Advances In Medical Informatics (ISSN 2819-8298) https://aimi.quantechquest.com
2025; 1: 7 https://doi.org/10.22034/AIMI.2025.12.7 Open Access

How neurological disorders affect acoustic features of speech? A data mining approach

Mohammadjavad Sayadi[1]*, Farhad Torabinezhad[2], Gholamreza Bayazian[3], Somayeh Abedian[4]

[1]Department of Computer Engineering, Faculty of Ilam, Technical and Vocational University, Iran
[2]Rehabilitation Research Center, Department of Speech Therapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran
[3]ENT and Head & Neck Research Center, The Five Sense Health Institute, Rasoul Akram Medical Complex, Iran University of Medical Sciences, Tehran, Iran
[4]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Canada

Article Info

Article type:
Research

Article History:
Received: 2025-04-25
Accepted: 2025-05-30
Published: 2025-06-25

**** Corresponding author:***
Mohammadjavad Sayadi

Department of Computer Engineering, Faculty of Ilam, Technical and Vocational University, Iran

Email: mjsayadi@nus.ac.ir

Keywords:
Neurological Disorders
Acoustic Features
Speech
Data Mining

A B S T R A C T

Introduction: Neurological disorders often manifest as alterations in speech production, affecting acoustic features such as pitch, rhythm, articulation, and voice quality. These speech changes can serve as non-invasive biomarkers for early diagnosis and monitoring. This study aims to investigate how various neurological disorders affect the acoustic features of speech across multiple speaking styles using a data mining approach.

Material and Methods: We collected speech recordings from 383 participants, including 264 individuals diagnosed with different neurological disorders and 119 healthy controls. Participants read a standardized text in four speaking styles: questioning, excited, angry, and happy. A comprehensive set of acoustic features, covering prosodic, spectral, voice quality, formant, and temporal parameters, was extracted. Deep learning models were trained separately for each speaking style to classify neurological status. Feature importance analyses were conducted to identify key acoustic indicators of neurological impairment.

Results: The deep learning models achieved classification accuracies ranging from 82.7% to 87.9% across speaking styles, with the highest performance observed in the angry speech condition. Prosodic features, particularly fundamental frequency and speaking rate, alongside voice quality measures such as jitter and shimmer, emerged as the most discriminative features. Distinct acoustic profiles were identified for different neurological disorders, and several features correlated significantly with clinical severity scores. Speaking style influenced the detectability of speech impairments, underscoring the value of analyzing diverse speech contexts.

Conclusion: Our findings demonstrate that neurological disorders induce characteristic alterations in acoustic speech features that can be effectively captured using data mining and deep learning techniques. Incorporating multiple speaking styles enhances diagnostic sensitivity. This approach holds promise for developing accessible, non-invasive speech-based biomarkers to support early diagnosis and monitoring of neurological diseases.

Cite this paper as:
Sayadi M, Torabinezhad F, Bayazian G, Abedian S. How neurological disorders affect acoustic features of speech? A data mining approach. Adv Med Inform. 2025; 1: 7.

INTRODUCTION

Neurological disorders are defined as conditions that affect the human nervous system and include neuro-infectious diseases, Alzheimer's disease and other dementias, encephalitis, central nervous system cancers, stroke, epilepsy, Parkinson's disease, tetanus, multiple sclerosis, meningitis, and other neurological disorders [1]. These conditions can lead to brain damage and impair cognitive, sensory, socio-

Copyright© 2025 Mohammadjavad Sayadi, Published by Advances in Medical Informatics. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY) License (http://creativecommons.org/).

emotional, motor, and other functions and behaviors [2]. As a result, neurological disorders represent a major and growing health challenge, placing a significant burden on society and healthcare systems worldwide [3]. Notably, they are the second leading cause of death globally [4].

With increasing life expectancy, neurological disorders have become especially common among older adults. A global review and comparison of data indicate that these disorders are more prevalent in low- and middle-income countries, contributing to significant health disparities [5]. According to statistics, neurological disorders caused approximately 10 million deaths worldwide in 2019, with conditions such as Parkinson's disease, dementias, meningitis, migraine, autoimmune disorders, and epilepsy each contributing substantially to the global loss of disability-adjusted life years (DALYs) [6, 7]. Given the immense impact and burden of neurological disorders on individuals, families, and societies, the World Health Organization established the Intersectoral Global Action Plan on epilepsy and other neurological disorders 2022–2031 (IGAP) to address these challenges at a global level [8].

Human speech serves as a primary means of communication and emotional expression, relying on the intricate coordination of cognitive, linguistic, and motor systems regulated by the brain [9]. The acoustic properties of speech, such as pitch, rhythm, articulation, and voice quality, are closely linked to the functional integrity of the nervous system. Consequently, neurological disorders often produce changes in speech that reflect deficits in cognitive-linguistic processing [10]. These changes may manifest as incorrect articulation, slowed speech, altered prosody, or reduced fluency. Moreover, evidence shows that an individual's speech style (e.g., questioning, emotional, fearful) and communicative context can further influence speech production [11]. This high degree of sensitivity and variability presents challenges for the accurate collection and analysis of speech data.

Previous studies have established that speech disorders are among the symptoms of neurological dysfunction. However, their interpretation in clinical practice is often subjective, relying on the clinician's perspective, and lacks the detailed granularity needed for precise diagnosis [12]. Furthermore, collecting accurate and reliable speech data is inherently challenging. Since initial speech changes may precede other clinical symptoms, speech analysis holds promise for early diagnosis and improved treatment management [13]. Human auditory perception is limited in its ability to detect subtle changes in speech, underscoring the need for objective, quantitative analysis of acoustic features to enhance clinical assessments [14].

Advances in digital technology and signal processing have now made it possible to collect high-quality voice data and analyze it rapidly and accurately [15]. Speech quality can be quantified using a range of devices, including sensors, smartphones, and audio-equipped systems, allowing for comprehensive and longitudinal monitoring of patients [16]. In parallel, artificial intelligence (AI), machine learning, and data mining have emerged as transformative tools in healthcare. Recent progress in these fields has enabled the systematic analysis of complex speech patterns, revealing acoustic features that may not be detectable with traditional statistical methods [17, 18]. Sophisticated speech processing tools and powerful machine learning algorithms are now capable of analyzing a wide array of acoustic features, facilitating the identification of subtle markers of neurological dysfunction [19].

Systematic reviews have demonstrated the effectiveness of machine learning algorithms for the early detection of mental, neurological, and laryngeal disorders using speech signals, highlighting their potential for non-invasive, rapid, and scalable diagnostics [20]. Data mining approaches have also proven valuable in distinguishing laryngeal disorders and predicting clinical outcomes based on acoustic speech features [21]. Despite these advances, most existing research has focused on neutral or read speech, with limited exploration of how different speaking styles—such as questioning, excited, angry, or happy tones—may interact with neurological impairment and affect acoustic features. Given that emotional and contextual variability can modulate speech production and potentially unmask deficits not apparent in neutral speech, there is a critical need for comprehensive, data-driven studies that examine the impact of neurological disorders on a wide array of acoustic features across diverse speaking styles.

The present study addresses this gap by systematically investigating how various neurological disorders affect the acoustic features of speech using a data mining and deep learning approach. We analyzed speech samples from individuals with and without neurological disorders, elicited in four distinct speaking styles: questioning, excited, angry, and happy. A comprehensive set of acoustic features was extracted, encompassing prosodic, spectral, voice quality, formant, and temporal parameters. Deep learning models were developed to classify neurological disorders based on these features, and feature importance analyses were conducted to identify the most informative acoustic markers. By integrating advanced analytics with nuanced speech data, our study aims to advance the field of speech-based biomarkers and contribute to improved diagnosis and care for individuals with neurological disorders.

MATERIAL AND METHODS

This research aimed to investigate the impact of various neurological disorders on speech acoustics using a data mining approach, employing deep learning models to analyze different speaking styles.

Study design and participants

We conducted a cross-sectional study involving a total of 383 participants. The study group comprised 264 individuals diagnosed with various neurological disorders, while the control group consisted of 119 healthy individuals without any known neurological conditions. Participants were recruited from multiple neurological clinics and community centers across the region. The study protocol was approved by the institutional review board, and all participants provided written informed consent before enrollment.

Inclusion criteria

- Age range: 18-80 years
- Ability to read and speak fluently in the language of the provided text
- For the neurological disorder group: confirmed diagnosis of a neurological condition by a licensed neurologist
- For the control group: absence of any known neurological disorders

Exclusion criteria:

- Presence of severe cognitive impairment that would interfere with the ability to follow instructions
- History of speech or language disorders unrelated to neurological conditions
- Acute illness or hospitalization within the past month

Neurological disorders

The 264 participants in the neurological disorder group represented a diverse range of conditions, including but not limited to:

1. Parkinson's disease
2. Alzheimer's disease and other dementias
3. Multiple sclerosis
4. Amyotrophic lateral sclerosis (ALS)
5. Huntington's disease
6. Stroke
7. Traumatic brain injury
8. Epilepsy
9. Cerebellar ataxia
10. Progressive supranuclear palsy (PSP)

Each participant's specific diagnosis was confirmed and documented by their treating neurologist. The severity and duration of the disorder were also recorded to account for potential variations in speech patterns across different stages of disease progression.

Speech data collection

Text selection

A carefully curated text passage was selected for the speech recording task. The passage was designed to include a variety of phonemes, stress patterns, and prosodic features representative of natural speech. The text was reviewed by a panel of linguists and speech pathologists to ensure its suitability for acoustic analysis across different speaking styles.

Recording procedure

All participants were recorded in a quiet room with minimal background noise. A high-quality directional microphone (Shure SM58) was used, connected to a digital audio interface (Focusrite Scarlett 2i2) to ensure consistent and clear recordings. The audio was captured at a sampling rate of 44.1 kHz with 16-bit depth.

Participants were seated comfortably and positioned approximately 20 cm from the microphone. They were given time to familiarize themselves with the text before the recording session began. Clear instructions were provided on how to perform each speaking style.

Speaking styles

Each participant was asked to read the same text passage in four distinct speaking styles:

1. Question: Participants were instructed to read the text as if they were asking questions, emphasizing rising intonation at appropriate points.
2. Excited: Participants were asked to read the text with an excited tone, conveying enthusiasm and heightened emotion.
3. Angry: Participants were instructed to read the text with an angry or frustrated tone, emphasizing intensity and sharp articulation.
4. Happy: Participants were asked to read the text with a happy and cheerful tone, focusing on positive emotion and upbeat prosody.

The order of speaking styles was randomized for each participant to minimize order effects. Participants

were given short breaks between each style to reset their vocal patterns and prevent fatigue.

Acoustic feature extraction

Following the data collection phase, we employed a comprehensive acoustic feature extraction process to capture a wide range of speech characteristics.

Preprocessing

Before feature extraction, all audio recordings underwent a preprocessing stage:

1. Noise reduction using spectral subtraction
2. Silence removal at the beginning and end of each recording
3. Amplitude normalization to ensure consistent volume across all samples

Feature categories

We extracted an extensive set of acoustic features, broadly categorized into the following groups:

Prosodic features:

- Fundamental frequency (F0) statistics (mean, median, standard deviation, range)
- Speaking rate (syllables per second)
- Articulation rate (excluding pauses)
- Rhythm metrics (Pairwise Variability Index, nPVI)

Spectral features:

1. Mel-Frequency Cepstral Coefficients (MFCCs) and their derivatives
2. Spectral centroid, flux, and rolloff
3. Harmonic-to-Noise Ratio (HNR)

Voice quality features:

- Jitter (cycle-to-cycle variation in fundamental frequency)
- Shimmer (cycle-to-cycle variation in amplitude)
- Harmonics-to-Noise Ratio (HNR)
- Cepstral Peak Prominence (CPP)

Formant features:

- First three formant frequencies (F1, F2, F3) and their bandwidths
- Formant dispersion and spacing

Energy and intensity features:

- Root Mean Square (RMS) energy
- Energy contour statistics

Temporal features:

- Voice Onset Time (VOT) for stop consonants
- Pause characteristics (number, duration, distribution)

Feature extraction tools

We utilized a combination of established speech processing libraries and custom scripts for feature extraction:

1. Praat: Used for extracting prosodic features, formants, and voice quality measures.
2. OpenSMILE: Employed for extracting a large set of low-level descriptors and functional.
3. Python libraries (librosa, pysptk): Used for spectral feature extraction and additional custom features.

Deep learning model architecture

For each speaking style (question, excited, angry, and happy), we developed a separate deep learning model to analyze the acoustic features and classify the presence and type of neurological disorder. The model architecture was designed to capture both temporal and spectral characteristics of speech.

Model structure

We implemented a hybrid model combining convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, inspired by recent advancements in speech recognition and analysis. The model structure for each speaking style was as follows:

1. Input layer: Accepts the time-series of extracted acoustic features.
2. Convolutional layers: Two 1D convolutional layers with 64 and 128 filters respectively, each followed by batch normalization, ReLU activation, and max pooling. These layers help in capturing local spectral patterns.
3. LSTM layers: Two bidirectional LSTM layers with 128 and 64 units respectively, to model temporal dependencies in the speech signal.
4. Attention mechanism: An attention layer to focus on the most relevant parts of the speech signal for classification.
5. Dense layers: Two fully connected layers with 128 and 64 units, using ReLU activation and dropout (0.5) for regularization.
6. Output layer: A softmax layer for multi-class

classification of neurological disorders.

Model training

The models were trained using the following parameters:

- Optimizer: Adam with learning rate of 0.001
- Loss function: Categorical cross-entropy
- Batch size: 32
- Epochs: 100 with early stopping based on validation loss
- Data split: 70% training, 15% validation, 15% testing

To address class imbalance, we employed weighted classes during training, assigning higher weights to underrepresented disorders.

Data analysis

Feature importance analysis

To determine which acoustic features were most affected by neurological disorders, we employed several techniques:

1. SHAP (SHapley Additive exPlanations) values: Calculated to understand the contribution of each feature to the model's predictions across different disorders and speaking styles3.
2. Permutation importance: Assessed the importance of each feature by measuring the decrease in model performance when the feature is randomly shuffled.
3. Gradient-based methods: Utilized integrated gradients to attribute the prediction of the deep learning models to their input features.

Statistical analysis

We conducted comprehensive statistical analyses to complement the machine learning approach:

1. ANOVA: To compare acoustic features across different neurological disorders and the control group.
2. Post-hoc tests: Tukey's HSD test for pairwise comparisons between groups when ANOVA showed significant differences.
3. Correlation analysis: Spearman's rank correlation to examine relationships between acoustic features and disorder severity.
4. Effect size calculation: Cohen's d to quantify the magnitude of differences in acoustic features between groups.

Ethical considerations

All participants provided informed consent, and their data was anonymized to ensure confidentiality. Participants were informed of their right to withdraw from the study at any time without consequence.

Data management and security

All audio recordings and extracted features were stored on secure, encrypted servers. Access to the data was restricted to authorized research team members only. Data processing and analysis were performed on dedicated workstations with appropriate security measures in place.

This comprehensive methodology allowed us to systematically investigate the impact of various neurological disorders on speech acoustics across different speaking styles, leveraging advanced data mining and deep learning techniques to uncover subtle yet significant patterns in speech production.

RESULTS

This section presents the detailed findings from our analysis of how neurological disorders affect acoustic features of speech across four different speaking styles: question, excited, angry, and happy. We report on the performance of the deep learning models, the differential impact of neurological disorders on acoustic features, and the feature importance results that highlight which acoustic parameters are most affected.

Participant demographics and data summary

The study included 383 participants, with 264 diagnosed with various neurological disorders and 119 healthy controls. The demographic characteristics were balanced across groups with no significant differences in age or gender distribution (Table 1).

Table 1: Participant demographics

Group	N	Mean Age (SD)	Gender (M/F)
Neurological	264	58.3 (12.4)	140/124
Control	119	56.7 (11.8)	62/57

Acoustic feature variations across groups

Statistical comparisons

Using ANOVA followed by Tukey's HSD post-hoc tests, we identified significant differences in multiple acoustic features between neurological disorder groups and controls across all speaking styles.

- Fundamental frequency (F0): Mean F0 was significantly lower in Parkinson's disease and ALS groups compared to controls, particularly pronounced in the angry and

question styles ($p<0.001$).

- Jitter and Shimmer: Voice quality measures such as jitter and shimmer were elevated in all neurological disorder groups, indicating increased vocal instability ($p<0.001$).
- MFCCs: Several MFCC coefficients showed significant alterations, reflecting changes in spectral properties of speech. These were most notable in Alzheimer's disease and stroke groups.
- Speaking rate: A marked reduction in speaking rate was observed in multiple sclerosis and stroke patients, especially in the happy and excited styles ($p<0.01$).
- Formant frequencies: Formant shifts, particularly F2, were significant in cerebellar ataxia and Huntington's disease groups, suggesting articulatory impairments.

Effect sizes

Cohen's d effect sizes indicated large effects ($d>0.8$) for jitter, shimmer, and F0 range in Parkinson's and ALS groups, moderate effects ($d=0.5$-0.8) for MFCC changes in Alzheimer's and stroke, and small to moderate effects for speaking rate and formants in other disorders.

Deep learning model performance

Separate deep learning models were trained for each speaking style to classify neurological disorders based on acoustic features.

Table 2: Deep learning model performance metrics by speaking style

Speaking Style	Accuracy	Precision	Recall	F1 Score
Question	85.2	0.84	0.83	0.83
Excited	82.7	0.81	0.80	0.80
Angry	87.9	0.86	0.88	0.87
Happy	83.5	0.82	0.81	0.81

According to Table 2, the Angry style model achieved the highest classification accuracy (87.9%), suggesting that speech produced with an angry tone reveals more distinctive acoustic markers of neurological disorders. The question style model also performed strongly, while excited and happy styles showed slightly lower but still robust performance.

Feature importance and interpretation

Using SHAP values and permutation importance, we identified the most influential acoustic features contributing to model predictions.

Notably, prosodic features such as mean fundamental frequency and speaking rate consistently ranked highest across all speaking styles, underscoring their sensitivity to neurological impairment. Voice quality parameters like jitter and shimmer were also critical, reflecting the vocal instability common in many neurological disorders (Table 3).

Table 3: Summary of top acoustic features influencing classification

Feature Category	Top Features Identified	Importance Rank (Average)
Prosodic	Mean F0, F0 range, Speaking rate	1
Voice Quality	Jitter, Shimmer, Harmonics-to-Noise Ratio (HNR)	2
Spectral	MFCC 1, MFCC 3, Spectral centroid	3
Formant	F2 frequency, F1 bandwidth	4
Temporal	Pause duration, Voice Onset Time	5

Disorder-specific acoustic profiles

Our models and feature analyses allowed us to delineate disorder-specific acoustic profiles:

- Parkinson's disease: Characterized by reduced mean F0, increased jitter and shimmer, and decreased speaking rate, particularly evident in angry and question styles.
- Alzheimer's disease: Marked by altered MFCC patterns and increased pause durations, reflecting cognitive-linguistic impairments impacting speech fluency.
- Multiple sclerosis: Showed slowed articulation rate and increased voice breaks, especially in excited and happy styles.
- Stroke: Exhibited significant formant frequency shifts and reduced spectral centroid values, indicating articulatory deficits.
- ALS: Presented with extreme voice quality degradation (high jitter/shimmer) and reduced F0 range.

Correlations with disease severity

Spearman correlation analyses revealed significant associations between acoustic features and disease severity scores (e.g., UPDRS for Parkinson's, EDSS for MS):

- Negative correlations between mean F0 and Parkinson's severity ($r= -0.62$, $p<0.001$)

- Positive correlations between jitter and ALS severity ($r= 0.58$, $p<0.001$)
- Increased pause duration correlated with Alzheimer's severity ($r=0.54$, $p<0.01$)

These findings support the potential of acoustic features as biomarkers for monitoring disease progression.

Cross-style comparisons

Comparing models and features across speaking styles revealed:

- The angry style elicited the most pronounced acoustic differences and highest classification accuracy.
- The question style also performed well, possibly due to its natural prosodic variability.
- Excited and happy styles showed more subtle differences but still contributed valuable information.

This suggests that emotional and interrogative speech styles may be particularly sensitive for detecting neurological speech impairments.

DISCUSSION

This study investigated the impact of neurological disorders on acoustic features of speech using a data mining approach, analyzing speech samples collected in four distinct speaking styles, question, excited, angry, and happy. By extracting a broad range of acoustic features and employing deep learning models tailored to each speaking style, we were able to identify characteristic speech alterations associated with different neurological conditions. Our findings contribute significant insights into the complex interplay between neurological impairment and speech production, with important implications for diagnosis, monitoring, and therapeutic interventions.

Summary of key findings

Our results demonstrated that neurological disorders profoundly affect multiple acoustic dimensions of speech, including prosody, voice quality, spectral characteristics, and temporal features. The deep learning models achieved high classification accuracies (up to 87.9%) in distinguishing individuals with neurological disorders from healthy controls, particularly when analyzing speech produced in the angry and question styles. Prosodic features such as fundamental frequency (F0) and speaking rate, along with voice quality measures including jitter and shimmer, emerged as the most informative predictors across all speaking styles. Additionally, we identified disorder-specific acoustic profiles, reflecting the unique pathophysiological mechanisms underlying speech impairments in conditions such as Parkinson's disease, Alzheimer's disease, multiple sclerosis, stroke, and amyotrophic lateral sclerosis (ALS). Importantly, several acoustic features showed significant correlations with clinical severity scores, underscoring their potential as non-invasive biomarkers for disease progression.

Interpretation of acoustic feature alterations

Prosodic features

Prosody, encompassing pitch, rhythm, and intonation, is critical for conveying linguistic and emotional information. Our finding of reduced mean F0 and F0 range in Parkinson's disease and ALS aligns with the well-documented hypophonia and monopitch observed in these disorders. These alterations likely reflect basal ganglia dysfunction and impaired motor control of the laryngeal muscles, resulting in diminished vocal fold vibration variability. The prominence of prosodic changes in the angry and question styles suggests that emotional and interrogative speech may exacerbate or reveal underlying deficits more clearly, possibly due to the increased demands on pitch modulation and intonation patterns in these styles.

The observed slowing of speaking and articulation rates in multiple sclerosis and stroke patients reflects motor and cognitive impairments affecting speech planning and execution. Reduced speech rate may also indicate compensatory strategies to maintain intelligibility in the presence of dysarthria or other motor speech disorders.

Voice quality features

Elevations in jitter and shimmer across neurological groups indicate increased cycle-to-cycle variability in frequency and amplitude, respectively, which are hallmarks of vocal instability and dysphonia. These findings are consistent with prior studies reporting breathiness, roughness, and hoarseness in neurological dysarthrias. The increased harmonic-to-noise ratio (HNR) in some groups may reflect compensatory changes or differences in vocal fold closure patterns. The sensitivity of these features to neurological impairment highlights their utility for early detection and monitoring.

Spectral and formant features

Alterations in Mel-frequency cepstral coefficients (MFCCs) and formant frequencies reflect changes in the spectral envelope and articulatory configuration of speech sounds. For example, shifts in F2 and F1 frequencies in cerebellar ataxia and Huntington's disease suggest impaired tongue and jaw movements, consistent with ataxic and chronic dysarthrias. These spectral changes contribute to

reduced speech clarity and intelligibility and may serve as objective markers of articulatory dysfunction.

Temporal features

Increased pause durations and altered voice onset times observed particularly in Alzheimer's disease and stroke patients likely reflect cognitive-linguistic deficits and impaired motor timing. These temporal disruptions can degrade speech fluency and naturalness, impacting communication effectiveness.

Speaking style as a modulating factor

Our approach of analyzing speech in multiple speaking styles revealed that emotional and interrogative speech (angry and question styles) provided richer acoustic cues for detecting neurological impairments. This may be due to the greater prosodic variability and expressive demands inherent in these styles, which unmask subtle deficits not as apparent in neutral or positive affective speech. The relatively lower classification performance for excited and happy styles suggests that positive emotions may mask or compensate for some speech impairments, or that these styles inherently involve less prosodic contrast. These findings underscore the importance of including diverse speaking styles in speech assessments to maximize diagnostic sensitivity.

Deep learning and feature importance

The use of hybrid CNN-LSTM architectures enabled effective modeling of both local spectral patterns and long-range temporal dependencies in acoustic features, contributing to the high classification accuracies achieved. The integration of attention mechanisms further enhanced model interpretability by highlighting the most relevant speech segments and features.

Feature importance analyses consistently identified prosodic and voice quality features as primary contributors to classification, corroborating their clinical relevance. The identification of disorder-specific acoustic signatures supports the potential for developing tailored diagnostic tools that can differentiate between neurological conditions based on speech patterns alone.

Clinical and research implications

Our findings have several important implications:

1. Non-invasive biomarkers: Acoustic features of speech, particularly prosody and voice quality, show promise as accessible, cost-effective biomarkers for early detection and monitoring of neurological disorders. This could facilitate timely intervention and improve patient outcomes.
2. Remote and continuous monitoring: Speech-based assessments can be integrated into telemedicine platforms and mobile applications, enabling remote monitoring of disease progression and treatment response without requiring frequent clinic visits.
3. Personalized therapy: Understanding disorder-specific speech impairments can guide the development of targeted speech therapy protocols, optimizing rehabilitation strategies.
4. Multimodal integration: Combining speech analysis with other biomarkers (e.g., imaging, genetics) could enhance diagnostic accuracy and provide a more comprehensive understanding of neurological disorders.

Comparison with Previous Studies

Several prior studies employing acoustic analysis and machine learning have reported high accuracy in detecting neurological disorders such as Parkinson's disease and ALS, with performance metrics comparable to or exceeding those found in our study [9, 18]. For example, recent work demonstrated classification accuracies up to 88.6% in differentiating progressive supranuclear palsy (PSP) and multiple system atrophy (MSA), closely matching our results in the angry and question styles. Other research focusing on disease progression monitoring in Parkinson's disease also highlights the importance of acoustic features and deep learning models, reinforcing the clinical utility of our approach. Our study's novelty lies in the comprehensive evaluation of multiple speaking styles and a broad range of acoustic features, providing deeper insight into how emotional and contextual variability modulates speech impairments in neurological disorders [11, 12].

Limitations

Despite promising results, limitations include sample size and focus on specific speaking styles. Future research should expand sample diversity, incorporate additional speaking styles, and explore multimodal data integration to enhance diagnostic accuracy and applicability. Longitudinal studies are also needed to assess changes in speech features over disease progression.

Future directions

Building on our results, future research should explore:

- Longitudinal speech monitoring: Tracking acoustic changes over time to identify early markers of disease onset and progression.
- Multilingual and cross-cultural validation:

Extending analyses to diverse languages and populations.

- Integration with other modalities: Combining speech data with neuroimaging, genetic, and clinical data for multimodal biomarker development.
- Real-world speech analysis: Analyzing spontaneous conversational speech and natural emotional expressions to enhance ecological validity.
- Development of clinical tools: Translating models into user-friendly applications for clinicians and patients.

CONCLUSION

In conclusion, our study demonstrates that neurological disorders exert significant and characteristic effects on the acoustic features of speech, which can be effectively captured and analyzed using advanced data mining and deep learning techniques. The differential impact of neurological conditions across speaking styles highlights the importance of incorporating diverse speech contexts in assessment protocols. Prosodic and voice quality features emerged as robust markers of neurological impairment, with potential applications in diagnosis, monitoring, and therapy. These findings advance the field of speech-based biomarkers and pave the way for innovative, non-invasive tools to improve neurological care.

AUTHOR'S CONTRIBUTION

All authors contributed to the literature review, design, data collection, drafting the manuscript, read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this study.

ETHICAL APPROVAL

This study was approved by the institute review board at Iran University of Medical Sciences (IR.IUMS.REC.1400.327).

FINANCIAL DISCLOSURE

No financial interests related to the material of this manuscript have been declared.

REFERENCES

1. Feigin VL, Nichols E, Alam T, Bannick MS, Beghi E, Blake N, et al. Global, regional, and national burden of neurological disorders, 1990–2016: A systematic analysis for the global burden of disease study 2016. Lancet Neurol. 2019; 18(5): 459-80. PMID: 30879893 DOI: 10.1016/S1474-4422(18)30499-X [PubMed]
2. Feigin VL, Vos T, Nichols E, Owolabi MO, Carroll WM, Dichgans M, et al. The global burden of neurological disorders: Translating evidence into policy. Lancet Neurol. 2020; 19(3): 255-65. PMID: 31813850 DOI: 10.1016/S1474-4422(19)30411-9 [PubMed]
3. Hecker P, Steckhan N, Eyben F, Schuller BW, Arnrich B. Voice analysis for neurological disorder recognition–a systematic review and perspective on emerging trends. Front Digit Health. 2022; 4: 842301. PMID: 35899034 DOI: 10.3389/fdgth.2022.842301 [PubMed]
4. Mayer G, Happe S, Evers S, Hermann W, Jansen S, Kallweit U, et al. Insomnia in neurological diseases. Neurol Res Pract. 2021; 3: 1-12. PMID: 33691803 DOI: 10.1186/s42466-021-00106-3 [PubMed]
5. Gilmour GS, Nielsen G, Teodoro T, Yogarajah M, Coebergh JA, Dilley MD, et al. Management of functional neurological disorder. J Neurol. 2020; 267: 2164-72. PMID: 32193596 DOI: 10.1007/s00415-020-09772-w [PubMed]
6. Ding C, Wu Y, Chen X, Chen Y, Wu Z, Lin Z, et al. Global, regional, and national burden and attributable risk factors of neurological disorders: The global burden of disease study 1990–2019. Front Public Health. 2022; 10: 952161. PMID: 36523572 DOI: 10.3389/fpubh.2022.952161 [PubMed]
7. Zhang R, Liu H, Pu L, Zhao T, Zhang S, Han K, et al. Global burden of ischemic stroke in young adults in 204 countries and territories. Neurology. 2023; 100(4): e422-34. PMID: 36307222 DOI: 10.1212/WNL.0000000000201467 [PubMed]
8. Steinmetz JD, Seeher KM, Schiess N, Nichols E, Cao B, Servili C, et al. Global, regional, and national burden of disorders affecting the nervous system, 1990–2021: A systematic analysis for the global burden of disease study 2021. Lancet Neurol. 2024; 23(4): 344-81. PMID: 38493795 DOI: 10.1016/S1474-4422(24)00038-3 [PubMed]
9. Saeedi S, Hetjens S, Grimm M, Latoszek BBv. Acoustic speech analysis in Alzheimer's disease: A systematic review and meta-analysis. J Prev Alzheimers Dis. 2024; 11(6): 1789-97. PMID: 39559890 DOI: 10.14283/jpad.2024.132 [PubMed]
10. Cho S, Cousins KAQ, Shellikeri S, Ash S, Irwin DJ, Liberman MY, et al. Lexical and acoustic speech features relating to Alzheimer disease pathology. Neurology. 2022; 99(4): e313-22. PMID: 35487701 DOI: 10.1212/WNL.0000000000200581 [PubMed]
11. Noffs G, Boonstra FM, Perera T, Kolbe SC, Stankovich J, Butzkueven H, et al. Acoustic speech analytics are predictive of cerebellar dysfunction in multiple sclerosis. Cerebellum. 2020; 19: 691-700. PMID: 32556973 DOI: 10.1007/s12311-020-01151-5 [PubMed]

12. He D, Feenaughty L, Wan Q. Global acoustic speech temporal characteristics for Mandarin speakers with Parkinson's disease during syllable repetition and passage reading. Am J Speech Lang Pathol. 2023; 32(5): 2232-44. PMID: 37625136 DOI: 10.1044/2023_AJSLP-23-00062 [PubMed]

13. Krýže P, Tykalová T, Růžička E, Rusz J. Effect of reading passage length on quantitative acoustic speech assessment in Czech-speaking individuals with Parkinson's disease treated with subthalamic nucleus deep brain stimulation. J Acoust Soc Am. 2021; 149(5): 3366-74. PMID: 34241103 DOI: 10.1121/10.0005050 [PubMed]

14. Kim KS, Wang H, Max L. It's about time: Minimizing hardware and software latencies in speech research with real-time auditory feedback. J Speech Lang Hear Res. 2020; 63(8): 2522-34. PMID: 32640180 DOI: 10.1044/2020_JSLHR-19-00419 [PubMed]

15. Rusz J, Krack P, Tripoliti E. From prodromal stages to clinical trials: The promise of digital speech biomarkers in Parkinson's disease. Neurosci Biobehav Rev. 2024; 167: 105922. PMID: 39424108 DOI: 10.1016/j.neubiorev.2024.105922 [PubMed]

16. Dorsey ER, Omberg L, Waddell E, Adams JL, Adams R, Ali MR, et al. Deep phenotyping of Parkinson's disease. J Parkinsons Dis. 2020; 10(3): 855-73. PMID: 32444562 DOI: 10.3233/JPD-202006 [PubMed]

17. Livezey JA, Glaser JI. Deep learning approaches for neural decoding across architectures and recording modalities. Brief Bioinform. 2021; 22(2): 1577-91. PMID: 33372958 DOI: 10.1093/bib/bbaa355 [PubMed]

18. Thies T, Mallick E, Tröger J, Baykara E, Mücke D, Barbe MT. Automatic speech analysis combined with machine learning reliably predicts the motor state in people with Parkinson's disease. NPJ Parkinsons Dis. 2025; 11(1): 105. PMID: 40316531 DOI: 10.1038/s41531-025-00959-4 [PubMed]

19. Huang Y-J, Lin Y-T, Liu C-C, Lee L-E, Hung S-H, Lo J-K, et al. Assessing schizophrenia patients through linguistic and acoustic features using deep learning techniques. IEEE Trans Neural Syst Rehabil Eng. 2022; 30: 947-56. PMID: 35358049 DOI: 10.1109/TNSRE.2022.3163777 [PubMed]

20. Sayadi M, Varadarajan V, Langarizadeh M, Bayazian G, Torabinezhad F. A systematic review on machine learning techniques for early detection of mental, neurological and laryngeal disorders using patient's speech. Electronics. 2022; 11(24): 4235.

21. Sayadi MJ, Langarizadeh M, Torabinezhad F, Bayazian G. Voice as an indicator for laryngeal disorders using data mining approach. Frontiers in Health Informatics. 2024; 13: 205.