**ISSN 2063-5346**

# Speaker Profiling: The Study of Supra Segmental Features of Different Malayalam Dialects for Forensic Speaker Identification

[1]Sherin Shaji, [2] Dr. Abhinav Sood

[1] M.Sc. Student, Department of Forensic Science, Chandigarh University, Gharuan, India,[2] Assistant Professor, Department of Forensic Science, Chandigarh University, Gharuan, India

[1] sherinshaji110@gmail.com, [2] abhinav.e13628@cumail.com

***Abstract— For forensic speaker identification, this study intends to analyze the acoustic properties of Malayalam dialects based on phonetic elements. The study includes the investigation of Malayalam native speakers from several locations in Kerala, India based on different dialects spoken, via voice samples. The voice samples were used to extract and analyze phonetic parameters such as vowel duration and formant frequencies. The analysis's findings showed that the Malayalam dialects differed significantly in terms of formant frequencies and vowel duration. These results can help distinguish between speakers of various Malayalam dialects, which is valuable for forensic speaker identification reasons. This can be especially pertinent when determining the source of a voice sample depends heavily on the dialect of the speaker. Overall, this work adds to the expanding body of knowledge on forensic speaker identification and emphasizes how crucial it is to consider dialectal variance.***

**Index Terms—** ***Dialects, Formants, Speaker profiling, Vowel duration.***

## I. INTRODUCTION

Speaker profiling is the process of identifying an individual's voice through their speech patterns and language use. This technique creates a voiceprint or "voice signature" that can be used to identify a person solely by their voice. Formant is a sound wave measured in dB (decibels) and as an indicator of how strong the sound is. The objective of this is to create a profile that can be used to identify an individual by their voice alone, even if they are not speaking in their native language or dialect [1]. Speaker profiling is commonly used in law enforcement and security to locate missing persons, fugitives, and terrorists, as well as to detect forgeries when someone attempts to impersonate another person's voice. Additionally, speaker profiling can be used to track changes in a person's voice over time and to detect emotional and psychological states reflected in speech patterns [2] However, factors such as background noise and interference caused by environmental noise, microphone quality, and network issues can affect the accuracy of voice identification. Technique such as Denoising can be used to filter out unwanted background noise and helps in increasing the accuracy and reliability of voice identification systems [3] Like any other biometric identification, speaker profiling can also be used as part of a larger, comprehensive approach to identification and security. Speaker profiling can also provide valuable insights into an individual's cultural background and regional dialect. It is an important tool in many fields and can provide valuable insights into speech patterns and language use [4][5].

In this study of forensic speaker identification, audio recordings are examined to determine speaker profiling based on dialect accent features. The study of acoustic characteristics based on phonetic elements of various dialects is a crucial component of this science [6]. Malayalam is a Dravidian language largely used in Kerala, an Indian state, and Lakshadweep, a union territory. It has many dialects, each with unique phonetic characteristics. In the present study, the acoustical and segmental features of 3 dialects (Kasaragod, Thrissur, and Pathanamthitta) of the Malayalam language, corollaries were made on phoneme production in voicing, identifying the Vowel Duration, and the main focus was to create a static profile of voice of the mentioned dialects to identify the native place of the sample voice through a comparison of formant frequency [7]. Formants values are largely used in forensics as they can be used to provide significant variations to describe accent variation caused by dialects and other linguistic characteristics [8]. By making it easier to identify between various speakers and dialects, the study of these dialects' acoustic characteristics based on phonetic aspects can help in forensic speaker identification [9]. Pitch, intensity, formants, and other acoustic measurements that may be retrieved from speech signals are only a few examples of acoustic characteristics based on phonetic elements that can be used [10]. Researchers can find trends and distinctions between various dialects and individual speakers by examining these qualities. Beyond forensic speaker identification, the study of acoustic characteristics based on phonetic properties of Malayalam dialects can also be useful in voice recognition and language acquisition [11]. Researchers can increase the efficacy of speech recognition systems and aid people in learning to

**ISSN 2063-5346**

speak and understand various dialects by comprehending the distinctive acoustic properties of different dialects. Overall, research into the acoustic properties of Malayalam dialects based on their phonetic qualities is useful and has many applications in fields like forensic science, speech, technology, and linguistics [12][13].

## II. MATERIALS AND METHOD

### *A. Subjects*

Speakers have been chosen from even areas of the native dialect as well as within a limited age group of 20–30, with at least higher secondary education and also can speak its pure native dialect. It was noted that none of subjects suffered from any speech defects.

### *B. Recordings*

The samples were collected on a digital recorder with 128 kbps recording speech & 16-bit mono recorder. It has taken care while recording of speech samples to make the free from distortions and noise.

For speech analysis, the Malayalam word 'PERU' was taken as a sample word means as 'Name' in English. From the test word 'PERU' the analysis was carried out in CVC Monosyllable for segment 'P-E-R'.

### *C. Software*

PRAAT Software is used for the analysis of phonetic features of speech [14]. It is an easy and open-source software that can be used for performing speech analysis. We have used it for analyzing vowel duration and formants.

## III. ANALYSIS AND DISCUSSION

The samples were converted to digital and then the word segments were marked and phonetic segments were determined, the segments thus marked constituted repeated segments of words. The window of the screenshot image of the analysis of the word 'peru' is shown below (Figure 1). The waveform of the segmented word is given in the first window. The second window is the spectrogram of speech signals determines the formant frequencies; red dots seen on the spectrogram are splits of formant tracks. Vowel duration is determined by analyzing the spectrograph by marking the dark segments of vowel voicing below.

Data represented in the tables are the average data of repeated segments of the target word so that maximum accuracy is maintained.

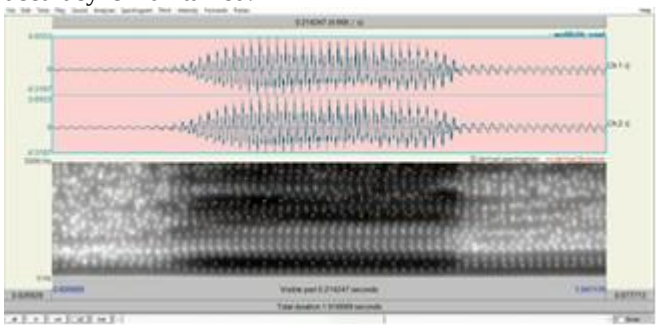

Figure 1. Praat's sound window

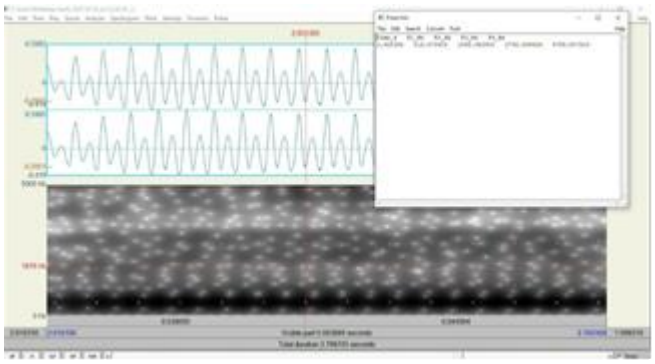

Fig 2: Analysis of formants from spectrogram and display of Formant values of F1, F2, F3, and F4. Measurement of the formant value can be obtained by clicking the Get Formant List menu on the formant menu.

20 Speech samples of male and female subjects analyzed over vowel /e/ –CVC syllabi of the Thrissur, Pathanamthitta, and Kasaragod dialects is shown in Table Ⅰ, Table Ⅱ, and Table Ⅲ.

Average of Vowel duration and formant frequencies of the repeated segments of 20 subjects of the Thrissur dialect of Vowel /e/

| Sample No. | Sex | Vowel Duration in CVC Syllable | F1(Hz) | F2(Hz) | F3(Hz) |
|---|---|---|---|---|---|
| T1 | M | 0.101 | 405.343 | 1901.652 | 2410.420 |
| T2 | M | 0.086 | 511.127 | 2147.313 | 2460.538 |
| T3 | M | 0.093 | 479.264 | 2019.610 | 2273.321 |
| T4 | M | 0.095 | 424.929 | 1737.062 | 2561.475 |
| T5 | M | 0.090 | 550.764 | 1865.913 | 2256.845 |
| T6 | M | 0.095 | 456.734 | 2094.914 | 2509.033 |
| T7 | M | 0.098 | 429.982 | 1798.091 | 2394.046 |
| T8 | M | 0.090 | 504.824 | 2094.736 | 2498.311 |
| T9 | M | 0.089 | 439.098 | 2133.604 | 2298.241 |
| T10 | M | 0.094 | 543.61 | 1948.83 | 2476.06 |

**ISSN 2063-5346**

| Sample Number | Sex | Vowel Duration Of CVC Syllable | F1 (Hz) | F2 (Hz) | F3 (Hz) |
|---|---|---|---|---|---|
| P1 | M | 0.099 | 415.727 | 1894.068 | 2458.943 |
| P2 | M | 0.103 | 481.274 | 2103.437 | 2588.015 |
| P3 | M | 0.105 | 454.625 | 1987.264 | 2511.761 |
| P4 | M | 0.097 | 437.274 | 1847.760 | 2436.117 |
| P5 | M | 0.102 | 439.995 | 2076.046 | 2545.981 |
| P6 | M | 0.096 | 498.665 | 1793.815 | 2675.876 |
| P7 | M | 0.098 | 502.873 | 1977.801 | 2488.726 |
| P8 | M | 0.109 | 423.845 | 2007.654 | 2511.723 |
| P9 | M | 0.110 | 485.906 | 1892.099 | 2599.066 |
| P10 | M | 0.099 | 524.908 | 2111.909 | 2600.526 |
| P11 | F | 0.089 | 475.481 | 2338.406 | 3060.335 |
| P12 | F | 0.085 | 537.334 | 2115.486 | 3464.884 |
| P13 | F | 0.097 | 640.166 | 2148.877 | 2904.094 |
| P14 | F | 0.087 | 550.322 | 2294.991 | 2991.020 |
| P15 | F | 0.093 | 597.765 | 2198.672 | 2967.912 |
| P16 | F | 0.095 | 498.803 | 2239.098 | 3020.432 |
| P17 | F | 0.096 | 563.723 | 2310.662 | 3340.991 |
| P18 | F | 0.094 | 610.712 | 2436.998 | 2945.091 |
| P19 | F | 0.093 | 544.609 | 2210.882 | 3411.229 |
| P20 | F | 0.094 | 587.732 | 2300.932 | 3043.206 |
| T11 | F | 0.092 | 411.534 | 1978.321 | 2784.258 |
| T12 | F | 0.087 | 510.865 | 2615.017 | 3111.067 |
| T13 | F | 0.101 | 443.686 | 2450.864 | 3107.269 |
| T14 | F | 0.098 | 585.531 | 2447.588 | 3203.769 |
| T15 | F | 0.095 | 489.274 | 2367.987 | 2887.543 |
| T16 | F | 0.089 | 600.876 | 2246.935 | 3123.711 |
| T17 | F | 0.099 | 573.364 | 2611.612 | 2995.243 |
| T18 | F | 0.089 | 511.725 | 2433.937 | 3209.872 |
| T19 | F | 0.100 | 493.625 | 2513.824 | 3009.336 |
| T20 | F | 0.094 | 592.711 | 2333.963 | 2993.207 |

I

Average of Vowel duration and formant frequencies of the repeated segments of 20 subjects of the Pathanamthittaa dialect of Vowel /e/

II

Average of Vowel duration and formant frequencies of the repeated segments of 20 subjects of the Kasaragod dialect of Vowel /e/

| Sample Number | Sex | Vowel Duration in CVC Syllable | F1(Hz) | F2(Hz) | F3(Hz) |
|---|---|---|---|---|---|
| K1 | M | 0.130 | 405.873 | 1956.805 | 2859.470 |
| K2 | M | 0.091 | 440.148 | 2023.819 | 2777.078 |
| K3 | M | 0.090 | 449.704 | 1877.881 | 2416.936 |
| K4 | M | 0.089 | 411.103 | 1620.979 | 2429.174 |
| K5 | M | 0.098 | 420.55 | 1968.55 | 2631.17 |

**ISSN 2063-5346**

| | | | | | |
|---|---|---|---|---|---|
| | | | 4 | 7 | 5 |
| K6 | M | 0.102 | 466.98 0 | 1786.54 3 | 2543.83 4 |
| K7 | M | 0.097 | 422.82 2 | 1832.32 3 | 2802.52 3 |
| K8 | M | 0.095 | 449.07 8 | 2087.67 3 | 2495.91 1 |
| K9 | M | 0.102 | 500.76 5 | 1709.77 3 | 2775.02 2 |
| K10 | M | 0.094 | 511.82 6 | 2095.65 6 | 2839.55 1 |
| K11 | F | 0.101 | 480.17 2 | 2374.34 3 | 2873.57 7 |
| K12 | F | 0.095 | 451.95 6 | 2881.14 9 | 3365.66 6 |
| K13 | F | 0.110 | 429.74 6 | 2287.14 9 | 3250.11 2 |
| K14 | F | 0.097 | 480.93 5 | 2456.82 3 | 3067.52 2 |
| K15 | F | 0.103 | 493.31 1 | 2654.98 7 | 3188.23 3 |
| K16 | F | 0.098 | 432.71 0 | 2567.87 6 | 3211.71 2 |
| K17 | F | 0.095 | 500.87 1 | 2802.76 2 | 2998.16 2 |
| K18 | F | 0.112 | 484.89 6 | 2455.87 2 | 2891.72 0 |
| K19 | F | 0.097 | 544.90 4 | 2564.90 8 | 3081.60 2 |
| K20 | F | 0.107 | 520.75 4 | 2765.31 1 | 3200.76 2 |

Ⅲ

## IV. RESULTS AND DISSCUSSIONS

Given below table contains Median of vowel duration and formant frequencies (F1,F2,F3) derived from the tables numbered Ⅰ, Ⅱ and Ⅲ

| Sample Name | Sex | Median of vowel duration | Median of F1 (Hz) | Median of F2 (Hz) | Median of F3 (Hz) |
|---|---|---|---|---|---|
| Pathanam thitta | M | 0.101 | 466.5 | 1969.1 | 2541.6 |
| | F | 0.092 | 560.6 | 2259.5 | 3114.9 |
| Thrissur | M | 0.093 | 474.5 | 1974.1 | 2413.8 |
| | F | 0.094 | 521.3 | 2403.0 | 3042.5 |
| Kasarago d | M | 0.098 | 447.8 | 1896.0 | 2657.0 |
| | F | 0.101 | 482.0 | 2581.1 | 3112.9 |

Ⅳ

Given below table contains the lowest and highest values of vowel duration and formant frequencies (F1, F2, F3) derived from the tables numbered Ⅰ, Ⅱ, and Ⅲ

| Sample Name | Sex | Lowest and highest recorded vowel duration (in secs) | Lowest and highest recorde dF1(Hz) | Lowest and highest recorded F2(Hz) | Lowest and highest recorded F3 (Hz) |
|---|---|---|---|---|---|
| Pathan-amthitta | M | 0.096-0.110 | 415-524 | 1793-2111 | 2436-2675 |
| | F | 0.085-0.097 | 475-640 | 2115-2436 | 2904-3464 |
| Thrissur | M | 0.086-0.101 | 405-550 | 1737-2133 | 2256-2561 |
| | F | 0.087-0.101 | 411-600 | 1978-2611 | 2784-3209 |
| Kasarag-od | M | 0.089-0.130 | 405-511 | 1620-2095 | 2416-2859 |
| | F | 0.095-0.112 | 429-544 | 2374-2881 | 2873-3365 |

V

This study reveals the correlation between sustained vowel /e/ –CVC syllabi of the Kasaragod, Pathanamthitta, and Thrissur dialects by means of analyzing 20 vocal samples of three different dialects. The results with the lowest and highest recorded vowel duration along with identifying the value of formant appropriately highlighted in Table Ⅳ and Ⅴ. To study the distinguishing characteristics of the vowels in speech samples, formant frequencies (F1, F2, and F3) have been considered for analysis. These vowel formants will help to identify the distinguishing features in dialectal variants. Vowels can be classified in sound by three properties: On the basis of length (tense/lax), position (high/low and front/back), and shape of lips (rounded, neural, and spread). The

**ISSN 2063-5346**

characteristics of vowels such as the back and front articulation are associated with the high and low values of F2 respectively [15]. Likewise, the high and low values of the F1 formants are related to the close and open articulation respectively. F3 shows the variation between rounded and unrounded vowels. Mostly, F3 formant analysis has been done to analyze singing voices. The significance of vowel formants can be appropriately established by F1 and F2 formant which deals with age, gender, and accent variability [16]. From the results shown in Tables 4 and 5, we can clearly see that the time duration needed to enunciate the chosen vowel showed more similarities between males and females, in the Thrissur dialect in comparison to Pathanamthitta and Kasaragod dialects. The highest recorded formant frequency (F1) among female subjects was recorded in the Pathanamthitta dialect (F1:640) While the female's formant frequency (F1) in the Kasaragod dialect was (F1: 544) and for formant frequency (F2) the highest recorded for females of Pathanamthitta dialect was (F2:2436) while in Kasaragod dialect it was (F2:2881). The highest recorded formant frequency among male subjects was recorded in the Thrissur dialect (F2:2133). While for the females the formant frequency recorded was among the Kasaragod dialect (F2: 2881). Intersex individuals say multiple words in a uniquely long manner which is different from how male and female subjects would usually pronounce words while speaking in normal dialect. This is another individualizing characteristic that could be found in this study. From the results obtained, the data showed that the three dialects have their own significant differences from each other on the basis of vowel duration as well as formant frequencies. From the previous studies [15][16]it is clear that formants F1 and F2 are forensically significant for dialect differentiation, this study also showed similar significant differences between F1 and F2 formants of Kasaragod and Thrissur dialects. The current study suggests a more in-depth analysis of dialects on the basis of suprasegmental features in various languages for creating reliable dialect profile data.

## REFERENCES

[1] H. Hollien, G. Didla, J. D. Harnsberger, and K. A. Hollien, "The case for aural perceptual speaker identification," *Forensic Sci Int*, vol. 269, pp. 8–20, 2016.

[2] M. Magdin, T. Sulka, J. Tomanová, and M. Vozár, "Voice Analysis Using PRAAT Software and Classification of User Emotional State," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5, no. 6, p. 33, 2019, doi: 10.9781/ijimai.2019.03.004.

[3] M. Coto-Jiménez, "Experimental Study on Transfer Learning in Denoising Autoencoders for Speech Enhancement," 2020, pp. 307–317. doi: 10.1007/978-3-030-49076-8_29.

[4] P. Angkititrakul, *Accent and speaker recognition for advanced automatic speech recognition*. University of Colorado at Boulder, 2004.

[5] S. Cenceschi, C. Meluzzi, and A. Trivilini, "The Variability of Vowels' Formants in Forensic Speech," *IEEE Instrum Meas Mag*, vol. 24, no. 1, pp. 38–41, 2021.

[6] H. Kwon, Y. Kim, H. Yoon, and D. Choi, "Selective Audio Adversarial Example in Evasion Attack on Speech Recognition System," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 526–538, 2020, doi: 10.1109/TIFS.2019.2925452.

[7] P. Ahuja and J. M. Vyas, "Forensic speaker profiling: the study of supra-segmental features of Gujarati dialects for text–independent speaker identification," *AustrAliAn JournAl of forensic sciences*, vol. 50, no. 2, pp. 152–165, 2018.

[8] P. Rose, *Forensic Speaker Identification*. CRC Press, 2002. doi: 10.1201/9780203166369.

[9] M. Jessen, "Forensic Phonetics," *Lang Linguist Compass*, vol. 2, no. 4, pp. 671–711, Jul. 2008, doi: 10.1111/j.1749-818X.2008.00066.x.

[10] N. B. Chittaragi, A. Limaye, N. T. Chandana, B. Annappa, and S. G. Koolagudi, "Automatic text-independent Kannada dialect identification system," in *Information Systems Design and Intelligent Applications: Proceedings of Fifth International Conference INDIA 2018 Volume 2*, Springer, 2019, pp. 79–87.

[11] G. Brown and J. Wormald, "Automatic sociophonetics: Exploring corpora with a forensic accent recognition system," *J Acoust Soc Am*, vol. 142, no. 1, pp. 422–433, Jul. 2017, doi: 10.1121/1.4991330.

[12] S. P. Todkar, S. S. Babar, R. U. Ambike, P. B. Suryakar, and J. R. Prasad, "Speaker Recognition Techniques: A Review," in *2018 3rd International Conference for Convergence in Technology (I2CT)*, 2018, pp. 1–5. doi: 10.1109/I2CT.2018.8529519.

[13] G. S. Didla and S. L. Niharika, "Voice Impersonation and Forensic Speaker Identification: A Review," *OSMANIA PAPERS IN LINGUISTICS*, p. 119.

[14] P. Boersma, D. J. M. Weenink, and V. Van Heuven, "PRAAT, a system for doing phonetics by computer Examining the contribution of statistical learning to grammar and literacy acquisition in children View project Lexical deficits in specific language impairment: the role of statistical learning View project Speak and unSpeak with PRAAT RAAT," 2001. [Online]. Available: https://www.researchgate.net/publication/208032992

[15] D. K. Phull and G. B. Kumar, "Vowel Analysis for Indian English," in *Procedia Computer Science*, Elsevier B.V., 2016, pp. 533–538. doi: 10.1016/j.procs.2016.07.264.

[16] E. Language, A. KumarC, and S. PrasadKM, ""Basic Acoustic Features Analysis of Vowels and C-V-C of Indian."