Supra Segmental features profile of Gamthi & Kathiyawadi dialects of Gujarati Language

Ms. Pooja Ahuja*, Dr. J. M. Vyas**

*(Faculty, Institute of Forensic Science, Gujarat Forensic Sciences University, India)
** (Director General, Gujarat Forensic Sciences University, India)

ABSTRACT: *In this present paper, Speech Samples of two major dialects of gujarati language i.e. Gamthi Gujarati (includes Surti, Charotari, Machi, Gramaya) & Kathiyawadi (includes Jhalawadi, Mer, Bhavanagari, Sorathi) were obtained. The speech includes 'Hello' as the test sample. The word 'Hello' used most frequently in conversations is a CV-monosyllable. The speech word was analyzed on supra segmental features viz. Stress, Length of segment, Vowel duration, VOT, Vowel consonant duration ratio, Consonant closure time, Fo frequency and formant frequencies. Significant results obtained for creating a reliable data profile for the language.*

Keywords – Dialects, Hello, Monosyllable, Supra segmental features, Vowel.

1. INTRODUCTION

Voice is a generally willful act of living beings, an act of verbal interaction between two or mass of living beings where in to and fro conversations of audible sounds[1]. Processes involving dynamics of sound: respiration, phonation, resonation and articulation leading to voiceless and voiced segments of speech [2].

In consideration of forensics, it related to verification of the human voices associated with personnel's in conversation. Speech analysis technique or voice comparison is typically a comparison of a person in general to prove his/her presence or involvement at an instance when the voice sample is being recorded [3]. Usually voice recordings have suspect and offender voice samples which are to be identified to fix the authenticity of voice.

Forensic Audio recordings for comparison traditionally are tape recordings, where in a forensic expert lends to analysis of the recordings to check its authenticity and reliability. In the present study, the aim is to advance the present analysis techniques by elaborating the search into phonemes which are basic units of voice. Phoneme strength can be analyzed basing on voicing, place of articulation and manner of production which further leads to spectrographically calculation of supra segmental features harmoniously a production of vowel and consonants. Indo- Aryan languages which include Gujarati language are descended from Sanskrit, depend variably on vowel production. During production, the lips, the tongue and velum altogether leads to vocal cord articulation. Determination and productivity of vowel depends on position of tongue, degree of stress/loudness depends on tongue and upper jaw spacing [3]. This pulsing are further modulated for voicing sounds. Aural graph of flow of vocal tract is plotted as X and Y-axis, pulsing time and amount of loudness in speech. Voice-graph is presentation of vowels and consonants in wave pattern plotted as frequency of sound over time [4] [5].

In the present paper, the acoustical and segmental features of two strained dialects (Gamthi & Kathiyawadi) of Gujarati language, corollaries where made on phoneme production in voicing, identifying the Consonant & Vowel ratio, Vowel Duration, Consonant closure time which leads to length of segment study [6]. Particular focus was to create a statically profile of voice of the fore mentioned dialects to identify the native place of the sample voice through a comparison of fundamental (F_0) and formant frequencies (F_1, F_2, F_3, F_4) data of the monosyllable elements selected [7] [8] [9].

2. SAMPLING & METHODOLOGY

2.1. Sampling

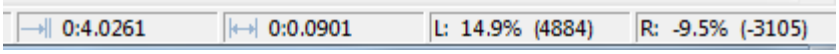
Speech samples of 20 subjects, Gujarat as their native place, age between 20-35 years, with at least higher secondary educational background were taken. 20 (10 male, 10 female) speech samples for Gamthi dialect & 20 (10 male, 10 female) speech samples for kathiyawadi dialect depending upon the region. The subject's speech sample selected for present study includes personals that are purely vegetarian, non-smoking without any ill-habits. It was noted that none of subject suffered from speech deformities. The samples were collected in a sound proof room on a digital recorder with 128 kbps recording speech & 16 bit mono recorder.

2.2. Materials & Procedure

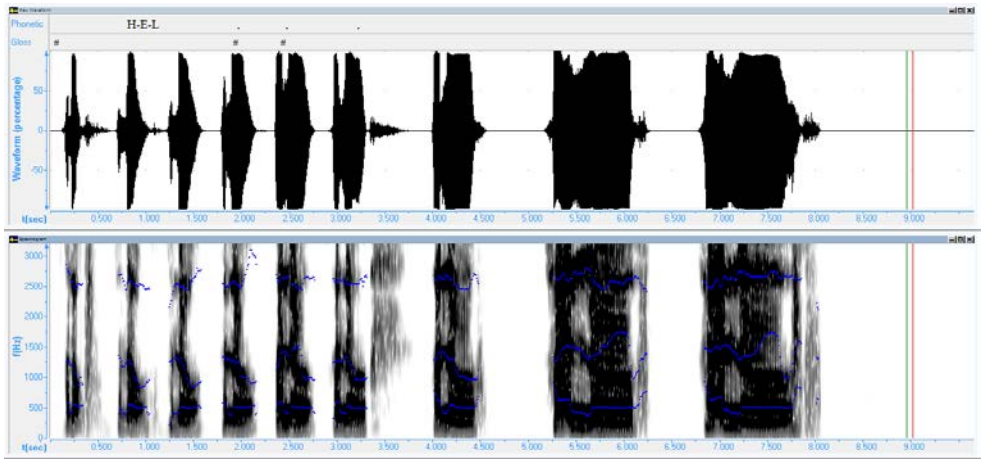
For the said study, the word 'Hello' was overtaken as a sample word. As the word is the preparatory of all exchange or is used as greeting address word. Two basic analysis were carried out from the test word 'Hello' i.e. CVC- Monosyllable analysis for the segment 'H-E-L' and Vowel analysis for Closed Vowel /E/ appears in the segment 'H-E-L' & CV analysis for the segment 'L-O and Vowel analysis for open vowel /O. The prompts for the word 'HELLO' were grouped which contained repetitions of word "HELLO'. For all the prompts collected, a measure was taken to keep the speech data at medial speech. During the recording session, subjects were asked to play with the target word 'HELLO" so that the alterations & fluctuations of voicing can be studied [10].

IJSER

3. ANALYSIS

The response speech samples recorded with a digital recorder which were lately transferred to the desktop system with USB-flash drive already inbuilt part of Digital recorder. The recordings by default were saved as WAV files at 48 KHz & 16 mono bytes. It was taken care of that speech samples were free from distortions & noise for which arrangements were made to record speech samples in sound proof room, integrated part of voice department. The samples converted to digital were marked a word segments and then phonetic segments were determined, the segments thus marked constitutes repeated segment of words. The window of the screen shot image of the analysis of word hello is shown below. The first window is the waveform of the segmented words, detailed zoom in and zoom out study minutely done over. Second window is the spectrogram of speech signals determines the formant frequencies, blue dots seen on spectrogram are splits of formant tracks. Over all voicing with split vowel onset and vowel duration is determined by placing the blue and red line on the onset and end of the phonetic vowel and lately is marked over. The value is flashed on the downside as shown 

on the lower toolbar of the spectrogram analysis [11].



4. RESULTS

Speech samples of male and female subjects analyzed over vowel /E/ –CVC syllabi and vowel /O – CV syllabi of the gamthi and kathiyawadi dialects is shown in the table below [12]. Data cited in the table are the average data of repeated segment of target word so that accuracy is maintained.

Table-I: Average of Duration and formant frequencies of the repeated segments 20 subjects of Gamthi dialect of Vowel /**E/**

IJSER

Sample No.	Sex	Vowel Onset Time in CVC syllable (in secs)	Vowel Duration (in secs)	Syllable length –CVC (in secs)	F1 (Hz)	F2 (Hz)	F3 (Hz)
G-1	M	0.1385	0.0918	0.4073	493.7	992.4	2424.7
G-2	M	0.1402	0.0874	0.3980	411.8	1201.4	2611.8
G-3	M	0.1466	0.0963	0.4163	509.2	1353.4	2666.5
G-4	M	0.1407	0.0862	0.4006	410.5	1386.7	2651.8
G-5	M	0.1398	0.0965	0.4674	465.8	1193.6	2687.4
G-6	M	0.1470	0.0900	0.3996	484.1	986.9	2557.3
G-7	M	0.1448	0.0913	0.4179	503.4	1117.7	2461.9
G-8	M	0.1376	0.0875	0.4080	468.8	1534.2	2659.0
G-9	M	0.1419	0.0927	0.4077	511.8	1368.2	2544.2
G-10	M	0.1475	0.0901	0.4026	525.6	1210.9	2513.7
G-11	F	0.1571	0.0999	0.3718	617.9	1835.9	2990.6
G-12	F	0.1662	0.1007	0.3896	609.5	1887.4	2684.3
G-13	F	0.1632	0.1075	0.3577	611.5	1833.2	2668.0

G-14	F	0.1599	0.0987	0.3553	603.5	1779.5	2876.4
G-15	F	0.1603	0.1021	0.3930	613.2	1837.4	2639.5
G-16	F	0.1598	0.0977	0.4052	643.2	1889.2	2744.1
G-17	F	0.1687	0.1056	0.3994	632.7	1853.7	2835.1
G-18	F	0.1618	0.0988	0.3990	630.1	1880.4	2690.2
G-19	F	0.1574	0.1072	0.4068	611.4	1845.2	2880.4
G-20	F	0.1493	0.0982	0.3885	625.3	1792.5	2773.6

Table-II: Average of Duration and formant frequencies of the repeated segments of 20 subjects of Gamthi dialect of Vowel /**O**

Sample No.	**Sex**	**Vowel Onset Time in CV syllable (in secs)**	**Vowel Duration (in secs)**	**Syllable length –CV (in secs)**	**F1 (Hz)**	**F2 (Hz)**	**F3 (Hz)**
G-1	M	0.1753	0.1610	0.3716	489.3	1273.4	2719.6
G-2	M	0.1703	0.1974	0.4067	491.9	1497.7	2583.7
G-3	M	0.1644	0.1606	0.3996	488.5	1242.9	2736.2
G-4	M	0.1703	0.1841	0.3861	470.5	1386.7	2651.8
G-5	M	0.1798	0.1947	0.4173	465.8	1493.6	2780.4
G-6	M	0.1644	0.1844	0.3922	484.1	1430.9	2557.3
G-7	M	0.1773	0.1907	0.4107	449.8	1274.8	2566.1
G-8	M	0.1637	0.1882	0.3992	490.8	1467.2	2706.3
G-9	M	0.1785	0.1690	0.3539	472.6	1509.3	2599.3
G-10	M	0.1608	0.1979	0.4029	495.2	1294.6	2770.8
G-11	F	0.1665	0.1705	0.3916	506.4	1807.3	2808.3
G-12	F	0.1799	0.1873	0.3832	519.3	1791.6	2905.7
G-13	F	0.1704	0.1892	0.3799	498.7	1798.5	2894.1
G-14	F	0.1691	0.1883	0.3819	520.6	1832.9	2876.4

IJSER

G-15	F	0.1701	0.1806	0.3783	536.4	1850.3	2893.5
G-16	F	0.1716	0.1842	0.3949	554.2	1864.2	2811.7
G-17	F	0.1653	0.1899	0.4118	498.4	1775.4	2907.1
G-18	F	0.1699	0.1902	0.4099	528.2	1711.8	2795.5
G-19	F	0.1701	0.1774	0.3993	544.7	1782.5	2799.3
G-20	F	0.1612	0.1780	0.3651	500.9	1694.3	2828.6

Table-III: Average of Duration and formant frequencies of the repeated segment of 20 subjects of Kathiyawadi dialect of Vowel /**E/**

Sample No.	**Sex**	**Vowel Onset Time in CVC syllable (in secs)**	**Vowel Duration (in secs)**	**Syllable length –CVC (in secs)**	**F1 (Hz)**	**F2 (Hz)**	**F3 (Hz)**
K-1	M	0.1826	0.1427	0.4705	552.3	1490.8	2452.1
K-2	M	0.1788	0.1393	0.4890	589.0	1139.7	2598.3
K-3	M	0.1803	0.1563	0.4261	522.2	1211.9	2749.9
K-4	M	0.1773	0.1432	0.4109	506.2	1414.0	2463.1
K-5	M	0.1697	0.1455	0.3999	591.2	1429.3	2482.2
K-6	M	0.1744	0.1369	0.4276	497.7	1490.3	2826.6
K-7	M	0.1791	0.1479	0.4130	574.3	1222.7	2256.2
K-8	M	0.1800	0.1581	0.3770	526.9	1411.9	2749.9
K-9	M	0.1699	0.1442	0.3905	460.2	1317.4	2435.9
K-10	M	0.1791	0.1377	0.3768	441.5	1324.4	2409.7
K-11	F	0.1871	0.1690	0.5361	689.4	1841.6	2694.2
K-12	F	0.1895	0.1596	0.5501	620.5	1476.6	2489.7
K-13	F	0.1904	0.1802	0.5807	525.6	1744.4	2499.4

IJSER

K-14	F	0.1822	0.1724	0.4708	570.6	1630.0	2662.7
K-15	F	0.1815	0.1655	0.4980	621.2	1838.3	2919.4
K-16	F	0.1908	0.1864	0.5106	547.3	1668.3	2527.6
K-17	F	0.1944	0.1756	0.4342	577.6	1699.9	2582.0
K-18	F	0.1863	0.1730	0.4655	638.4	1803.5	2730.4
K-19	F	0.1832	0.1504	0.4122	595.0	1796.4	2367.1
K-20	F	0.1575	0.1733	0.4009	642.7	1877.2	2630.7

Table-IV: Average of Duration and formant frequencies of the repeated segment of 20 subjects of Kathiyawadi dialect of Vowel /**O.**

IJSER

Sample No.	**Sex**	**Vowel Onset Time in CV syllable (in secs)**	**Vowel Duration (in secs)**	**Syllable length –CV (in secs)**	**F1 (Hz)**	**F2 (Hz)**	**F3 (Hz)**
K-1	M	0.1834	0.2130	0.3716	601.5	1026.1	2563.8
K-2	M	0.1816	0.2506	0.4561	578.3	982.5	2500.1
K-3	M	0.1906	0.2116	0.3922	596.1	1023.8	2584.6
K-4	M	0.1911	0.2477	0.4821	595.5	1005.9	2402.5
K-5	M	0.1835	0.2511	0.4923	571.2	960.2	2497.6
K-6	M	0.1962	0.3070	0.5147	612.9	1106.7	2023.9
K-7	M	0.1894	0.2419	0.4973	586.6	1163.9	2121.8
K-8	M	0.1905	0.2671	0.4298	603.7	1031.9	2566.7
K-9	M	0.1942	0.2840	0.3917	586.7	1182.6	2446.5
K-10	M	0.1899	0.2509	0.3994	578.9	978.6	2491.6
K-11	F	0.2061	0.2607	0.4905	593.2	1158.9	2682.5
K-12	F	0.1987	0.2615	0.4792	619.4	1205.6	2611.7

K-13	F	0.2009	0.2593	0.4674	604.0	1334.5	2612.9
K-14	F	0.1904	0.2640	0.4719	628.2	1273.2	2664.3
K-15	F	0.1972	0.2493	0.4439	611.2	1393.4	2678.0
K-16	F	0.2103	0.2342	0.4507	590.7	1105.4	2570.9
K-17	F	0.2065	0.2516	0.4475	632.3	1309.2	2616.5
K-18	F	0.1992	0.2409	0.4446	587.3	1296.4	2529.4
K-19	F	0.2019	0.2511	0.4602	653.2	1179.4	2517.5
K-20	F	0.1991	0.2337	0.4385	608.8	1239.4	2679.2

From the cited data, it is clearly shows dialect differentiation on supra segmental features, for accuracy of results the obtained results where statistically tested, it is observed that the p-value is less than 0.05. Hence the results were considered practically opt for data profiling [13] [14].

IJSER

5. CONCLUSION

The experimentally modules data on supra segmental features of vowel voicing featuring of vowel /E/ and /O of gamthi and kathiyawadi dialects depends on acoustic correlates. Significant results were obtained & statistical data in co-operated showed that two dialects showed a fixed effect difference of 0.06 seconds in vowel duration and 0.03 seconds in vowel onset time. Notable deviation was recorded on segment length. As far as frequencies are concerned no detrimental data for dialect differentiation can be pooled out as the data showed almost similar pattern of frequencies i.e. F_1: 400-600 Hz, F_2: 1200-1400 Hz and F_3: 2200-2700 in both the dialects. The results revealed that the gamthi speakers concerning to vowel /E/ & /O have smaller on burst voicing and smaller vowel duration as compared to kathiyawadi speakers. The study implies that the native dialects have a pure association with language segments. The present findings suggest that in depth analysis of dialects on the basis of segmental features can be useful to create a reliable dialect profile of the language.

6. REFERENCES

1. D. H. Klatt and L. C. Klatt. “Analysis, Synthesis, and Perception of Voice Quality Variations among Female and Male Talkers”, J. Acoust. Soc. Am. 87(2), pp. 820-857, February 1990.
2. C. Gobl and A. N. Chasaide. “Acoustic Characteristics of Voice Quality”, Speech Communication, v 11, pp. 481-490, 1992.
3. D. G. Childers. “Vocal Quality Factors: Analysis, Synthesis, and Perception”, J. Acoust. Soc. Am. 90(5), pp. 2394-2409.
4. Diehl R. and D. Rosenberg 1977. Acoustic feature analysis in the perception of voicing contrasts. *Perception and Psychophysics* 21(5). 418-422.

5. Lisker L. 1986. 'Voicing' in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29, 1. 3-11.
6. Denes P. 1955. Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America* 27. 761-764.
7. Ladefoged P. and I. Maddieson. 1996. *The Sounds of the World's Languages*. Malden: Blackwell Publishing.
8. Haggard M., Q. Summerfield and M. Roberts. 1991. Psychoacoustical and cultural detriments of phoneme boundaries: Evidence from trading F0 cues in the voiced voiceless distinction. *Journal of Phonetics* 9. 49-62.
9. Whalen D., A. Abramson, L. Lisker and M. Mody. 1993. F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America* 93 (4). 2152-2159.
10. Hoole P., C. Gobl and A. Chasaide. 1999. Laryngeal coarticulation. In Hardcastle W. and N. Hewlett (eds.). *Coarticulation: Theory, Data and Techniques*. Cambridge: Cambridge University Press. 105-143.
11. Sjölander K. and J. Beskow. 2000. WaveSurfer - An open source speech tool. In Yuan B., T. Huang and X. Tang (eds.), *Proceedings of ICSLP 2000, 6th International Conference on Spoken Language Processing*. Beijing. 464-467.
12. Ladefoged P. 2003. *Phonetic Data Analysis*. Malden: Blackwell Publishing.
13. Riede T., B. Mitchell, I. Tokuda and M. Owren. 2005. Characterizing noise in nonhuman vocalizations: Acoustic analysis and human perception of barks by coyotes and dogs. *Journal of the Acoustical Society of America* 118 (1). 514-522.
14. Harrington J. and S. Cassidy. 1999. *Techniques in Speech Acoustics*. Dordrecht: Kluwer Academic Publishers.

IJSER