

COMPARISON OF MALE AND FEMALE VOWEL FORMANTS BY NATIVE SPEAKERS OF MARATHI LANGUAGE

D. K. Shinde*, Y.N. More, A. S. Sonawane, V. S. Pawade, Dr. K. V. Kulkarni

Directorate of Forensic Science Laboratories, Kalina Santacruz (E), Mumbai 400098, Home Dept.

Govt. of Maharashtra

E-mail: *dipakshinde888@gmail.com

Abstract: -Now a day's many studies are being conducted on speaker identification between male and female voice. In our present study, we compared and found formant frequencies of male and female voice samples in a number of 12 Marathi vowels produced by 10 Marathi native speakers (Male and Female). Day by day voice related crimes are increasing rapidly across the world, e.g. Bribery, Ransom call, threatening and terrorist activities cases, etc. Hence there are lots of applications available on voice and speaker identification and as we know that, voice's uniqueness in the resonant frequencies proves that voice is unique. So we focused on Marathi native speakers (Male and Female), the voice comparison is made on the basis of resonant frequencies, Linear Predictive Coding (LPC) and spectrographic analysis using Multi-Speech application. This technique uses Linear Predictive Coding (LPC) and we compared formant frequencies of Marathi native speakers (Male and Female).

Keywords: -Marathi native speakers, vowels, gender, voice analysis, formant frequency measurement, sound and pitch.

Introduction: -Many research papers and studies on voice comparison (Male and Female) have been done in the field of major global language for example English, Spanish and French etc. But a lot is still to be explored in native languages that are inherited in our country. With a view to study and learn more about Marathi language vowels, we intended to conduct a research and prepare a paper on the same. Since DFSL is located in Mumbai, Maharashtra and receive the cases in native language i.e. Marathi, this study would be helpful.

When audio cases are received, the disciplined involved while solving those voice cases are not confined to one particular scientific field solely i.e. ^[1] not only do we consider the fundamental frequencies and resonant frequencies of voice but also take into consideration the state of mind involved, behavioral pattern portrayed while committing a crime i.e. acting in his/her natural senses etc. and compare this natural behavioral pattern with learned behavioral pattern. The paper intends to study the issues related to the above clauses.

As per studies that were being conducted till now did not gave us particular frequency values for male and female in Marathi native language. Our motto is to conduct a research and work on the same so as to come out with certain set of frequency values that could be certainly considered as standard values.

Participants: -Five male and five female native speakers of Maharashtra were well versed in Marathi language and Marathi was their first language, took part in this study. All were university students with normal hearing capability.

Table 1:- Marking of Female and Male voice samples.

<i>Sr. No.</i>	<i>Female voice Samples</i>	<i>Marked</i>	<i>Male voice samples</i>	<i>Marked</i>
1	Marathi Female Voice Sample1	MFVS1	Marathi Male Voice Sample1	MMVS1
2	Marathi Female Voice Sample2	MFVS2	Marathi Male Voice Sample2	MMVS2
3	Marathi Female Voice Sample3	MFVS3	Marathi Male Voice Sample3	MMVS3
4	Marathi Female Voice Sample4	MFVS4	Marathi Male Voice Sample4	MMVS4
5	Marathi Female Voice Sample5	MFVS5	Marathi Male Voice Sample5	MMVS5

Recorded samples: -12 Marathi vowels ^[5] (a/ə/, ā/a/, I/i/, Ī/i:/, U/u/, Ū/u:/, e/e/, ai/əi/, o/o/, au/əu/, aṃ/əṃ/, aḥ/əḥə/) were recorded from the participants (5 male and 5 female speakers).

Devices used: -To record the samples we used the following devices:

- I. Digital Voice Recorder named – Sony company, sampling rate is 44100Hz.
- II. Laptop named –Lenovo Company (G-50-80).
- III. Mobile phone named - One plus one Company.

From the above recording devices, we considered the samples recorded from Digital Voice Recorder for analysis purpose.

Procedure: -^[2]To go ahead with the approach, we called our participants to record the sample voice. To do so, we needed an ideal room so as to cut down the external disturbances interfering the recording procedure i.e.^[1]noise from AC & fans, other conversations, vehicle horn, ticking of watch etc. We did this so that the natural quality of the voice recorded remains intact and to get precise results. Hence we chose to record the sample voice in a soundproof room.

Every speaker was instructed to repeat one vowel thrice and in all we considered 12 such vowels for analysis. These samples were recorded using mobile, laptop and digital voice recorder. Of all the recorded samples, the sample recorded from digital voice recorder was considered for our further analysis.

Basic information: - ^[2]Most of the sound generated by the human voice organ can be understood by the following human. Actually a spoken word has more meaning than its explanation in the dictionary. We produce sound to express emotions, mental imagery, and language. Human voice is the most natural form of communication involved with environmental and situational reality. The formant frequencies of female are higher than that of male. Voice tract of the female is shorter than that of male, which causes sharp and clear voice generation in females.

^[4]The formant is term applied to describe the resonant frequencies of the voice tract. The formants are frequency information characteristics of the voice. The formants change the Centre frequency and bandwidth during speech. Adult men and women have different vocal folds sizes; reflecting the male – female differences in larynx size. Adult male voices are usually lower pitched and have larger folds. The male vocal folds are between 17mm and 25mm in length. The female vocal folds are between 12.5mm and 17.5mm in length. Male and female voices differ in the fundamental frequency mainly due to length of the vocal voice tract because formant frequencies are inversely proportional to the length of the voice tract.^[1] Individual speakers differ from larynx mechanism and the vocal tract, but it has been seen that differences in the vocal fold vibration play a major role in enabling listeners to identify the individual voice.

Analysis:-Words were first extracted from the frame sentence. Since the sample voice was recorded thrice, only the most acoustically satisfactory occurrence was selected, making up a total of 360 samples (10 speakers, 12 vowels). We then segregated and labeled voice samples in phones. This task was performed with the help of Gold wave application.

The file format of the sample was converted from .mp3 format to .wav (PCM sine 16-bit mono) format with the help of Gold wave software. Our study revolves around comparative study of 12 vowels between male and female Marathi native speakers (a/ə/, ā/a/, I/i/, Ī/i:/, U/u/, Ū/u:/, e/e/, ai/əi/, o/o/, au/əu/, aṃ/əṃ/, aḥ/əḥə/). We used Multi-Speech application for the analysis purpose. We opened the respective file and ran it in Multi-Speech one at a time. We then down sampled the sample voice at 11025Hz, to make digital audio signal smaller by lowering its sampling rate or sample size (bits per sample). Down sampling is done to decrease the bit when transmitting over a limited bandwidth or to convert to a more limited audio format. Further we conducted spectrograph test for

detailed analysis. [3]The spectrographic test is divided into three parts: 1) Amplitude vs. Time analysis (waveform analysis), 2) Frequency vs. Time analysis (formant analysis),3) Energy vs. Frequency .Once this done, we focused on formant frequencies (F1 & F2) of male and female samples with the help of LPC technique. We did this because of the role it plays in the recognition and differentiation of speech sounds.

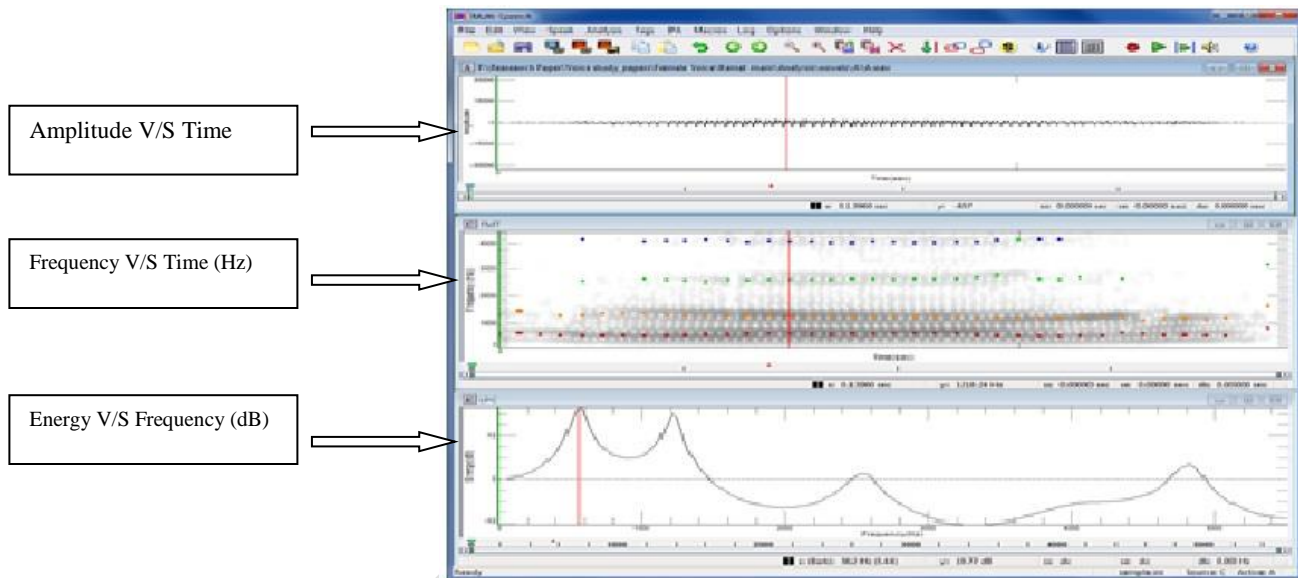


Figure: -1: Spectrographic result of one of the native speaker.

[4]It must be stressed that the formant pattern of particular sound is the outcome of the acoustic character of the whole tract working as one resonant system. Hence is not justifiable to assign any one formant to a particular part of the vocal tract. The frequency F1 and F2 are interdependent, since in general the lengthening of one section of the tract implies the shortening of the other. It is true, however, that the vertical section is longer than the horizontal and is therefore responsible for the wavelength of the lowest formant, F1 while the shorter section tends to determine F2. This results in an F1 of comparatively low frequency together with an F2 of high frequency.

The following observations are based on 5 female native speakers that were our participants. Formant frequency (F1)values are obtained after a detailed analysis of the sample voice taken. After obtaining these values for every vowel, we have taken out their mean values for the same vowel.

Table 2:-F1 Values of MFVS1 to MFVS5.

Sr. No	Name	Marathi Vowels											
		a /ə/	Ā /a/	I /i/	Ī /i:/	U /u/	ū /u:/	e /e/	ai /əi/	o /o/	au /əu/	aṃ /əṃ/	aḥ /əḥə/
1	MFVS1	604.0 2	936.33	404.33	378	435.17	436.81	431.22	431.13	438.4	440.52	651.67	787.11
2	MFVS2	563.5 1	964.54	417.51	426.77	433.88	441.24	435.43	422.24	470.15	564.74	509.51	602.44
3	MFVS3	751.2 2	991.66	478.86	464.07	370.28	400.95	546.09	521.92	524.89	807.86	794.95	995.77
4	MFVS4	639	992.62	321.35	362.51	321.97	346.09	524.81	740.99	535.76	485.89	768.15	813.43
5	MFVS5	539.8 1	930.88	299.14	350.13	453.01	355.29	527.91	502.77	495.25	558.8	720.32	894.79
Total		3097. 5	4816.0 3	1921.19	1981.4 8	2014.3 1	1980.3 8	2465.4 6	2619.0 5	2464.4 5	2857.8 1	3444.6	4093.5
Average		619.5 1	963.20	384.238	396.29	402.86	396.07	493.09	523.81	492.89	571.56	688.92	818.70

The following observations are based on 5 female native speakers that were our participants. Formant frequency (F2) values are obtained after a detailed analysis of the sample voice taken. After obtaining these values for every vowel, we have taken out their mean values for the same vowel.

Table 3:-F2 Values of MFVS1 to MFVS5.

Sr. No	Name	Marathi Vowels											
		A /ə/	ā /a/	I /i/	Ī /i:/	U /u/	Ū /u:/	e /e/	ai /əi/	o /o/	au /əu/	aṃ /əṃ/	aḥ /əḥə/
1	FVS1	1321.1 2	1491.2 4	1485.59	1564.8 7	701.71	701.71	2799.5 3	2800.2 3	859.34	861.54	1284.3 8	1471.1 7
2	FVS2	1218.9 2	1462.9 6	2588.38	2660.3 5	899.87	899.87	2535.0 2	2522.1 1	919.8	1008.7 7	1229.3 7	1078.0 3
3	FVS3	1306.0 8	1483.2 4	2950.87	2940.0 5	1043.8 1	1043.8	2727.4 9	2521.0 6	1067.0 4	1264.8	1311.4 9	1409.6 5
4	FVS4	1300.1 3	1514.8 4	2881.65	2885.5 3	805.21	805.21	2173.6 1	1742.8 2	990.72	1013.3 5	1249.0 7	1365.8 1
5	FVS5	1261.0 9	1483.8	3014.66	2812.0 6	884.02	884.02	2570.7 9	2823.6 3	967.84	974.6	1245.8 1	1543.4 1
Total		6407.3 4	7436.0 8	12921.1 5	12862. 8	4334.6 2	4334.6	12806. 4	12409. 8	4804.7 4	5123.0 6	6320.1 2	6868.0 7
Average		1281.4 6	1487.2 1	2584.23	2572.5 7	866.92 4	866.92	2561.2 8	2481.9 7	960.94 8	1024.6 1	1264.0 2	1373.6 1

The following observations are based on 5 male native speakers that were our participants. Formant frequency (F1) values are obtained after a detailed analysis of the sample voice taken. After obtaining these values for every vowel, we have taken out their mean values for the same vowel.

Table 4:-F1 Values of MMVS1 to MMVS5.

Sr. No	Name	Marathi Vowels											
		A /ə/	ā /a/	I /i/	Ī /i:/	U /u/	Ū /u:/	e /e/	ai /əi/	o /o/	au /əu/	aṃ /əṃ/	aḥ /əḥə/
1	MVS1	663.13 2	1001.4 2	423.81	451.58	451.06	459.14	485.17	611.01	498.94	638	780.53	788.11
2	MVS2	550.15	627.62	348.54	302.32	352.41	363.11	416.27	467.12	428.96	484.15	626.19	753.3
3	MVS3	550.57	767.98	301.99	275.7	328.31	335.65	470.67	455.45	505.9	498.99	575.78	665.98
4	MVS4	539.74	750.77	330.65	329.18	361.96	379.31	432.95	351.32	491.8	478.23	592.73	686.88
5	MVS5	492.2	841.37	333.36	333.31	369.38	369.62	385.89	402.42	545.14	520.92	412.25	778.75
Total		2795.7 9	3989.1 6	1738.35	1692.0 9	1863.1 2	1906.8 3	1774.6 8	2287.3 2	2470.7 4	2620.2 9	2987.4 8	3673.0 2
Average		559.15 8	797.83 2	347.67	338.41 8	372.62 4	381.36 6	443.67	457.46 4	494.14 8	524.05 8	597.49 6	734.60 4

The following observations are based on 5 male native speakers that were our participants. Formant frequency (F2) values are obtained after a detailed analysis of the sample voice taken. After obtaining these values for every vowel, we have taken out their mean values for the same vowel.

Table 5:-F2 values of MMVS1 to MMVS5

Sr. No	Name	Marathi Vowels											
		A /ə/	ā /a/	I /i/	Ī /i:/	U /u/	Ū /u:/	e /e/	ai /əi/	o /o/	au /əu/	aṃ /əṃ/	aḥ /əḥə/
1	MVS1	1346.5 3	1456.8 7	2448.22	2250.6 5	731.15	752.62	2275.5 1	2227.8 5	814.67	1081.4 5	1237.6 3	1314.8 3
2	MVS2	1137.8 1	1224.3 6	2150.73	2146.6 8	673.69	669.58	2044.1 1	1753.8 9	758.92	846.05	1130.8 7	1271.0 3
3	MVS3	1283.1 1	1287.0 4	2399.54	2385.4 5	743.71	765.43	2137.9 7	2067.5 4	798.02	996.05	1276.7	1308.1 5

4	MVS4	1125.7 5	1146.3 7	2314.88	2351.4 5	659.93	681.08	2052	2141.5 6	818.99	826.49	1275.1	1146.1 9
5	MVS5	1294.7 3	1308.7 8	2719.25	1479.3 3	792.45	788.92	2353.7 1	2414.7 4	1202.3 1	1158.8 1	1118.5 7	1367.4
Total		6187.9 3	6423.4 2	12032.6 2	10613. 5	3600.9 3	3657.6 3	10863. 3	10605. 5	4392.9 1	4908.8 5	6038.8 7	6407.6
Average		1237.5 8	1284.6 8	2406.52 4	2122.7 1	720.18 6	731.52 6	2172.6 6	2121.1 1	878.58 2	981.77	1207.7 7	1281.5 2

Observation: -The comparative study of male and female speakers is tabulated as follows:

The mean values of F1 and F2 for 5 female native speakers are as follows:

Table 6:- F1 and F2 Values of MFVS1 to MFVS5.

Sr. No	F.F.	Marathi Vowels											
		A /ə/	ā /a/	I /i/	Ī /i:/	U /u/	Ū /u:/	e /e/	ai /əi/	o /o/	au /əu/	aṃ /əṃ/	aḥ /əḥə/
1	F1	619.51 8	963.20 2	384.238	396.29 8	402.86 4	396.07 6	493.09	523.81 4	492.89 8	571.56 8	688.92 6	818.708
2	F2	1281.4 6	1487.2 1	2584.23	2572.5 7	866.92 4	866.92	2561.2 8	2481.9 7	960.94 8	1024.6 1	1264.0 2	1373.61 4

Note: - Formant Frequencies (F.F.)

The mean values of F1 and F2 for 5 male native speakers are as follows:

Table 7:-F1 and F2 Values of MMVS1 to MMVS5.

Sr. No	F.F.	Marathi Vowels											
		A /ə/	ā /a/	I /i/	Ī /i:/	U /u/	Ū /u:/	e /e/	ai /əi/	o /o/	au /əu/	aṃ /əṃ/	aḥ /əḥə/
1	F1	559.15 8	797.83 2	347.67	338.41 8	372.62 4	381.36 6	443.67	457.46 4	494.14 8	524.05 8	597.49 6	734.604
2	F2	1237.5 8	1284.6 8	2406.52 4	2122.7 1	720.18 6	731.52 6	2172.6 6	2121.1 1	878.58 2	981.77	1207.7 7	1281.52

Note: - Formant Frequencies (F.F.)

The above table: - 6 and 7 gives us mean values of F1 and F2 formants for pure vowels of Marathi based on data obtained from samples of Marathi native speakers.

Result and discussion: - The formant frequencies of all the native speakers have been observed.

Fundamental frequency is measure of how high or low the frequency of voice sound is of the person. A person's voice typically varies over a range of fundamental frequencies. This acoustic analysis has given interesting results concentrated on fundamental frequencies and harmonic frequencies. We obtained the mean of F1 and F2 formant frequencies. We significantly saw that the formant frequencies in Marathi vowels are higher for female than male.

Comparison of mean values of formant frequencies (F1 and F2) of female and male native speakers.

Table 8:-F1 and F2 Values Male and Female voice samples.

Sr. No	F.F	Marathi Vowels											
		a /ə/	ā /a/	I /i/	Ī /i:/	U /u/	ū /u:/	e /e/	ai /əi/	o /o/	au /əu/	aṃ /əṃ/	aḥ /əḥə/
1	Femal eF1	619.51	963.20	384.238	396.29	402.86	396.07	493.09	523.81	492.89	571.56	688.92	818.70
2	Male F1	559.158	797.83 2	347.67	338.41 8	372.62 4	381.36 6	443.67	457.46 4	494.14 8	524.05 8	597.49 6	734.60
3	Femal e F2	1281.46	1487.2 1	2584.23	2572.5 7	866.92 4	866.92	2561.2 8	2481.9 7	960.94 8	1024.6 1	1264.0 2	1373.6

4	Male	1237.58	1284.6	2406.52	2122.7	720.18	731.52	2172.6	2121.1	878.58	981.77	1207.7	1281.5
	F2		8	4	1	6	6	6	1	2		7	

Note: - Formant Frequencies (F.F.)

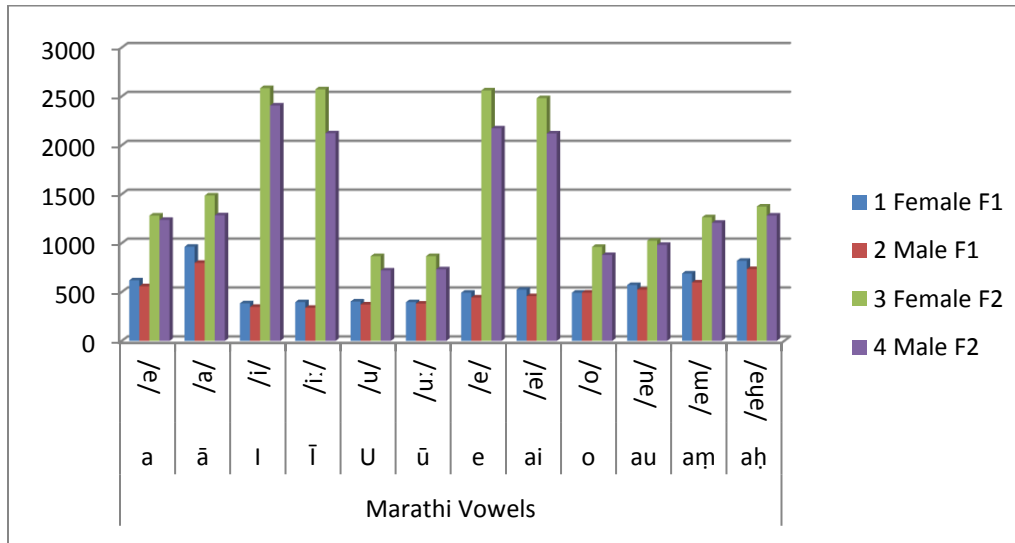


Figure 2: All mean values of Formant Frequencies (F1 and F2) of Female (F) and Male (M) Marathi native speakers

Conclusion: -Thus in this study we compared the formant frequencies of female and male speakers using Multi-Speech application. It is found that the formant frequencies of female sample voice is higher than male sample voice for certain vowels (a/a/, ā/a/, I/i/, Ī/i:/, U/u/, Ū/u:/, e/e/, ai/ai/,am/am/, aḥ/aḥ) but for certain vowels (o/o/, au/au/)formant frequencies (F1 and F2) are found to be overlapping. Also from this study we tried to obtain standard values for F1 and F2 for native speakers for Marathi language.

This study would prove as a useful reference and a resource for future applications related to voice analysis.

Acknowledgement: -We are thankful to the authors and all the participants of the Directorate of Forensic Science Laboratories, Mumbai, Maharashtra state for their support for providing the research environment and tools required for voice analysis.

REFERENCES:

[1] D.B. Fry, "The Physics of Speech" CAMBRIDGE UNIVERSITY PRESS.
 [2] Harry Hollien, "Forensic Voice Identification" Institute for Advance Study of the communication process University of Florida, USA.
 [3] Lee, Chine- hui, Soong, Frank k, Paliwal, K.K. (Kuldip K.) "Automatic Speech And Speaker Recognition" IV Series, USA 1996.
 [4] Keshet, Joseph, Bengio, Samy "Automatic Speech And Speaker Recognition" UK 2009.
 [5] Balasaheb Shinde, "Marathi Language" Maharashtra, India.