Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 20 May 2026
Revised 29 June 2026
Accepted 21 July 2026

Natural Resources for Human Health 2026, Vol. 6 (8s): 1170–1174

KEYWORDS:
Short-Time Fourier Transform (STFT), Linear Predictive Coding (LPC),Formants, Phoneme, articulation disorder,Intensity,Pitch,Jitter,shimmer .

Comparative Acoustic Analysis in Dental Plosives Using Linear Predictive Coding for Articulation Disorder Assessment

[1]Renuka Chandrakanth, [2]Sunil Yaraguntappa, [3]Manjula G. Hegde, [4]Shashidhara K S, [5]Dr. V. Shashidhar, [6]Malaya Kumar Nath

[1]Department of Electronics and Communication Engineering, Maharaja Institute of Technology Mysore, Mandya 571477, Karnataka, India | Department of Electronics and Communication Engineering, M.S, Ramaiah University of Applied Sciences, Peenya, Bengaluru 560058, Karnataka, India
[2]Department of Electronics and Communication Engineering, M.S. Ramaiah University of Applied Sciences, Peenya, Bengaluru 560058, Karnataka, India
[3]Dept. of ECE, R. R. Institute of Technology, Bengaluru, Karnataka, India
[4]Nitte (Deemed to be University), Nitte Meenakshi Institute of Technology, Department of Electronics Engineering(VLSI Design and Technology), Bengaluru, India. (Corresponding Author)
[5]Assoc. Prof. &HoD, CSE (IoT CS BCT), Rajarajeswari College of Engineering, Bangalore, India
6Department of ECE, National Institute of Technology Puducherry, Karaikal 609609, India

ABSTRACT: Dental plosives /d/ and /t/ in Kannada (Mysore dialect) share nearly identical place and manner of articulation, making them acoustically similar and particularly difficult for articulation-disordered individuals to learn. This study presents a comparative acoustic analysis establishing a baseline profile from a normal native speaker, with each phoneme iterated five times and analysed using Praat for spectrogram generation and Origin 2026b for statistical visualization. Formant frequencies (F1–F4) were compared across three estimation approaches—direct measurement, Linear Predictive Coding (LPC), and Short-Time Fourier Transform (STFT) envelope peaks—alongside intensity, pitch, jitter, and shimmer. Results show jitter and shimmer remain nearly identical between the two phonemes, while intensity and pitch maxima differ significantly, offering a reliable acoustic marker for distinguishing them. LPC yielded smoother, more accurate resonance tracking than the comparatively noisy STFT envelope, especially at lower frequencies. These baseline findings provide a quantitative acoustic reference supporting diagnostic and therapeutic strategies in speech-language pathology.

INTRODUCTION

Speech is a complex acoustic signal shaped by the coordinated movement of the vocal tract, and any deviation in articulatory precision manifests as measurable changes in its acoustic structure [1], [2]. Among the various classes of speech sounds, dental plosives such as /d/ and /t/ are particularly challenging for individuals with articulation disorders, as both consonants share nearly identical place and manner of articulation, differing primarily in voicing and subtle temporal-spectral cues [3].

Natural Resources for Human Health
eISSN: 2583-1194

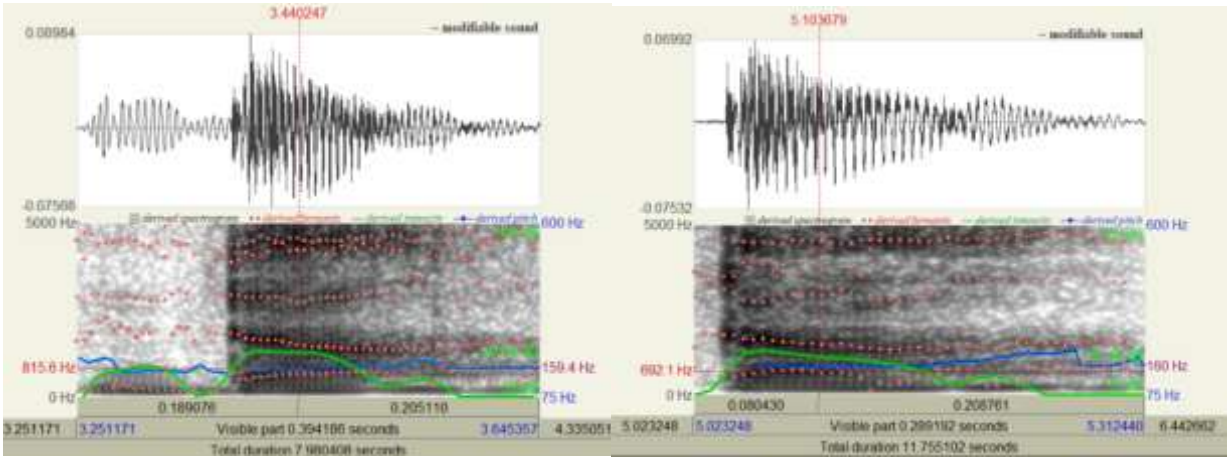
This acoustic closeness makes /d/ and /t/ a frequent source of confusion during speech therapy, especially in under-resourced regional languages such as Kannada, where standardized acoustic baselines are limited [4].

Acoustic analysis using tools such as Praat has become a widely accepted, non-invasive method for objectively characterizing normal and disordered speech through parameters including fundamental frequency, formant structure, jitter, and shimmer [5], [6]. Formant frequencies (F1–F4), representing vocal tract resonances, remain among the most reliable indicators for distinguishing phonemes with similar articulatory postures [7]. Complementing formant-based analysis, digital signal processing (DSP) techniques such as the Short-Time Fourier Transform (STFT) and Linear Predictive Coding (LPC) provide two contrasting approaches to spectral estimation: STFT offers a direct time-frequency representation of the raw signal [8],[10], while LPC models the vocal tract as an all-pole filter, yielding smoother and often more accurate resonance tracking, particularly at lower frequencies [11], [12].

Beyond phoneme-level formant analysis, acoustic markers such as jitter and shimmer have proven valuable in characterizing voice quality and articulatory stability across a range of speech and voice disorders, including dysarthria, stuttering, and dysphonia [13],[15]. However, comparative studies applying these DSP-based techniques specifically to confusable dental consonants in Indian regional languages remain scarce. This study addresses that gap by presenting a comparative acoustic analysis of Kannada /d/ and /t/, combining spectrographic, LPC, and STFT-based formant estimation to establish a quantitative baseline that can support diagnostic and therapeutic decision-making for articulation-disordered individuals.

METHODOLOGY

Phoneme articulation in Kannada language (Mysore dialect) by a normal native speaker is captured using microphone (Cirrus high definition) of Dell laptop. The client is asked to iterate phoneme ADA and ATA for five times and the signal is captured for acoustic analysis. Using Praat(version 7.0) an opensource software, spectrogram is obtained and analysed based on the formants (F1-Fundamental frequency). The spectrograms of phonemes for consonants /d/ and /t/ are sketched along with all the formants in the form of speckles shown as red dots in the Figure 1./d/ and /t/ are the dental consonant and very difficult for the articulation disordered patient to learn and articulate.



/d/ /t/

Figure 1: Spectrogram of Phoneme articulated /d/ and /t/

Though the sound has almost similar acoustic characteristics they differ in the intensity as shown in the comparison chart for both /d/ and /t/ in Figure 2. Thesedentalplosives on perseverance and manner of articulation appears to be identical. A bar graph that includes intensity maximum,pitchmaximum,jitter and shimmer are plotted for both /d/ and /t/ phoneme to highlight the acoustical significant difference while articulating. Jitter and shimmer are almost identical for both the phonemes except the intensity and pitch for corresponding formants. The graph is platted using Origin 2026b software tool for analysing the acoustic characteristics in consonant.

Natural Resources for Human Health
eISSN: 2583-1194

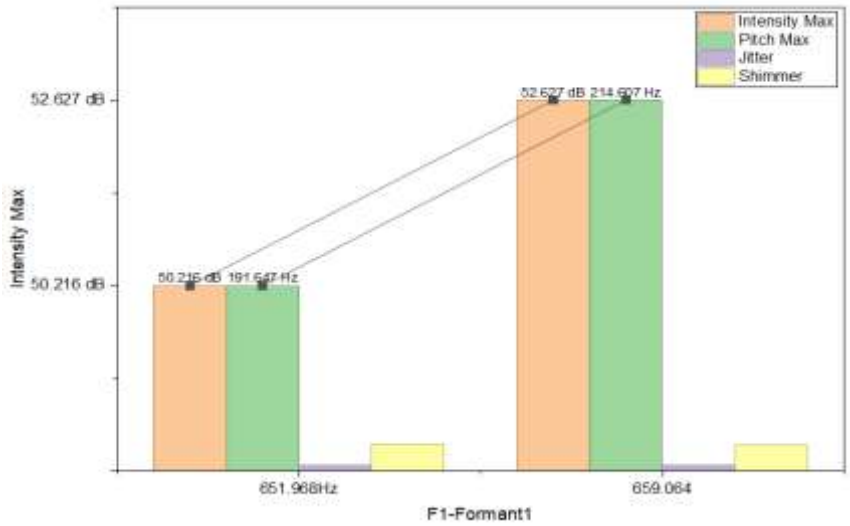


Figure 2: comparison bar graph depicting significance difference in intensity and pitch of articulated phonemes /d/ and /t/ respectively

In order to analyse acoustic difference in the articulated phonemes in Kannada (/d/ and /t/) a comparative analysis is carried out on the spectrogram using STFT (Digital Signal Processing) obtained using Praat and linear predictive coding using Origin 2026b.The formant frequency, LPC and STFT (to detect envelop peak) for both /d/ and /t/ is captured in Table1 and Table 2 respectively.

Table 1: Formant Frequency Comparison for phoneme ADA-Kannada language

Formants	**Frequency**	**LPC**	**STFT Spectrogram (Envelop peak)**
F1	651.97 Hz	785.5 Hz	215.3 Hz
F2	1444.17 Hz	1481.1 Hz	581.4 Hz
F3	2868.49 Hz	2905.7 Hz	818.3 Hz
F4	4136.34 Hz	4352.2 Hz	1432.0 Hz

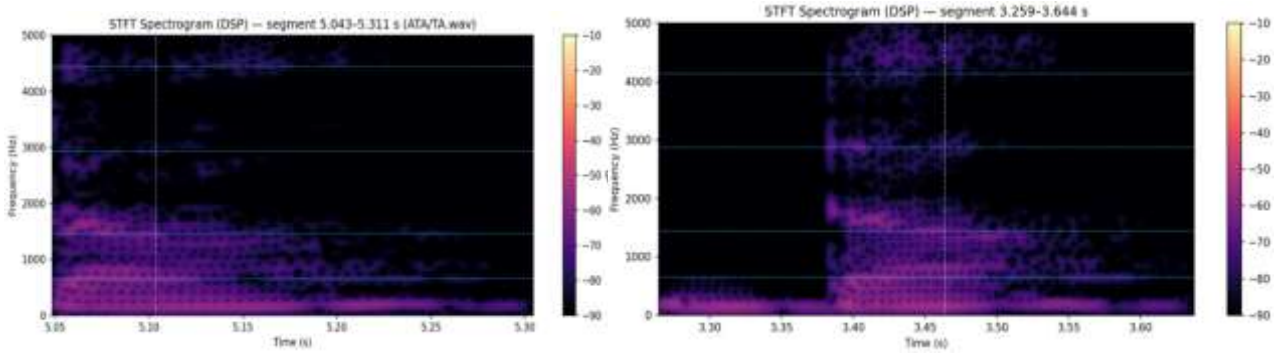
Table 2: Formant Frequency Comparison for phoneme ATA-Kannada language

Formants	**Frequency**	**LPC**	**STFT Spectrogram (Envelop peak)**
F1	659.06 Hz	695.5 Hz	247.6 Hz
F2	1460.79 Hz	1453.7 Hz	656.8 Hz
F3	2923.96 Hz	3031.0 Hz	1335.1 Hz
F4	4435.47 Hz	4427.8 Hz	4489.7 Hz

Natural Resources for Human Health
eISSN: 2583-1194

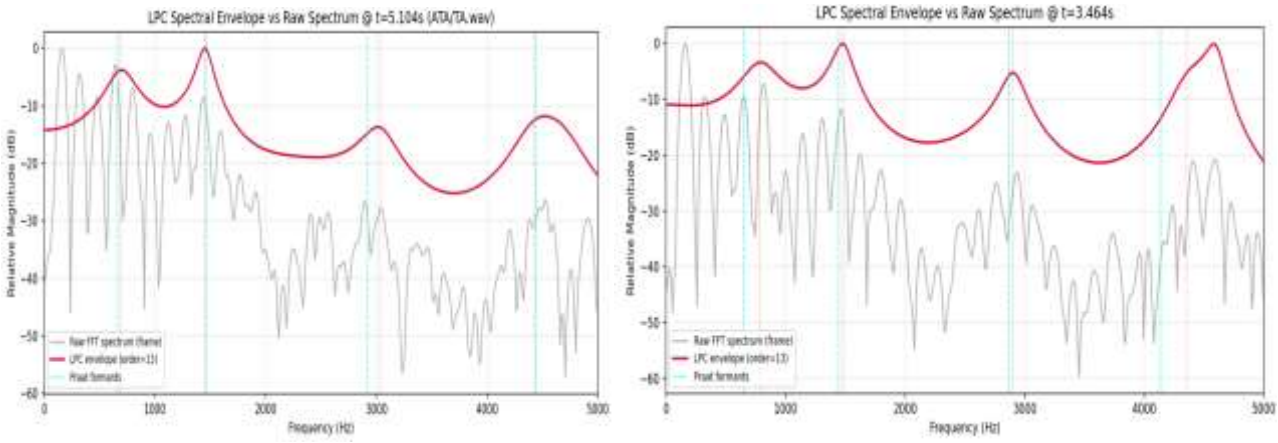
RESULT:

To study the phoneme articulation manner based on two approaches one being spectrogram a raw voice print and LPC for more accurate analysis methods that operates on directly on time-series blocks that can predict error and minimize it before converting it to frequency domain. Spectrum showing the envelop peak for high frequencies for both /d/ and /t/ phonemes in Figure3.



/d/ /t/

Figure 3: Short Time Fourier Time (STFT) spectrogram of phoneme/d/ and /t/ with frequency (Hz) in Y axis and Time (s) on X-axis



/d/ /t/

Figure 4: Linear predictive coding (LPC) providing smooth curve of phoneme spectrum articulated for /d/ and /t/ respectively

On comparing the STFT spectrogram and LPC envelope for tracking the vocal-tract resonance the former is noisier than LPC. LPC directly plots all the pole that corresponds to the vocal tract resonances, providing more accuracy.STFT doesn't provide continues resonance peaks for low frequency resulting in noisy true formants. From Figure 4 it's clearly evident that on increasing the window we can smear closely spaced low frequency signals.Accuracy of LPC is approximately near to the original and STFT underperforms.

CONCLUSION:

This study established a baseline acoustic comparison of the Kannada dental plosives /d/ (ADA) and /t/ (ATA) from a normal native speaker of the Mysore dialect, using spectrographic, LPC, and STFT-based formant analysis. While the two phonemes are perceptually and articulatorily similar, the results demonstrate that intensity and pitch maxima provide a consistent and measurable point of acoustic divergence, whereas jitter and shimmer remain largely indistinguishable between them. Comparison of formant-tracking methods further confirmed that LPC yields smoother, more reliable resonance

estimates than the STFT envelope, particularly at lower frequencies where spectral smearing degrades STFT accuracy. These findings offer a quantitative acoustic reference that can help clinicians and speech-language pathologists differentiate /d/ and /t/ production more objectively, supporting targeted diagnostic assessment and therapy planning for Kannada-speaking individuals with articulation disorders. Future work will extend this baseline to a larger cohort of both normal and disordered speakers to validate its clinical applicability.

DECLARATIONS

Renuka Chandrakanth: methodology, formal analysis, data curation,writing original draft preparation,writing-review and editing; Sunil Yaraguntappa: methodology, formal analysis, data curation, reviewing supervision; Malaya Kumar Nath: writing-review, editing and review. All authors have read and agreed to the publication of the manuscript.

Data availability

The datasets generated and analysed during the current study are not publicly available because they contain identifiable speech data and are subject to institutional ethics, participant consent, and privacy restrictions. The data are available from the corresponding author upon reasonable request and applicable ethical guidelines.

REFERENCES:

[1] B. A. Al-Qatab and M. B. Mustafa, "Classification of dysarthric speech according to the severity of impairment: an analysis of acoustic features," *IEEE Access*, vol. 9, pp. 18183–18194, 2021.

[2] Slis, N. Lévêque, C. Fougeron, M. Pernon, F. Assal, and L. Lancia, "Analysing spectral changes over time to identify articulatory impairments in dysarthria," J. Acoust. Soc. Am., vol. 149, no. 2, pp. 758–769, Feb. 2021.

[3] A. Joshy and R. Rajan, "Automated dysarthria severity classification: A study on acoustic features and deep learning techniques," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1147–1157, 2022.

[4] W. Xue, R. van Hout, C. Cucchiarini, and H. Strik, "Assessing speech intelligibility of pathological speech in sentences and word lists: The contribution of phoneme-level measures," *J. Commun. Disord.*, vol. 102, art. 106301, 2023.

[5] P. Hippargekar, S. Bhise, S. Kothule, and S. Shelke, "Acoustic voice analysis of normal and pathological voices in Indian population using Praat software," *Indian J. Otolaryngol. Head Neck Surg.*, vol. 74, suppl. 3, pp. 5069–5074, Dec. 2022.

[6] G. Li, Q. Hou, C. Zhang, Z. Jiang, and S. Gong, "Acoustic parameters for assessing voice quality in patients with voice disorders," *Ann. Palliat. Med.*, vol. 10, no. 1, pp. 130–136, 2021.

[7] M. Elsherbeny, H. Baz, and O. Afsah, "Acoustic characteristics of voice and speech in Arabic-speaking stuttering children," *Egyptian J. Otolaryngology*, vol. 38, art. 8, 2022.

[8] Elfaki, A. L. Asnawi, A. Z. Jusoh, A. F. Ismail, S. N. Ibrahim, and N. F. Mohamed Azmin, "Using the Short-Time Fourier Transform and ResNet to diagnose depression from speech data," in Proc. 2021 IEEE Int. Conf. Computing (ICOCO), 2021, pp. 372–376.

[9] T. Kaneko, K. Tanaka, H. Kameoka, and S. Seki, "ISTFTNET: Fast and lightweight mel-spectrogram vocoder incorporating inverse short-time Fourier transform," in *Proc. ICASSP 2022*, 2022, pp. 6207–6211.

[10] K. Radha, D. V. Rao, K. V. K. Sai, R. T. Krishna, and A. Muneera, "Detecting autism spectrum disorder from raw speech in children using STFT layered CNN model," in *Proc. 2024 Int. Conf. Green Energy, Computing and Sustainable Technology (GECOST)*, 2024, pp. 437–441.

[11] L. Jiashen and Z. Xianwu, "Extracting speech spectrogram of speech signal based on generalized S-transform," *PLoS ONE*, vol. 20, no. 1, art. e0317362, Jan. 2025.

[12] D. T. Phan, "Comparison performance of spectrogram and scalogram as input of acoustic recognition task," in *Advances in Information and Communication (FICC 2025)*, Lecture Notes in Networks and Systems, vol. 1283, Springer, Cham, 2025.

[13] S. Kadiri, P. Alku, and B. Yegnanarayana, "Extraction and utilization of excitation information of speech: a review," *Proc. IEEE*, vol. 109, no. 12, pp. 1920–1941, 2021.

[14] M. Khan, W. Gueaieb, A. El Saddik, and S. Kwon, "MSER: Multimodal speech emotion recognition using cross-attention with deep fusion," *Expert Syst. Appl.*, vol. 245, art. 122946, 2024.

[15] E. Yeo, S. Kim, and M. Chung, "Automatic severity classification of Korean dysarthric speech using phoneme-level pronunciation features," in *Proc. Interspeech 2021*, 2021.