Identifying the Acoustics Features of Malayalam Vowels

Nasla Abdul Azeez
PhD Research Scholar, Department of Linguistics
Thunchath Ezhuthachan Malayalam University
naslaabdulazeez@gmail.com
and
Dr. Saidalavi.C
Principal
Malabar College of Advanced Studies, Vengara
saidmutty@gmail.com

Abstract

The study of vowel acoustics is fundamental to understanding the phonetic and phonological structure of a language. This research focuses on identifying and analysing the acoustics of vowel system in Malayalam by examining key acoustic parameters, particularly formant frequencies (F1, F2, and F3). Formant frequencies serve as critical indicators of vowel quality, influencing speech perception and linguistic classification. Through acoustic analysis, this study explores the distribution and variation of these formants across different vowel categories in Malayalam. The results provide valuable insights into the phonetic properties of the language, contributing to a deeper understanding of its vowel space. Furthermore, this research has potential applications in speech recognition, linguistic analysis, and language preservation efforts. By documenting and analysing the acoustic characteristics of Malayalam vowels, the study aims to support advancements in phonetic studies and computational linguistics while preserving the linguistic diversity of the language.

Keywords: Malayalam Vowels, Acoustic Analysis, Formant Frequencies F1, F2, F3.

Introduction

Vowel acoustics play a crucial role in the study of phonetics and phonology, as vowels serve as the fundamental building blocks of spoken language. The acoustic properties of vowels, particularly formant frequencies, provide valuable insights into the articulation, phonetic structure, and perceptual characteristics of a language. The fundamental frequency of the voice is determined by the rate of vocal cord vibrations per second of the voice produced, which is thus controlled by the speaker

through adjusting the tension in the vocal cords. It accompanies preferred resonating frequencies of the vocal tract referred to as formants. They are usually referred to as F1, F2, F3 etc. F1 is the first formant frequency, resulting from resonance in the first chamber, which extends from the glottis to the points of articulation or approximation. F2 is the second formant frequency, caused by resonance in the second chamber, which runs from the point of articulation or the area where the tongue reaches a particular height, to the front part of the tongue or the dental ridge. Finally, F3 is the third formant, occurring in the third chamber, a smaller air column between the tongue and the lips.

The first formant (F1) is primarily associated with vowel height, the second formant (F2) is linked to vowel frontness or backness, and the third formant (F3) can provide additional information about vocal tract characteristics, including speaker identity and certain phonetic features. Analysing these formant frequencies is essential for understanding how vowels are produced and perceived across different languages.

Languages vary in their vowel systems, and their acoustic characteristics are influenced by multiple factors, including the shape and length of the vocal tract, articulatory settings, and coarticulatory effects. While numerous acoustic studies have been conducted on well-documented languages such as English, French, and German, many lesser-studied languages, including Malayalam, lack detailed acoustic analysis. Malayalam, a Dravidian language spoken primarily in the Indian state of Kerala and by diaspora communities worldwide, has a distinct phonetic inventory that sets it apart from Indo-Aryan and other Dravidian languages. The vowel system of Malayalam consists of both short and long vowel pairs, each with unique articulatory and acoustic properties. Malayalam is a five vowel system language. /ɪ/, /e/, /a/, /u/, /ɔ/ with length contrast are the pure vowels in Malayalam and diphthongs /aɪ/, /au/ are also present in the vowel system of Malayalam. However, despite its linguistic richness, there has been relatively little acoustic research dedicated to the vowel system of Malayalam, particularly in terms of formant frequency analysis.

Given this gap in research, the present study aims to systematically analyse the vowels of Malayalam using acoustic parameters, specifically focusing on formant frequencies (F1, F2, and F3). By measuring and evaluating these formants, this study seeks to determine the distribution, variation, and acoustic characteristics of Malayalam vowels. This research will provide a detailed phonetic profile of Malayalam vowels, contributing to a more comprehensive understanding of the language's sound system.

The significance of this study extends beyond linguistic description. The findings will be beneficial for various applications, including speech synthesis, automatic speech recognition (ASR) systems, forensic linguistics, and language documentation efforts. A precise acoustic characterization of Malayalam vowels can improve speech technology applications by refining pronunciation models

and enhancing voice-based systems. Additionally, given the increasing focus on language preservation, documenting the acoustic properties of Malayalam contributes to linguistic conservation efforts and aids future research in comparative phonetics and phonology.

This paper is structured as follows: Section 2 presents the methodology used for data collection and acoustic analysis, including details on participant selection, recording procedures, and analytical techniques. Section 3 discusses the results of the formant frequency measurements, highlighting patterns and variations within the Malayalam vowel system. Section 4 interprets the findings in a broader linguistic context, comparing them with established phonetic principles. Finally, Section 5 concludes the study by summarizing key insights, discussing practical implications, and suggesting directions for future research.

Literature Review

There are several studies available focusing on the acoustic features of vowels in English and how they differ across dialects of English and speaker groups. There are also significant studies on the acoustic properties of vowels in languages other than English. Choi (1991) analyzed the vowels of Kabardian, a Circassian language spoken in the northwest Caucasus, measuring formant frequencies using wide-band FFT power spectra with a filter bandwidth of 300 Hz. The mean formant frequencies obtained from the FFT measurements were considered in the study. Choi highlighted the impact of different consonant types on the vowels of Kabardian. Another notable study focused on the vowels of Hidatsa, a language from the Missouri Valley branch of the Siouan language family, spoken primarily around the Fort Berthold Indian Reservation in North Dakota. Additionally, Boyle et al. (2024) and Rai (2017) examined the acoustic features of Koyee, a language spoken in Nepal.

Significant research has also been conducted on the acoustic properties of vowels in Indian languages. Basu et al. (2017) investigated the acoustics of vowels in five North East Indian languages from Nagaland: Nagamese, Ao, Lotha, Sumi, and Angami. The vowels studied were /u/, /o/, /ə/, /a/, /e/, and /i/. A study by Chittaragi & Koolagudi (2019) identified certain features necessary for recognizing vowels in five different dialects of the Kannada language. In addition to formant frequencies (F1-F3), prosodic features were also considered for dialect identification. Other studies include research on the vowels of Lambada, a minor language spoken in various parts of India, with a special focus on the speakers from Telangana (Kumar & Duli, 2014); vowels of Bagri, a dialect of the Rajasthani language; vowels of Dhundhari, another Rajasthani language; vowels of Kashmiri; and a comparative study of the acoustic features of vowels in Maithili and Angika.

There is extensive research on the acoustic properties of vowels in various languages, the study of Malayalam vowels remains relatively under explored. Pallath (2023) examined the formant frequecies of Malayalam vowels in a fundamental level. This study aims to address these gaps by

conducting a detailed acoustic analysis of Malayalam vowels. By leveraging modern signal processing technique and exploring inter speaker variability, this research seeks to provide a more comprehensive understanding of Malayalam acoustics

Objective of the study

To identify and analyse the vowel system in Malayalam by examining key acoustic parameters, particularly formant frequencies (F1, F2 and F3), to gain insights into their distribution and variations in the language.

Methodology

Participant: A total of 30 participants took part in this study, including 15 males and 15 females aged between 25 and 35 years. They were from various regions of Kerala and were native speakers of Malayalam. The informants are devoid of any structural or functional speech abnormalities.

Method: The speakers' utterances were recorded in a semi-controlled environment. The analysis focused on the Malayalam vowels /ɪ/, /e/, /a/, /u/, and /ɔ/.The speakers' utterances were recorded in a semi-controlled environment. The selected words contained vowels appearing in all three-word positions: initial, medial, and final. The analysis focused on the Malayalam vowels /ɪ/, /e/, /a/, /u/, and /ɔ/. Specifically, six words each were chosen for /ɪ/ and /e/, twelve for /a/, nine for /u/, and seven for /ɔ/. Each word was uttered twice. Each recording was made using a high-quality IOS device, resulting in minimal background noise. The recorded voice samples were analysed using Praat Speech Recognition software.

Analysis and Discussion

Vowel identification of Malayalam is performed by analysing acoustic features, specifically by measuring the first three formant frequencies (F1, F2, and F3) of each vowel. The collected data is prepared carefully, maintaining its homogeneity. Trimming off the irrelevant pieces, collected sound were loaded to PRAAT. The first phase of analysis was to identify and record the three formant frequencies F1, F2 and F3 of the vowels /ɪ/, /e/, /a/, /u/ and /ɔ/. As we have considered repeated utterance of same instances of every vowel, the readings expected to be identical to almost all influences of practical situations on the production of vowels. The choice of words with vowels being at different positions too contributed to this.The mean formant frequencies of the vowels /ɪ/, /e/, /a/, /u/ and /ɔ/ are listed in the table below. As each entry is subjected to critical analysis, it was evident that the same vowel produced by the same candidate in the same conditions, almost instantly have some variations in the readings. This was observed in all of F1, F2 and F3. This suggested that taking

mean would not be an appropriate method to develop the identification of vowels. As the prevailing literature considers mean to determine vowels with F1, F2 and F3, we suggest a novel approach by considering a range of frequencies for each formant frequencies. Instead of looking for a precise value, this method will look in to a range of values which is expected to be more inclusive.

The range is taken around the mean.. With the readings, ranges are constructed for F1, F2 and F3 of each vowel of consideration. The left end point is taken as $mean - Std.Deviation$ and the right end point is taken as $mean + Std.Deviation$

Table 1

The *Table 1* depicts the means, standard deviations and ranges of F1, F2 and F3 frequencies of /ɪ/, /e/, /a/, /u/, /ɔ/ in Malayalam.

Vowel	F1 (Hz)			F2(Hz)			F3(Hz)		
	Mean	Std. Deviation	Range	Mean	Std. Deviation	Range	Mean	Std. Deviation	Range
/ɪ/	350.3	40.916	309.4 – 391.2	2281.55	254.65	2026.9 _ 2536.2	2971.2	213.49	2757.7 – 3184.7
/e/	478.32	54.586	423.7 – 532.9	2140.5	196.22	1944.3 – 2336.8	2759.1	195.15	2563.9 – 2962.2
/a/	744.375	126.937	622.4 – 876.31	1474.35	121.012	1353.3 – 1595.4	2712.68	276.534	2436.2 – 2989.2
/u/	398.08	40.186	357.9 – 438.3	1247.96	402.63	845.33 – 1650.6	2749.304	244.44	2504.9 – 2993.7
/ɔ/	496.54	85.053	411.5 – 581.6	1050.5	240.43	809.6 – 1290.9	2749.8	287.48	2462.3 – 3037.3

Formant Frequencies of Front vowel /ɪ/:

The mean first formant frequency (F1) of the front vowel /ɪ/ is 350.3 Hz, while the second formant (F2) is 2281.55 Hz, and the third formant (F3) is 2971.2 Hz which is tabled below. The standard deviation of the first formant frequency (F1) for the vowel /ɪ/ is 40.916 Hz, while the standard deviations of the second (F2) and third (F3) formants are 254.65 Hz and 213.49 Hz, respectively. Therefore, the formant frequency ranges for the high front vowel /ɪ/ in Malayalam are as follows: F1 ranges from 309.4 to 391.2 Hz, F2 ranges from 2026.9 to 2536.2 Hz, and F3 ranges from 2757.7 to 3184.7 Hz.

Formant frequencies of front vowel /e/:

The mean formant frequencies for the front vowel /e/ are 478.32 Hz for F1, 2140.5 Hz for F2, and 2759.1 Hz for F3. The mean F1 of the vowel /e/ is 128.02 Hz higher than that of /ɪ/, while the mean F2 and F3 of vowel /e/ are slightly lower compared to /ɪ/. The standard deviations of formant frequencies F1, F2, and F3 for the vowel /e/ are 54.586 Hz, 196.22 Hz, and 195.15 Hz, respectively. Thus, the formant frequency ranges for the vowel /e/ are as follows: F1 ranges from 423.7 to 532.9 Hz, F2 ranges from 1944.3 to 2336.8 Hz, and F3 ranges from 2563.9 to 2962.2 Hz. The F2 and F3 ranges of the vowels /ɪ/ and /e/ overlap significantly, but they can be distinguished based on their F1 range.

Formant frequencies of front vowel /a/:

The mean formant frequencies for the front vowel /a/ are 744.375 Hz for F1, 1474.35 Hz for F2, and 2712.68 Hz for F3.The mean F1 of the vowel /a/ is 266.055 Hz greater than the mean F1 of the vowel /e/. Additionally, the F1 of /a/ is greater than that of /e/, which in turn is higher than that of /ɪ/. The mean F2 of the vowel /a/ is the lowest compared to the F2 of vowels /e/ and /ɪ/.The mean F3 is slightly higher for the vowel /ɪ/, followed by /e/, and then /a/. The standard deviations of F1, F2, and F3 for the vowel /a/ are 126.937 Hz, 121.012 Hz, and 276.534 Hz, respectively. Therefore, the formant frequency ranges for the vowel /a/ are: F1 from 622.4 to 876.31 Hz, F2 from 1353.3 to 1595.4 Hz, and F3 from 2436.2 to 2989.2 Hz. The F1 ranges of the front vowels /ɪ/, /e/, and /a/ are distinct and do not overlap. In contrast, the F2 ranges of vowels /ɪ/ and /e/ overlap, while the F2 range of vowel /a/ is lower than the other two and does not overlap with them. The F3 range overlaps across all three vowels: /ɪ/, /e/, and /a/.

Formant Frequencies of back vowel /u/:

The mean formant frequencies for the back vowel /u/ are 398.08 Hz for F1, 1247.96 Hz for F2, and 2749.304 Hz for F3. The mean F1of vowel /u/ is lower than F1 of vowel /a/ and /e/ but higher than F1 of vowel /ɪ/. The mean F1 of the vowel /u/ is lower than that of vowels /a/ and /e/ but higher than that of vowel /ɪ/. However, the mean F2 of the vowel /u/ is the lowest among all the vowels /ɪ/, /e/, and /a/. F3 is almost similar in all cases. The F3 values are nearly the same across all vowels. The standard deviations of F1, F2, and F3 for the back vowel /u/ are 40.186 Hz, 402.63 Hz, and 244.44 Hz, respectively. Hence, the formant frequency ranges for the back vowel /u/ are: F1 from 357.9 to 438.3 Hz, F2 from 845.33 to 1650.6 Hz, and F3 from 2504.9 to 2993.7 Hz. The vowels /u/ and /a/ do not overlap in terms of range of F1. However, /ɪ/ and /u/ show significant overlap in the range of F1, while /u/ and /e/ overlap only slightly in the range of F1. So vowels /ɪ/ and /u/ cannot be distinguished in terms of range of F1. In the case of F2 range, the vowels /ɪ/ and /e/ do not overlap with /u/, but there is a slight overlap between the F2 ranges of vowels /u/ and /a/. all the vowels overlaps in the case of range of F3.

Formant frequencies of back vowel /ɔ/:

The mean formant frequencies for the back vowel /ɔ/ are 496.54 Hz for F1, 1050.5 Hz for F2, and 2749.8 Hz for F3. The mean of F1 of vowel /ɔ/ is higher than that of /e/, /u/ and /ɪ/ but lower than that of vowel /a/. the vowel /ɔ/ has the lowest mean F2 when compared to Malayalam vowels /ɪ, e, a, u/. Mean F3 is almost similar in all vowels in Malayalam. The standard deviations of F1, F2, and F3 for the back vowel /ɔ/ are 85.053 Hz, 240.43 Hz, and 287.48 Hz, respectively. Therefore, the formant frequency ranges for the back vowel /ɔ/ in Malayalam are: F1 from 411.5 to 581.6 Hz, F2 from 809.6 to 1290.9 Hz, and F3 from 2462.3 to 3037.3 Hz. The range of F1 of back vowel /ɔ/ do not overlap with the range of F1 of vowels /a/ and /ɪ/. The range of F1 of /ɔ/ overlaps heavily with /e/ and slightly with /u/. Meanwhile, the F2 range of /ɔ/ overlaps with /u/ but does not overlap with /ɪ/, /a/, or /e/.

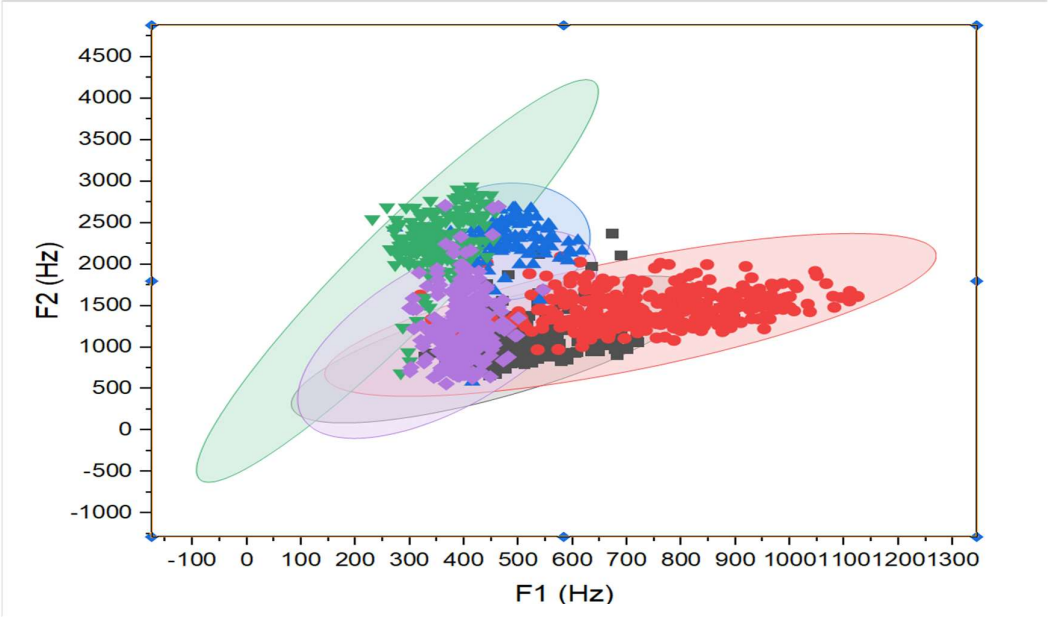


Figure 1

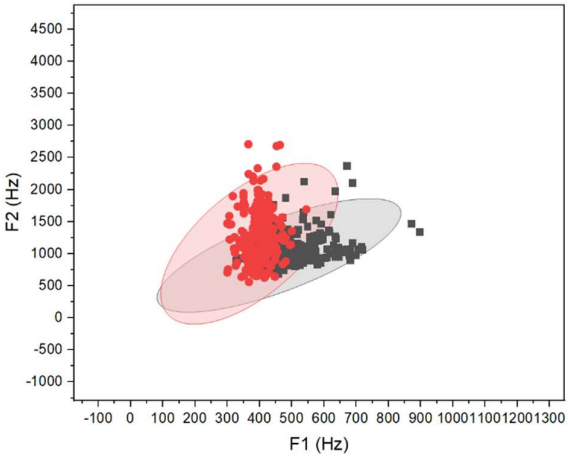


Figure 2

The low

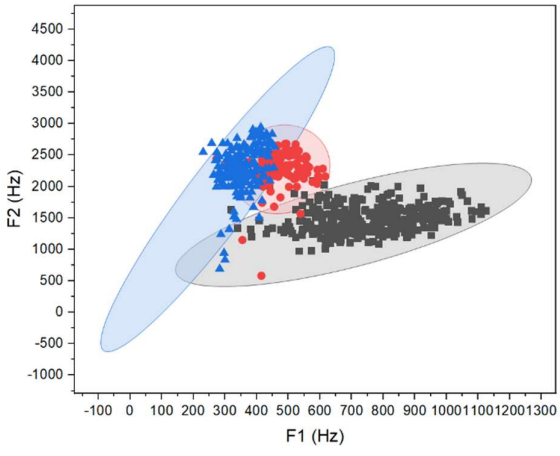
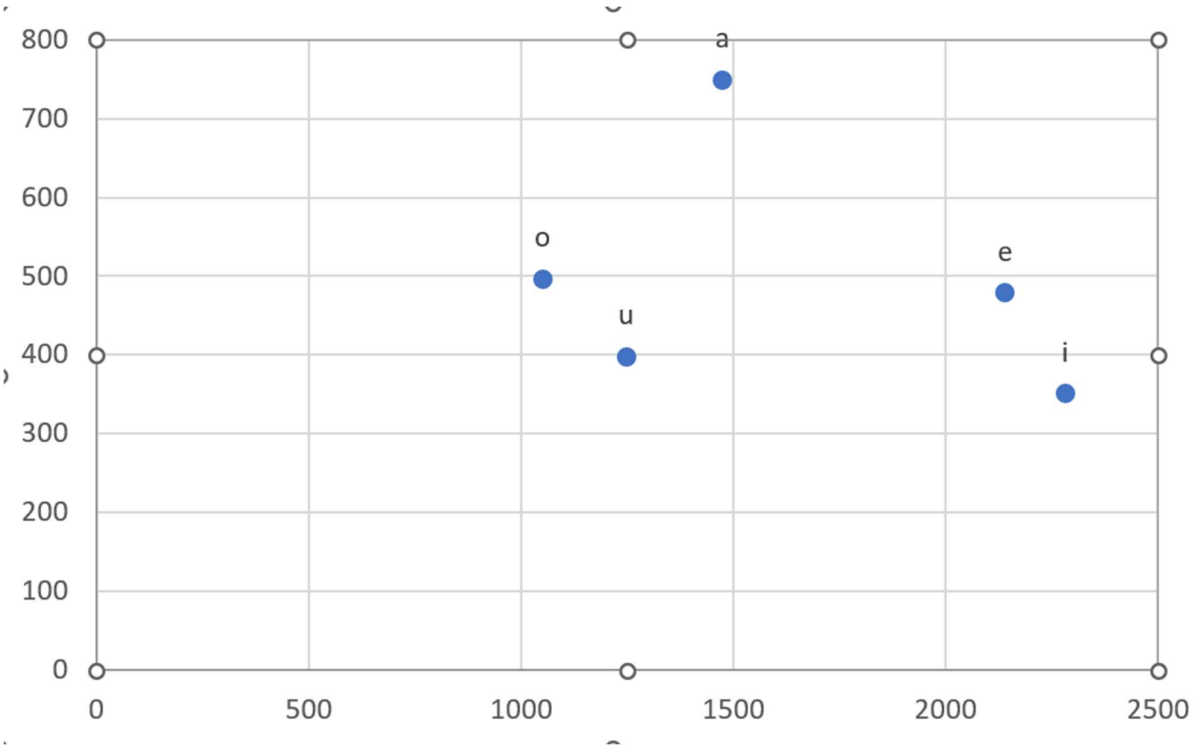


Figure 2

Front vowel /a/ has the highest F1 frequency range, spanning from 622.4 to 876.31 Hz, while the high front vowel /ɪ/ has the lowest F1 frequency range, ranging from 309.4 to 391.2 Hz in Malayalam vowels. Low vowels exhibit a higher F1 frequency range, while high vowels have a lower F1 frequency range. This indicates a pattern of increasing F1 frequency range from high to mid to low vowels. The pattern follows the order: /a/ > /e/ > /ɔ/ > /u/ > /ɪ/ in terms of F1 frequency range.

For the F2 frequency range, the front vowel /ɪ/ has the highest frequency, ranging from 2026.9 to 2536.2 Hz, while the back vowel /ɔ/ has the lowest frequency range, spanning from 809.6 to 1290.9 Hz in Malayalam vowels. Front vowels have a higher F2 frequency range compared to back vowels. This shows an increasing F2 frequency range from back vowels to front vowels, following the pattern: /ɪ/ > /e/ > /a/ > /u/ > /ɔ/.



Conclusion

This study aimed to identify and analyse the vowel system in Malayalam by examining key acoustic parameters, particularly formant frequencies (F1, F2, and F3). Through acoustic analysis, the study provided valuable insights into the distribution and variation of these formants, contributing to a deeper understanding of the phonetic characteristics of Malayalam vowels. The findings confirm that low vowels exhibit higher F1 frequencies, while high vowels have lower F1 frequencies, consistent with general phonetic principles. Similarly, front vowels show higher F2 frequencies than back vowels, aligning with established vowel classification patterns. The range of first formant frequency

of vowel in Malayalam exhibit an increase from high to mid to low vowels. The pattern follows the order: /a/ > /e/ > /ɔ/ > /u/ > /ɪ/. the range of second formant frequency of vowels in Malayalam shows an increasing F2 frequency range from back vowels to front vowels following the pattern: /ɪ/ > /e/ > /a/ > /u/ > /ɔ/. These results highlight the systematic nature of vowel production in Malayalam and provide empirical data for future linguistic research.

The study's findings have significant implications for various fields, including phonetics, speech technology, and language documentation. Understanding the acoustic properties of Malayalam vowels can aid in improving speech recognition systems, text-to-speech synthesis, and linguistic education. Additionally, the research contributes to the preservation and documentation of Malayalam phonetics, which is crucial for future comparative linguistic studies.

Despite its contributions, this study has certain limitations. The analysis was based on a specific set of vowel samples, and factors such as speaker variation, dialectal differences, and coarticulatory effects were not extensively explored. Future research could expand on these aspects by incorporating a larger dataset, analysing dynamic speech contexts, and comparing the acoustic properties of Malayalam vowels with those of other languages.

In conclusion, this study provides a foundational understanding of the acoustic characteristics of Malayalam vowels, contributing to both theoretical and applied linguistics. The insights gained from this research can serve as a valuable resource for further investigations into the phonetics and phonology of Malayalam and other Dravidian languages.

Reference

Basu, Joyanta, et al. "Acoustic Analysis of Vowels in Five Low Resource North East Indian Languages of Nagaland." *2017 20th Conference of the Oriental Chapter of the International Coordinating Committee on Speech Databases and Speech I/O Systems and Assessment (O-COCOSDA)*, IEEE, 2017, pp. 1–6.

Chittaragi, Nagaratna B., and Shashidhar G. Koolagudi. "Acoustic-Phonetic Feature Based Kannada Dialect Identification from Vowel Sounds." *International Journal of Speech Technology*, vol. 22, no. 4, Dec. 2019, pp. 1099–113.

Choi, John. "An Acoustic Study of Kabardian Vowels." *Journal of the International Phonetic Association*, vol. 21, 1991, pp. 4–12.

Johnson, Keith. *Acoustic and Auditory Phonetics*. 3rd ed., Wiley-Blackwell, 2011

Kewley-Port, Diane, and Yoko Atsuko Tatsumi. "Critical Spectral Regions for Vowel Identification." *The Journal of the Acoustical Society of America*, vol. 120, no. 1, 2006, pp. 405–410.

Kumar, R. Ramachandra, and Babi Duli. "An Acoustic Analysis of the Vowels of Lambada Language." *International Journal of Science and Research (IJSR)*, vol. 3, no. 12, Dec. 2014, pp. 1928–1932.

Ladefoged, Peter, and Ian Maddieson. "Vowels of the World's Languages."*Journal of Phonetics*, vol. 18, 1990, pp. 93–122.

Ladefoged, Peter, and Keith Johnson. *A Course in Phonetics*. 7th ed., Cengage Learning, 2014.

Ladefoged, Peter. *Vowels and Consonants: An Introduction to the Sounds of Languages*. 2nd ed., Blackwell, 2005.

Pallath, Nobert Thomas. "Formant Analysis of Malayalam Phonemes for Cleft Lip". *International Journal of Scientific Development and Research*, vol. 8, no. 1, Jan. 2023, pp. 731–739

Paul, Elizebath P., and Satish Kumaraswami. "Acoustic Characteristics of Vowels in Clear and Conversational Speaking Style in Malayalam." *International Journal of Scientific Development and Research*, vol. 5, no. 8, Aug. 2020, pp. 350–351.

Rai, Tara Mani. "Koyee Vowels: An Acoustic Analysis."*Gipan*, vol. 3, no.2, 2017, pp.136-143.