Voice Quality in Multiple Sclerosis: A Clinical Perspective on Cepstral Peak Prominence-Smoothed and Harmonic-to-Noise Ratio

Voice Quality in Multiple Sclerosis: A Clinical Perspective on Cepstral Peak Prominence-Smoothed and Harmonic-to-Noise Ratio

*,§Giulia Fusari, *Alessandro Torchio, *Cinzia Baldanzi, *Sara Isernia, †Valeria Crispiatico, *Marco Rovaris, ‡Davide Sgrò, *Chiara Vitali, ‡Alessio Carullo, and *,§Davide Cattaneo, *†§*Milan, and* ‡*Turin, Italy*

SUMMARY: Objectives/hypothesis. This study aims to evaluate voice quality in people with multiple sclerosis (pwMS) with moderate-to-severe disability using objective acoustic measures: Cepstral peak prominence-smoothed (CPPS) and harmonic-to-noise ratio (HNR). The objectives are to compare these measures between pwMS and healthy subjects (HS) and assess their diagnostic accuracy in identifying inadequate voice quality and the need for voice rehabilitation. We hypothesize that pwMS will exhibit reduced voice quality compared to HS, with CPPS being a more reliable marker than HNR for assessing vocal impairments and guiding rehabilitation decisions.

Study design. This cross-sectional study involved 34 pwMS with moderate-to-severe disability and 40 sex-matched HS.

Methods. Participants' voice recordings were collected through a 1-minute monologue and analyzed using *PRAAT* software (University of Amsterdam, Amsterdam, The Netherlands). Acoustic measures, CPPS and HNR, were extracted and compared between groups. Clinical perceptual judgment was employed as the gold standard. Diagnostic accuracy of CPPS and HNR was evaluated through receiver operating characteristic curve analysis.

Results. CPPS and HNR values were significantly lower in pwMS compared to HS, indicating impaired voice quality due to reduced harmonicity and increased aperiodicity. CPPS (threshold = 11.07 dB) demonstrated high sensitivity (0.76) and specificity (0.82) for detecting inadequate voice quality. HNR (threshold = 11.29 dB) exhibited lower sensitivity (0.73) and specificity (0.73). CPPS proved to be a more reliable measure than HNR for identifying vocal impairments and determining the need for rehabilitation.

Conclusions. Our findings confirm significant voice alterations in pwMS compared with a HS group in all collected measures. Although the data showed that the CPPS can be a valuable tool for voice assessment in multiple sclerosis, further studies should explore its applicability in broader populations and its integration into clinical protocols for voice assessment in pwMS.

Key Words: Voice quality–Multiple sclerosis–Acoustical analysis–Rehabilitation–Dysphonia–Cepstral peak prominence-smoothed.

INTRODUCTION

Multiple sclerosis (MS) is a chronic, immune-mediated disease of the central nervous system, characterized by the disruption or loss of axonal myelin driven by both inflammatory and degenerative mechanisms.[1] In addition to the commonly observed motor symptoms, MS can also affect communication, particularly speech and voice, impacting all speech subsystems, including respiration, phonation, resonance, articulation, and prosody.[2] Among these, articulation and phonation subsystems are the most frequently impaired, leading to common speech disorders in MS: dysarthria and dysphonia.[3–5] These impairments frequently co-occur due to shared neuromuscular dysfunctions, significantly reducing speech intelligibility and voice quality. These voice alterations manifest in the perception of pitch instability, asthenia, hoarseness, and breathiness, particularly during prolonged phonation.[2,4,6] Such changes negatively impact quality of life, limiting social, professional, and emotional communication. Numerous studies have reported that vocal changes in people with multiple sclerosis (pwMS) are mainly due to glottic inefficiency with incomplete glottal closure and posterior glottal chink,[7] impaired respiratory support, weakness of respiratory muscles and fatigability.[4,8–10] Furthermore, Feijò and collegues demonstrated that dysphonic symptoms in MS stem from demyelinating lesions in the periaqueductal gray matter, which disrupt the concentrated neuronal projections impairing the synchronous activation of the respiratory and phonatory musculature.[11]

Assessing voice quality in MS is particularly complex due to its multidimensional nature resulting from the interaction between all speech subsystems. Traditionally, voice quality has been evaluated through auditory-perceptual assessments using scales such as the GRBAS or

Accepted for publication May 22, 2025.

This work was supported and funded by the Italian Ministry of Health—Current Research 2023 and by the FISM—Fondazione Italiana Sclerosi Multipla—cod. 2022/PR-Single/016.

From the *IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; †Department of Psychology, University of Milano-Bicocca, Milan, Italy; ‡Department of Electronics and Telecommunications, Politecnico of Turin, Turin, Italy; and the §Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy.

Address correspondence and reprint requests to Chiara Vitali, IRCCS Fondazione Don Carlo Gnocchi, Via Capecelatro 66, Milan, Italy. E-mail: cvitali@dongnocchi.it

Journal of Voice, Vol xx, No xx, pp. xxx–xxx

0892-1997

https://doi.org/10.1016/j.jvoice.2025.05.026

GIRBAS scales.[12,13] In this type of assessment, clinicians assess voice quality based on listening and rating. While considered the gold standard,[14,15] perceptual evaluations are inherently subjective, influenced by listener bias, such as training and experience, and inter-rater variability. Furthermore, these assessments do not fully capture underlying physiological deficits, making it challenging to quantify subtle voice impairments in MS.[16–19] Therefore, to complement and enhance the robustness of perceptual evaluations, acoustic analysis can be used.[20,21] Acoustical analysis provides objective and quantifiable measures of voice quality, reducing the subjectivity of clinical evaluations.[20] In particular, changes in voice quality correspond to measurable alterations in the harmonic structure and periodicity of the signal.[22]

Among the various parameters used to assess voice quality and harmonic structure, harmonic-to-noise ratio (HNR) and cepstral peak prominence-smoothed (CPPS) are two of the most frequently metrics used in clinic.[20,23,24] HNR (dB) evaluates the overall voice quality through the evaluation of the periodicity of successive glottal cycles and the ratio between the harmonic energy (periodic components) and the noise energy (aperiodic components).[25] The periodic component arises from vocal-fold vibration, while the non-periodic component reflects glottal turbulence due to incomplete closure. Higher HNR values indicate a better voice quality.

Among cepstral-based measures, CPPS (dB) is recognized as one of the most reliable dysphonia parameters by the American Speech-Language-Hearing.[21] In particular, given a signal, its Cepstrum is equal to the Fourier transform of the logarithm of its power spectrum. Its robustness comes from the fact that it does not need previous pitch detection and tracking.[26] Therefore, CPPS quantifies harmonic dominance in the spectrum and outperforms HNR in voice analysis, especially in connected speech and dysphonic or dysarthric voices, as it is less influenced by fundamental frequency estimation errors.[20,27–29] Maryn et al,[20] conducted a meta-analysis showing CPPS's superior performance over traditional perturbation measures (jitter, shimmer, HNR) across multiple clinical populations and speaking tasks. As a result, CPPS has been widely incorporated into objective dysphonia indices, underscoring its versatility and clinical relevance, with higher CPPS values stand for a better voice quality.[30,31]

HNR has been studied in the MS population, with mixed results, some showing no significant differences from healthy subjects (HS) and others reporting significant differences. Conversely, the use of CPPS in this population remains unexplored.[5–7] Given the growing clinical interest in CPPS as a robust measure of voice quality, further investigation into its diagnostic and clinical relevance in MS is warranted. This is especially relevant when studying continuous speech. In fact, clinicians with experience in MS rehabilitation know that most speech and voice symptoms are most often signaled by the subjects themselves in continuous speech. In connected speech (eg, monologue), the higher accuracy of CPPS compared to HNR seems to be proven, as well as its utility in clinical practice as an outcome to measure the effects of rehabilitation treatments.[26,31–34] Although continuous speech introduces greater variability and presents methodological challenges in segmentation and standardization, it offers a more ecologically valid assessment of voice function by capturing the dynamic and complex nature of everyday communication in MS.[35] For this reason, its inclusion is increasingly recommended, particularly when assessing individuals with neurological disorders, in whom voice impairments may be more apparent during connected speech tasks.[16,35]

Therefore, this cross-sectional study aims to:

1. Describe and compare voice quality during monologue, in HS and pwMS with moderate-to-severe disability combining MS-related clinical measures (Expanded Disability Status Scale [EDSS], MS-type and disease duration) and acoustic parameters (CPPS and HNR).
2. Assess the ability of CPPS and HNR to distinguish between individuals with adequate and inadequate voice quality and identify those who require voice rehabilitation, using clinical perceptual judgment as the gold standard.

By addressing these objectives, this study seeks to enhance the understanding of voice quality in pwMS, emphasize the clinical utility of acoustic measures in monitoring voice changes, and guide the development of personalized rehabilitation strategies. We hypothesize that the pwMS will exhibit reduced voice quality compared to HS, with CPPS emerging as a more reliable marker than HNR for assessing voice impairments in MS and guiding rehabilitation decisions.

MATERIAL AND METHODS

Study design

This cross-sectional study was conducted at the Fondazione Don Carlo Gnocchi IRCCS Santa Maria Nascente (Milan) between September 2023 and August 2024 and was approved by the local Ethics Committee at Fondazione Don Carlo Gnocchi in Milan. All participants provided informed consent and voluntarily agreed to take part in the study.

Participants

Thirty-four native Italian speakers with MS who were referred to our Center for routine neurological visits were recruited with the following inclusion criteria: diagnosis of MS according to revisited McDonald's criteria[36]; > 18 years of age; Mini-Mental State Examination score > 21 points[37]; absence of severe dysarthria (as confirmed by two speech-language therapists), and ability to understand the Italian language and the aims of the study.

Participants were excluded if they had other neurological disorders or clinical history of laryngeal cancer, chemotherapy, radiotherapy, head and neck trauma or endotracheal intubation, visual/hearing impairments hindering the rehabilitation program, relapse, or sudden changes in MS symptoms within the previous 3 months.

A sample of 40 euphonic native Italian speakers HS, was also recruited to provide normative data for CPPS and HNR, being attentive to the female-to-male ratio. All the HS underwent instrumental voice examinations (video laryngoscopy with stroboscopic examination) and were found to be free of larynx and vocal fold pathologies.

PROCEDURE

Sample characteristics

Demographic and clinical characteristics

Age, sex, and smoking habits were collected for the whole sample. Type of MS, disability level, disease duration, were collected for the MS group. Disability level was rated using the EDSS.[38]

Perceptual characteristics of voice

The GIRBAS scale,[12] was used to assess voice quality through a 1-minute monologue recorded by a speech therapist (see "Recording setup and audio selection" for details). In addition, voice intensity (dB sound pressure level (SPL)) was extracted.

GIRBAS scale consists of six parameters: G (overall grade of hoarseness), I (instability), R (roughness), B (breathiness), A (asthenicity), and S (strained quality), each rated on a four-point equal-appearing interval scale from 0 (no problem) to 3 (severe problem). In our study, three speech therapists with experience in voice evaluation conducted a blind evaluation of each parameter based on the recorded monologue. The median of the three ratings from each parameter was reported.

Recording setup and audio selection

For the MS group, 1 minute of monologue was recorded using a portable high-quality recorder (24 bit/96 kHz) with an integrated in-air microphone (Roland, R-05) (Roland, 2036-1 Nakagawa, Hosoe-cho, Hamana-ku, Hamamatsu, Shizuoka 431-1304, Giappone) placed 30 cm from the mouth. Participants were asked to speak at their comfortable frequency and intensity level. A specific task identical for each subject was given: Please speak for at least a minute, explaining: "how to make coffee," "how you feel today," and "what is scheduled for the day."

Audio tracks were assessed to remove incomplete, saturated, or noisy background files manually.

The same procedure was followed for the HS group, but they were asked to talk about their day.

Acoustic features extraction

The available voice recordings (34 pwMS and 40 HS) were analyzed using a *PRAAT* script (version 6.4.06). Before extracting the parameters for the vocal analysis, an initial preprocessing step was performed to remove low frequency noise and potential artifacts by means of a high-pass filter with a cut-off frequency of 34 Hz. Then, the silences in each recording were identified and removed using the *PRAAT* TextGrid object, with silence defined as a group of samples within intervals of 46 ms (corresponding to typical inter-syllabic pauses in the Italian language) and intensity below −25 dB.

For each recording, the remaining samples, representing vocal content, were concatenated into a new wave file to subsequently compute the vocal parameters. The percentage of voiced signal was calculated using Praat's "Voice Report" function.

The HNR was calculated using the PRAAT command To Harmonicity, which performs short-term HNR analysis and generates a harmonicity object representing the degree of acoustic periodicity, by employing the autocorrelation method described by Paul Boersma.[39] The parameters for the HNR analysis were set as follows: a time step of 10 ms (the duration of each frame), a pitch floor of 75 Hz (defining the length of the analysis window), a silence threshold of 0.1 (frames with amplitudes below 10% of the global maximum amplitude were considered silent), and an average of 4.5 pitch periods per analysis window. The Power Cepstrogram was computed using the To PowerCepstrogram function, which downsamples the digital sequence to 22,050 Sa/second with frames collecting 1024 samples, each corresponding to a time interval of 46.4 ms. Concerning the spectral calculation by means of the Fourier transform, a Gaussian weighting window (spanning three periods of the pitch floor) was applied to each frame to mitigate the leakage distortion caused by the noncoherent sampling of the signal. Then, the Power Cepstrum was derived by applying a second Fourier transform to the logarithm of the power spectrum. The mean value of CPPS was calculated using the Get CPPS command, which determines the average CPP across individual frames after smoothing the Power Cepstrogram with a time-averaging window of 14 ms and a quefrency-averaging window of 1 ms. This method of CPPS computation in PRAAT follows the algorithms implemented and described by Maryn and Weenink,[40] which are based on the original concept introduced by Hillenbrand and Houde,[41] that enhances the assessment of the voice signal's periodicity, reflecting the harmonicity of the vocal fold vibrations.

In this smoothed Power Cepstrogram, each cepstral value represents the average of the cepstral values within the respective averaging window, symmetrically centered around each frame in both the time and quefrency domains. The peak search is performed within a pitch range of 60–330 Hz, with a 5% tolerance and parabolic interpolation; once the cepstral peak is identified, it is compared to a regression line calculated by fitting a straight line over the cepstrum, starting from a quefrency of 1 ms to minimize the influence of very low quefrency data on the fit.[42] The CPPS is then obtained by subtracting the peak value to the one on the regression line at the same quefrency.

Clinician's perceptive judgment

To integrate objective measures with clinical and perceptual assessment, we asked for clinical judgment in distinguishing between individuals with adequate and inadequate voice quality and in identifying those who require voice rehabilitation. We asked three speech therapists experienced in voice evaluation to make a blind perceptual judgment of both HS and pwMS by answering the following two questions: question 1—"Based on perceived voice quality and considering all subsystems, do you judge this voice as adequate?", and question 2—"Based on the perceived voice quality and considering all subsystems, would you recommend speech therapy treatment for this person?". The possible answers were limited to "yes" or "no."

STATISTICAL ANALYSIS

Demographic and clinical data were normally distributed and reported as mean ± standard deviation or number (percentage, %). The GIRBAS scale was presented as the number (percentage, %) of participants for each score, showing how participants were distributed across different severity levels for each of the six parameters.

For each group, the percentage of voiced signal was reported as median and interquartile range (IQR), and a Wilcoxon test was performed to compare groups. Acoustical parameters, CPPS and HNR, were reported as median (5th, 25th, 75th, 95th percentile), and differences between HS and pwMS were tested using Wilcoxon test. Higher CPPS and HNR values indicate better voice quality, while lower values reflect poorer voice quality due to increased aperiodicity. Since HS and pwMS had different ages, we also performed the between-group analyses, including age as a covariate with no statistical associations found with acoustical parameters; since age was not significant, we opted for the Wilcoxon test for simplicity and clarity in comparing the groups. From the HS group, we used the 5th percentile as a preliminary reference threshold and reported which pwMS presented CPPS and HNR values falling below this cutoff. Thus, we reported the frequency (n, %) of pwMS with scores below the threshold, along with their median (5th, 25th, 75th, 95th percentile) values for both CPPS and HNR.

In addition, a stepwise linear regression analysis was conducted to explore the relationship between disease duration (years), EDSS, SPL, type of MS (relapsing-remitting [RR]; secondary progressive [SP]; primary progressive [PP]), and the values of HNR and CPPS. Each model was appropriately specified to meet the assumptions of linearity, homoscedasticity, and normality. The goodness of fit statistics, including R^2 and adjusted R^2, were reported to assess the model's explanatory power.

Finally, to assess the accuracy of CPPS and HNR in distinguishing between participants with adequate versus inadequate voice quality (question 1) and in determining the need for speech therapy (question 2), receiver operating characteristic (ROC) curve analysis were performed separately for each question. For each acoustic parameter, a separate optimal threshold was identified for question 1 and question 2 using Youden's index to maximize the combined sensitivity and specificity. For each threshold, performance metrics (ie, sensitivity, specificity, positive predictive value [PPV], and negative predictive value [NPV]) were computed.

All the statistical analyses were conducted using *R* (version 4.1), and graphical representations of ROC curves were generated using pROC package in R. All statistical tests were 2-tailed, and a P-value < 0.05 was considered statistically significant.

RESULTS

Sample characteristics

Demographic and clinical characteristics

In the MS group (N = 34), 20 participants (58.8%) were female, with a mean age of 56.7 ± 11.36 years. In contrast, the HS group (N = 40) had 27 female participants (67.5%), with a considerably lower mean age of 29.5 ± 10.83 years. Although the proportion of females did not differ significantly between the 2 groups (P = 0.60), the significant difference in mean age (P < 0.001) was accounted for in the analysis. Moreover, while no HS reported smoking habits, 12 out of 34 pwMS were active smokers.

Regarding clinical characteristics of pwMS, 19 participants (57.6%) presented RR, while 14 participants (42.4%) presented PP or SP MS. The mean EDSS and disease duration were 6.5 ± 1.3 points and 22.9 ± 11.9 years, respectively.

Perceptual characteristics of voice

In the MS group, the median percentage of extracted voiced signal from the 1-minute monologue was 59.97%, with an IQR of 13.8%, while in the HS group the median was 64.8% (8.83%) (P-value = 0.59). The mean voice intensity was 71.7 ± 4.0 dB SPL for the MS group and 72.8 ± 5.5 dB SPL for the HS group. In the HS group, all subjects scored GIRBAS G = 0, indicating no perceptible voice alteration. In the MS group, the GIRBAS G score (general degree of voice alteration) and the sub-scores medians (Instability, Roughness, Breathiness, Asthenia, and Strain) are reported in Table 1. We also calculated the mean GIRBAS G scores across raters, which were 0.79, 1.06, and 1.06, points respectively, to facilitate comparison with previous studies using average values. Twenty-eight pwMS were evaluated with abnormal GIRBAS G scores (≥1 points, mild-to-moderate voice alteration).

Aim 1: To describe and compare voice quality in pwMS and HS using acoustic measures

CPPS

The median value of CPPS (dB) of a monologue (5th, 25th, 75th, 95th percentile) was 11.71 (9.47, 11.39, 12.56, 13.51)

TABLE 1.
GIRBAS Scores for MS Group (N = 34)

	GIRBAS Scores			
	0	1	2	3
Global (G)	6/34 (17.7%)	21/34 (61.8%)	7/34 (20.5%)	0/34 (0%)
Instability (I)	19/34 (55.9%)	12/34 (35.3%)	3/34 (8.8%)	0/34 (0%)
Roughness (R)	8/34 (23.5%)	17/34 (50%)	7/34 (20.6%)	2/34 (5.9%)
Breathiness (B)	29/34 (85.3%)	5/34 (14.7%)	0/34 (0%)	0/34 (0%)
Asthenia (A)	16/34 (47.1%)	13/34 (38.2%)	5/34 (14.7%)	0/34 (0%)
Strain (S)	16/34 (47.1%)	13/34 (38.2%)	5/34 (14.7%)	0/34 (0%)

Note: Data are reported as the number (percentage) of pwMS ho achieved the specified score.

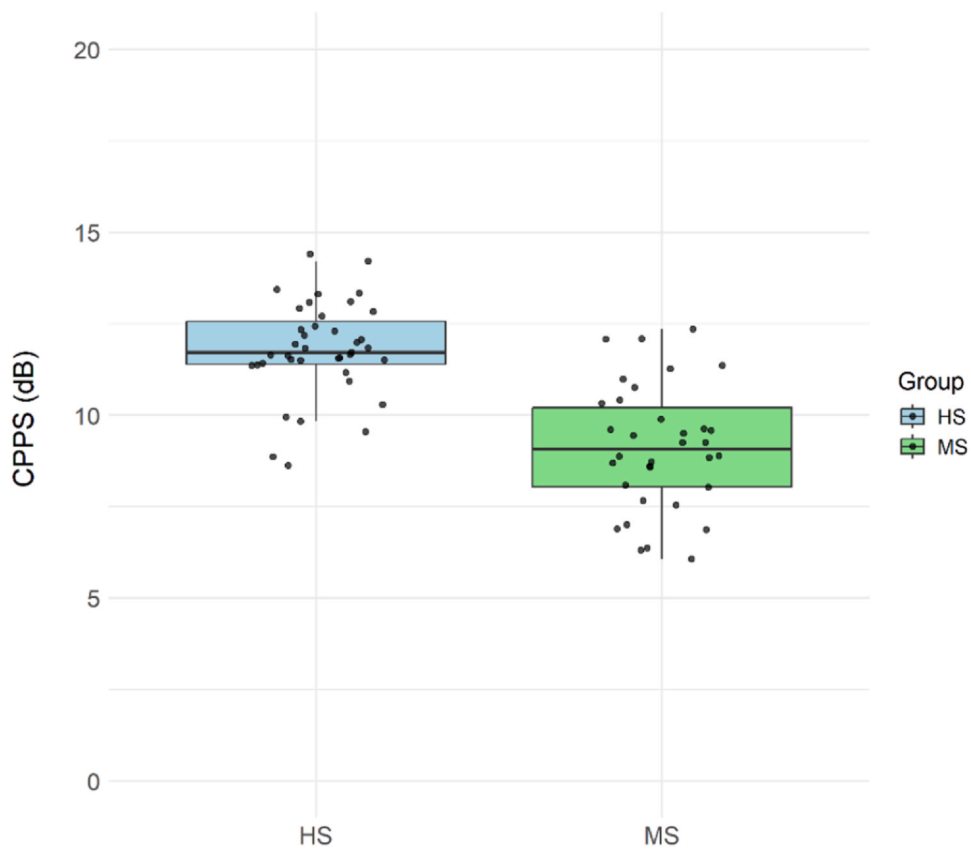

FIGURE 1. Comparison of CPPS values (in dB) between healthy subjects and people with multiple sclerosis. The boxplots represent the median, interquartile range, and variability of CPPS for each group, with individual data points overlaid.

dB for HS group and 9.07 (6.35, 8.04, 10.21, 12.08) dB for the MS group (W = 1178.5, $P < 0.001$) (Figure 1). In our MS sample, 20 subjects (58.82%) had CPPS values below the HS group 5th percentile with a median value of 8.33 dB (6.30, 6.97, 8.84, 9.26).

The linear model assessing CPPS, with independent variables selected through stepwise analysis, was statistically significant overall (F = 2.77, $P = 0.02$), explaining 62% of the variance ($R^2 = 0.62$, adjusted $R^2 = 0.40$). Among the predictors, only voice intensity (dB SPL) (estimate = 0.17, SE = 0.066, $P = 0.017$) and disease duration (years) (estimate = 0.076, SE = 0.03, $P = 0.01$) emerged as significant factors. In contrast, disability level (EDSS) did not significantly predict CPPS but remained as a predictor in the stepwise analysis.

HNR

The median value of HNR (dB) of a monologue was 12.88 (9.67, 11.51, 14.02, 16.05) dB for the HS group and 9.20 (5.85, 7.62, 10.70, 13.76) dB for the MS group (W = 1153.5, $P < 0.001$) (Figure 2).

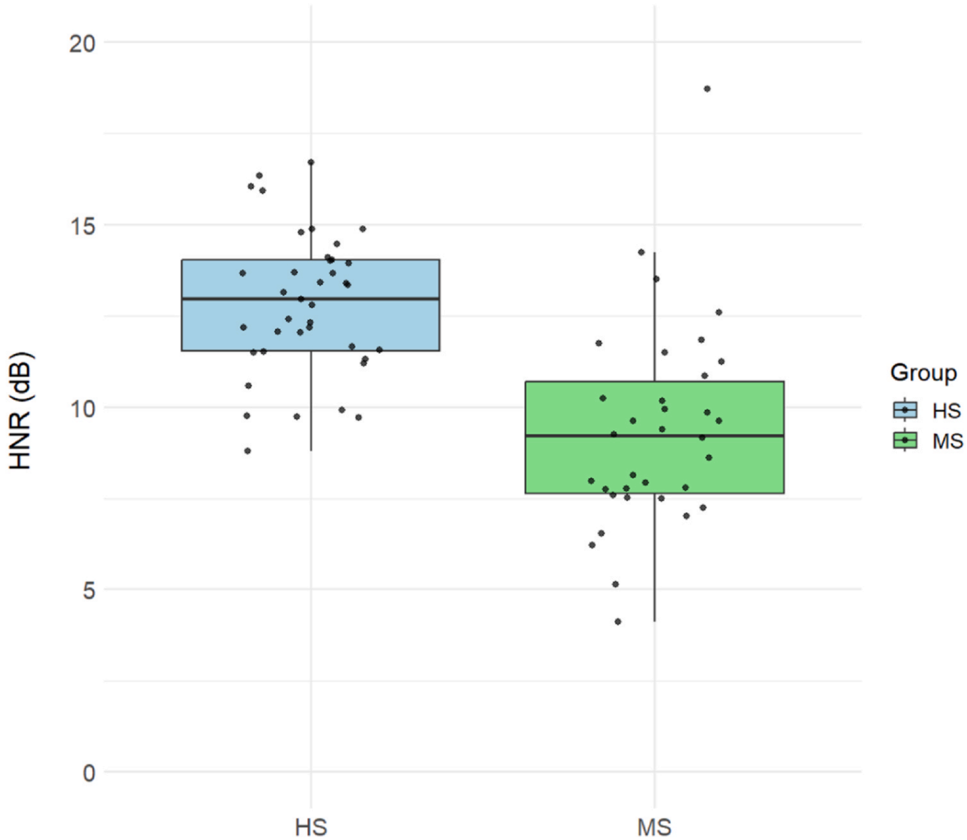

FIGURE 2. Comparison of CPPS values (in dB) between healthy subjects and people with multiple sclerosis. The boxplots represent the median, interquartile range, and variability of CPPS for each group, with individual data points.

In our MS sample, 21 subjects (61.76%) had HNR values below the HS group 5th percentile with a median value of 7.77 (5.15, 7.25, 8.61, 9.62) dB.

The linear regression model examining HNR, with independent variables selected through stepwise selection, was not statistically significant overall ($F = 0.69$, $P = 0.66$). Additionally, none of the predictors emerged as significant factors in explaining the variance in HNR.

Aim 2: Diagnostic accuracy of acoustic parameters

The ROC curve analysis was conducted to evaluate the diagnostic accuracy of CPPS and HNR in differentiating between participants with inadequate voice quality (*question 1*, Figure 3A), and in determining the need for speech therapy (*question 2*, Figure 3B).

For question 1, 33 participants out of 74 (4 HS and 29 MS) were judged to have inadequate voice quality, while for *question 2*, 29 participants out of 74 (0 HS and 29 MS) were judged to be candidates for voice rehabilitation.

CPPS

For CPPS, the optimal threshold for identifying inadequate voice quality (question 1) and determining the need for treatment (question 2) was consistently 11.07 dB.

The threshold for identifying inadequate voice quality (question 1) demonstrated a sensitivity of 0.76 and a specificity of 0.82. The PPV was 0.84, while the NPV was 0.73. For determining the need for treatment (question 2), the same threshold yielded a sensitivity of 0.64, and a specificity of 0.86. The PPV was 0.92, while the NPV was 0.49.

HNR

The optimal thresholds for HNR to identify inadequate voice quality (question 1) was 11.29 dB and to determine the need for treatment (question 2) was 11.94 dB. The question 1 threshold resulted in a sensitivity of 0.73 and specificity of 0.73. The PPV was 0.77, while the NPV was 0.69.

The question 2 threshold resulted in a sensitivity of 0.53 and a specificity of 0.86. The PPV was 0.90, while the NPV was 0.42.

DISCUSSION

This study is the first to examine voice quality during monologue in pwMS with moderate-to-severe disability, focusing on objective acoustic measures, specifically CPPS and HNR.

Voice quality

Our findings confirm that pwMS present significant voice alterations compared to HS. This was evident both

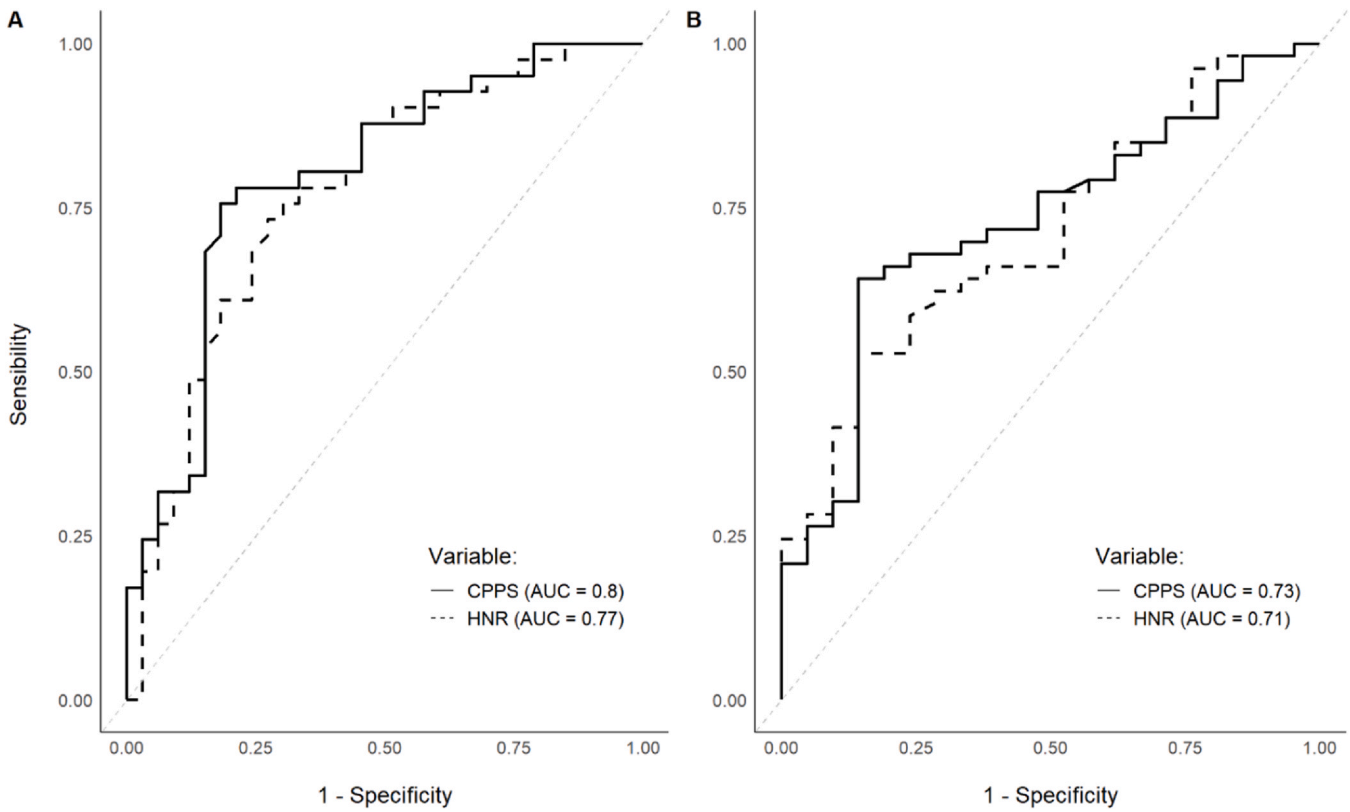

FIGURE 3. ROC curve for CPPS and HNR in response to question 1 (**A**) and question 2 (**B**).

perceptually, with 82.3% of pwMS exhibiting mild-to-moderate voice alterations (GIRBAS), and acoustically, with significantly lower CPPS and HNR values.

These results are in line with the literature, with studies reporting that pwMS are perceptually mildly impaired, both in the overall grade of dysphonia (GRBAS G) and in all the subscales.[5,7] In our sample, Breathiness emerged as the least affected component, with most participants scoring 0 points. In contrast, a significant percentage of pwMS exhibited mild Roughness, followed by Asthenia and Strain. Consistent with our findings, Yaşar et al[5] identified Roughness as the most prominent feature in pwMS. Moreover, in agreement with our results, other studies have reported significantly higher Asthenia (A) and Strain (S) scores in MS patients compared to controls.[5]

Considering acoustical assessment, mixed results exist regarding HNR, with studies reporting normal or low values in pwMS,[7,43] while no prior research has specifically examined CPPS in this population. Our results indicate that pwMS exhibit voice quality impairments, as reflected in low CPPS values. Furthermore, our regression analysis revealed that pwMS with higher voice intensity tend to show better voice quality. This is in line with Brockmann-Bauser and colleagues,[44] who demonstrated a strong correlation between increased voice intensity and higher CPPS in both dysphonic patients and normative groups. The relation between voice intensity and the level of harmonics in the signal may stem from increased muscle tone and medial compression of the vocal folds, which may result in improved glottal closure and signal periodicity, as indicated by a higher CPPS.[44] Clinically, this finding underscores the importance of maintaining adequate voice intensity for improved voice quality, highlighting a potential target for speech therapy intervention. In addition, CPPS was found to be associated with the disease duration in our group, suggesting that a longer disease duration is related to a slight increment in harmonicity. This finding may seem counterintuitive, as a progressive neurodegenerative disease like MS would typically be expected to degrade voice quality over time. However, several factors may contribute to this relationship, including the presence of compensatory mechanisms or individual adaptation strategies, due to past voice rehabilitation experiences. Unfortunately, no previous studies have directly investigated the relation between CPPS and disease duration, highlighting the need for further research to investigate potential factors associated with mechanisms underlying voice changes in pwMS.

Regarding HNR, our results showed a lower HNR in pwMS compared to HS, possibly reflecting inadequate vocal fold closure, and resulting in excessive airflow through the glottis. Differently from CPPS, our linear regression did not identify significant predictors for HNR, suggesting that this parameter might be influenced by factors beyond those considered in our current model.

Interestingly, EDSS did not significantly predict voice quality for either CPPS or HNR, suggesting that voice impairment in pwMS may not be directly associated with overall disability level. Alternatively, this could indicate that the EDSS, which primarily evaluates motor dysfunction, does not specifically assess voice-related aspects. This finding aligns with studies reporting no significant association between EDSS and HNR, while the relationship between CPPS and EDSS remains unexplored.[44]

Diagnostic accuracy

In the second part of our study, we evaluated the ability of CPPS and HNR to discriminate between individuals with inadequate voice quality and those who need voice treatment, using clinical assessment as the gold standard.

Considering CPPS, the results of our study suggest a unique threshold (11.07 dB) for identifying inadequate voice quality and determining the need for treatment. This threshold allowed us to correctly classify most participants, distinguishing between those with and without adequate voice quality. Notably, when CPPS indicated an inadequate voice quality, it was accurate in most cases, confirming that most individuals flagged as having voice alterations truly had them. Similarly, when CPPS suggested adequate voice quality, this was generally reliable, though with a slightly lower certainty. These findings suggest that CPPS can be a valuable tool for clinical assessment in MS.

In contrast, HNR demonstrated a lower yet still meaningful ability to distinguish between adequate and inadequate voice quality, with its optimal threshold set at 11.29 dB. While HNR was able to correctly classify a substantial portion of participants, its accuracy was slightly lower than CPPS. This means that, although HNR can help identify voice impairments, it may be more prone to misclassifying individuals, either by flagging some adequate voices as inadequate or failing to detect subtle voice issues.

To date, numerous studies have explored the ability of acoustic parameters to differentiate pathological from nonpathological voices, emphasizing the importance of integrating objective measures with clinical assessments for improved diagnostic accuracy. Considering CPPS, when connected speech is studied, threshold values for distinguishing dysphonic from non-dysphonic speakers range from 4.0 to 19.10 dB,[45] depending on factors such as the extraction algorithm, speech task, and the speaker's native language.[26,30,46] To our knowledge, no studies have specifically investigated the discriminative ability of CPPS in MS, including only people with MS in the sample. However, a recent study on Brazilian Portuguese speakers, including a small percentage of pwMS,[45] identified a threshold of 11.30 dB (ROC: 0.872; sensitivity: 76.29%; specificity: 78.57%), similar to our result. Other recent studies from Cantor-Cutiva and Murton and collegues,[30] found CPPS to be a strong predictor of voice disorders, though methodological differences may explain discrepancies with our results. In fact, Cantor-Cutiva and colleagues conducted a retrospective analysis incorporating various vocal tasks,[47] including sustained vowel phonation and the six sentences from the consensus auditory-perceptual evaluation of voice, while Murton et al[30] analyzed a standardized reading passage from English speakers using *PRAAT*. Differences in study populations, disease distribution, voice characteristics, and parameters estimation methods likely contributed to this variation.

Although HNR has been less frequently studied as a diagnostic measure for dysphonia, Nguyen and colleagues[48] found no significant differences in HNR between dysphonic and non-dysphonic subjects. Similarly, Shama et al[49] reported that the time-based HNR had lower discrimination ability in classifying normal and pathologic voices compared to the spectral-based HNR.

Another way acoustic analysis can support clinical judgment is to identify individuals needing speech therapy. In this context, our results suggest that despite CPPS and HNR presenting similar threshold values, CPPS demonstrated a stronger ability in detecting pwMS who need voice rehabilitation, compared to HNR. Specifically, CPPS demonstrated a stronger ability to correctly identify those requiring treatment, ensuring that most individuals classified as needing therapy truly required it. Additionally, while CPPS was less effective in ruling out those who did not need intervention, its overall predictive value remained high. These findings reinforce the idea that CPPS is a more reliable tool for determining the need for voice rehabilitation, while HNR may still be useful as a supplementary measure. One possible explanation for the consistent underperformance of HNR compared to CPPS may lie in the f0-dependent nature of HNR. Since accurate measurement of HNR requires reliable cycle boundary detection, its precision decreases in cases of dysphonic voices, where irregular vocal fold vibration complicates this detection. In contrast, CPPS does not rely on cycle detection, making it more robust in voices with high variability in vocal fold vibration.[42,50]

Clinical implications

The findings from Aim 2 highlight CPPS as the most reliable acoustic parameter for identifying voice impairment. However, when evaluating whether these parameters could guide treatment decisions, our results suggest a limitation. Specifically, the optimal thresholds for identifying inadequate voice quality and determining treatment needs were nearly identical for both CPPS and HNR. Consequently, all subjects identified as having deviant voice quality were also considered in need of treatment, rendering the additional threshold for treatment decisions unnecessary. This suggests that when voice quality is evaluated in terms of the interactions within all speech subsystems (question 1), it captures a level of dysfunction significant enough to warrant intervention.

In fact, in clinical practice with neurological voice disorders, treatment decisions are not solely based on impairments related to the phonatory system (eg, roughness or breathiness). Rather, they are influenced by alterations in respiratory control, articulatory precision, and resonance (eg, reduction in voice intensity or the presence of dysarthria). Furthermore, the clinical assessment of neurological voice disorders integrates professional competence based on clinical expertise, often referred to as the "silent know-how" in vocal rehabilitation, as well as the patient's specific needs.[51] Therefore, in our opinion, a threshold based solely on acoustic parameters may not be sufficient to guide the decision to initiate or withhold speech therapy.

Since voice quality is influenced by all these factors, the slight superiority of CPPS over HNR in identifying pwMS in need of voice rehabilitation could stem from the fact that CPPS better reflects the interaction between the phonatory, respiratory, and resonant systems.[52]

In conclusion, while CPPS and HNR are valuable measures, they represent only two components within a comprehensive framework of clinical, objective, and subjective assessments recommended for clinical voice evaluation. Importantly, the interpretation of these measures should be integrated with additional acoustic parameters, aerodynamic assessments, perceptual listener ratings, medical examination findings (including laryngeal endoscopic imaging), and patient self-reports.[21]

LIMITS

As with all studies, this work has certain limitations, which also present opportunities for further research. Firstly, the study included only pwMS with overall moderate-to-severe disability, limiting the generalizability of the findings to broader populations. Additionally, the cohort consisted of participants with a mild-to-moderate degree of dysphonia, as assessed by the GIRBAS scale, and some presented mild dysarthria as perceived by clinicians. Future studies should include within-group analyses comparing individuals with and without dysarthria or other speech impairments. Age differences between HS and pwMS could have introduced bias in comparisons. However, no statistically significant correlations were found between age and voice quality in our study, aligning with findings reported in the literature related to the effect of aging on CPPS and HNR.[53,54]

Furthermore, variables that could influence voice, such as respiratory, pulmonary, cognitive, or mood disorders, and a stroboscopic examination, were not available for all subjects. Another limitation is that only free speech samples (monologues) were recorded. While this approach reflects an ecological and naturalistic context, it does not allow for uniform speech material across participants, which may limit the comparability of results. Additionally, implementing more standardized tasks, such as sentence repetition or reading exercises, could enhance consistency and comparability across subjects. Finally, the higher proportion of smokers in the pwMS group compared to the HS group may have influenced some of the acoustic measures, as smoking can affect voice quality and related parameters. Future studies should consider controlling for smoking status during participant recruitment to minimize its potential confounding effect.

CONCLUSIONS

This study examines voice quality during monologue in pwMS with moderate-to-severe disability, using CPPS and HNR as objective acoustic measures. Our findings confirm significant voice alterations in pwMS compared with a HS group in all collected measures.

Our findings reinforce CPPS as a valuable tool in MS voice assessment. Moreover, our results underscore the importance of integrating acoustic analysis into routine assessments for pwMS, especially for guiding treatment decision-making. Including acoustical analysis in clinical practice may contribute to earlier identification of voice impairments, facilitate timely interventions, and potentially improve communication and quality of life.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Kevin Russell Bateman and Dr Elisa Gervasoni for helping revise the manuscript. GF thanks the 40th cycle of the PhD program in translational medicine for its valuable support.

References

1. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology*. 2014;83:278–286. https://doi.org/10.1212/WNL.0000000000000560.
2. Plotas P, Nanousi V, Kantanis A, et al. Speech deficits in multiple sclerosis: a narrative review of the existing literature. *Eur J Med Res*. 2023;28:252. https://doi.org/10.1186/s40001-023-01230-3.
3. Bauer V, Aleric Z, Jancic E, et al. Subjective and perceptual analysis of voice quality and relationship with neurological disfunction in multiple sclerosis patients. *Clin Neurol Neurosurg*. 2013;115:S17–S20. https://doi.org/10.1016/j.clineuro.2013.09.015.
4. Noffs G, Perera T, Kolbe SC, et al. What speech can tell us: a systematic review of dysarthria characteristics in Multiple Sclerosis. *Autoimmun Rev*. 2018;17:1202–1209. https://doi.org/10.1016/j.autrev.2018.06.010.
5. Yaşar Ö, Tahir E, Erensoy I, et al. Comparing dysphonia severity index, objective, subjective, and perceptual analysis of voice in patients with multiple sclerosis and healthy controls. *Mult Scler Relat Disord*. 2024;82:105378. https://doi.org/10.1016/j.msard.2023.105378.
6. Duranovic M, Salihovic N, Ibrahimagic A, et al. Characteristics of voice in individuals with multiple sclerosis. *Mater Socio Med*. 2011;23:23–27.
7. Dogan M, Midi I, Yazici MA, et al. Objective and subjective evaluation of voice quality in multiple sclerosis. *J Voice*. 2007;21:735–740. https://doi.org/10.1016/j.jvoice.2006.05.006.
8. Hamdan AL, Farhat S, Saadeh R, et al. Voice-related quality of life in patients with multiple sclerosis. *Autoimmune Dis*. 2012;2012:143813. https://doi.org/10.1155/2012/143813.
9. Nordio S, Bernitsas E, Meneghello F, et al. Expiratory and phonation times as measures of disease severity in patients with Multiple Sclerosis. A case-control study. *Mult Scler Relat Disord*. 2018;23:27–32. https://doi.org/10.1016/j.msard.2018.04.010.
10. Taveira FM, Teixeira AL, Domingues RB. Early respiratory evaluation should be carried out systematically in patients with multiple

sclerosis. *Arq Neuropsiquiatr.* 2013;71:142–145. https://doi.org/10.1590/s0004-282×2013000300003.
11. Feijó AV, Parente MA, Behlau M, et al. F. Acoustic analysis of voice in multiple sclerosis patients. *J Voice.* 2004;18:341–347. https://doi.org/10.1016/j.jvoice.2003.05.004.
12. Dejonckere PH, Remacle M, Fresnel-Elbaz E, et al. Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. *Rev Laryngol Otol Rhinol ((Bord)).* 1996;117:219–224.
13. Hirano M, McCormick KR. Clinical examination of voice. *J Acoust Soc Am.* 1986;80:1273. https://doi.org/10.1121/1.393788.
14. Lee Y, Kim G, Kwon S. The usefulness of auditory perceptual assessment and acoustic analysis for classifying the voice severity. *J Voice.* 2020;34:884–893. https://doi.org/10.1016/j.jvoice.2019.04.013.
15. Kreiman J, Gerratt BR, Kempster GB, et al. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *J Speech Hear Res.* 1993;36:21–40. https://doi.org/10.1044/jshr.3601.21.
16. Kent R. Hearing and believing: some limits to the auditory-perceptual assessment of speech and voice disorders. *Am J Speech Lang Pathol.* 1996;5:7–23.
17. Kreiman J, Gerratt BR. Sources of listener disagreement in voice quality assessment. *J Acoust Soc Am.* 2000;108:1867–1876. https://doi.org/10.1121/1.1289362.
18. Kreiman J, Gerratt BR, Precoda K. Listener experience and perception of voice quality. *J Speech Hear Res.* 1990;33:103–115. https://doi.org/10.1044/jshr.3301.103.
19. Kreiman J, Gerratt BR, Precoda K, et al. Individual differences in voice quality perception. *J Speech Hear Res.* 1992;35:512–520. https://doi.org/10.1044/jshr.3503.512.
20. Maryn Y, Roy N, De Bodt M, et al. Acoustic measurement of overall voice quality: a meta-analysis. *J Acoust Soc Am.* 2009;126:2619–2634. https://doi.org/10.1121/1.3224706.
21. Patel RR, Awan SN, Barkmeier-Kraemer J, et al. Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association Expert Panel to develop a protocol for instrumental assessment of vocal function. *Am J Speech Lang Pathol.* 2018;27:887–905. https://doi.org/10.1044/2018_AJSLP-17-0009.
22. Heman-Ackah YD, Michael DD, Goding GS. The relationship between cepstral peak prominence and selected parameters of dysphonia. *J Voice.* 2002;16:20–27. https://doi.org/10.1016/s0892-1997(02)00067-x.
23. Park Y, Stepp CE. Test-retest reliability of relative fundamental frequency and conventional acoustic, aerodynamic, and perceptual measures in individuals with healthy voices. *J Speech Lang Hear Res.* 2019;62:1707–1718. https://doi.org/10.1044/2019_JSLHR-S-18-0507.
24. Eskenazi L, Childers DG, Hicks DM. Acoustic correlates of vocal quality. *J Speech Hear Res.* 1990;33:298–306. https://doi.org/10.1044/jshr.3302.298.
25. Awan SN, Frenkel ML. Improvements in estimating the harmonics-to-noise ratio of the voice. *J Voice.* 1994;8:255–262. https://doi.org/10.1016/s0892-1997(05)80297-8.
26. Fraile R, Godino-Llorente JI. Cepstral peak prominence: a comprehensive analysis. *Biomed Signal Process Control.* 2014;14:42–54. https://doi.org/10.1016/j.bspc.2014.07.001.
27. Heman-Ackah YD, Sataloff RT, Laureyns G, et al. Quantifying the cepstral peak prominence, a measure of dysphonia. *J Voice.* 2014;28:783–788. https://doi.org/10.1016/j.jvoice.2014.05.005.
28. Heman-Ackah YD, Heuer RJ, Michael DD, et al. Cepstral peak prominence: a more reliable measure of dysphonia. *Ann Otol Rhinol Laryngol.* 2003;112:324–333. https://doi.org/10.1177/000348940311200406.
29. Lowell SY, Colton RH, Kelley RT, et al. Spectral- and cepstral-based measures during continuous speech: capacity to distinguish dysphonia and consistency within a speaker. *J Voice.* 2011;25:e223–e232. https://doi.org/10.1016/j.jvoice.2010.06.007.
30. Murton O, Hillman R, Mehta D. Cepstral peak prominence values for clinical voice evaluation. *Am J Speech Lang Pathol.* 2020;29:1596–1607. https://doi.org/10.1044/2020_AJSLP-20-00001.
31. Alharbi GG, Cannito MP, Buder EH, et al. Spectral/cepstral analyses of phonation in parkinson's disease before and after voice treatment: a preliminary study. *Folia Phoniatr Logop.* 2019;71:275–285. https://doi.org/10.1159/000495837.
32. Hartl DM, Hans S, Vaissière J, et al. Objective voice quality analysis before and after onset of unilateral vocal fold paralysis. *J Voice.* 2001;15:351–361. https://doi.org/10.1016/S0892-1997(01)00037-6.
33. Solomon NP, Awan SN, Helou LB, et al. Acoustic analyses of thyroidectomy-related changes in vowel phonation. *J Voice.* 2012;26:711–720. https://doi.org/10.1016/j.jvoice.2012.06.006.
34. Thijs Z, Knickerbocker K, Watts CR. The degree of change and relationship in self-perceived handicap and acoustic voice quality associated with voice therapy. *J Voice.* 2024;38:1352–1358. https://doi.org/10.1016/j.jvoice.2022.04.021.
35. Maryn Y, Corthals P, Van Cauwenberge P, et al. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice.* 2010;24:540–555. https://doi.org/10.1016/j.jvoice.2008.12.014.
36. McDonald WM. Overview of neurocognitive disorders. *Focus (Am Psychiatr Publ).* 2017;15:4–12. https://doi.org/10.1176/appi.focus.20160030.
37. Folstein MF, Folstein SE, McHugh PR. Mini-mental state examination. *PsycTESTS Dataset.* 2014. https://doi.org/10.1037/t07757-000.
38. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology.* 1983;33:1444–1452. https://doi.org/10.1212/WNL.33.11.1444.
39. Boersma P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proc Inst Phonetic Sci.* 2000;17:97–110.
40. Maryn Y, Weenink D. Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and acoustic voice quality index. *J Voice.* 2015;29:35–43. https://doi.org/10.1016/j.jvoice.2014.06.015.
41. Hillenbrand J, Houde RA. Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *J Speech Hear Res.* 1996;39:311–321. https://doi.org/10.1044/jshr.3902.311.
42. Hillenbrand J, Cleveland RA, Erickson RL. Acoustic correlates of breathy vocal quality. *J Speech Hear Res.* 1994;37:769–778. https://doi.org/10.1044/jshr.3704.769.
43. Yamout B, Fuleihan N, Hajj T, et al. Vocal symptoms and acoustic changes in relation to the expanded disability status scale, duration and stage of disease in patients with multiple sclerosis. *Eur Arch Otorhinolaryngol.* 2009;266:1759–1765. https://doi.org/10.1007/s00405-009-1003-y.
44. Brockmann-Bauser M, Van Stan JH, Carvalho Sampaio M, et al. Effects of vocal intensity and fundamental frequency on cepstral peak prominence in patients with voice disorders and vocally healthy controls. *J Voice.* 2021;35:411–417. https://doi.org/10.1016/j.jvoice.2019.11.015.
45. Lopes LW, de Abreu SR. Accuracy and cut-off values of cepstral measures in the clinical evaluation of Brazilian Portuguese speakers. *J Voice.* 2024. https://doi.org/10.1016/j.jvoice.2024.04.021.Published online May 8.
46. Delgado-Hernández J, León-Gómez N, Jiménez-Álvarez A. Diagnostic accuracy of the Smoothed Cepstral Peak Prominence (CPPS) in the detection of dysphonia in the Spanish language. *Loquens.* 2019;6:e058. https://doi.org/10.3989/loquens.2019.058.
47. Cantor-Cutiva LC, Ramani SA, Walden PR, et al. Screening of voice pathologies: identifying the predictive value of voice acoustic parameters for common voice pathologies. *J Voice.* 2023. https://doi.org/10.1016/j.jvoice.2023.12.005.Published online December 23.
48. Nguyen T, Kim M, Gwak J, et al. Investigation of brain functional connectivity in patients with mild cognitive impairment: a functional near-infrared spectroscopy (fNIRS) study. *J Biophotonics.* 2019;12:e201800298. https://doi.org/10.1002/jbio.201800298.
49. Shama K, krishna A, Cholayya NU. Study of harmonics-to-noise ratio and critical-band energy spectrum of speech as acoustic

indicators of laryngeal and voice pathology. *EURASIP J Adv Signal Process.* 2006;2007(1):085286. https://doi.org/10.1155/2007/85286.

50. Brinca LF, Batista APF, Tavares AI, et al. Use of cepstral analyses for differentiating normal from dysphonic voices: a comparative study of connected speech versus sustained vowel in European Portuguese female speakers. *J Voice.* 2014;28:282–286. https://doi.org/10.1016/j.jvoice.2013.10.001.
51. Iwarsson J. Reflections on clinical expertise and silent know-how in voice therapy. *Logop Phoniatr Vocol.* 2015;40:66–71. https://doi.org/10.3109/14015439.2014.949302.
52. Madill C, Nguyen DD, Yick-Ning Cham K, et al. The impact of Nasalance on cepstral peak prominence and harmonics-to-noise ratio. *Laryngoscope.* 2019;129:E299–E304. https://doi.org/10.1002/lary.27685.
53. Taylor S, Dromey C, Nissen SL, et al. Age-related changes in speech and voice: spectral and cepstral measures. *J Speech Lang Hear Res.* 2020;63:647–660. https://doi.org/10.1044/2019_JSLHR-19-00028.
54. S. K, Shetty R. An acoustic analysis on voice changes in adults and geriatrics. *Int J Health Sci Res.* 2023;13:106–118. https://doi.org/10.52403/ijhsr.20230513.