200	4.9455	0.9891
19	0.190	6	2.8	4.02	166	11.2560	1.8760
20	0.196	6	4.9	8.05	166	39.4450	6.5740
21	0.202	5	3.15	5.42	200	17.0730	3.4146
22	0.207	6	1.75	4.02	166	7.0350	1.1725
23	0.213	5	2.8	2.27	200	6.3560	1.2712
24	0.218	5	2.10	4.02	200	8.4420	1.6884
25	0.223	5	3.67	4.90	200	17.9830	3.5966
26	0.228	5	1.92	4.2	200	8.0640	1.6128
27	0.233	5	1.4	3.1	200	4.3400	0.8680
28	0.238	6	3.5	1.57	166	5.4950	0.9158
29	0.244	5	0.525	2.97	200	1.5590	0.3118
30	0.249	6	1.92	1.75	166	3.3600	0.5600
31	0.254	5	0.575	1.22	200	1.0675	0.2135
32	0.259	5	0.575	1.75	200	1.5313	0.3063
33	0.264	4	1.22	2.10	250	2.5620	0.6405
34	0.268	4	0.175	0.700	250	0.1225	0.0306
35	0.272	5	0.875	1.92	200	1.6800	0.3360
36	0.276	5	1.75	1.05	200	1.8375	0.3675
37	0.282	6	1.75	1.92	166	3.3600	0.5600
38	0.288	6	0.875	2.62	166	2.2925	0.3820
39	0.294	6	2.8	1.92	166	5.3760	0.8960
40	0.300	6	1.22	1.22	166	1.4884	0.2480
41	0.306	6	1.75	1.57	166	2.7475	0.4579
42	0.312	5	1.4	0.875	200	1.2250	0.2450
43	0.317	5	1.4	0.875	200	1.2250	0.2450
44	0.322	5	1.22	0.700	200	0.8450	0.1708
45	0.327	6	0.350	0.700	166	0.2450	0.0408
46	0.333	5	1.4	0.350	200	0.4900	0.0980



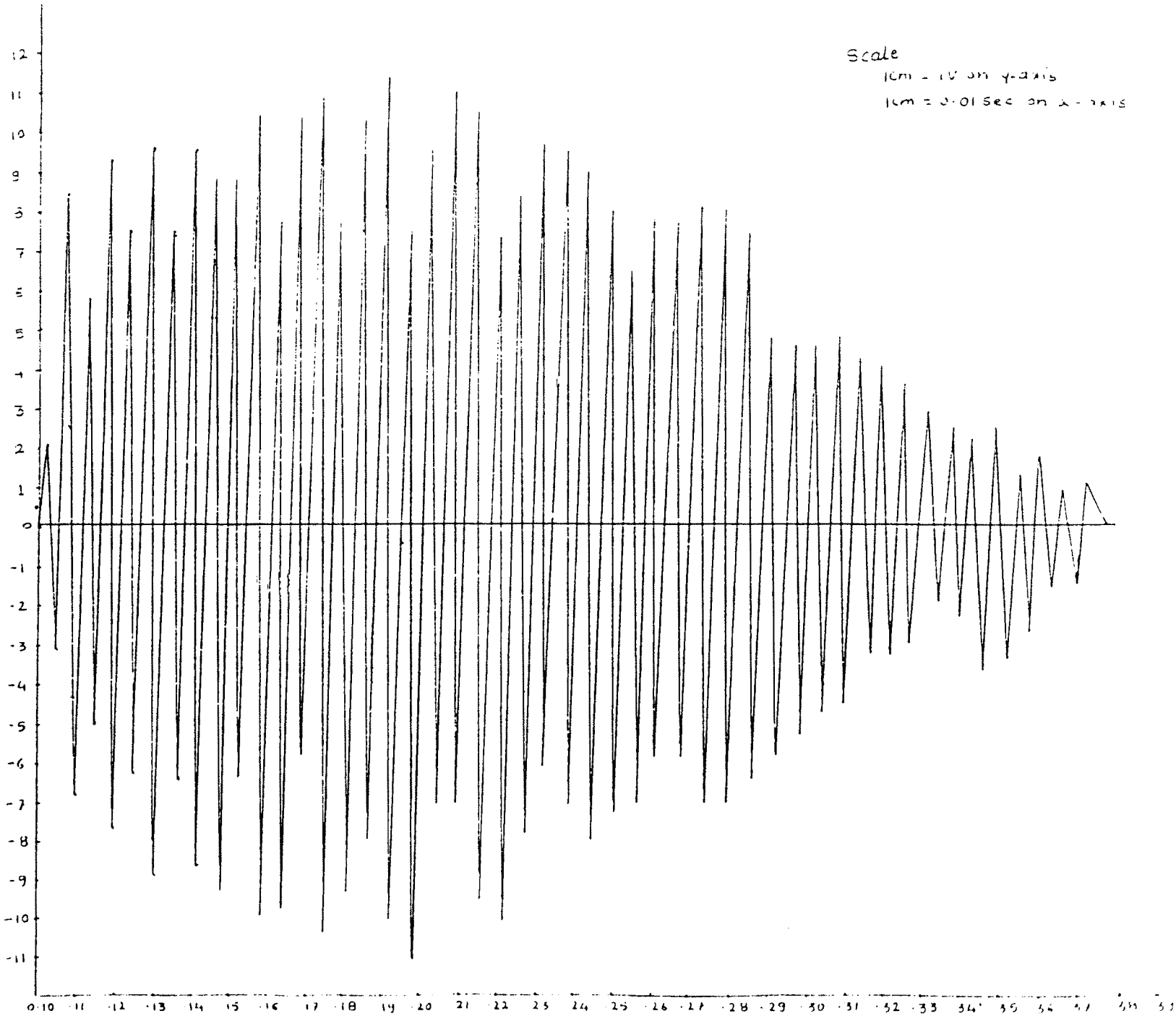
Amplitude Vs time graph for speech sound "u"



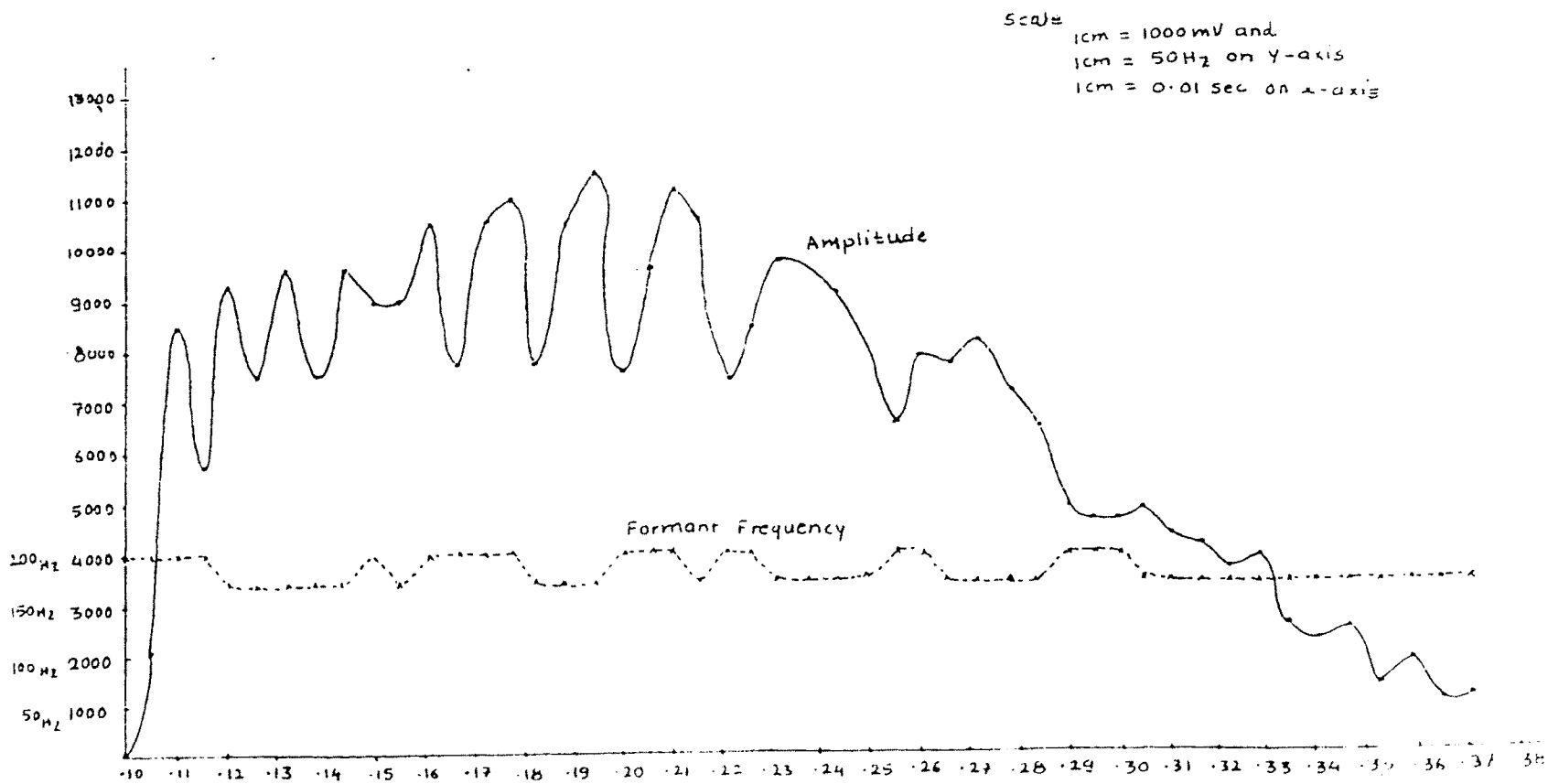
Amplitude and Formant frequency for "u"

4.4.7 Cycle analysis for 'o'and representation

cycle No.	Time	Period	Amplitude +ve	Amplitude -ve	Formant frequency	Energy	power
1	0.1	5	0.0	0.70	200	0.490	0.0980
2	0.105	5	2.1	3.15	200	6.615	1.3230
3	0.110	5	8.4	6.82	200	57.288	11.4576
4	0.115	5	5.77	5.07	200	29.254	5.8508
5	0.120	6	9.27	7.70	166	71.379	11.8965
6	0.126	6	7.52	6.30	166	47.376	7.8960
7	0.132	6	9.62	8.92	166	85.810	14.3000
8	0.138	6	7.52	6.47	166	48.654	8.1090
9	0.144	6	9.62	8.57	166	82.443	13.7400
10	0.150	5	8.92	9.27	200	82.688	16.5376
11	0.155	6	8.92	6.3	166	56.196	9.3660
12	0.161	5	10.5	9.8	200	102.9	20.5800
13	0.167	5	7.7	9.72	200	74.844	14.9688
14	0.172	5	10.5	5.77	200	60.585	12.1170
15	0.177	5	11.0	10.3	200	113.3	22.6600
16	0.182	6	7.70	9.27	166	71.379	11.8965
17	0.187	6	10.5	7.87	166	82.635	13.7725
18	0.194	6	11.5	9.97	166	114.655	19.1092
19	0.200	5	7.52	11.0	200	82.720	16.5440
20	0.205	5	9.57	7.0	200	66.990	13.3980
21	0.210	5	11.2	7.0	200	78.400	15.6800
22	0.215	6	10.6	9.45	166	100.170	16.6950
23	0.221	5	7.35	10.0	200	73.500	14.7000
24	0.226	5	8.40	7.70	200	64.680	12.9360
25	0.231	6	9.8	6.12	166	59.976	9.9960
26	0.237	6	9.62	7.0	166	67.34	11.2233
27	0.243	6	9.1	8.05	166	73.255	12.2090
28	0.249	6	8.22	7.17	166	58.937	9.8229
29	0.255	5	6.47	7.0	200	44.290	9.0580
30	0.260	6	7.87	5.77	166	45.409	7.5683
31	0.265	6	7.70	5.77	166	44.429	7.4048
32	0.272	6	8.22	7.0	166	57.540	9.5900
33	0.278	6	7.17	7.0	166	50.19	8.3650
34	0.284	6	6.47	6.47	166	41.860	6.9768
35	0.290	5	4.90	5.77	200	28.273	5.6546
36	0.295	5	4.72	5.25	200	24.78	4.9560
37	0.301	5	4.72	4.72	200	22.278	4.4556
38	0.305	6	4.9	3.50	166	17.15	2.8580
39	0.311	6	4.37	3.15	166	13.765	2.2942
40	0.317	6	4.20	3.32	166	13.944	2.3240
41	0.323	6	3.67	2.97	166	10.899	1.8165
42	0.329	6	2.97	1.92	166	5.702	0.9504
43	0.335	6	2.62	2.27	166	5.947	0.9912
44	0.341	6	2.27	3.67	166	8.330	1.3884
45	0.347	6	2.62	3.32	166	8.698	1.4447
46	0.353	6	1.4	2.62	166	0.366	0.6113



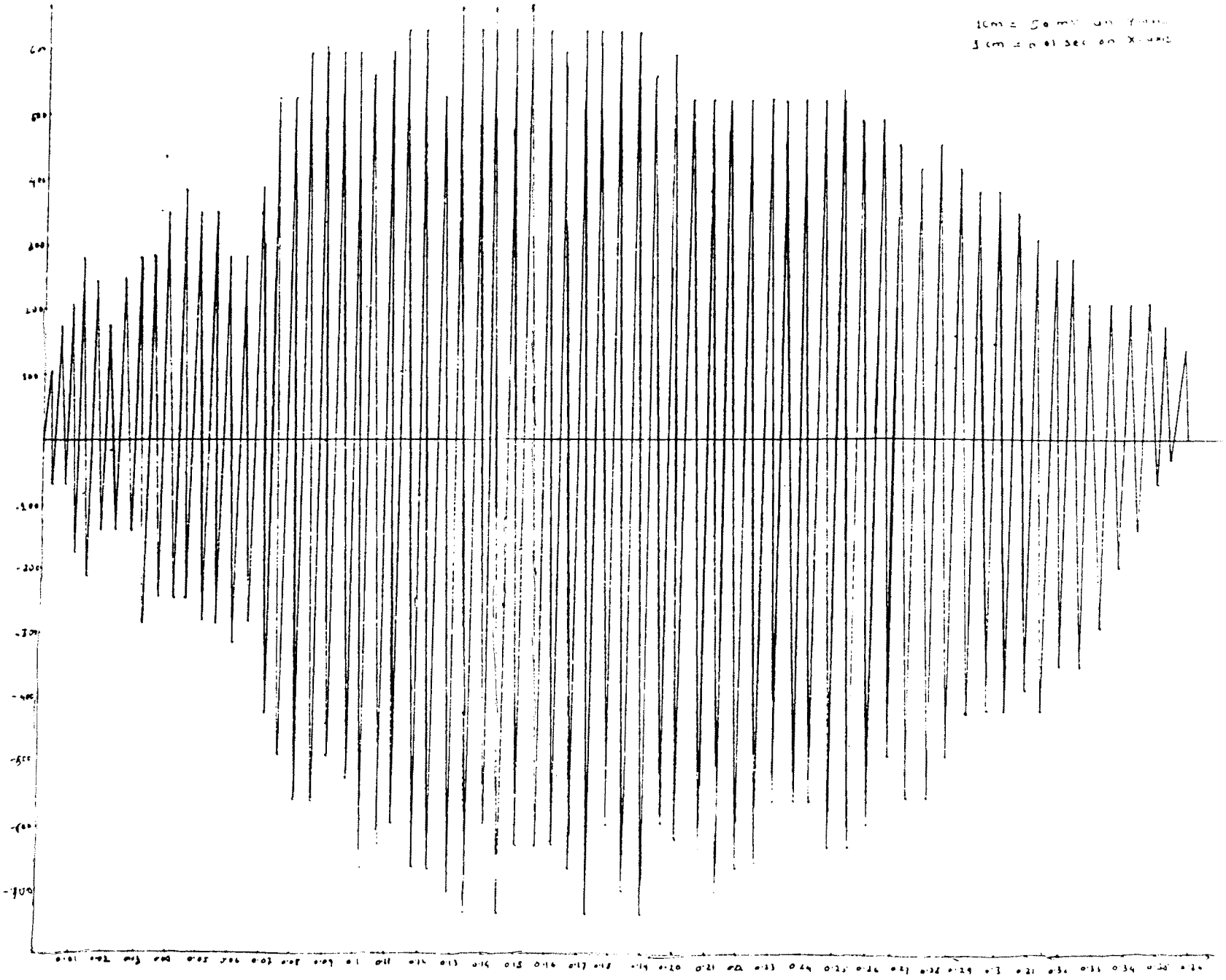
Amplitude Vs time graph for speech sound "o"



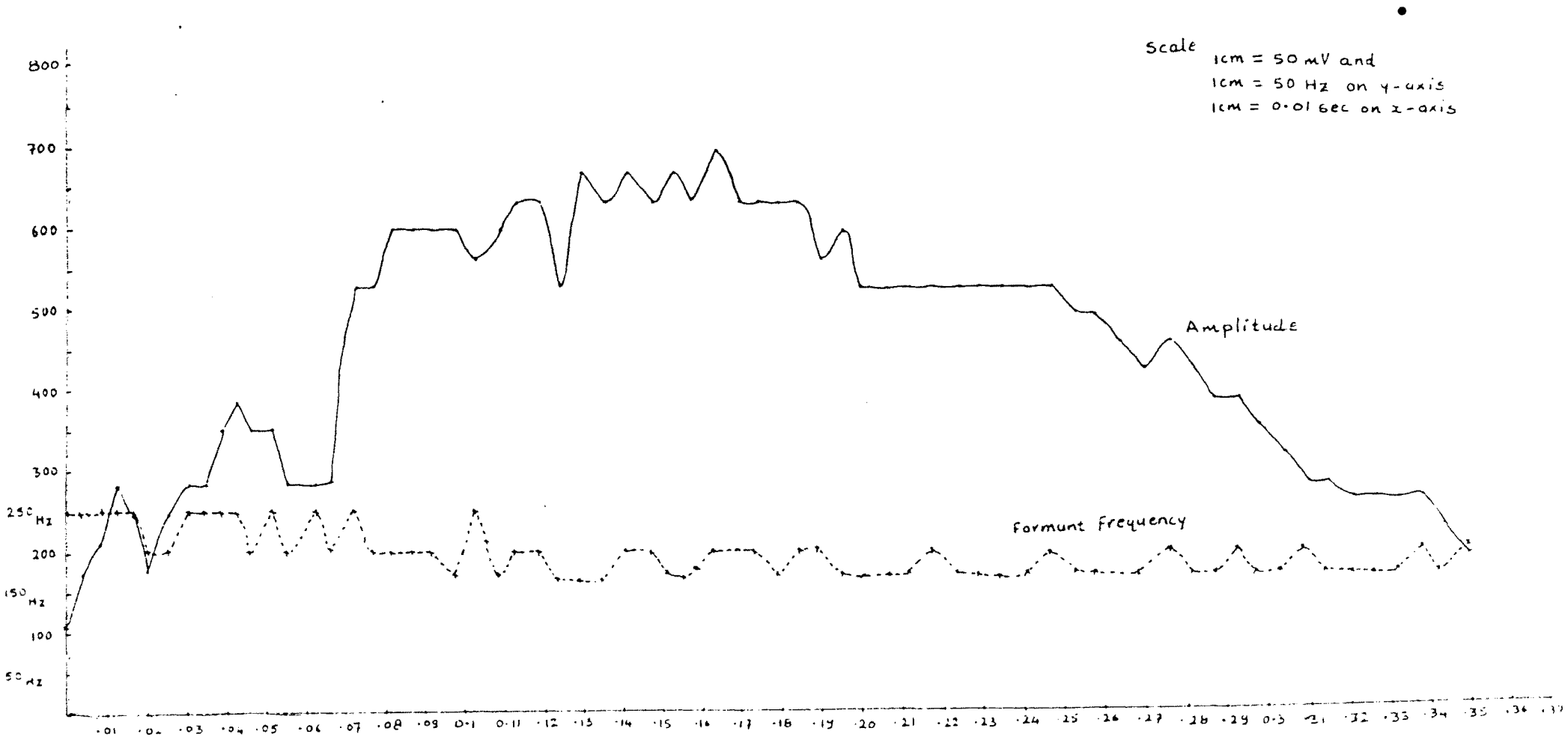
Amplitude and Formant frequency for "o"

4.4.8 Cycle analysis for 'E' and representation

cycle No.	Time	Period	Amplitude +ve	Amplitude -ve	Formant frequency	Energy	power
1	0.000	4	105	70	250	0.0735	0.0018
2	0.004	4	175	70	250	0.0122	0.0030
3	0.008	4	210	175	250	0.0367	0.0091
4	0.012	4	280	210	250	0.0588	0.0147
5	0.016	4	245	140	250	0.0343	0.0087
6	0.020	5	175	140	200	0.0245	0.0049
7	0.025	5	245	140	200	0.0343	0.0068
8	0.030	4	280	280	250	0.0784	0.0196
9	0.034	4	280	245	250	0.0686	0.0172
10	0.038	4	350	245	250	0.0857	0.0214
11	0.042	4	385	245	250	0.0857	0.0214
12	0.046	5	350	280	200	0.0980	0.0196
13	0.051	4	350	385	250	0.1347	0.0330
14	0.055	5	280	315	200	0.0882	0.0176
15	0.062	4	280	280	250	0.0784	0.0196
16	0.066	5	385	420	200	0.1617	0.0323
17	0.072	4	525	490	250	0.2572	0.0643
18	0.076	5	525	560	200	0.2940	0.0588
19	0.081	5	595	560	200	0.3332	0.0666
20	0.086	5	595	490	200	0.2915	0.0583
21	0.091	6	595	525	166	0.3123	0.0520
22	0.097	5	595	665	200	0.3956	0.0791
23	0.102	6	560	630	166	0.3528	0.0588
24	0.108	4	595	595	250	0.3540	0.0885
25	0.112	6	630	665	166	0.4189	0.0698
26	0.118	5	630	665	200	0.4189	0.0837
27	0.123	5	525	700	200	0.3675	0.0735
28	0.128	5	665	735	200	0.4887	0.0977
29	0.134	6	630	595	166	0.3748	0.0624
30	0.140	6	665	735	166	0.4887	0.0814
31	0.146	5	630	630	200	0.3969	0.0794
32	0.151	5	665	630	200	0.4189	0.0837
33	0.156	6	630	630	166	0.3969	0.0661
34	0.162	6	595	665	166	0.3969	0.0659
35	0.168	5	630	735	200	0.4630	0.0926
36	0.173	5	630	595	200	0.3748	0.0749
37	0.178	5	630	700	200	0.4410	0.0882
38	0.183	6	630	735	166	0.4630	0.0772
39	0.189	5	560	595	200	0.3332	0.0666
40	0.194	5	595	630	200	0.3748	0.0749
41	0.199	6	525	630	166	0.3307	0.0551
42	0.205	6	525	700	166	0.3675	0.0612
43	0.211	6	525	665	166	0.3491	0.0582
44	0.217	6	525	630	166	0.3307	0.0551
45	0.223	5	525	560	200	0.2940	0.0588
46	0.228	6	525	566	166	0.2940	0.0490



Amplitude Vs time graph for speech sound "E"



Amplitude and Formant frequency for "E"

4.4.8 (Continued.....)

cycle No.	Time	Period	Amplitude +ve	Amplitude -ve	Formant frequency	Energy	power
47	0.234	6	525	560	166	0.2940	0.4900
48	0.240	6	525	630	166	0.3307	0.0511
49	0.246	6	525	630	166	0.3307	0.0551
50	0.252	5	490	595	200	0.2915	0.0583
51	0.257	6	490	490	166	0.2401	0.0400
52	0.263	6	455	560	166	0.2548	0.0424
53	0.269	6	420	560	166	0.2352	0.0392
54	0.275	6	455	490	166	0.2229	0.0372
55	0.281	5	420	420	200	0.1764	0.0353
56	0.286	6	385	420	166	0.1617	0.0269
57	0.292	5	385	420	200	0.1617	0.0323
58	0.297	6	350	385	166	0.1347	0.0224
59	0.303	6	315	420	166	0.1323	0.0220
60	0.309	5	280	350	200	0.0980	0.0196
61	0.314	6	280	350	166	0.0980	0.0163
62	0.320	6	210	280	166	0.0588	0.0098
63	0.326	6	210	210	166	0.0441	0.0074
64	0.332	6	210	140	166	0.0294	0.0049
65	0.338	5	210	70	200	0.0147	0.0029
66	0.343	6	175	35	166	0.0061	0.0010
67	0.349	5	140	0	200	0.0196	0.0039

4.5 RESULTS AND DISCUSSION:

1. Analysis of waveform of 'ONE', 'ON', 'e', 'o' and 'u' reveals that there are no spikes superimposed on the waveform of vowels e, o and u. While there are spikes superimposed on the formant frequencies of the waveform of 'ONE' and 'ON'.

2. The formant of 'e' initially varies between 250-200 Hz and then varies between 200-150 Hz. The variation range frequency is 70 Hz or 200 ± 35 Hz.

3. The formant frequency of 'o' varies very little between 200 - 175 Hz.

4. The formant frequency for 'u' drops from 250 Hz to 200 Hz initially and remains constant at 200 Hz. Thus, it can be concluded that as you go from e-o-u the variation of formant tends to decrease.

5. The variation in amplitude with time is almost smooth for 'e'. The maximum variation is 600 mv. The average amplitude increases with time almost linearly in the period 0.80 ms. It then tends to level off and decreases linearly in the range 24 ms - 35 ms. The total span of 'e' is 300 ms.

6. The amplitude pattern for 'o' shows sudden increase of amplitude at 0.11 sec. and then there is sinusoidal variation of amplitude with maximum level = 11.5 V and minimum level = 1 V for t > 0.26 sec. the amplitude decreases almost linearly with time. The total span of 'o' is 0.37 sec.

7. The amplitude variation of 'u' is very random with V max = SV and V min. = 0.2 V. The total spread of 'u' is 0.33 sec.

The total spread of waveform for 'ONE' is 330 ms. The formant frequency varies between 200 Hz - 125 Hz. The variation is more in the period 0 - 160 ms. There are spikes superimposed on the waveforms. The details are given in Table 1.

Table 1 Spike analysis for 'one'

Spike No.	Amplitude Max. (v)	Amplitude Min. (v)	Spread in time (ms)	Variation type
1	0.099	0.021	2.0	e^{-at} Sin θ
5	0.550	0.015	1.5	e^{-at} Sin θ
10	0.056	0.046	1.0	e^{-at}
15	0.055	0.050	1.0	e^{-at}
25	0.041	0.038	1.0	e^{-at}
30	0.077	0.070	1.0	e^{-at}

The amplitude variation of 'ONE' shows that there is a large increase in the amplitude initially V max = 2800 mv = 2.8 V. The variation is not regular. However, for t > 160 ms the amplitude decreases with time almost linearly. Vmin = 100 mv = 0.1 V.

The total spread of waveform for 'ON' in 355 m sec. The formant frequency varies between 200 Hz - 143 Hz. The variation is more in the period 0 -80 ms. There are spikes superimposed on the waveforms. The details of spikes analysis are given in Table 2.

Table 2 Spike analysis for 'on'

Spike No.	Amplitude Max. (v)	Amplitude Min. (v)	Spread in time (ms)	Variation type
1	0.046	0.042	1.0	e^{-at}
15	0.038	0.035	1.0	e^{-at}
20	0.053	0.049	1.0	e^{-at}
25	0.061	0.056	1.0	e^{-at}
30	0.027	0.024	1.0	e^{-at}

4.6 SPEECH SYNTHESIS- A PROPOSED APPROACH:

Synthesis by Analysis System

We have the stored waveform by Single Cycle Analysis Method. A formant frequency has been found out. It is a main source in the synthesis. On this frequency, there are number of synthesis on this frequency harmonies and other amplitude variation superimposed.

Amplitude Variation

These variations are present in the amplitude of the formant frequency. The time - amplitude graph can be fitted in the 2^{nd} order or 1^{st} order damped waveform (with step input) by splitting it into proper window. This can be explained as follows.

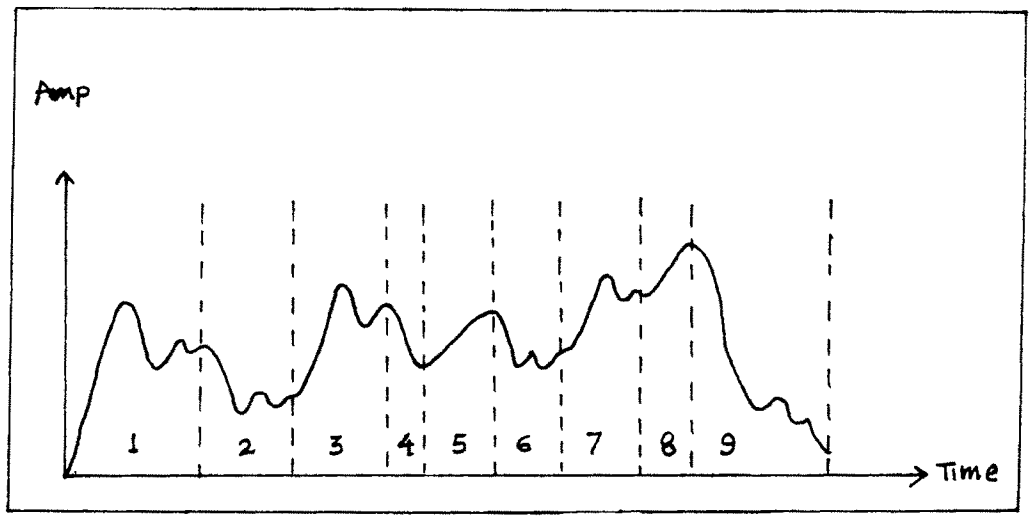


Fig. 4.2 Typical waveform

Now 1, 2, 3, 6, 7, and 9 can be viewed as the response of some second order systems step input. From the waveform its damping ratio & further R, L, C can be calculated. Also 4,5,6 can be viewed as 1st order system's step input responses having same R and C values.

Now we can generate these systems by programmable R and C by keeping some L constant)as in 2nd order system) and can get the required waveform of amplitude with the help of + Ve or _ Ve step excitation. This amplitude (response) becomes the limit of the amplitude of formant frequency. The block diagram of the proposed system can be as shown.

Harmonic Addition

The harmonic present in the waveform of the speech are in the form of voltage spikes. These can be mixed in the waveform by the help of microprocessor and one differentiator - FIR system. It can be designed as Fig 4.4.

Thus, one RC network if excited with a proper valued square impulse, can generate a spike.

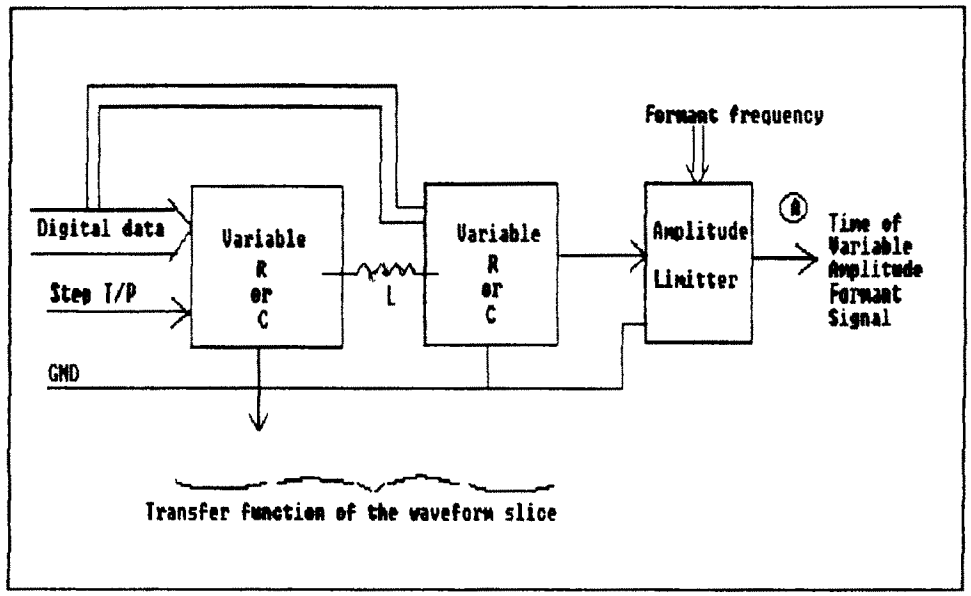


Fig. 4.3 Block diagram of the proposed system

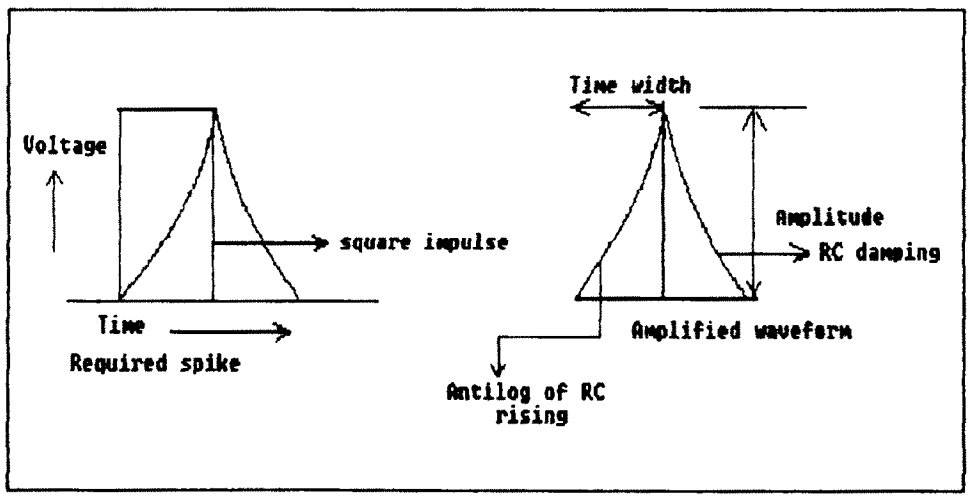


Fig. 4.4 Harmonic addition (Spike)

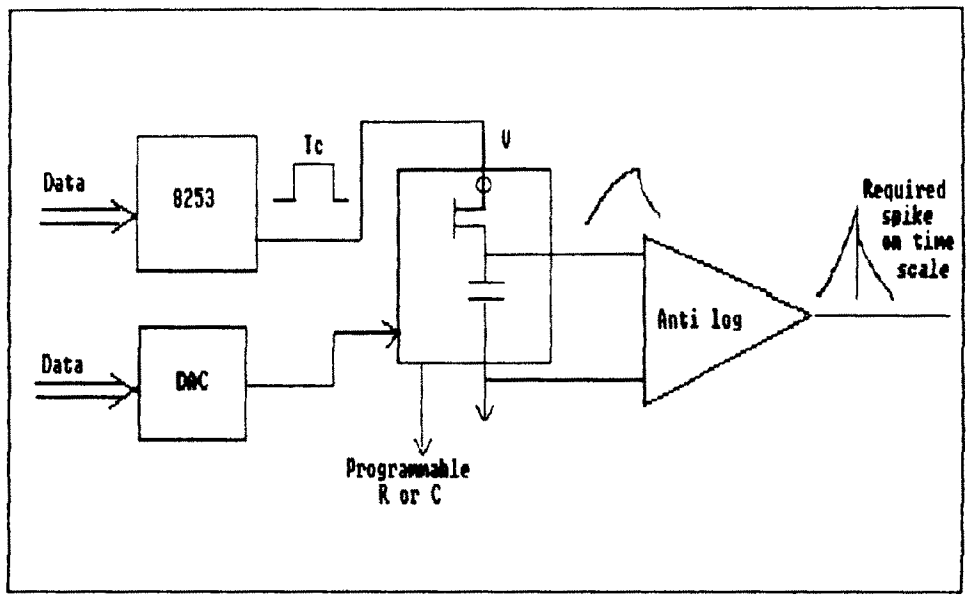


Fig. 4.5 Addition of Amplitude variation and Harmonics to get required spike

The required no. of spikes can be generated and fed from microprocessor and the dominant harmonics can be added in 1st formant signal structure.

Mixing 1 and 2 we get the required waveform of the speech in its electrical equivalent signal sending it through speaker we can get the required synthesized sound of the required formant frequency.

REFERENCES

1. R.K. Balani. "Proceedings of National Symposium on recent advances in DSP." BARC - SPS - 1 (1990).
2. A.M. Nolly, J. Acoust. Soc. America Vol.- 41, pp 293-309 (1967).
3. M.M. Sondhi. IEEE Irans Andia Electroacoust. Vol. Av-16, pp. 262-266 (1968).
4. J.D. Marked and A.N. Grey. "Linear prediction of speech." pp. 2066-211, N.. Springer (1976).
5. M.J. Ross, M.J. Shaffer, A. Cohen, R. Froudberg and H.J. Manley. IEEE Irans ASSP. Vol. A SSP-22, pp. 353-362 (1974).
6. J.J. Dubnowski, R.K. Schafer and L.R. Rabinor. IEEE Irans Aconst. Vol. ASSP-24, pp. 2-8 (1976).
7. M.R. Kesheorey and S.K. Bhatia. "Proceedings of the National Symposium on recent advances in DSP." BARC SP 3-2 (1990).