CHAPTER 5

ESTIMATION OF FORMANT FOR BODO VOWELS AND BODO WORDS OF DIFFERENT TYPES

5.1 Introduction-Formant Frequency

The process of artificial machine generated acoustic signal that is similar to the speech sound of humans is called speech synthesis. Formants frequency can be considered as one of the most important distinguishing features of human speech. The specific resonance frequency of the human vocal tract where there is maximum energy concentration during vowel utterance is called Formant frequency. In a sound spectrum, formants are the spectral peak of the sound spectrum [43]. In general, there are three formant frequencies namely -F1, F2 and F3 that are taken into consideration for discrimination and perception and discrimination [9].

5.2 Estimation of Formants

During the last few decades, many approaches have been developed for the analysis and synthesis of speech sound signal. Among those, the basic method used for research in speech recognition and synthesis is estimation of formants [12]. In the present study, the Formant Estimation technique is applied for estimating the frequencies for various types of Bodo words and vowels. Applying this technique, the entire range of frequencies are divided into a definite number of segments with a specified boundary, say 'x'. A predictor polynomial which is defined as the FFT of the Corresponding 2nd order predictor and is formulated as [46] :

$$A_x\left(e^{j\omega}\right) = 1 - \alpha_x e^{j\omega} - \beta_x e^{-j2\omega} \quad \ldots (5.1)$$

Here, the predictor co-efficients are real valued and denoted by α_x and β_x.

Thus, the equation given above can be elaborated as:

$$\begin{aligned}\left|A_x\left(e^{j\omega}\right)\right|^2 &= 1 + \propto_x{}^2 + \beta_x{}^2 - 2\alpha_x(1-\beta_x)\cos(\omega) - 2\beta_x\cos(2\omega) \\ &= (1+\beta_x)^2 + \alpha_x{}^2 + [\alpha_x{}^2((1-\beta_x)^2)/4\beta_x] \\ &\quad -4\beta_x[cos\omega + \{\alpha_x(1-\beta_x)\}/4\beta_x]^2\end{aligned} \quad \ldots (5.2)$$

Where β_x determines the BW of the resonator. It is defined as negative logarithm of $(-\beta_x)\left[-A_x\left(e^{j\omega}\right)\right]^2$.

The Formant Frequency is:

$$P_f = across[-\alpha_x(1-\beta_x)/4\beta_x] \quad \ldots (5.3)$$

The predictor error E (Using equation (5.2)) is:

$$E\left(\omega_{x-1}, \frac{\omega_x}{\alpha_x.\beta_x}\right) = \left(1 + \alpha_x{}^2 + \beta_x{}^2\right)r_x(0) - 2\alpha_x(1-\beta_x)r_x(1) - 2\beta_x r_x(2) \quad \ldots (5.4)$$

The autocorrelation coefficients is given by $r_x(\gamma)$ for any segment x.

Hence, the value for the predictor co-efficients (α_x and β_x) are given as,

$\alpha_x < 2,$

$$-1 < \beta_x < [-|\alpha_x|/(4-|\alpha_x|)] \quad \ldots (5.5)$$

5.3 Formant Frequencies of different Bodo Vowels and Words

Now, with the help using the equation (5.5), the Formant Frequencies of the Bodo vowels and different type of Bodo words of different structures (CV, CVC and VC; C:Consoant, V:Vowel,) are extracted for the two different genders of speakers-

male and female. These frequencies are extracted with the help of the following algorithm:

Algorithm

I. Y=speech wave ("path of .wav file")

II. F_{sa}=16000; (frequency for sampling)

III. X=fft(Y); (apply Fourier Transformation to the signal)

IV. Y=abs (X); (taking the absolute value)

V. Freq=1:4000; (range of frequency)

VI. Plot (Freq,8*log10(y(Freq)), 'b-');
(plot the waveform)

VII. p=lpc(x,10); (formant frequency by using LPC)

VIII. p_k=a(2); (predictor coefficent)

IX. q_k=a(3); (predictor coefficent)

X. q=1;

XI. [h,ws]=freq(p,q,100);

XII. W1=ws*F_s / W1(2*Pi);

XIII. Wait;

XIV. Plot (ws, %*lof(abs(0.4*h), 'k - -');
(plots the frequency response)

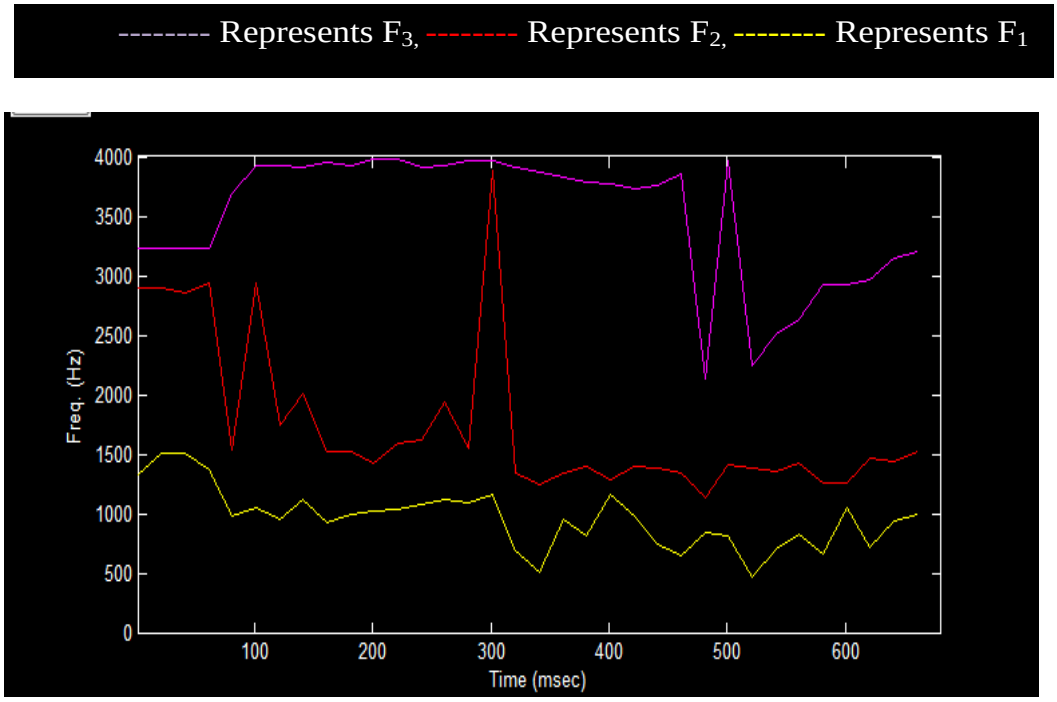
The algorithm given above has been used to find the formant frequencies (F1, F2 and F3) for the different structured words (CV, CVC and VC) and vowels of the Bodo language and plotted. The variation of Formant frequencies Bodo vowels for male and female has been shown in **Table 5.1 and Table 5.2**; and spectral representations are depicted in the **Figure 5.1, 5.2, 5.3, 5.4, 5.5 and 5.6.**

Table 5.1: Variations in formant frequencies (F1, F2 and F3) of Bodo vowels for Male

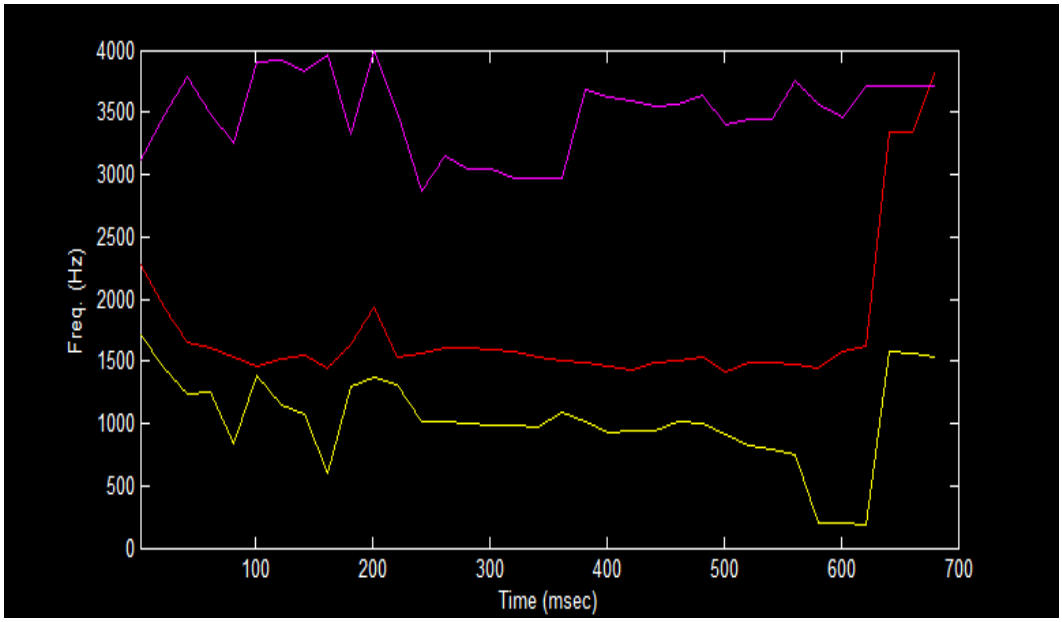
Vowel	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
/a/	Max	1.504952	3.886769	3.981763
	Min	0.468134	1.129728	2.138753
/e/	Max	0.466512	2.451997	3.993118
	Min	0.220518	2.160765	2.621805
/i/	Max	2.370138	3.131034	3.998523
	Min	0.26408	2.398267	3.068158
/o/	Max	2.023203	3.780806	3.279644
	Min	0.325972	0.655055	2.473697
/u/	Max	1.915692	3.972639	3.794718
	Min	0.277038	0.898572	3.074773
/ ɯ /	Max	2.898747	3.362456	3.999982
	Min	0.233757	1.060256	2.782809

Table 5.2: Variations in formant frequencies (F1, F2 and F3) of Bodo vowels for Female

Vowel	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
/a/	Max	1.704159	3.828826	3.997937
	Min	0.179824	1.418355	2.861879
/e/	Max	0.001198	0.002757	0.003927
	Min	0.00035	0.002493	0.003393
/i/	Max	0.338206	3.414298	3.988886
	Min	0.25403	2.701021	3.50482
/o/	Max	0.623018	1.189958	3.999093
	Min	0.282194	0.821005	3.28823
/u/	Max	0.795688	2.671362	3.722507
	Min	0.300116	0.610074	2.871546
/ ɯ /	Max	2.301972	3.753257	3.930727
	Min	0.322836	0.85328	3.103658



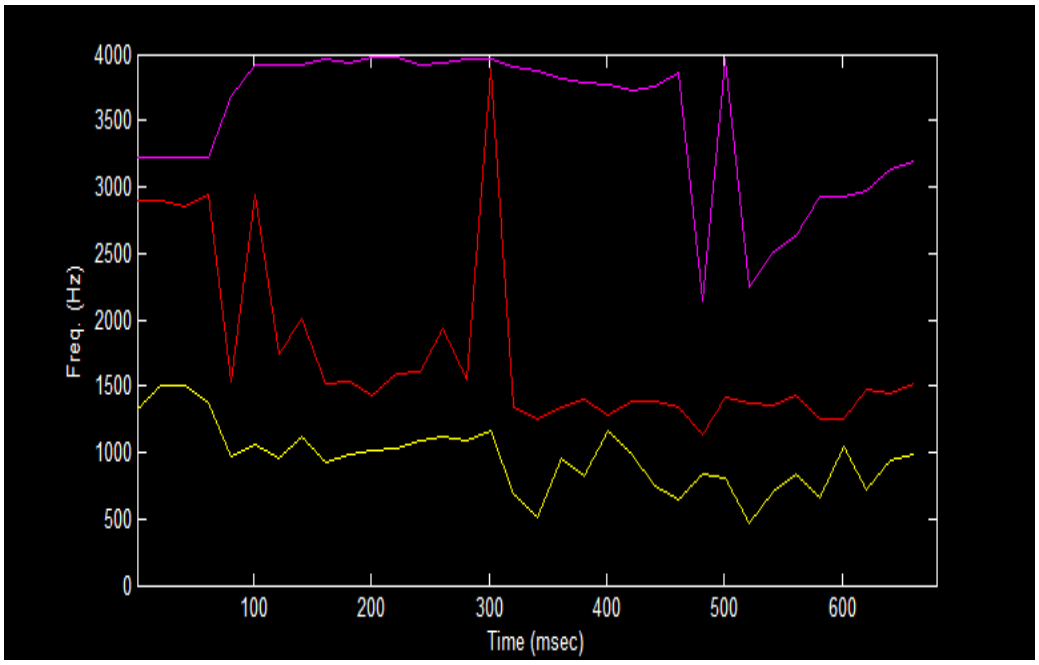
(Male)



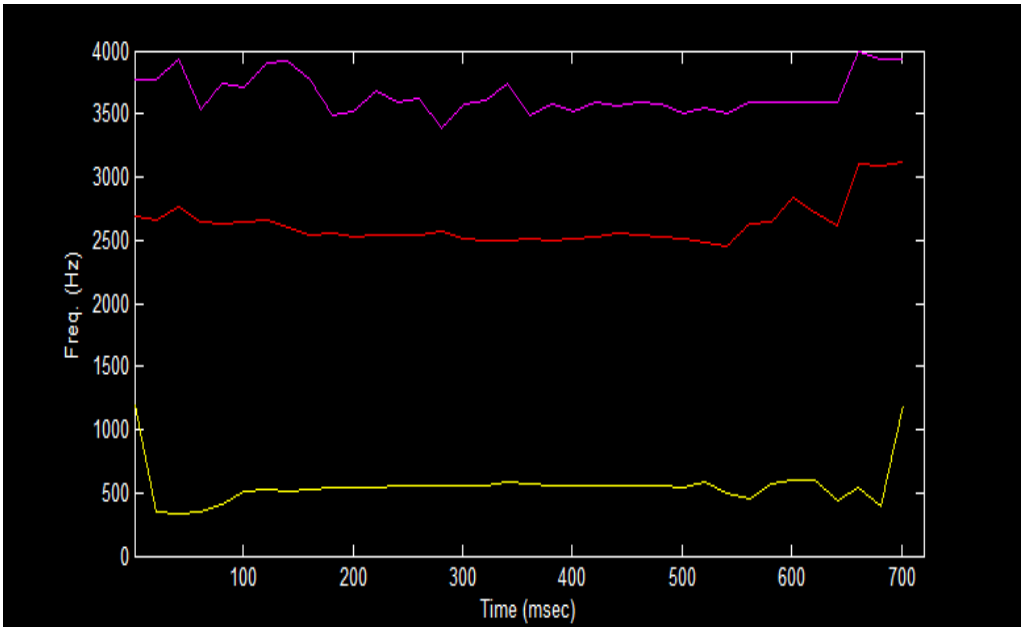
Female

Fig 5.1: Formant Frequencies (F1, F2 and F3) of Bodo vowels /a/ for Male and Female

-------- Represents F_3, -------- Represents F_2, -------- Represents F_1



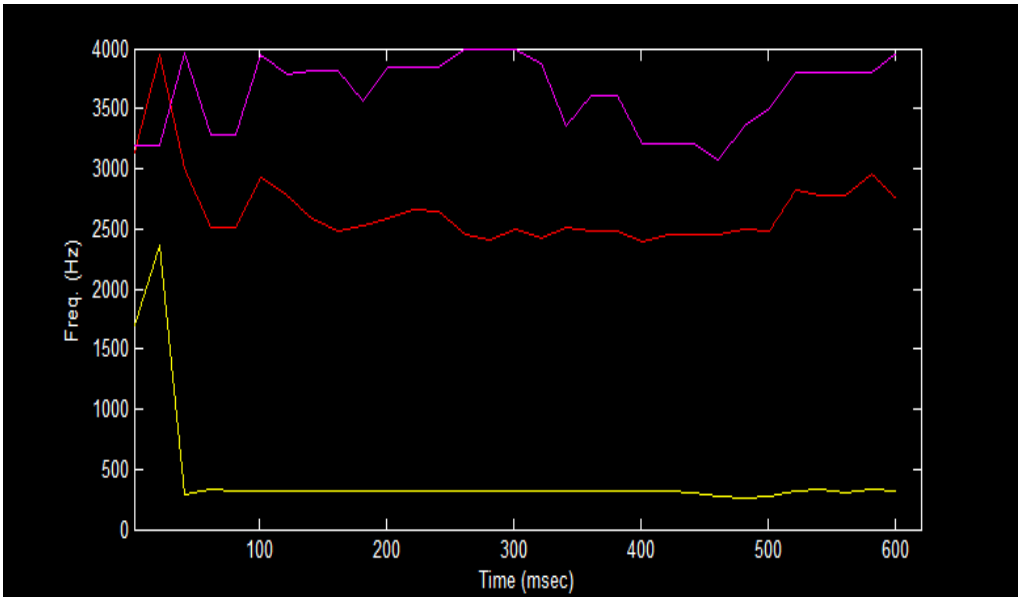
Male



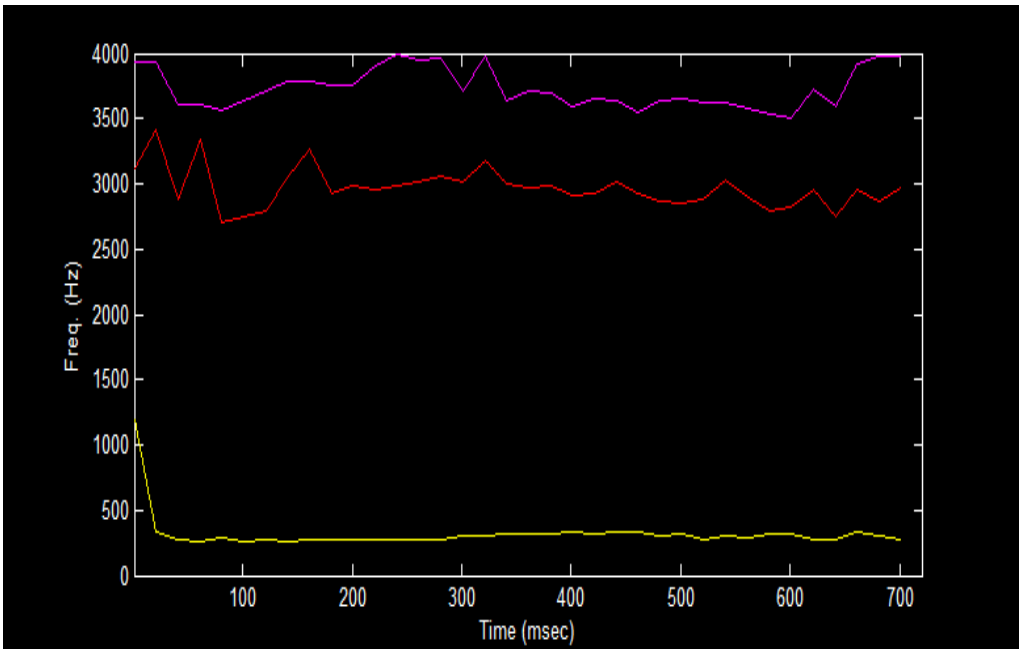
Female

Fig 5.2: Formant Frequencies(F1, F2 and F3) of Bodo vowels /e/ for Male and Female

-------- Represents F_3, -------- Represents F_2, -------- Represents F_1

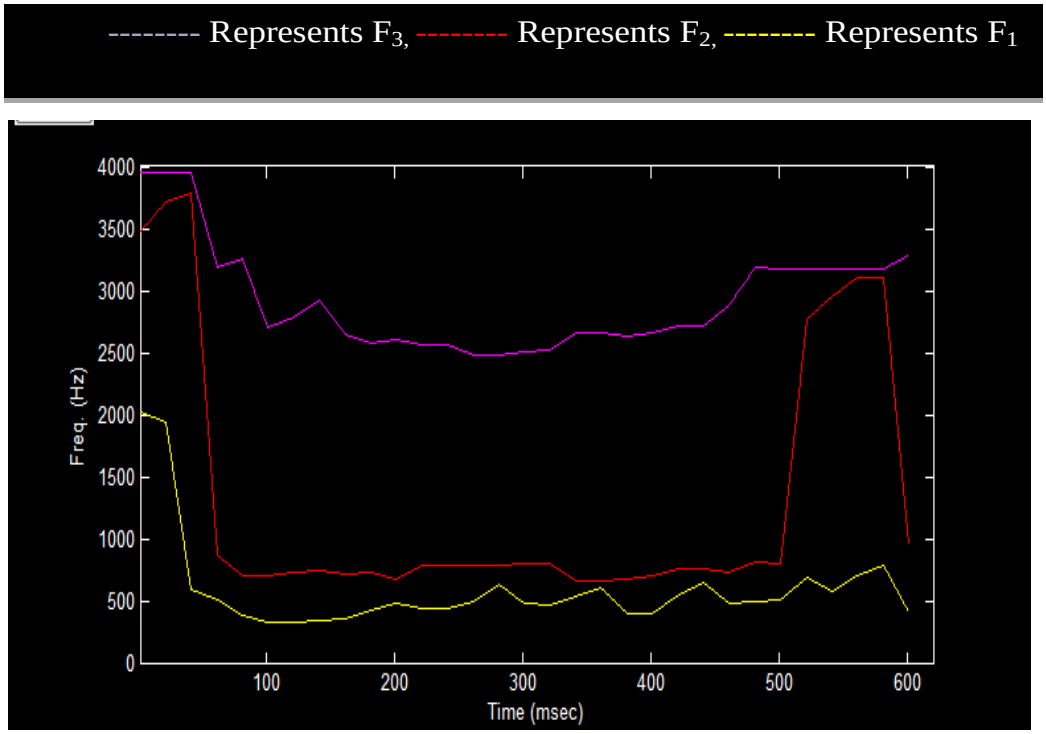


Male

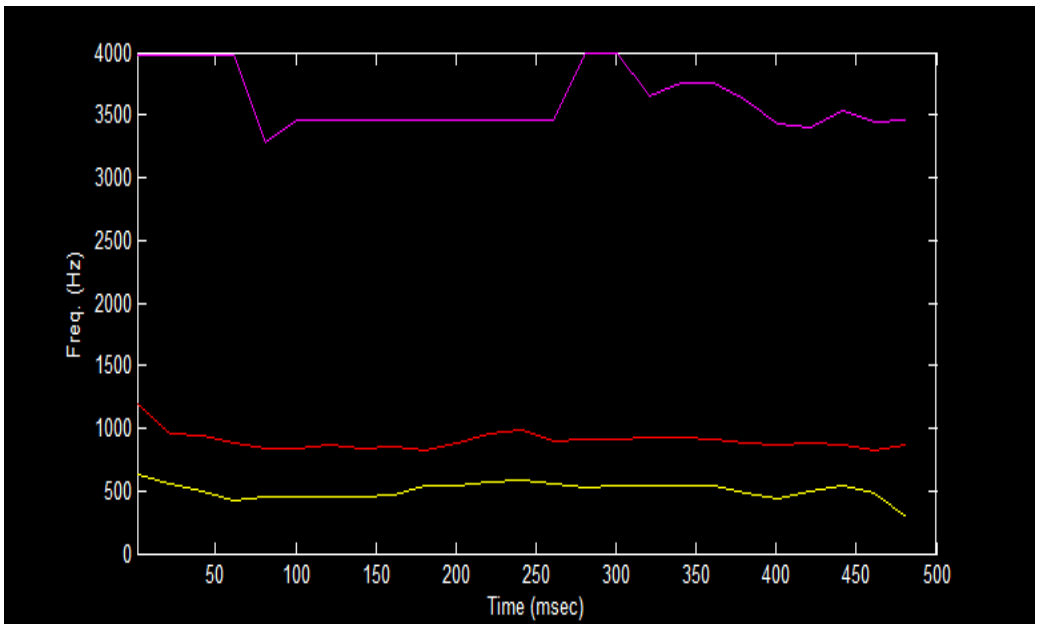


Female

Fig 5.3: Formant Frequencies (F1, F2 and F3) of Bodo vowels /i/ for Male and Female

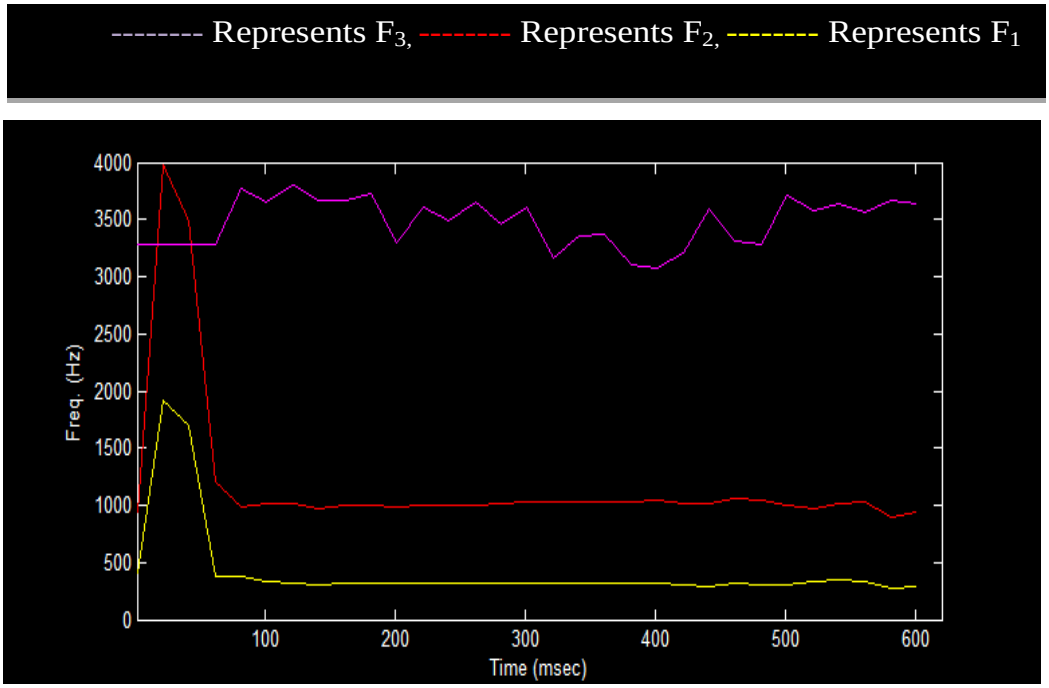


Male

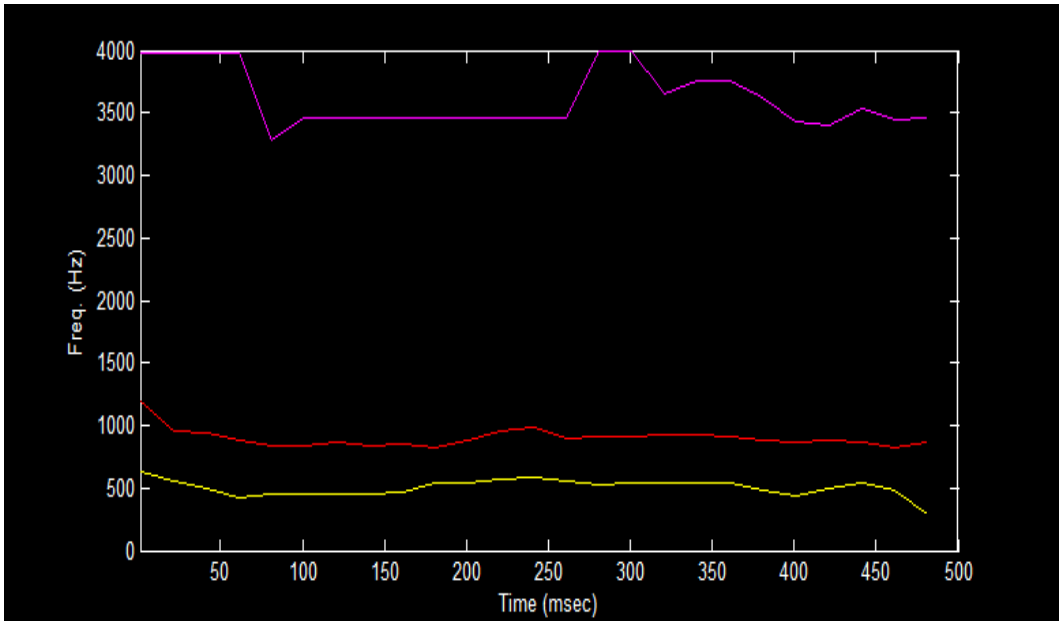


Female

Fig 5.4: Formant Frequencies (F1, F2 and F3) of Bodo vowels /o/ for Male and Female

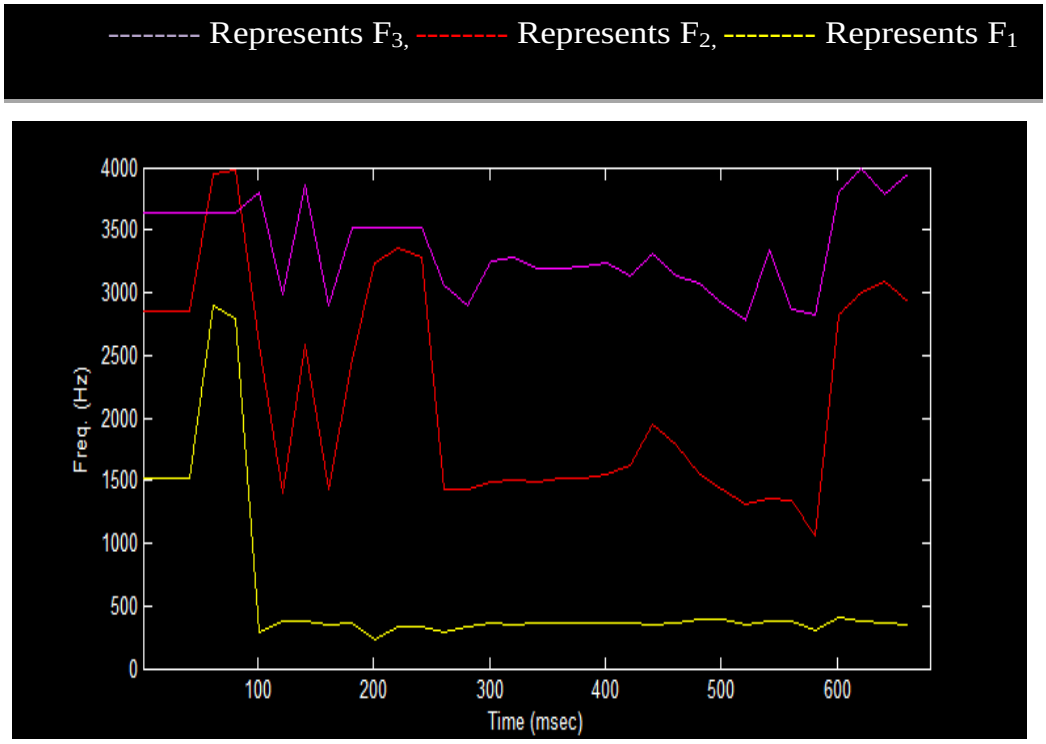


Male

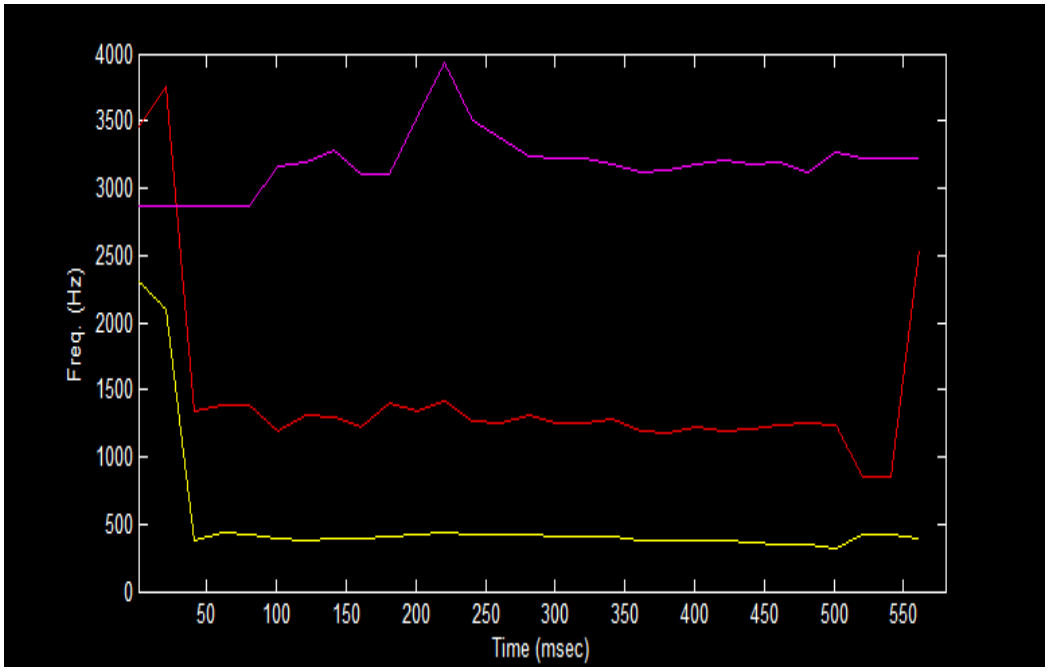


Female

Fig 5.5: Formant Frequencies(F1, F2 and F3) of Bodo vowels /u/ for Male and Female



Male



Female

Fig 5.6: Formant Frequencies(F1, F2 and F3) of Bodo vowels /ɯ /for Male and Female

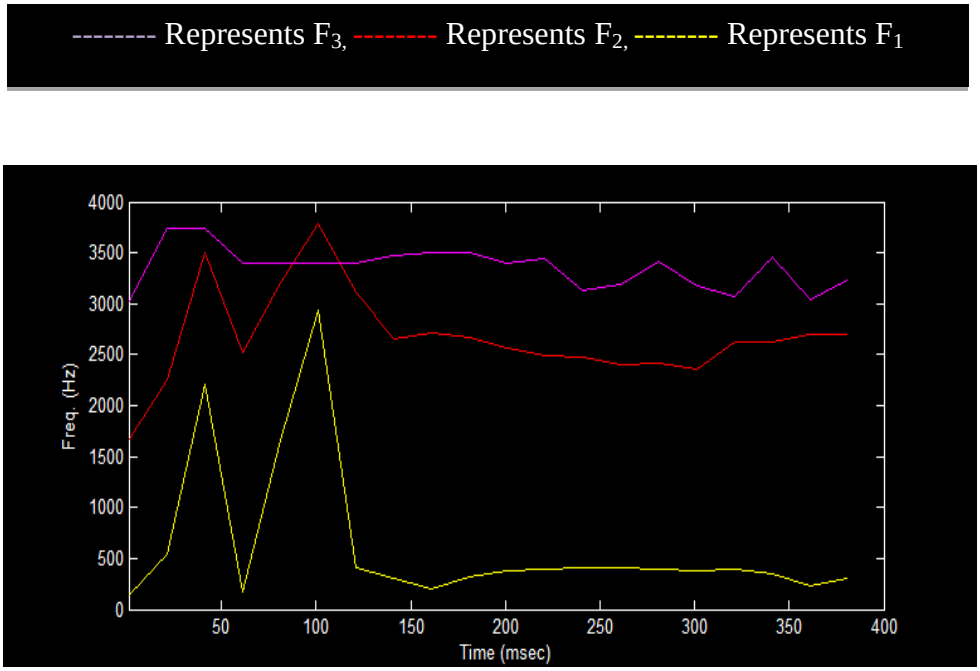
Similarly, for CV type of Bodo words, results for the male informants are given in **Table 5.3** and results for the female informants are given in **Table 5.4** respectively and their spectral representations are depicted in the **Figure 5.7, 5.8 and 5.9**; for the Bodo word of structure CVC, results are shown in **Table 5.5** and **Table 5.6** for male and female respectively, corresponding spectral representations are displayed in the **Figures- 5.10, 5.11 and 5.12** and for VC type of Bodo words for male informants, is given in **Table 5.7** and for female informants, it is given in **Table 5.8** and their spectral representations are depicted in **Figures - 5.13, 5.14 and 5.15.**

Table 5.3 Variations in formant frequencies (F1, F2 and F3) for Bodo words of structure CV (Male)

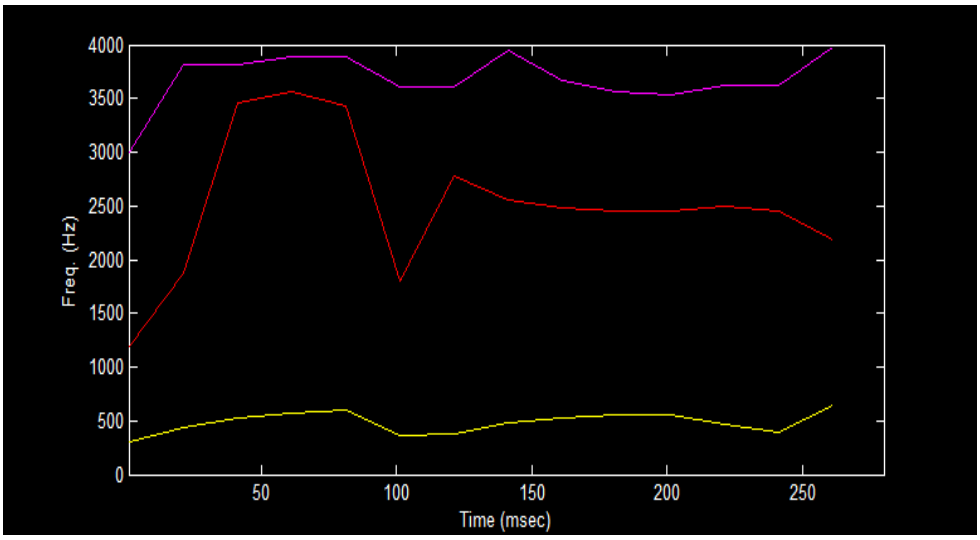
Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
k^he	Max	2.940092	3.778961	3.732705
	Min	.138524	1.577268	3.028903
Saa	Max	1.206797	2.890950	3.417528
	Min	.194924	.972059	2.306451
t^haa	Max	1.505742	3.729189	3.894283
	Min	.092787	1.207206	2.716378

Table 5.4 Variations in formant frequencies (F1, F2 and F3) for Bodo words of structure CV (Female)

Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
k^he	Max	539.608	3556.938	3970.419
	Min	360.13	1794.776	3540.252
Saa	Max	1773.655	3354.128	3834.984
	Min	180.198	1255.276	2927.229
t^haa	Max	1.092158	2.951778	3.861136
	Min	.259971	1.555119	2.797289

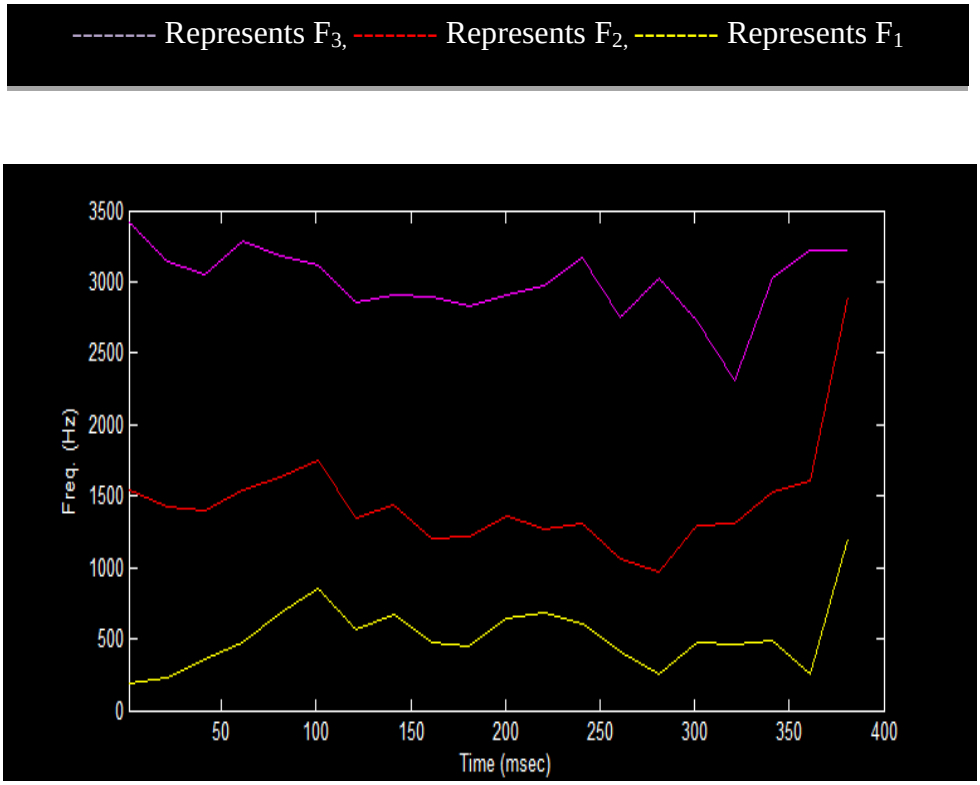


Male

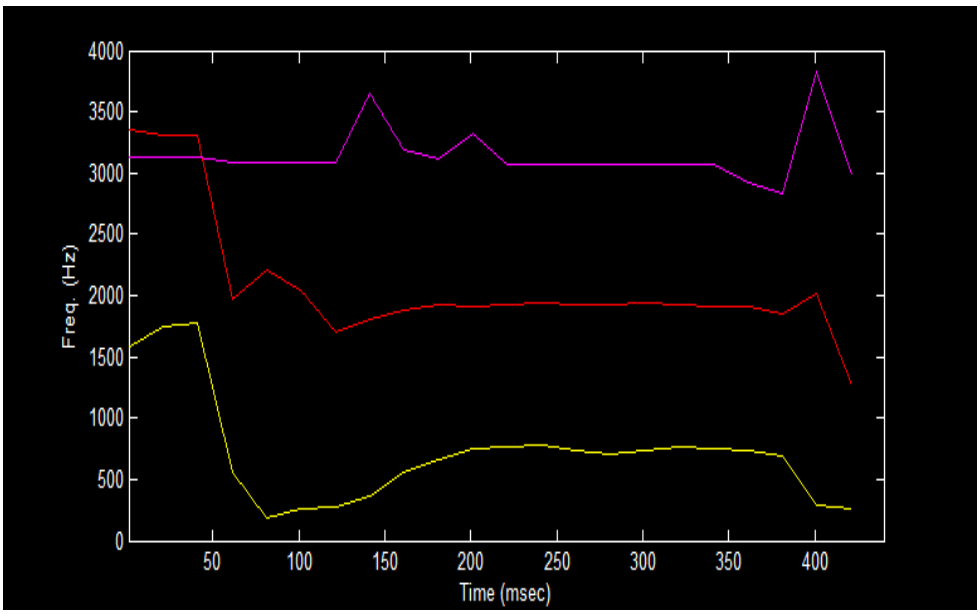


Female

Fig 5.7: Formant frequency (F1, F2 and F3) of CV type word /k^{h}e/ for Male and Female



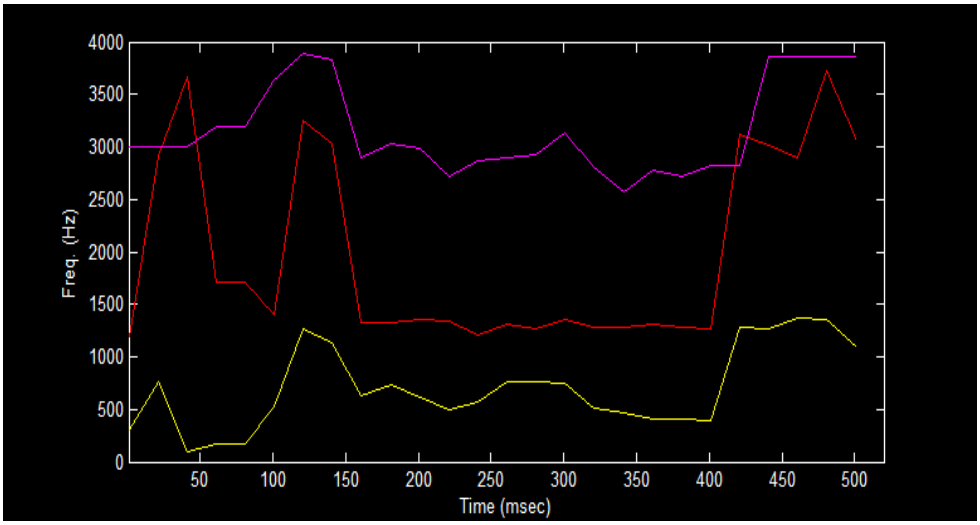
Male



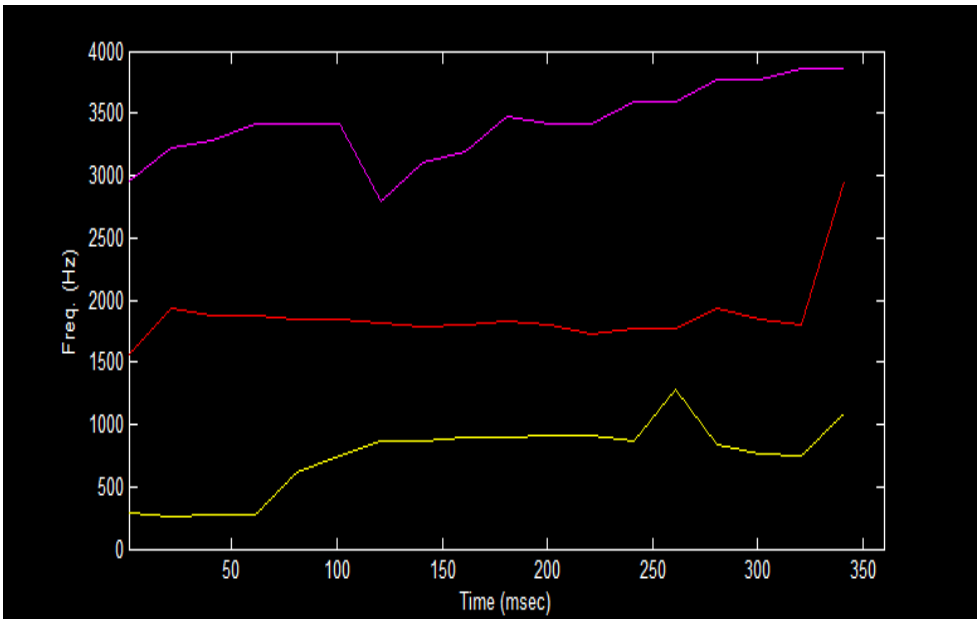
Female

Fig 5.8: Formant frequency (F1, F2 and F3) of CV type word /Saa/ for Male and Female

-------- Represents F_3, -------- Represents F_2, -------- Represents F_1



Male



Female

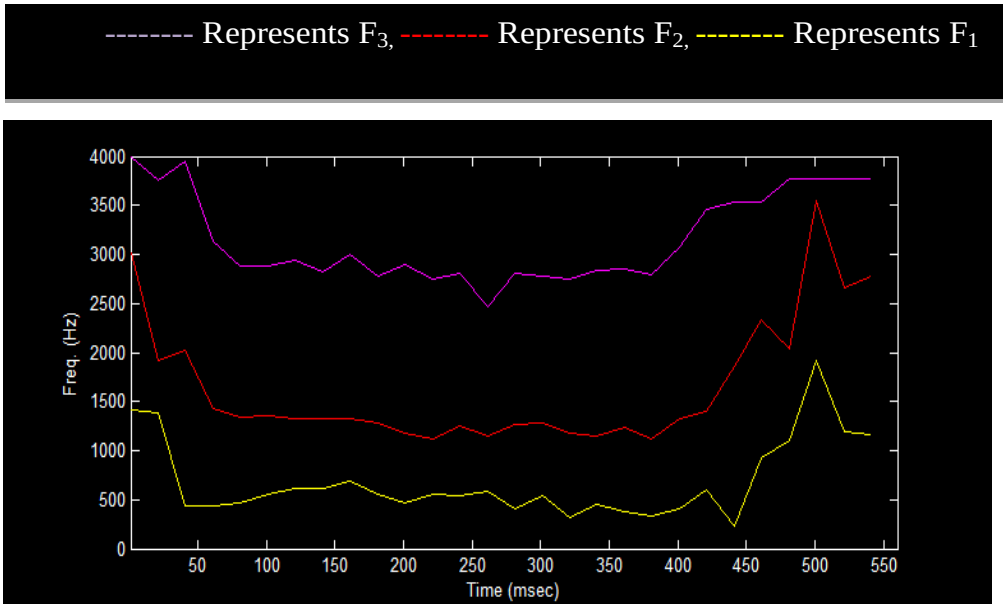
Fig 5.9: Formant frequency (F1,F2 and F3) of CV type word /t^haa/ for Male and Female

Table 5.5 Variations in formant frequencies (F1, F2 and F3) of some Bodo words of CVC Type for Male

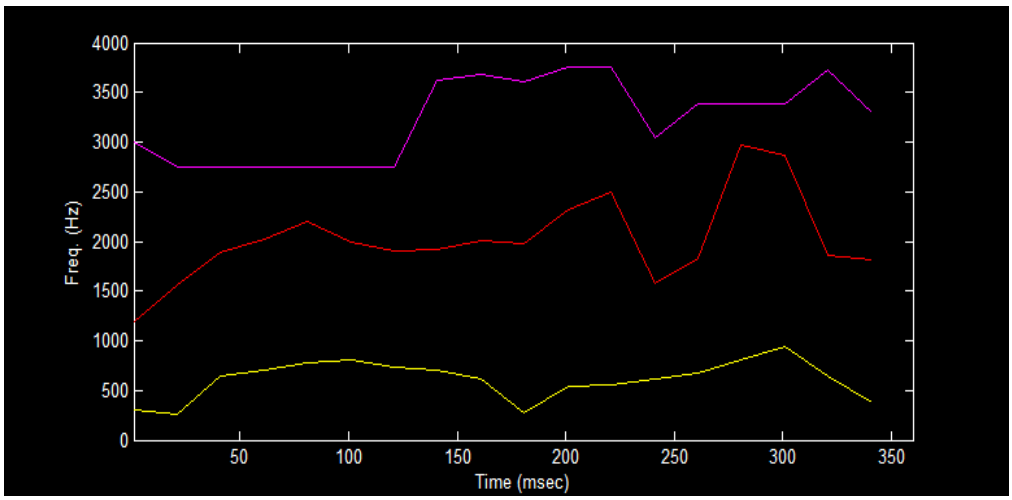
Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
gaab	Max	1.918941	3.545219	3.941283
	Min	.235756	1.118598	2.472513
Lir	Max	1.533586	3.288515	3.838438
	Min	.202258	1.250258	2.910778
naay	Max	1.455426	3.245492	3.755698
	Min	.157458	1.248238	2.535539

Table 5.6 Variations in formant frequencies (F1, F2 and F3) of some Bodo words of CVC Type for Female

Word	Frtequency	F1(KHz)	F2(KHz)	F3(KHz)
gaab	Max	.811582	2.955215	3.755503
	Min	.258005	1.550864	2.749731
Lir	Max	1.799738	2.687556	3.882493
	Min	.163174	.864491	2.599785
naay	Max	1.158539	3.745989	3.500589
	Min	.153293	1.065905	2.524303



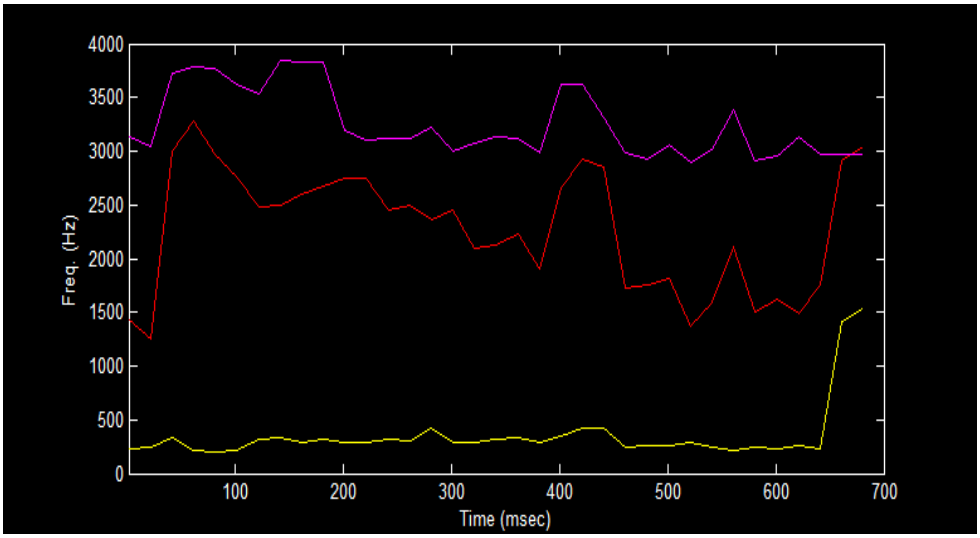
Male



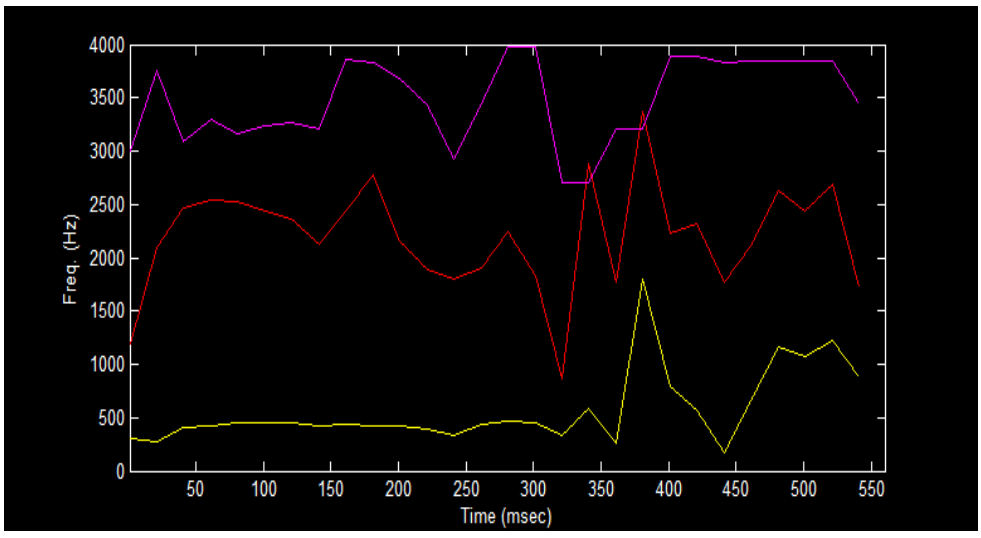
Female

Fig 5.10: Formant frequency (F1,F2 and F3) of CVC type word /gaab/ for Male and Female

-------- Represents F_3, -------- Represents F_2, -------- Represents F_1

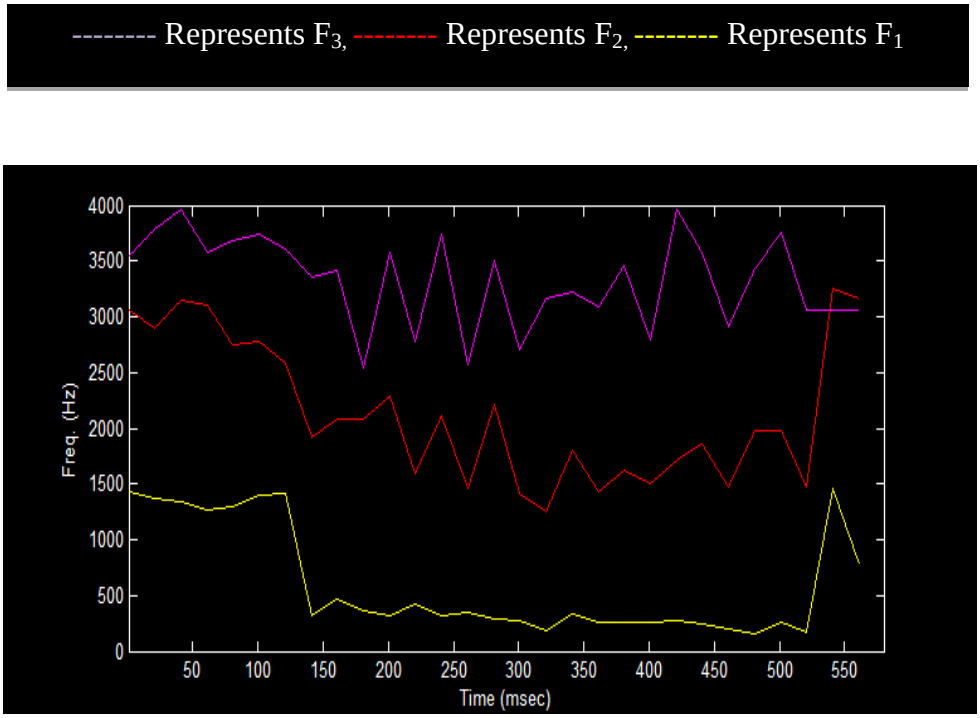


Male

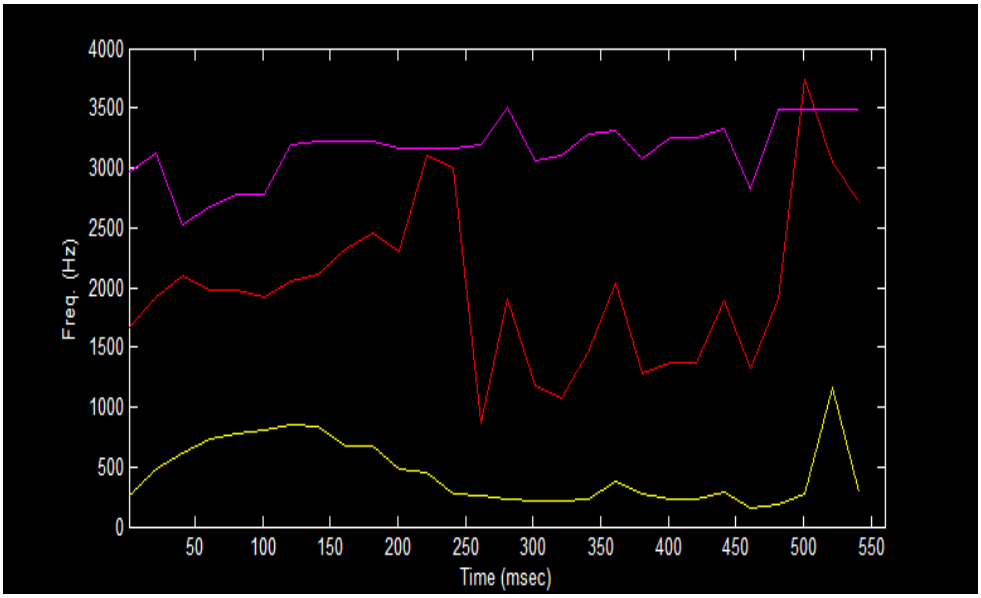


Female

Fig 5.11: Formant frequency (F1, F2 and F3) of CVC type word /lir/ for Male and Female



Male



Female

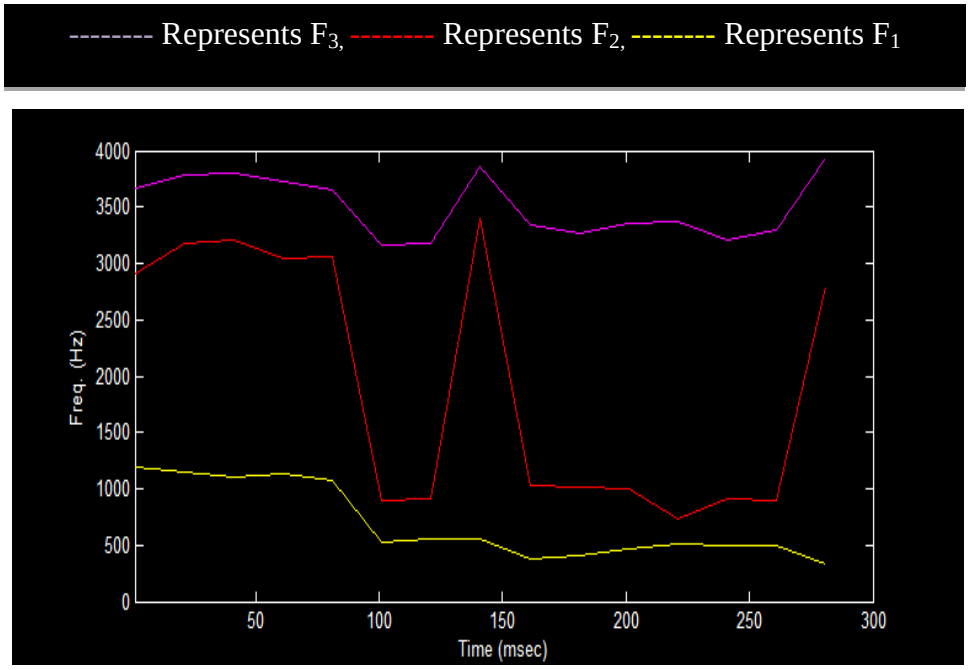
Fig 5.12: Formant frequency (F1, F2 and F3) of CVC type word /naay/ for Male and Female

Table 5.7 Variations in formant frequencies (F1, F2 and F3) of some Bodo words of VC Type for Male

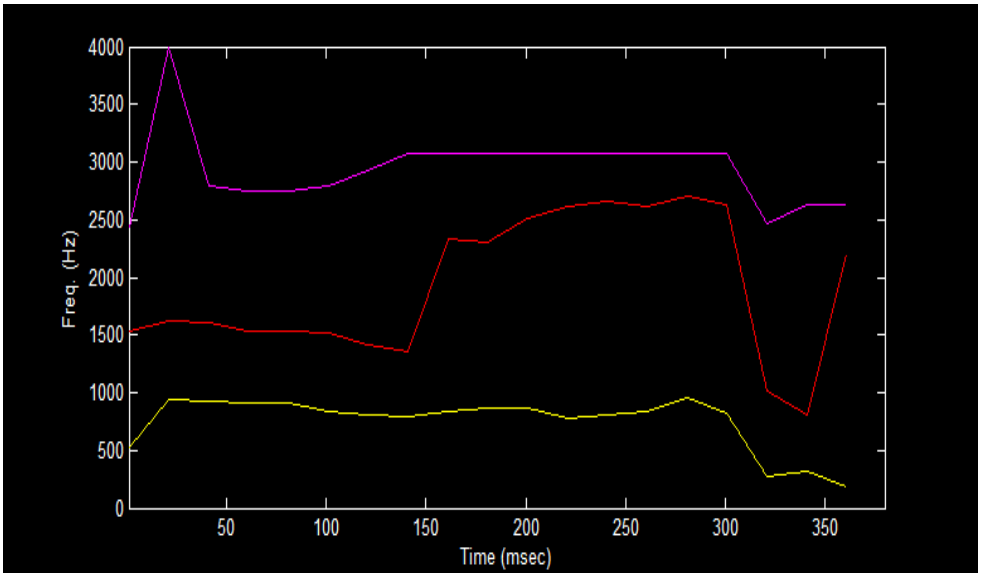
Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
aaw	Max	1.119256	3.208623	3.938411
	Min	.330325	.736218	3.155440
on	Max	1.954625	3.439314	3.985615
	Min	.190579	.785831	2.898596
un	Max	.429009	3.044419	3.858840
	Min	.229748	.885991	2.935936

Table 5.8 Variations in formant frequencies (F1, F2 and F3) of some Bodo words of VC Type for Female

Word		F1(KHz)	F2(KHz)	F3(KHz)
aaw	Max	.939718	2.705352	3.986321
	Min	.188443	.800056	2.433698
on	Max	.485250	2.820344	3.553849
	Min	.173752	.722737	2.500963
un	Max	3.43755	1.597975	3.084086
	Min	2.91155	.731739	2.751246

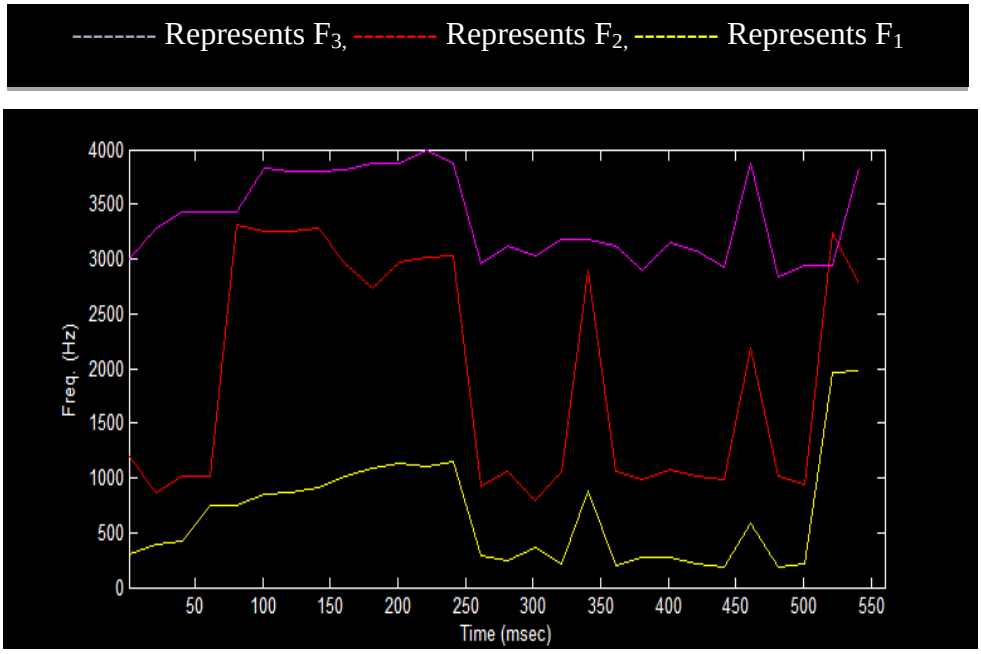


Male

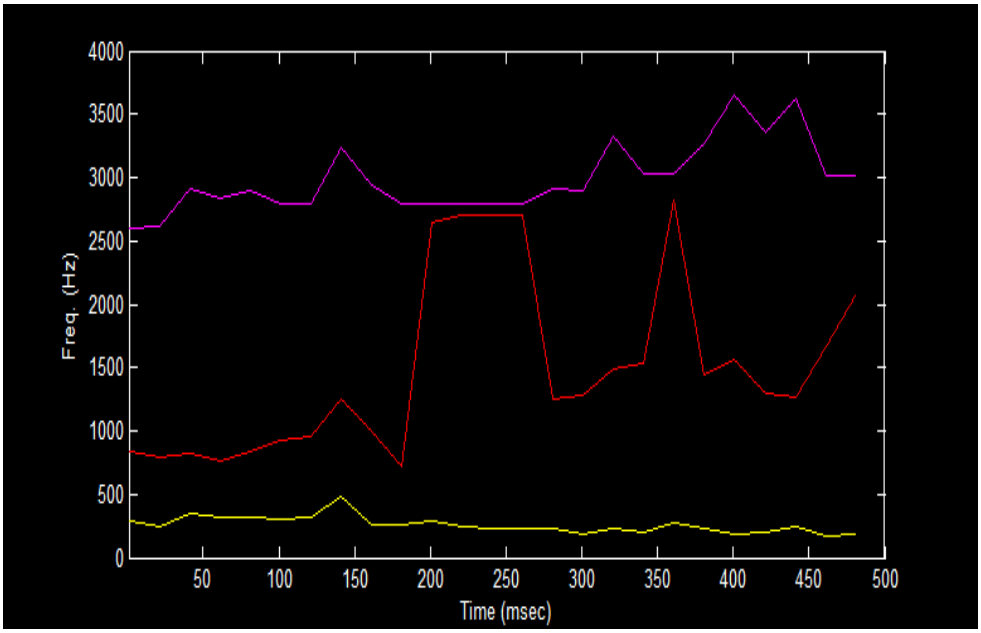


Female

Fig 5.13: Formant frequency (F1, F2 and F3) of VC type word /aaw/ for Male and Female

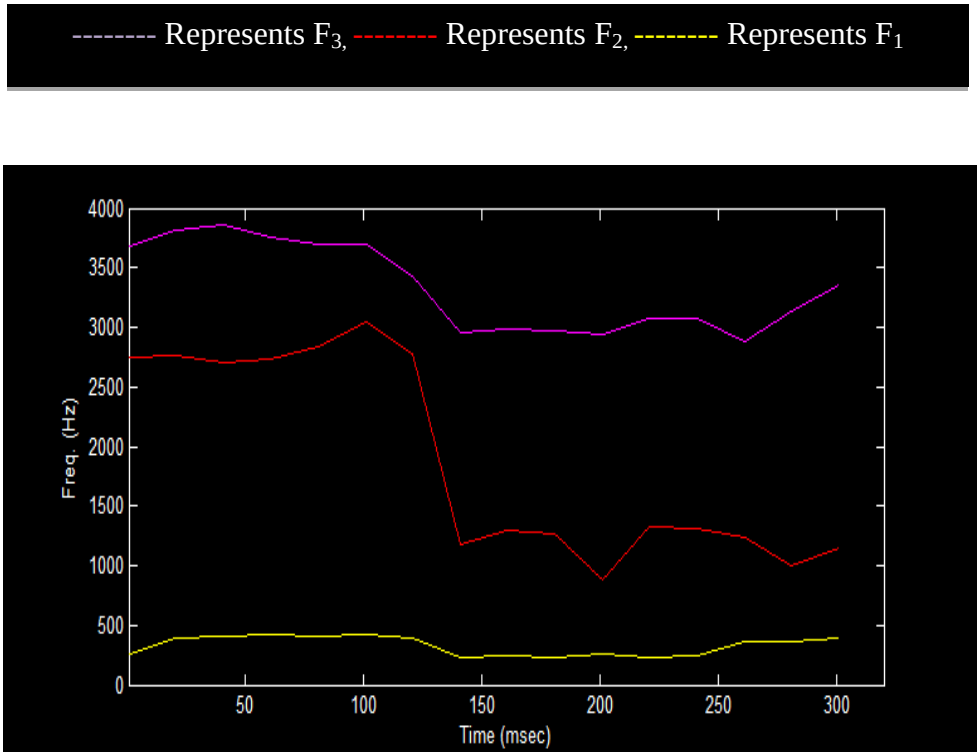


Male

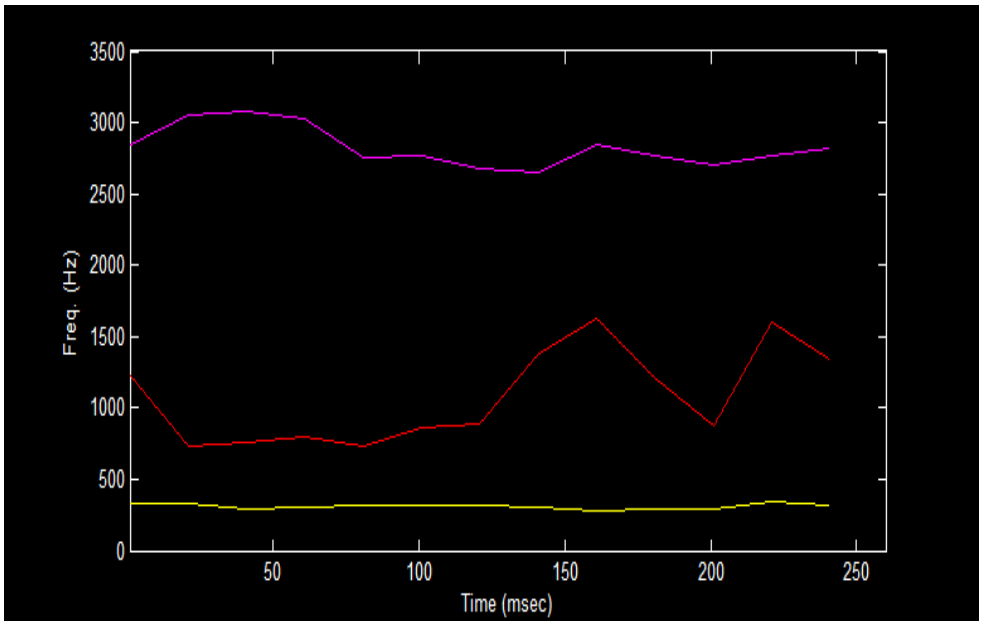


Female

Fig 5.14: Formant frequency (F1, F2 and F3) of VC type word /on/ for Male and Female



Male



Female

Fig 5.15: Formant frequency (F1, F2 and F3) of VC type word /un/ for Male and Female

5.4 Variation in Formant frequencies for Bodo Words in different Emotions

The variation in Formant Frequencies of the different type of Bodo Words CV (Consonant-Vowel), VC (Vowel- Consonant) and CVC (Consonant-Vowel- Consonant) for the three emotional moods under consideration, i.e Angry, Surprise and Normal moods is shown below:

Table 5.9 Variations in formant frequencies (F1, F2and F3) of CV type words in different emotional moods for male informant

Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
/ Saa / Angry mode	Max	374.2707	379.8961	363.2687
	Min	142.8642	176.7255	302.8882
/ Saa / Surprise mode	Max	400.2092	387.8572	383.2755
	Min	135.8634	165.7288	312.8712
/ Saa / Normal mode	Max	263.0183	356.8772	344.2611
	Min	125.7684	148.7257	332.8711

Table 5.10 Variations in formant frequencies (F1, F2 and F3) of CV type words in different emotional moods for female informant

Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
/ Saa / Angry	Max	568.627	267.6911	387.1418
	Min	370.10	289.3786	374.2270
/ Saa / Surprise	Max	639.614	378.6927	398.1417
	Min	371.22	179.4716	374.0272
/ Saa / Normal	Max	649.636	357.8984	387.9416
	Min	379.21	189.4773	364.1362

Table 5.11 Variations in formant frequencies (F1, F2 and F3) of VC type words in different emotional moods for male

Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
/ un / Angry	Max	126.937	341.072	399.051
	Min	30.9339	76.2229	35.6254
/ un / Surprise	Max	121.837	330.96	379.94
	Min	40.0336	78.631	32.5545
/ un / Normal	Max	112.935	324.773	368.852
	Min	37.2336	75.7226	32.3546

Table 5.12 Variations in formant frequencies (F1,F2 and F3) of VC type words in different emotional moods for female

Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
/ un / Angry	Max	948.715	285.641	379.633
	Min	167.853	180.725	223.637
/ un / Surprise	Max	849.716	279.645	359.73
	Min	327.852	183.415	125.347
/ un / Normal	Max	789.89	269.642	338.644
	Min	168.434	182.145	233.577

Table 5.13 Variations in formant frequencies (F1, F2 and F3) of CVC type words in different emotional moods for male

Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
/ K^har / Angry	Max	381.994	365.432	384.148
	Min	325.678	112.879	267.273
/ K^har / Surprise	Max	181.804	374.542	385.137
	Min	135.577	121.97	257.281
/ K^har / Normal	Max	189.89	369.832	380.773
	Min	140.67	121.878	274.36

Table 5.14 Variations in formant frequencies (F1, F2 and F3) of CVC type words in different emotional moods for female

Word	Frequency	F1(KHz)	F2(KHz)	F3(KHz)
/ K^har / Angry	Max	195.768	296.622	357.672
	Min	152.814	160.28	260.947
/ K^har / Surprise	Max	192.171	287.671	375.84
	Min	135.711	167.098	269.083
/ K^har / Normal	Max	190.248	269.631	380.74
	Min	152.817	157.095	267.948

5.5 Observation

An analytical study has been carried out for the Formant Frequencies of different structured words (CV, CVC and VC) and also the Bodo vowels. The following facts have been derived from the study.

i) From the pictorial representation of the Formant Frequencies (F1,F2 and F3), an observation is made that in case of both male and female spoken Bodo vowels the variation of F1 and upto some extend the F2, with respect to different vowel is quite distinct.

ii) F2 of /a/, /e/, /o/ and /w/ play important role (Fig 5.1, 5.2, 5.4 and 5.6) in sex determination of Bodo informants.

iii) It is also observed that the third formant frequency F3, does not play any significant role as far as Speech Recognition is concerned.

iv) In case of VC type of words, the variation of Format Frequency F2 is seen to be the highest as compared to the other structured words.

v) From all the above observations, we can come to the conclusion that the format frequency F2 is an important feature for speech recognition and synthesis.

vi) In case of CV type of word for Male, the variation in F1=400.2092, F2=387.8572 and F3=383.2755 is observed to be the highest in Surprise mood, than the other moods in the Bodo language.

vii) In case of CV type word for Male, the variation in F1=125.7684 and F2=148.7257 in Normal and F3=302.8882 in the Angry mood is observed to be the lowest as compared to the other moods in the Bodo language.

viii) In case of CV type word for female, the variation in F1=639.614, F2=378.6927 and F3=398.1417 is observed to be the highest in Surprise mood, than the other moods in the Bodo language.

ix) In case of CV type of word for Female, the variation in F1=370.10 in Angry, F2=179.4716 in Surprise and F3=364.1362 in Normal mood are observed as lowest as compared to the other moods in the Bodo language.

x) In case of VC type of word for Male, the variation in the F1=126.937, F2=341.072 and F3=399.051 are observed to be highest in Angry mood, than the other moods in the Bodo language.

xi) In case of Male VC type word, the variation in the F1=30.9339 in Angry, F2=75.7226 and F3=32.3546 in Normal moods are observed to be lowest as compared to the other moods in the Bodo language.

xii) In case of Female VC type word, the variation in the F1=167.853 in Angry, F2=180.725 in Angry and F3=125.347 in Surprise mood are observed to be lowest as compared to the other moods in the Bodo language.

xiv) In case of CVC type word for Male, the variation in the F1=381.994 in Angry mood, F2=374.542 and F3=385.137 in Surprise mood are observed to be highest as compared to the other moods in the Bodo language.

xv) In case of CVC type word for Male, the variation in the F1=135.577 and F3=257.281 in Surprise mood and F2=112.879 in Angry mood are observed to be lowest as compared to the other moods in the Bodo language.

xvi) In case of CVC type word for Female, the variation in the F1=195.768 and F2=296.622 in Angry mood and F3=380.74 in Normal mood are observed to be highest as compared to the other moods in the Bodo language.

xvii) In case of CVC type word for Female, the variation in the F1=135.711 in Surprise mood, F2=157.095 in Normal Angry mood and F3=260.947 in Angry mood are observed to be lowest as compared to the other moods in the Bodo language.

5.6 Conclusions

From the present study, it can be concluded that the Formant Frequency values - F1, F2 and F3 for the different Bodo words of VC, CV and CVC type can play a very important role in the determination of the type of emotional content by a Female or Male in the Bodo Language. The physical significance of the frequencies F1, F2 and F3 for the three different type of emotions are studied from the results obtained so far from the analysis of different Bodo words of CV, VC, and CVC types. The determination of three formant frequencies- F1, F2, F3 gives the description and correlation of different type of vocal tract characteristics in all the three emotion types under consideration. Moreover, the formant frequencies so obtained were compared and evaluated statistically. In this study, the speech used has been collected in the form of Bodo words and sentences which has been spoken by both male and female informants in different emotional moods like angry, surprise and normal. The analysis of the formant frequencies gives a good discrimination to study the comparison between neutral and emotional modes for both male and female voices. The spectral parameters, thus obtained together with the other characteristics of speech, are used to build the feature vector set. This vector set is used to build the emotional speech classifier