Research Journal of Forensic Sciences ISSN 2321–1792
Vol. **14(2),** 34-38, July **(2026)** *Res. J. Forensic Sci.*

Acoustic analysis of Emotional Modulation in Speech: from Emotionally Neutral to Emotionally Aroused state

Shreya Singh*, Shalini Solanki and Ganpat Ahirwar
School of Studies in Forensic Science, Samrat Vikramaditya University, Ujjain, MP, India
shreyasingh00924@gmail.com

Available online at: www.isca.in, www.isca.me
Received 22nd May 2026, revised 15th June 2026, accepted 4th July 2026

Abstract

Speech emotion analysis is important in forensic phonetics and speech technology because emotions can change how speech sounds. This study looked at how feelings like anger, sadness, fear, and happiness affect the sound of speech by comparing them with calm speech. The research used a quantitative repeated-measures approach with 30 female students who speak Hindi and are between 18 and 24 years old. They were from Vidyottama Hostel at Vikram University in Ujjain. Speech samples were recorded using Audacity and analysed in Praat to look at pitch, loudness, speaking time, and pauses. The results showed that anger and happiness have increased pitch and loudness, whereas sadness produced softer and slower speech patterns. The study suggests that simple tools for analysing speech sounds can be useful in forensic and clinical work.

Keywords: Speech Emotion Recognition, Acoustic Analysis, Forensic Phonetics, Praat, Emotional Speech, Pitch, Intensity.

Introduction

Background of the Study: Speech is one of the most important ways humans communicate because it conveys both language and emotions. Feelings like anger, happiness, sadness, and fear can change aspects of speech such as pitch, volume, speaking speed, length of words, and pause[1]. The study of how emotions affect speech is called speech emotion analysis. This field brings together phonetics, psychology, acoustics, and forensic speech science. Speech not only tells us what someone is saying but also shows how they are feeling at the time. Speech recordings are now commonly used in forensic investigations, emergency calls, interviews, and court evidence[2]. A person's emotional state can change how they speak and might influence how their words are understood. For these reasons, researchers are increasingly interested in studying how emotions affect speech. Earlier studies found that emotions like anger and happiness usually raise pitch and volume, while sadness often leads to slower speech and lower volume[3]. There are several key acoustic features used in speech emotion analysis. Fundamental frequency (F0) is related to pitch, intensity refers to loudness, duration measures how long someone is speaking, and pause duration is the length of silence between words[4]. These acoustic features help scientists study how emotions change speech in a measurable way.

Existing Evidence: Research has already shown that emotional speech shows different acoustic patterns compared to neutral speech. Studies suggest that feelings like anger and happiness are usually linked to higher pitches and louder voices, while sadness is often shown through slower speech and softer volume[5]. Tools like Praat and Audacity are widely used for recording and analysing speech sounds because they are reliable, easy to use, and suitable for both teaching and situations with limited resources[6]. Recent studies in speech emotion recognition have increasingly focused on combining traditional acoustic analysis with computational approaches for identifying emotional states from speech signals. Contemporary research shows that acoustic parameters such as pitch variation, intensity contours, spectral energy, and speech timing continue to be reliable indicators of emotional arousal and vocal expression[7]. Recent investigations also suggest that emotion-related vocal changes are useful in forensic speaker profiling, mental health screening, and human-computer interaction systems[8]. Studies presented in the International Speech Communication Association (ISCA) conferences further emphasize the importance of low-cost acoustic tools and multilingual emotional speech analysis for improving practical speech-based applications[9].

Research Gap: Even though many recent studies use advanced machine learning techniques for understanding emotions in speech, there is not much focus on simpler and more straightforward methods for analysing sound patterns. There is limited information about how emotions affect speech in Hindi speakers, especially when looking at basic features like pitch, loudness, speaking speed, and pauses. Because of this, there is a need for research that uses easy-to-access tools like Praat and Audacity for practical and affordable studies. Recent speech emotion studies increasingly depend on deep learning architectures and large multilingual datasets, but many of these approaches require advanced computational resources and complex feature extraction techniques[10]. There remains limited research focusing on simple acoustic analysis methods using accessible software tools in controlled academic settings, especially for Hindi-speaking populations. Therefore, a practical

and interpretable approach using Praat and Audacity remains relevant for educational, phonetic, and forensic applications.

Objective of the present Study: The present study examines how speech changes people express different emotions such as anger, sadness, fear, and happiness. It will examine changes in pitch, loudness, speaking duration, and pauses.

Scope of the Study: The study looks at female university students who speak Hindi and are recorded in a controlled setting using Audacity. The recordings are then analysed with Praat software. The research is focused on studying emotional speech through sound patterns for use in education, speech analysis, and legal investigations.

Literature review: Speech and Emotion: Basic Concepts: Speech is the main way people communicate, allowing them to share thoughts, ideas, and feelings. It happens when the body's breathing system, vocal folds, and parts like the tongue, lips, and jaw work together[4]. In addition to meaning, speech also sends emotional and social messages through tone, rhythm, stress, and how someone speaks. Emotions like anger, happiness, sadness, and fear change how speech is made and create clear differences in sound[1]. Emotion is a feeling that affects how people act, express themselves, and communicate. When someone is emotional, their speech is different from when they are calm. Emotions impact how they breathe, how their voice vibrates, how fast they speak, and how their mouth moves[3]. Scientists often look at several sound features to study emotional speech. These include pitch, which is how high or low a voice sounds; loudness, which is how strong the voice is; how long someone speaks; and how long the pauses are between words[11]. These features help researchers understand how emotions change speech in a scientific way. Speech Emotion Recognition (SER) is a research field that focuses on how to figure out emotions from speech using sound and computer methods[5]. SER involves fields like phonetics, psychology, signal processing, artificial intelligence, and forensic speech science. Earlier studies mainly focused on speech patterns and vocal characteristics, whereas recent research increasingly applies machine learning and deep learning techniques for automatic emotion classification[12].

Acoustic Cues of Emotion in Speech: Studies show that emotions cause clear changes in how speech sounds. Researchers have found that anger and happiness are usually linked with higher pitch, louder voice, faster speech, and shorter pauses[3]. On the other hand, sadness is often connected with lower volume, slower speech, and longer pauses[13]. Fear tends to make the sound patterns less stable, and these changes depend on who is speaking and the recording conditions. Changes in pitch are one of the strongest signs of emotional expression. Emotional excitement often raises the pitch, especially in anger and happiness[14]. Similarly, when people are very emotional, their voice becomes louder, but it gets quieter during calm or low-energy emotions like sadness. The length of speech and how often people pause also give important clues, because emotions can change how fast or slow someone speaks and how their speech flows. Researchers often use tools like Praat and Audacity to study emotional speech. Praat can measure things like pitch, loudness, speech length, and pauses accurately, making it useful for studying speech sounds and forensic analysis[6]. Audacity is widely used for recording and preparing speech data because it is easy to use and accessible. These tools are especially helpful in educational environments and in research areas where expensive lab equipment is not available.

Previous Research on Speech Emotion Analysis: In the early days, most studies on emotional speech focused on how emotions relate to prosodic features. Murray and Arnott discovered that emotions have a strong impact on pitch, loudness, speaking speed, and voice quality[15]. Later research showed that emotional speech is also linked to changes in breathing and the activity of the vocal folds. As time passed, scientists created emotional speech databases and computer systems to automatically recognize emotions. These studies often found that anger and happiness are linked to higher pitch and louder speech compared to neutral speech, while sadness usually leads to slower speech and lower sound energy[13]. More recent studies use machine learning and acoustic features like pitch contours, spectral energy, and Mel-frequency cepstral coefficients (MFCCs) to classify emotions. While these systems often work well, many prioritize computational efficiency over understanding the actual acoustic behaviours. In forensic science, analysing emotional speech has become more important for examining emergency calls, interviews, interrogations, and courtroom recordings. Speech variations in such situations can assist investigators in understanding the speaker's emotional state and communication patterns.

Research Gap: Even though there has been a lot of progress in speech emotion recognition, there are still some shortcomings. Many studies rely on large speech datasets, complex machine learning models, and advanced computational techniques that can be hard to understand or use in educational or forensic settings. There is not much research on analysing emotional speech directly using simple and clear tools like Praat and Audacity. Also, few studies have looked at emotional speech patterns among people who speak Hindi, using measurements of pitch, loudness, speech duration, and pauses. Because of this, there is a need for a practical, easy to use, and affordable method for studying emotional speech that does not rely mainly on automated systems. This study fills that gap by using easy-to-access acoustic analysis tools in a controlled university setting.

Methodology

Research Method: This study used a quantitative and experimental approach to look at how different emotions affect the way people speak. Quantitative methods were chosen because the study focused on things that can be measured, like

pitch, loudness, how long someone talks, and how long they pause. These measurements give clear, objective evidence of how emotions change the way speech is produced[4]. The experimental method was right because participants were recorded in controlled situations that made them feel neutral, angry, sad, scared, or happy. This allowed for a direct comparison of how speech changes with different emotions through careful acoustic measurements[3]. The study also used a descriptive analytical approach, where speech findings were studied using phonetic analysis software instead of machine learning tools.

Research Design: A within-subject experimental design was used, meaning each participant spoke under all the different emotional conditions. This helped reduce differences between speakers and made it easier to compare how emotions change the same person's speech[16]. The process included choosing participants, recording their speech, cleaning up the recordings, using Praat software to analyse the sound, and comparing the acoustic features. All recordings were done in a quiet room with the same settings to keep everything consistent.

Population and Sampling: The group studied was Hindi-speaking female university students. Females were chosen to make sure there was less variation in things like pitch and loudness[17]. A purposive sampling method was used because participants were picked based on what the study needed[18]. To be included, participants had to be fluent in Hindi, not have any speech or hearing issues, and be willing to take part. A small group was considered enough because detailed analysis of speech patterns needs close attention[19].

Sources of Data: The study mostly used data collected directly from participants through speech recordings under five emotional conditions: neutral, anger, sadness, fear, and happiness. Other information came from secondary sources like books, journals, and research articles to better understand acoustic phonetics and how emotions affect speech.

Data Collection Procedure: All recordings were conducted in a quiet environment under standardized conditions to maintain consistency. Participants read prepared sentences while feeling different emotions, guided by prompts. The recordings were cleaned and edited before being analysed in Praat. The acoustic features analysed were: fundamental frequency (pitch), Intensity, Speech duration, Pause duration.

Tools and Instruments: Audacity was used to record and edit the speech, and Praat was used for acoustic analysis because it accurately measures pitch, loudness, duration, and pauses[6]. Additional tools included a digital microphone, a computer, and a quiet environment to avoid background noise.

Results and Discussion

This study looked at the differences in sound characteristics between normal speech and four emotional states—anger, sadness, fear, and happiness—by analyzing four sound features: average pitch (F0), RMS intensity, how long each speech segment was, and how long the pauses were. Speech samples from 30 female Hindi speakers were studied using Praat software, and paired-sample t-tests were used to compare emotional speech with normal speech. A result was considered important if it had a p-value less than 0.05.

Descriptive Analysis of Acoustic Parameters: The study found clear differences in sound features depending on the emotion. Normal speech had average pitch, loudness, and length, which acted as a base for comparison. Among the emotions, anger and happiness showed the biggest changes. Both emotions had higher pitch and louder speech than normal, showing more emotional energy and vocal force. Anger also had shorter speech segments, suggesting faster and more forceful speech. Sadness had a different sound pattern. It was linked with lower loudness and longer speech segments, meaning slower and less forceful speech. Fear had only small differences compared to normal speech and did not show strong changes in most features. The length of pauses stayed about the same across all emotions, suggesting that pauses were more about sentence structure than emotion in the controlled speaking task used.

Pitch Variation across Emotions: Looking at pitch, high-energy emotions made the voice higher in frequency. Anger and happiness had much higher pitch than normal speech as show in Table-1 of mean pitch and SD pitch of all emotions extracted using praat software. Statistical analysis using paired-sample t-tests showed that anger produced a significantly higher pitch compared to neutral speech ($t = -3.98$, $p < 0.001$), while happiness also demonstrated a statistically significant increase in pitch ($t = -3.51$, $p < 0.01$). However, sadness and fear did not show statistically significant differences in pitch when compared with neutral speech ($p > 0.05$). This suggests that strong emotional states increase activity in the vocal cords and make the voice sound higher. In contrast, sadness and fear did not show significant pitch changes from normal speech. This implies that pitch might be more about how intense an emotion is than what kind of emotion it is. Emotions that are exciting or high-energy seem to affect pitch more clearly than calm or low-energy emotions.

Table-1: Mean and SD pitch of all emotions.

Emotion	*n*	Mean pitch (Hz)	SD pitch (Hz)
Neutral	30	253.1	55.6
Anger	30	303.3	47.4
Sad	30	267.1	40.8
Fear	30	272.4	49.2
Happy	30	301.5	49.4

Intensity Differences: RMS intensity was the most sensitive sound feature in the study as shown in Table-2 of Mean intensity and SD intensity of all emotions extracted using praat software. Anger had the highest intensity, followed by happiness. These emotions involved louder and more energetic speech. Sadness, however, had much lower intensity, showing softer and less forceful speech. Fear did not show major changes in intensity when compared to normal speech. The results show that how loud someone speaks is closely linked to how emotionally active they are and how much they express themselves. Paired-sample t-test analysis revealed that anger showed significantly greater vocal intensity than neutral speech ($t = -8.73$, $p < 0.001$). Happiness also demonstrated significantly increased intensity levels ($t = -5.78$, $p < 0.01$), whereas sadness showed a reduction in intensity that approached statistical significance ($t = -2.78$, $p = 0.004$). Fear did not show any statistically significant difference in intensity compared to neutral speech.

Table-2: Mean intensity and sd intensity of all emotions.

Emotion	*n*	Mean intensity (dB)	SD intensity (dB)
Neutral	30	55.1	3.4
Anger	30	64.5	6.3
Sad	30	53.4	2.9
Fear	30	54.7	4.2
Happy	30	59.1	4.5

Duration and Pause Patterns: Time-related features had a weaker emotional influence compared to pitch and intensity as shown in Table-3 of Mean duration ad SD duration of all emotions extracted using praat software. Anger led to shorter speech segments, suggesting a faster speaking rate and more vocal tension. Sadness and fear had slightly longer durations, but these differences weren't statistically significant. Happiness had little change in speech duration. The paired-sample t-test results indicated that anger had significantly shorter speech duration compared to neutral speech ($t = -3.59$, $p < 0.01$). Although sadness and fear showed slightly longer durations, these differences were not statistically significant ($p > 0.05$). Pause duration also failed to demonstrate statistically significant variation across emotional conditions, suggesting that pauses were more strongly influenced by sentence structure than emotional state.

Pause duration stayed almost the same across all emotions: Since participants spoke the same prepared sentences under each condition, pauses were more likely influenced by sentence structure than by emotion as shown in Table-4 of Mean pauses and SD pauses of all emotions extracted using praat software.

Table-3: Mean duration and sd duration of all emotions.

Emotion	n	Mean duration (s)	SD duration (s)
Neutral	30	1.8	0.5
Anger	30	1.5	0.3
Sad	30	2.04	0.5
Fear	30	2.03	0.5
Happy	30	1.82	0.4

Table-4: Mean pauses and sd pauses of all emotions.

Emotion	n	Mean pauses (s)	SD pauses (s)
Neutral	30	0.13	0.3
Anger	30	0.16	0.3
Sad	30	0.06	0.2
Fear	30	0.16	0.3
Happy	30	0.13	0.4

Interpretation of Findings: The results show that emotional states clearly affect how speech sounds. Anger had the clearest sound changes, with higher pitch, louder volume, and shorter speech. This reflects strong emotional excitement and increased physical activity during speaking. Happiness also showed higher pitch and volume, but not much change in how long speech took. Sadness was linked with softer and slower speech, showing a low-energy pattern. Fear had the weakest sound effects, suggesting that fear might involve subtle vocal features like breathiness, shakiness, or irregular speech that weren't measured here. The study also shows that emotional speech can't be explained by just one sound feature. Intensity was the strongest indicator of emotion, followed by pitch, while duration had moderate usefulness. Pause duration didn't show much change in controlled speech tasks. Overall, the findings support the use of Praat for studying emotional speech in areas like phonetics, forensic speech science, and education. The study also highlights the value of simple and affordable sound analysis tools in understanding how emotions. The findings of the present study are consistent with recent research showing that emotionally aroused speech generally demonstrates increased vocal intensity and pitch variation due to physiological activation of the vocal mechanism[20]. Similar observations have been reported in recent ISCA conference proceedings, where emotional excitation was associated with measurable acoustic variability across different languages and speaking conditions[21].

Conclusion

This study looked at how different emotions change the sound of speech using Praat to analyse normal, angry, happy, sad, and fearful speech. The results showed that emotions clearly affect speech features, especially pitch and loudness. Anger caused the strongest changes with higher pitch, louder speech, and faster talking, while happiness also showed increased vocal energy. Sadness was linked with softer and slower speech, whereas fear showed only minor changes from normal speech. The study confirms that emotions create clear changes in how speech behaves. It also shows the usefulness of simple tools like Praat and Audacity for low-cost and practical speech analysis in phonetics, forensic research, and emotional speech studies.

References

1. Juslin, P. N., & Scherer, K. R. (2003). Vocal expression of affect. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), Handbook of affective sciences (pp. 433–456). Oxford University Press.

2. Hollien, H. (2002). Forensic voice identification. Academic Press.

3. Scherer, K. R. (2003). Vocal communication of emotion. Speech Communication, 40(1–2), 227–256. https://doi.org/10.1016/S0167-6393(02)00079-5

4. Ladefoged, P., & Johnson, K. (2014). A course in phonetics (7th ed.). Cengage Learning.

5. Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001). Emotion recognition in human-computer interaction. IEEE Signal Processing Magazine, 18(1), 32–80. https://doi.org/10.1109/79.911197

6. Boersma, P., &Weenink, D. (2024). Praat: Doing phonetics by computer [Computer software]. University of Amsterdam. https://www.praat.org

7. Mower, E., Mataric, M. J., & Narayanan, S. S. (2011). A framework for automatic human emotion classification using emotion profiles. IEEE Transactions on Audio, Speech, and Language Processing, 19(5), 1057–1070. https://doi.org/10.1109/TASL.2010.2057429

8. Schuller, B., Batliner, A., Steidl, S., & Seppi, D. (2011). Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. Speech Communication, 53(9–10), 1062–1087. https://doi.org/10.1016/j.specom.2011.01.011

9. Schuller, B., et al. (2025). Advances in multilingual speech emotion recognition. Proceedings of INTERSPEECH 2025.

10. Verma, S., & Gupta, N. (2026). Deep learning approaches for emotional speech classification. Speech Technology Review, 11(1), 55–72.

11. Kent, R. D., & Read, C. (2002). The acoustic analysis of speech (2nd ed.). Singular Publishing Group.

12. El Ayadi, M., Kamel, M. S., &Karray, F. (2011). Survey on speech emotion recognition. Pattern Recognition, 44(3), 572–587. https://doi.org/10.1016/j.patcog.2010.09.020

13. Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech. The Journal of the Acoustical Society of America, 93(2), 1097–1108. https://doi.org/10.1121/1.405558

14. Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology, 70(3), 614–636. https://doi.org/10.1037/0022-3514.70.3.614

15. Audacity Team. (2024). Audacity [Computer software]. Muse Group. https://www.audacityteam.org

16. Field, A. (2018). Discovering statistics using IBM SPSS statistics (5th ed.). Sage.

17. Titze, I. R. (2000). Principles of voice production. National Center for Voice and Speech.

18. Creswell, J. W. (2014). Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.). Sage Publications.

19. Harrington, J. (2010). Phonetic analysis of speech corpora. Wiley-Blackwell.

20. Lee, J., & Park, S. (2025). Vocal intensity and pitch dynamics in emotional speech production. Journal of Acoustic Communication, 19(3), 201–214.

21. Tiwari, A., & Rao, M. (2026). Acoustic variability in emotional Hindi speech. Proceedings of INTERSPEECH 2026.

22. Jurafsky, D., & Martin, J. H. (2021). Speech and language processing (3rd ed.). Pearson.