Research Article

Acoustic Measures of Dysphonia in Amyotrophic Lateral Sclerosis

Marc F. Maffei,[a] Jordan R. Green,[a,b] Olivia Murton,[a] Yana Yunusova,[c,d,e]
Hannah P. Rowe,[a] Farah Wehbe,[c,d] Kathleen Diana,[f] Katharine Nicholson,[f]
James D. Berry,[f] and Kathryn P. Connaghan[a]

[a] Department of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, MA [b] Speech and Hearing Bioscience and Technology Program, Harvard University, Cambridge, MA [c] Department of Speech-Language Pathology, University of Toronto, Ontario, Canada [d] Hurvitz Brain Sciences Program, Sunnybrook Research Institute, Toronto, Ontario, Canada [e] Toronto Rehabilitation Institute, University Health Network, Ontario, Canada [f] Department of Neurology, Neurological Clinical Research Institute, Massachusetts General Hospital, Boston

ABSTRACT

Purpose: Identifying efficacious measures to characterize dysphonia in complex neurodegenerative diseases is key to optimal assessment and intervention. This study evaluates the validity and sensitivity of acoustic features of phonatory disruption in amyotrophic lateral sclerosis (ALS).
Method: Forty-nine individuals with ALS (40–79 years old) were audio-recorded while producing a sustained vowel and continuous speech. Perturbation/noise-based (jitter, shimmer, and harmonics-to-noise ratio) and cepstral/spectral (cepstral peak prominence, low–high spectral ratio, and related features) acoustic measures were extracted. The criterion validity of each measure was assessed using correlations with perceptual voice ratings provided by three speech-language pathologists. Diagnostic accuracy of the acoustic features was evaluated using area-under-the-curve analysis.
Results: Perturbation/noise-based and cepstral/spectral features extracted from /a/ were significantly correlated with listener ratings of roughness, breathiness, strain, and overall dysphonia. Fewer and smaller correlations between cepstral/spectral measures and perceptual ratings were observed for the continuous speech task, although post hoc analyses revealed stronger correlations in speakers with less perceptually impaired speech. Area-under-the-curve analyses revealed that multiple acoustic features, particularly from the sustained vowel task, adequately differentiated between individuals with ALS with and without perceptually dysphonic voices.
Conclusions: Our findings support using both perturbation/noise-based and cepstral/spectral measures of sustained /a/ to assess phonatory quality in ALS. Results from the continuous speech task suggest that multisubsystem involvement impacts cepstral/spectral analyses in complex motor speech disorders such as ALS. Further investigation of the validity and sensitivity of cepstral/spectral measures during continuous speech in ALS is warranted.

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease characterized by the progressive loss of muscle strength and function. Over the course of the disease, most individuals with ALS will experience bulbar symptoms affecting speech, feeding, and swallowing (Haverkamp et al.,

1995). Over 80% of individuals with ALS develop dysarthria (B. Tomik & Guiloff, 2010), primarily a mixed spastic–flaccid subtype resulting from the deterioration of both upper motor neurons (UMNs) and lower motor neurons (LMNs) involved in speech production (Darley et al., 1969).

Dysphonia in ALS

Dysphonia, or an impairment of voice production, is a commonly reported sign of dysarthria in individuals

Correspondence to Kathryn P. Connaghan: kconnaghan@mghihp.edu.
Disclosure: The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

with ALS due to abnormal motor function of the larynx. Chen and Garrett (2005) conducted a medical record review of 44 patients with bulbar-onset ALS and reported that 48% presented with dysphonia. Detectable changes in voice quality or loudness are observed in both patients with spinal-onset symptoms and those with bulbar-onset symptoms (Robert et al., 1999). Commonly reported phonatory characteristics associated with dysphonia in ALS include hoarseness, roughness, strain, and breathiness (Carrow et al., 1974; Chen & Garrett, 2005; J. Tomik et al., 2015). In a study of 69 patients with ALS, 80% presented with a harsh voice quality; 67% with a breathy quality; and 59% with a strained–strangled vocal quality (Carrow et al., 1974). J. Tomik et al. (2015) reported hoarseness in 47% of their sample at initial evaluation, which increased to 55% at reevaluation 6 months later.

Measuring Dysphonia in ALS

Valid, reliable, and easily implemented measures of dysphonia are critical to efficacious assessment and treatment for individuals with ALS. Various quantitative acoustic measures have been used in prior literature to assess phonatory function in ALS, particularly to identify bulbar involvement, track bulbar decline over time, and monitor progress during clinical trials (see Chiaramonte & Bonfiglio, 2020, for a review).

Traditional (Perturbation/Noise-Based) Acoustic Measures

Among the most commonly used acoustic measures for quantifying dysphonia in ALS have been jitter (a measure of cycle-to-cycle variation in frequency) and shimmer (a measure of cycle-to-cycle variation in amplitude; J. Kent et al., 1992; Ramig et al., 1990; Robert et al., 1999; Silbergleit et al., 1997). These two measures are considered acoustic perturbation measures since they reflect the amount of involuntary variation (or *perturbation*) in the vocal signal. In addition, a noise measure—the harmonics-to-noise ratio (HNR; a measure of turbulent noise present in the voice signal)—has also been used. These features are often considered acoustic correlates of perceptual measures of voice. Both jitter and shimmer have primarily been associated with perceived roughness and overall voice quality (Arends et al., 1990; Hillenbrand, 1988; Lopes et al., 2012), whereas HNR has been associated with perceived breathiness and roughness (Bhuta et al., 2004; Hillenbrand, 1988).

Jitter, shimmer, and HNR have been used to identify aberrant voice production in subjects with ALS. For example, a study of 10 women with ALS found that five participants had abnormal levels of jitter and six had abnormal levels of shimmer relative to healthy controls (J. Kent et al., 1992). Significantly higher jitter values have been found in individuals with ALS compared with healthy controls even when the speakers with ALS present with perceptually normal voices (Silbergleit et al., 1997). A longitudinal case study examining acoustic voice features of a 69-year-old man with ALS revealed a trajectory of decreasing vocal quality over 6 months, from no initial vocal symptoms to vocal dysfunction acoustically characterized by increased shimmer and jitter and abnormal HNR (Ramig et al., 1990). Another longitudinal case study followed a 53-year-old woman with ALS over 2 years and documented highly variable values of shimmer, jitter, and HNR over time, rather than steady changes (R. D. Kent et al., 1991).

Although these studies suggest the efficacy of traditional perturbation- and noise-based measures of dysphonia in ALS, existing literature related to acoustic characterization of dysphonia in ALS is limited by several methodological factors. First, many studies utilizing these acoustic features in ALS include small sample sizes (e.g., J. Kent et al., 1992; Ramig et al., 1990; Strand et al., 1994), which precludes generalization to the heterogeneous population of individuals with ALS. Additionally, concerns regarding the psychometric limitations of these measures have been raised (Brockmann-Bauser & Drinnan, 2011; Carding et al., 2004; Martin et al., 1995; Yiu, 1999); for instance, the reliability of jitter values may decrease as the severity of dysphonia increases (Rabinov et al., 1995). Finally, perturbation analysis is limited to sustained vowels produced at a steady pitch (Garrett, 2013), since it relies on accurate time-based detection of cycle boundaries. Specifically, characteristics of continuous speech such as short vowel durations, fundamental frequency variation, pauses, and voiceless consonants may significantly impact perturbation measures (Awan et al., 2010).

Cepstral/Spectral Measures of Dysphonia

A growing body of literature has demonstrated the utility of cepstral and spectral approaches (e.g., cepstral peak prominence [CPP], low–high spectral ratio [L/H ratio], and related features) as an alternative to traditional acoustic measures for objectively measuring dysphonia (Awan et al., 2010; Brinca et al., 2014; Heman-Ackah et al., 2003, 2014; Yu et al., 2018). Significant advantages of cepstral/spectral measures over traditional acoustic measures include that they can be extracted from continuous speech (Heman-Ackah et al., 2002) and that they may be more reliable across the range of impairment severity than traditional features (Heman-Ackah et al., 2003).

CPP values are obtained from a signal's *cepstrum* (or, more formally, the real part of the cepstrum), which is calculated by applying an inverse Fourier transformation to the magnitude of the frequency spectrum. For signals with strong harmonic components, the cepstrum contains a tall peak at a location corresponding to the period of the harmonic frequencies, and the CPP is defined as the distance between that peak and a regression line through

the cepstrum as a whole. CPP values are generally higher in voices with strong periodicity and lower in voices characterized by aperiodic noise and have been shown to correlate with listener perception of overall dysphonia (Eadie & Baylor, 2006; Heman-Ackah et al., 2002; Jannetts & Lowit, 2014; Murton et al., 2020) and breathiness (Hillenbrand & Houde, 1996; Jannetts & Lowit, 2014; Schultz et al., 2021). There is less agreement regarding the correlation between cepstral measures and the perceptual characteristic of roughness, with some studies showing a significant relationship (da Silva Antonetti et al., 2020) and others showing no correlation (Heman-Ackah et al., 2002). The standard deviation of CPP (CPP *SD*) has also shown potential as an acoustic measure of dysphonia (Awan et al., 2014; Carson et al., 2016; Watts & Awan, 2011), likely in part because it is reflective of frequency and amplitude variations in normal speech intonation patterns (Watts & Awan, 2011).

The utility of cepstral/spectral analyses for discrimination and/or progress monitoring has been explored in clinical populations including individuals with Parkinson's disease (Behrman et al., 2020; Benba et al., 2017; Orozco-Arroyave et al., 2015), hypokinetic and ataxic dysarthria (Byeon, 2021; Jannetts & Lowit, 2014), and Friedreich ataxia (Carson et al., 2016), with results providing promising evidence that cepstral measures are associated with perceptual voice characteristics in dysarthric speakers. Indeed, the American Speech-Language-Hearing Association expert panel tasked with developing a protocol for instrumental voice evaluation (Patel et al., 2018) recommended CPP as a global measure of dysphonia severity. However, given the complex nature of voice and speech impairment in ALS including breathy, strained, and/or rough voice quality, we extended our evaluation of acoustic measures to include CPP *SD*, L/H ratio and its variance (the standard deviation of L/H ratio [L/H ratio *SD*]), and the Cepstral Spectral Index of Dysphonia (CSID).

L/H ratio, a spectral measure used in voice evaluation, is typically calculated as the ratio of spectral energy below 4 kHz and spectral energy at or above 4 kHz. L/H ratio has been shown to differentiate typical from dysphonic voices (Yu et al., 2018) and is correlated with the perceptual features of breathiness and overall dysphonia severity (Awan & Roy, 2006; Schultz et al., 2021), although its usefulness for differentiating dysphonic from nondysphonic voices may be limited (Lowell et al., 2012). L/H ratio *SD* reflects its variability across the duration of a voice sample, and a higher value on this measure may be interpreted as corresponding to better laryngeal support for dynamic adjustments of the larynx particularly in continuous speech (Watts, 2015). The CSID is a multivariate estimate of dysphonia severity that incorporates CPP, CPP *SD*, L/H ratio, and L/H ratio *SD* to approximate a rating from a 100-point visual analog scale of overall dysphonia severity (Awan et al., 2016).

To date, the application of cepstral/spectral measures to quantify phonatory function in individuals with ALS has been limited. Chiaramonte and Bonfiglio (2020) conducted a systematic review of 26 studies published between 1990 and 2019 that used acoustic methods to analyze voice in bulbar ALS; of note, no studies meeting their inclusion criteria used cepstral or spectral measures. Byeon et al. (2016) reported significant differences in values of CPP and L/H ratio between a group of eight women with ALS and 20 healthy controls. More recently, Eshghi et al. (2021) compared four groups of individuals with ALS: predominantly hypernasal, predominantly dysphonic, mixed hypernasal and dysphonic, and voice/hypernasality asymptomatic. They showed that values of CPP differentiated all groups except asymptomatic versus mixed and that the L/H ratio did not differ among groups.

The Effect of Task on Cepstral/Spectral Measures of Dysphonia

The nature of a speech stimulus can affect cepstral/spectral values. Zraick et al. (2005) found that CPP values differed significantly when measured during a sustained vowel or during continuous speech from the same speaker. Different sentences included in the protocol for the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) have different levels of correlation with the CSID, and none of those correlations are as strong as the correlation with a sustained vowel (Awan et al., 2010). Supporting this finding, Watts (2015) reported a significant main effect of sentence type on measurements of CPP, L/H ratio, and L/H ratio *SD* and recommended the use of a variety of stimuli during clinical voice assessment to overcome the effect of glottal and supraglottal differences on acoustic measurements.

Measuring Dysphonia in Complex Neurodegenerative Diseases Such as ALS

The complexity of neurodegenerative diseases such as ALS, in which multiple speech subsystems (i.e., respiratory, phonatory, resonatory, and/or articulatory) may be affected, complicates the direct assessment of single subsystems since each may have a significant confounding effect on the measurement of the others. For example, the co-occurrence of dysphonia and hypernasality (impairments of the phonatory and resonatory subsystems, respectively) in ALS has been shown to have contradictory effects on spectral energy (Eshghi et al., 2021), potentially reducing the validity of acoustic measures in populations in which both

subsystems are impacted. Foundational research validating acoustic voice measures has often focused on individuals with dysphonia at the exclusion of individuals with multiple subsystem impairment (e.g., Awan et al., 2010; Heman-Ackah et al., 2003; Kreiman et al., 2002; Rabinov et al., 1995). Voice analysis in ALS is also complicated by the variable manifestation of the disease across individuals (Swinnen & Robberecht, 2014) and within an individual across time (Ravits & La Spada, 2009). The diagnosis of ALS requires evidence of clinical signs of co-occurring UMN and LMN involvement. Lesions to these distinct neural pathways are thought to result in vastly different phonatory disorders, with UMN dysfunction producing a spastic dysphonia and LMN dysfunction producing a flaccid dysphonia (Colton et al., 2011).

The involvement of multiple speech subsystems also appears to impact auditory-perceptual voice evaluations, in which listeners make subjective judgments of dysphonic features such as roughness, breathiness, or strain. Auditory-perceptual voice assessments are very common and constitute the current gold standard for voice evaluation in ALS. However, they require specialized training, are prone to reliability constraints (Kreiman & Gerratt, 2010; Lu & Matteson, 2014; Webb et al., 2004), and may be complicated by the co-occurrence of subsystem impairments. For example, Imatomi et al. (2003) synthesized voices with varying degrees of roughness and hypernasality and found that severely rough voices were rated as less hypernasal than normal or moderately rough voices, especially in the case of severe hypernasality. Taken together, the above evidence suggests that the interaction of voice and speech impairments—which manifest specifically during continuous speech and will be less noticeable or absent during a task such as a sustained vowel—is likely to have a significant effect on the acoustic assessment of voice and the correlations between acoustic features and perceptual voice ratings.

Given the lack of validation of measures of dysphonia in ALS, the overarching aim of this research is to identify valid, reliable, and sensitive acoustic features for monitoring bulbar impairment in ALS. In the long term, this research aims to facilitate assessment and intervention to mitigate the devastating impacts of the disease. As an initial step toward that goal, we sought (a) to establish the criterion validity of acoustic features for evaluating dysphonia in ALS based on their relationship to the current gold standard (i.e., auditory-perceptual assessment) and (b) to examine the diagnostic accuracy of acoustic features for identifying dysphonia in this population. Specifically, our research questions are as follows:

Research Question 1. Are perturbation/noise-based and cepstral/spectral acoustic measures valid indicators of dysphonia severity in speakers with ALS?

To comprehensively address this question, we also examined the following:

Research Question 1.1. Does the correlation between cepstral/spectral measures and perceptual voice features differ by task?

Research Question 1.2. Is the correlation between cepstral/spectral measures and perceptual voice features impacted by speech impairment severity?

Research Question 2. What is the diagnostic accuracy of perturbation/noise-based and cepstral/spectral measures for detecting dysphonia among individuals with ALS?

Method

Data for this study were collected from participants as part of a longitudinal, multisite project conducted at the University of Nebraska–Lincoln and the University of Toronto. Additional data from four participants were collected through the Dominant Inherited ALS (DIALS) Network, a multicenter study at Massachusetts General Hospital and Washington University. All study procedures were approved by the institutional review boards of all institutions involved in data collection.

Participants

The participants included 49 individuals (ages 40–79 years, $M = 60.2$, $SD = 9.5$; 22 women) diagnosed with ALS by a neurologist using the revised El Escorial criteria (Brooks et al., 2000). Inclusion criteria for the study were being a native speaker of English and demonstrating sufficient cognitive, hearing, vision, and literacy skills to complete the research tasks. Participants were excluded if they had a history of other neurological disorders (e.g., stroke). Participants passed hearing screenings at 25 dB for both ears at 500, 1000, 2000, and 4000 Hz, except participants from the DIALS study, who reported no hearing concerns. Participants demonstrated a range of dysarthria severity on a variety of assessments, from no impairment to severely affected. For participants whose scores on the Speech Intelligibility Test (SIT; Yorkston et al., 2007) were available ($n = 45$), scores ranged from 6.4% to 100% intelligibility ($M = 87.9$, $SD = 21.5$), and their speaking rate derived from the SIT ranged from 37.8 to 251.9 words per minute ($n = 44$, $M = 138.8$, $SD = 53.5$). Table 1 presents relevant demographic and clinical information about the sample.

Procedure

Participants were audio-recorded using a high-quality lapel microphone (Audio-Technica AT831R) located

Table 1. Participant demographics.

Variable	n	M	SD	Min	Max
Age (years)	49	60.2	9.5	40.7	79.4
Time since symptom onset (months)	42	36.6	34.0	7.0	183.0
Intelligibility (% words understood)	45	87.9	21.5	6.4	100.0
Speaking rate (words per minute)	44	138.8	53.5	37.8	251.9
Sex					
Female	22				
Male	27				
Site of onset					
Bulbar	7				
Spinal	31				
Mixed	6				
Unavailable	5				

Note. Min = minimum; Max = maximum.

approximately 15 cm from the mouth, with a sampling rate of 44.1 kHz and 16-bit resolution. In one session, participants were recorded producing a sustained vowel and reading a passage aloud. For the sustained vowel task, participants were instructed to produce /a/ for as long as possible on one breath at their typical pitch and loudness following a model from the researcher. For the passage reading task, participants read aloud the Bamboo Passage, a 97-word paragraph written at a fifth-grade reading level initially designed to aid in automatic pause boundary detection (Green et al., 2004). This study used existing data from projects with aims such as quantifying the rate of subsystem decline in ALS using pause frequency and duration; thus, the most commonly available continuous speech sample in our sample was the Bamboo Passage.

Acoustic Processing

Acoustic analysis consisted of a multistep process. First, all files were down-sampled to 22050 Hz. For sustained vowels, a steady-state 2-s portion of the vowel beginning at least 1 s after voicing onset was extracted in Praat (Boersma & Weenink, 2022). For the Bamboo Passage recordings, files were trimmed to leave approximately 500 ms before the beginning of the first word and after the end of the last word. The Loudness Normalization function in Audacity (Audacity Team, 2021) was used to standardize root-mean-square

(RMS) amplitude to −20 dB for sustained vowels and to −30 dB for continuous speech, facilitating perceptual ratings by generating signals of comparable perceived loudness. Continuous speech was standardized to a lower RMS amplitude than sustained vowels to avoid clipping related to plosive bursts.

Perturbation- and Noise-Based Measures

The Voice Report command in Praat was implemented on the trimmed sustained /a/ recordings to extract the following features: local jitter (i.e., frequency perturbation calculated as the average absolute difference between consecutive periods, divided by the average period), local shimmer (i.e., amplitude perturbation calculated as the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude), and HNR (i.e., the ratio of acoustic periodicity to noise, expressed in decibels).

Cepstral/Spectral Measures

Cepstral/spectral measures were extracted from the sustained vowel and continuous speech samples using Analysis of Dysphonia in Speech and Voice (ADSV; PENTAX Medical, 2011). The cepstral/spectral variables of interest and their definitions (as provided in ADSV documentation) are provided in Table 2. Participant sex was accounted for when calculating CSID values derived from sustained /a/; a single calculation

Table 2. Cepstral/spectral variables with definitions and interpretations.

Variable	Description
CPP	Cepstral peak prominence. The mean difference between the actual and linear regression predicted cepstral peaks for the selected voiced data frames.
CPP SD	CPP standard deviation. The standard deviation of the mean differences between the actual and linear regression predicted cepstral peaks for the selected voiced data frames.
L/H ratio	Low–high spectral ratio. The mean ratio of signal energy below 4 kHz to the energy above 4 kHz for the selected voiced data frames.
L/H ratio SD	L/H ratio standard deviation. The standard deviation of the ratio of signal energy below 4 kHz to the energy above 4 kHz for the selected voiced data frames.
CSID	Cepstral Spectral Index of Dysphonia. An acoustic estimate of dysphonia severity incorporating multiple cepstral- and spectral-based measures.

was used for all participants during the continuous speech task (Awan et al., 2016; Peterson et al., 2013).

Perceptual Analysis

Three of the authors (M.F.M., O.M., and H.P.R.), who are speech-language pathologists with expertise in voice/dysarthria research and perceptual voice evaluation (M = 5.5 years), independently rated the phonatory quality of all participants using the sustained /a/ and, in a separate rating session, the Bamboo Passage recordings. For each participant, *roughness*, *breathiness*, *strain*, and *overall dysphonia* were rated on 100-point visual analog scales adapted from the CAPE-V (Kempster et al., 2009) with descriptive labels of "normal" corresponding to 0 and "severe" corresponding to 100. For the Bamboo Passage recordings, listeners additionally rated each participant's *dysarthria severity* and *resonance impairment*, also on 100-point visual analog scales, in order to aid our interpretation of perceptual ratings within the context of ALS-related dysarthria. Data were collected via REDCap (Harris et al., 2009), a browser-based electronic data capture system. Each rater heard the files in a different randomized order. Raters were permitted, but not required, to listen to each audio file up to 3 times. Operational definitions for each voice feature were discussed with each rater before their rating sessions and presented on the screen during all ratings. *Roughness* was defined as perceived irregularity in voice; *breathiness*, as excess air escape during voicing; *strain*, as excess vocal effort/hyperfunction; *overall dysphonia*, as a global estimate of voice quality during speech; *hypernasality*, as the degree of altered resonance due to increased resonance in the nasal cavity; and *overall speech impairment*, as the overall speech impairment severity considering all speech subsystems. Each participant received a single rating for each perceptual feature per task, calculated as the average of the three listeners' ratings.

Inter- and Intrarater Reliability

Reliability was assessed within (intrarater) and across (interrater) the three listeners. To assess interrater reliability, ratings for each perceptual feature (i.e., roughness, breathiness, strain, overall dysphonia, hypernasality, and overall speech impairment) by the three raters were compared. To assess intrarater reliability, a randomly selected 30% of the recordings (15 recordings: nine men, six women) was rated a second time by each of the three raters, resulting in 64 total ratings (49 + 15) per task completed by each listener. Interrater reliability was measured using two-way random effects intraclass correlation coefficients (ICCs; absolute agreement, mean of k raters type)–ICC(2,k)–calculated with the *irr* package in R (Gamer et al., 2012). Intrarater reliability was assessed using two-way mixed-effects ICCs (consistency, mean of k raters type)–ICC(3, k)–for each perceptual feature for each rater using the *irr* package in R. ICCs were then averaged across raters to generate an intrarater reliability score for each feature.

Statistical Analysis

Criterion Validation of Acoustic Features (Research Question 1)

The validity of acoustic measures was evaluated using correlation coefficients between each acoustic feature and the perceptual voice measures. The Kolmogorov–Smirnov test of normality was conducted to assess the normality of model residuals for each pair of perceptual/acoustic features (e.g., roughness and CPP derived from the sustained vowel task). For pairs that did not violate the assumption of normality (p > .05; n = 11, 21% of correlations), Pearson correlation coefficients and associated p values were calculated; for pairs that did violate the assumption of normality (p ≤ .05; n = 41), Spearman's rank correlation coefficients and associated p values were calculated. To examine differences in correlations across tasks, the correlations between perceptual ratings and acoustic features were analyzed during both the sustained vowel task and the continuous speech task due to known effects of speech stimulus on cepstral/spectral values.

To examine the impact of speech impairment severity on cepstral/spectral values, three separate post hoc analyses were conducted on data from the continuous speech task. To examine the effect of overall speech impairment on the correlations, the sample was stratified into two groups: a "no/low speech impairment" group of subjects who received a perceptual rating of 10 or less on *overall speech impairment* and a "high speech impairment" group of subjects who were rated higher than 10. To examine the effect of intelligibility, the sample was divided into a "high intelligibility" group of subjects who were above 94% intelligible on the SIT and a "low intelligibility" group of those who were 94% intelligible or less. This cutoff was chosen because an intelligibility level above 94% is the recommended severity grouping corresponding to clinical severity ratings of "normal" (Stipancic et al., 2021). To examine the effect of hypernasality, the sample was divided into a "no/low hypernasality" group who were rated as 10 or less on *hypernasality* and a "high hypernasality" group who received a rating above 10. While standards for visual analog scale ratings of hypernasality severity are not currently established, Baylis et al. (2015) found that most subjects rated as 0 on a 0–5 equal-appearing interval scale of hypernasality were also rated below 10 on a visual analog scale of hypernasality. Correlation coefficients and associated p values calculated for each acoustic feature/perceptual rating pair were calculated for these six groups using the same methods described above for the entire pooled sample.

Diagnostic Accuracy of Acoustic Features (Research Question 2)

The area under the receiver operating characteristic (ROC) curve (AUC) was then calculated and used to

quantify the overall performance of each acoustic variable at differentiating perceptually normal and dysphonic voices (i.e., its diagnostic accuracy). The AUC can range from 0 to 1, with .5 indicating that a feature is no better than chance at discriminating between participants, a .7–.8 being considered adequate, and a .8–.9 being considered excellent (Hosmer et al., 2013). Given known voice changes related to normal aging of the larynx (Kendall, 2007) and the age range of the participants in this study, as well as the strong but inherently imperfect reliability of the CAPE-V (Karnell et al., 2007), a cutoff of 5 or below on any of the perceptual features was considered to be "normal" or "non-dysphonic" for the AUC analysis (i.e., mean ratings of 0–4 on the 100-point visual analog scales were considered normal, and mean ratings of 5–100 were considered abnormal). Additionally, the following values were calculated: the optimal threshold of each feature for classifying dysphonic versus perceptually normal voices, the specificity of each feature (i.e., the proportion of participants with a perceptually normal voice who were correctly classified), the sensitivity of each feature (i.e., the proportion of participants with a perceptually dysphonic voice who were correctly classified), and the accuracy of each feature (i.e., the number of participants correctly classified). These analyses were conducted using the *pROC* (Robin et al., 2011) and *ROCR* (Sing et al., 2005) packages in R.

Results

Reliability of Listener Ratings

The inter- and intrarater reliability of the perceptual voice and speech features are presented in Table 3. All values in Table 3 were significant at the level of $p < .001$. Using the guidelines provided by Koo and Li (2016), ICCs revealed good interrater reliability (i.e., .75–.90 averaged across three raters) for all perceptual features, except hypernasality and speech impairment in the continuous speech task, which were both excellent (i.e., > .90), and strain in the sustained vowel task, which was moderate (i.e., .50–.75). Intrarater reliability for all perceptual features ranged from good to excellent, except breathiness in the continuous speech task, which was moderate.

Listener Ratings and Acoustic Features

Means, standard deviations, and ranges for the ratings of each perceptual voice feature are included in Table 4. In the sustained vowel task, overall dysphonia was rated as the most impaired ($M = 18.80$, $SD = 17.15$), whereas breathiness was rated as the least impaired ($M = 11.95$, $SD = 14.09$); in the continuous speech task, strain was rated as the most impaired ($M = 17.64$, $SD = 19.61$), whereas breathiness was rated as the least impaired ($M = 8.36$, $SD = 10.82$). Table 5 shows the means, standard deviations, and ranges of values for all acoustic features included in this study.

Criterion Validity of Acoustic Features (Research Question 1)

A summary of the correlations between perceptual ratings and values of each acoustic variable is presented in Table 6. For the sustained vowel task, all three perturbation/noise-based features (i.e., jitter, shimmer, and HNR) demonstrated moderate to very strong (correlation coefficient > .5 and > .8, respectively) and statistically significant ($p < .001$) correlations with all four perceptual features of dysphonia. Cepstral/spectral acoustic measures, especially those related to CPP (i.e., CPP, CPP *SD*, and CSID), were also significantly correlated with perceptual

Table 3. Inter- and intrarater reliability scores by three skilled raters for perceptual ratings of voice quality, hypernasality, and speech impairment during (a) a sustained vowel and (b) continuous speech.

Feature	Interrater reliability ICC(2,k)	Intrarater reliability ICC(3,k), M (range)
Sustained /a/		
Roughness	.81	.93 (0.91–0.95)
Breathiness	.86	.84 (0.76–0.90)
Strain	.65	.85 (0.78–0.96)
Overall dysphonia	.89	.93 (0.89–0.96)
Hypernasality	.81	.94 (0.88–0.97)
Continuous speech		
Roughness	.80	.85 (0.82–0.90)
Breathiness	.87	.73 (0.50–0.89)
Strain	.89	.93 (0.89–0.98)
Overall dysphonia	.81	.83 (0.67–0.95)
Hypernasality	.91	.93 (0.91–0.94)
Speech impairment	.97	.95 (0.94–0.96)

Note. All values were significant at the level of $p < .001$. ICC = intraclass correlation coefficient.

Table 4. Descriptive statistics of mean perceptual voice ratings across subjects.

Feature	M	SD	Min	Max
Sustained /a/				
Roughness	15.18	16.86	0.00	67.33
Breathiness	11.95	14.09	0.00	64.33
Strain	13.04	12.84	0.00	43.00
Overall dysphonia	18.80	17.15	0.00	61.67
Continuous speech				
Roughness	13.95	14.12	0.00	74.33
Breathiness	8.36	10.82	0.00	50.00
Strain	17.64	19.61	0.00	69.67
Overall dysphonia	17.37	16.20	0.00	71.67

Note. N = 49, range = 0–100. Min = minimum; Max = maximum.

Table 5. Descriptive statistics of acoustic features.

Feature	Sustained vowel				Continuous speech			
	M	*SD*	Min	Max	*M*	*SD*	Min	Max
Jitter	0.007	0.010	0.001	0.072	—	—	—	—
Shimmer	0.050	0.034	0.008	0.201	—	—	—	—
HNR	19.473	5.862	0.132	32.591	—	—	—	—
CPP	11.364	2.576	1.926	15.437	5.240	1.412	1.952	9.223
CPP *SD*	1.037	0.731	0.367	3.040	3.525	0.652	1.851	5.043
L/H ratio	35.308	5.372	23.312	44.819	32.735	5.781	19.950	42.022
L/H ratio *SD*	1.734	0.722	0.796	4.944	9.421	1.743	4.873	13.798
CSID	16.540	18.870	−13.669	81.335	6.340	15.390	−24.052	42.160

Note. Em dashes indicate that there are no values to report for these features (i.e., jitter, shimmer, and HNR) for the continuous speech task because they cannot be calculated from continuous speech. Min = minimum; Max = maximum; HNR = harmonics-to-noise ratio; CPP = cepstral peak prominence; CPP SD = CPP standard deviation; L/H ratio = low–high spectral ratio; L/H ratio SD = L/H ratio standard deviation; CSID = Cepstral Spectral Index of Dysphonia.

measures—in several cases, with stronger correlations than traditional perturbation/noise measures (i.e., CPP × Breathiness, CSID × Strain, and CSID × Overall Severity). Correlations between cepstral/spectral measures and perceptual ratings during the continuous speech task were fewer and smaller, although CPP, CPP *SD*, and L/H ratio were significantly correlated with perceptual ratings. Additional analysis of the correlations among severity groups based on overall speech impairment, intelligibility, and hypernasality is presented in Table 7, demonstrating that, in general, correlations during continuous speech are stronger in subjects with less severe speech impairment.

Diagnostic Accuracy of Acoustic Features for Differentiating Normal and Dysphonic Voices (Research Question 2)

The AUC, threshold, specificity, sensitivity, and overall classification accuracy for each acoustic feature are presented in Table 8. All perturbation/noise-based acoustic

features (i.e., jitter, shimmer, and HNR) extracted from the sustained /a/ demonstrated outstanding AUCs (i.e., > .90; Hosmer et al., 2013). CSID had an excellent AUC (i.e., .80–.90), and CPP, CPP *SD*, and L/H ratio *SD* each had AUCs in the acceptable range (i.e., .70–.90). L/H ratio did not demonstrate an adequate ability to differentiate these groups (i.e., < .70). For the continuous speech task, none of the assessed acoustic features (i.e., CPP, CPP *SD*, L/H ratio, L/H ratio *SD*, or CSID) demonstrated an adequate ability to differentiate between groups.

Discussion

Consistent with previous literature reporting dysphonia among individuals with ALS, most participants in our sample were perceptually rated as presenting with dysphonia characterized by roughness, strain, and/or breathiness. However, the complex manifestation of UMN and LMN dysfunction in individuals with ALS, particularly during

Table 6. Correlations between perceptual voice ratings and acoustic measures.

Feature	Sustained /a/				Continuous speech			
	Roughness	Breathiness	Strain	Severity	Roughness	Breathiness	Strain	Severity
Jitter	.79***	.59***	.66***	.71***	—	—	—	—
Shimmer	.77***	.65***	.60***	.71***	—	—	—	—
HNR	−.81***	−.59***	−.64***	−.69***	—	—	—	—
CPP	−.53***	−.74***	−.45**	−.67***	.01	−.19	.33*	.15
CPP *SD*	.68***	.37**	.64***	.61***	−.26	−.40**	−.03	−.13
L/H ratio	−.17	.01	−.22	−.09	.15	.31*	.37**	.32*
L/H ratio *SD*	.32*	.20	.40**	.33*	−.20	−.07	−.27	−.20
CSID	.78***	.63***	.69***	.75***	.24	.03	−.19	−.05

Note. Em dashes indicate that there are no values to report for these features (i.e., jitter, shimmer, and HNR) for the continuous speech task because they cannot be calculated from continuous speech. HNR = harmonics-to-noise ratio; CPP = cepstral peak prominence; CPP *SD* = CPP standard deviation; L/H ratio = low–high spectral ratio; L/H ratio *SD* = L/H ratio standard deviation; CSID: Cepstral Spectral Index of Dysphonia.

*p ≤ .05. **p ≤ .01. ***p ≤ .001.

Table 7. Correlation coefficients and significance levels between perceptual ratings and acoustic features during continuous speech, separated by speech impairment, intelligibility, and hypernasality severity levels.

Correlation pair	Pooled $N = 49$	Speech impairment		Intelligibility		Hypernasality	
		No/low $n = 23$	High $n = 26$	High $n = 36$	Low $n = 9$	No/low $n = 34$	High $n = 15$
CPP ×							
Roughness	.01	−.47*	.01	−.32	.04	−.64***	−.26
Breathiness	−.19	−.42*	−.27	−.46*	−.13	−.49**	−.43
Strain	.33*	−.19	.42*	−.22	.60*	−.33	.24
Voice severity	.15	−.27	.17	−.30	.31	−.46**	−.16
CPP SD ×							
Roughness	.26	−.51*	−.19	−.42*	−.47	−.61***	−.45
Breathiness	−.40**	−.59**	−.40*	−.59***	−.37	−.64***	−.35
Strain	−.03	−.31	.13	−.36*	.01	−.46**	−.03
Voice severity	−.13	−.40	−.02	−.44*	−.12	−.56***	−.17
L/H ratio ×							
Roughness	.15	.19	−.35	.30	.23	.42*	−.29
Breathiness	.31*	.24	−.08	.34	.26	.48**	−.15
Strain	.37**	.45*	−.10	.47**	.35	.57***	.00
Voice severity	.32*	.20	−.09	.35	.49	.48**	−.01
L/H ratio SD ×							
Roughness	−.20	−.07	−.11	−.05	−.02	−.22	−.06
Breathiness	−.07	.22	−.04	−.10	.08	−.07	.04
Strain	−.27	−.22	−.26	−.16	−.17	−.25	−.33
Voice severity	−.20	−.04	−.17	−.16	−.02	−.23	−.10
CSID ×							
Roughness	.24	.45*	.29	.47**	.00	.51**	.53*
Breathiness	.03	.15	.10	.23	−.13	.21	.35
Strain	−.19	−.11	−.28	.07	−.47	.09	.02
Voice severity	−.05	.15	.00	.22	−.37	.29	.21

Note. CPP = cepstral peak prominence; CPP SD = CPP standard deviation; L/H ratio = low–high spectral ratio; L/H ratio SD = L/H ratio standard deviation; CSID = Cepstral Spectral Index of Dysphonia.

*$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.

---

continuous speech, appears to complicate the relationship between acoustic measures and perceptual ratings of dysphonia. This study was designed to examine the efficacy of various acoustic measures for the assessment of dysphonia in ALS, including both traditional perturbation/ noise-based measures (i.e., jitter, shimmer, and HNR) and cepstral/spectral measures of voice quality (i.e., CPP, L/H ratio, CSID, and the variability of CPP and L/H ratio), which are relatively understudied among individuals with ALS. An optimal acoustic measure of voice quality should

Table 8. Classification accuracy values for each acoustic feature.

Statistic	Jitter	Shimmer	HNR	CPP	CPP SD	L/H ratio	L/H ratio SD	CSID
Sustained vowel								
AUC	**.908**	**.914**	**.910**	**.764**	**.733**	.585	**.731**	**.826**
Threshold	0.004	0.034	22.858	10.770	0.790	33.617	1.173	6.670
Specificity	1.000	1.000	.900	1.000	.900	.800	.600	.800
Sensitivity	.718	.769	.846	.487	.513	.462	.923	.795
Accuracy	.776	.816	.857	.592	.592	.531	.857	.796
Continuous speech								
AUC	—	—	—	.562	.690	.660	.564	.562
Threshold	—	—	—	4.763	3.545	30.184	9.299	3.666
Specificity	—	—	—	.846	.769	.615	.769	.692
Sensitivity	—	—	—	.472	.667	.722	.500	.611
Accuracy	—	—	—	.571	.694	.694	.571	.633

Note. Em dashes indicate that there are no values to report for these features (i.e., jitter, shimmer, and HNR) for the continuous speech task because they cannot be calculated from continuous speech. HNR = harmonics-to-noise ratio; CPP = cepstral peak prominence; CPP SD = CPP standard deviation; L/H ratio = low–high spectral ratio; L/H ratio SD = L/H ratio standard deviation; CSID = Cepstral Spectral Index of Dysphonia; AUC = area under the curve (boldface indicates an AUC above 0.7).

be reliable, reproducible, and correlated with dysphonia severity (Heman-Ackah et al., 2003), and with these criteria in mind, we examined (a) the criterion validity of each acoustic measure via their correlations with perceptual ratings and (b) the diagnostic accuracy of each acoustic measure for differentiating speakers with and without perceptually dysphonic voices.

Summary of Findings

Sustained Vowel Task

For the sustained /a/ task, both analyses confirmed that jitter, shimmer, and HNR are robust acoustic measures of dysphonia, which were strongly correlated with four of the most common auditory-perceptual features of vocal dysfunction: roughness, breathiness, strain, and overall dysphonia severity. These findings confirm previous literature demonstrating associations between jitter and shimmer with roughness and overall voice quality (Arends et al., 1990; Hillenbrand, 1988; Lopes et al., 2012) and between HNR and breathiness and roughness (Bhuta et al., 2004; Hillenbrand, 1988). Additionally, our findings provide evidence supporting jitter and shimmer as markers of breathiness and strain and HNR as a marker of strain and overall dysphonia severity. Our findings also add to the literature reporting abnormal levels of jitter, shimmer, and HNR in individuals with ALS (e.g., J. Kent et al., 1992; R. D. Kent et al., 1991; Ramig et al., 1990; Silbergleit et al., 1997).

Strong and significant correlations were also found between cepstral/spectral acoustic features and perceptual ratings in the sustained vowel task. In fact, the strongest acoustic correlate of breathiness was CPP, and the strongest acoustic correlate of both strain and overall dysphonia severity was CSID. In particular, it appears that acoustic features related to CPP (i.e., CPP, CPP *SD*, and CSID) correlate well with perceptual ratings in the sustained vowel task, confirming prior findings (Heman-Ackah et al., 2002; Hillenbrand & Houde, 1996). Interestingly, despite literature linking L/H ratio to perceptual ratings of dysphonia (e.g., Awan & Roy, 2006; Schultz et al., 2021; Yu et al., 2018, although see Eshghi et al., 2021; Lowell et al., 2012), we did not observe significant correlations between L/H ratio and any of the perceptual features. However, L/H ratio *SD* demonstrated several significant correlations of fair strength (i.e., correlation coefficient > .3) with ratings of roughness, strain, and overall dysphonia severity.

AUCs are an effective way to summarize the overall diagnostic accuracy of an assessment (Mandrekar, 2010). For the sustained vowel task, all perturbation/noise-based acoustic features (i.e., jitter, shimmer, and HNR) demonstrated outstanding discrimination between dysphonic and nondysphonic voices (i.e., AUCs > .90; Hosmer et al., 2013). Cepstral/spectral measures also discriminated the groups with excellent (i.e., CSID) or acceptable (i.e., CPP, CPP *SD*, and L/H ratio *SD*) diagnostic accuracy, defined as an AUC > .8 and > .7, respectively, supporting the notion that CPP can effectively discriminate between perceptually dysphonic and nondysphonic voices (Awan & Roy, 2005; Sauder et al., 2017). Of note, while perceptual ratings of dysphonia varied widely among our sample, the mean values of each rating suggest that our sample was skewed toward less dysphonic voices. The fact that most acoustic features demonstrated solid diagnostic accuracy despite our relatively mild sample supports the robustness of these features, specifically that they perform well even with low levels of dysphonia.

Continuous Speech (Reading Passage Task)

In the continuous speech task, CPP and CPP *SD* were each significantly correlated with only one perceptual feature (i.e., CPP × Strain and CPP *SD* × Breathiness), and CSID was not significantly correlated with any perceptual features. The negative correlation between CPP *SD* and breathiness was expected since CPP *SD* is sensitive to changes in frequency and amplitude that are likely diminished in the presence of breathiness (Watts & Awan, 2011). A potential explanation for the lack of correlations with CSID is that this measure—an acoustic estimate of dysphonia severity based on an algorithm incorporating several cepstral- and spectral-based measures related to CPP and L/H ratio (Awan et al., 2016; Peterson et al., 2013)—has only been validated for use with sustained /a/ and with the protocol of the CAPE-V (Awan et al., 2016).

Some of the correlations between cepstral/spectral measures and perceptual ratings in the continuous speech task were unexpected. First, L/H ratio was significantly correlated with breathiness, strain, and overall dysphonia severity during the continuous speech task but was not correlated with any perceptual features in the sustained vowel task. Of further interest, these significant correlations in the continuous speech task were all positive. Typically, a higher L/H ratio—indicating a larger concentration of spectral energy in the fundamental frequency and lower formats—is associated with perceptually nondysphonic voices (Watts & Awan, 2011). However, our results indicated that as ratings of breathiness, strain, and overall dysphonia severity increased in our sample, so did L/H ratio. Because the perceptual feature with the strongest positive correlation with L/H ratio was strain, it may be that speakers with ALS who demonstrate a strained vocal quality are producing less high-frequency energy, because laryngeal hyperfunction may dampen even a normal, nondysphonic amount of high-frequency noise.

Second, the only significant correlation with CPP in the continuous speech task (i.e., CPP × Strain) suggested that, in our sample, higher CPP values—typically

associated with a greater periodicity and stability of phonation—were associated with ratings of increased strain. This result was unexpected since CPP and strain have been shown to be strongly negatively correlated in speakers with dysphonia (Lowell et al., 2012). However, a relatively high CPP value associated with strain is not unprecedented. Anand et al. (2019) found a high positive correlation (r = .80, p < .001) between CPP and strain during a sustained vowel task. Wolfe and Martin (1997) found higher values of CPP in strained voices than in hoarse or breathy voices. Kapsner-Smith et al. (2022) did not find a significant difference in CPP between speakers with vocal hyperfunction (typically associated with strain) and nondysphonic controls. They not only attributed the absence of group differences primarily to the relatively low overall dysphonia severity ratings of their hyperfunctional group but also acknowledged the potential role of unexplored differences between hyperfunctional speakers and other populations. Indeed, cepstral/spectral analyses are similarly underinvestigated in ALS, and such unexpected relationships may be idiosyncratic to the disease.

The Effect of Task on Acoustic Voice Assessment in ALS

In the current investigation, the correlations between cepstral/spectral measures and ratings of voice quality during continuous speech differed in meaningful ways from the same correlations derived from the sustained vowel task, in line with previous findings (Phadke et al., 2020; Watts & Awan, 2011). During continuous speech, cepstral measures are affected by fluctuations in a variety of laryngeal and articulatory factors including vocal intensity, fundamental frequency, sound pressure level, syllable stress, vowel context, and vowel type (Awan et al., 2012; Phadke et al., 2020; Sampaio et al., 2020).

During sustained vowel production, several cepstral/spectral measures—particularly those related to CPP—were strongly correlated with perceptual ratings of vocal quality, matching or outperforming the traditional perturbation/noise measures in the correlation analyses. Sustained vowels (typically, as in this study, /a/) offer a standardized voice sample with several practical advantages in a clinical setting, including ease of elicitation and production. They are also less confounded by co-occurring articulatory impairments that may disrupt a listener's ability to focus on the voice signal (de Krom, 1994). However, there is evidence that acoustic measures derived from sustained vowels are not valid clinical indices of the severity of dysphonia in continuous speech (Qi & Hillman, 1997; Wolfe et al., 1995). Furthermore, the requirement that commonly used acoustic measures of voice quality (e.g., jitter and shimmer) must be obtained during sustained vowel production limits the applicability of these measures to more ecologically valid assessment tasks such as

continuous speech. Of note, shimmer and jitter are similarly impacted by vocal intensity and the specific vowel being produced, as well as by fundamental frequency (Brockmann et al., 2011) and vowel type (Akif Kiliç et al., 2004).

Continuous speech, on the other hand, is more representative of habitual voice use patterns (de Krom, 1995; Eadie & Baylor, 2006), containing pitch and loudness variations that serve as important perceptual indicators of vocal dysfunction (Askenfelt & Hammarberg, 1986). CPP has been shown to correlate with perceptual measures obtained from continuous speech samples (Heman-Ackah et al., 2002), and CPP SD is sensitive to features of continuous speech that impact the variability in the voice signal's periodicity, such as changes in the vowel spectrum and changes in the frequency spectrum due to intonation patterns (Watts & Awan, 2011). Importantly, in our current investigation, the correlations during continuous speech improved when our correlation analysis was limited to speakers with perceptually normal or low levels of overall speech severity and hypernasality, suggesting that the acoustic assessment of continuous speech is complicated by various factors related to dysarthric speech.

The Effect of Speech Impairment on Voice Assessment in ALS

The assessment of dysphonia in complex neurological diseases such as ALS requires careful consideration of the influences of multisubsystem impairment on both perceptual ratings and acoustic features. In this study, the lack of expected correlations between cepstral/spectral features and roughness, breathiness, and overall severity in the continuous speech task was largely unexpected, given that CPP has been shown to be a valid indicator of vocal quality across dysphonia severity levels (Murton et al., 2020) and is even more strongly correlated with breathiness and roughness during continuous speech than during a sustained vowel task (Heman-Ackah et al., 2002). It may be that particular characteristics of motor speech disorders (e.g., reduced articulation rate, reduced vowel space, and increased frequency and length of pausing) confound cepstral/spectral acoustic measures when applied to individuals with complex neurodegenerative diseases. Additionally, the mixed spastic–flaccid dysarthria associated with ALS, which may manifest as variable dysphonia profiles within and across speakers, has the potential to "cancel out" straightforward associations among individuals with complex neurodegenerative diseases.

To address the lack of significant correlations in the continuous speech task, we conducted post hoc correlation analyses within groups of individuals with and without perceptual hypernasality, reduced intelligibility, and overall speech impairment. These analyses revealed compelling findings revealing that the associations of cepstral/spectral

features and perceptual ratings of vocal quality are not consistent across dysarthria severity levels in individuals with ALS. For instance, in the impaired-intelligibility group (i.e., speakers with < 94% words intelligible), only a single correlation was significant: CPP × Strain. Of course, this division resulted in a small number of speakers in the impaired-intelligibility group ($n = 9$), reducing the statistical power of the analysis and making it less likely that a significant correlation, if present, could be identified. However, when we examined the impaired-speech group, which had a larger number of speakers ($n = 26$), there was a similar decrease in the number of significant correlations, with only CPP × Strain and CPP *SD* × Breathiness remaining significant. Of particular note is that, within the no/low speech impairment group—which is conceptually more similar to a sample with pure dysphonia in the absence of dysarthria—CPP was significantly correlated with roughness and breathiness in continuous speech, as expected from prior literature (da Silva Antonetti et al., 2020; Heman-Ackah et al., 2002; Hillenbrand & Houde, 1996; Schultz et al., 2021).

However, perhaps the most striking post hoc analysis was the division of our sample by hypernasality severity. In addition to the acoustic features of dysphonia addressed in this study, there has also been substantial progress made regarding acoustic measures of hypernasality, including one-third octave spectra of isolated vowels (Eshghi et al., 2021; Kataoka et al., 2001; Lee et al., 2003). However, acoustic features quantifying vocal and resonatory quality have been shown to correlate with perceptual ratings primarily when a single subsystem (e.g., phonation or resonance) is impacted. The effect of co-occurring dysphonia and hypernasality—a common feature of dysarthria due to ALS—on these features remains largely unknown. Eshghi et al. (2021) examined this issue and found that CPP values were significantly different among hypernasal-only, dysphonic-only, and mixed hypernasal–dysphonic speaker groups. When we examined correlations among the no/low hypernasality group—which we again posit is more similar to a group of speakers with pure dysphonia without the confounding effects of co-occurring dysarthria—CPP, CPP *SD*, and L/H ratio all became significantly correlated with all four perceptual measures, with the sole exception of CPP × Strain. Additionally, the correlation between CSID and roughness became significant, as it did in the no/low speech severity and high intelligibility groups. These findings, in conjunction with the high reliability achieved by the raters in this study, provide compelling evidence for a confounding effect of speech impairment (particularly hypernasality) on acoustic features of dysphonia. Interestingly, the only significant correlation in the hypernasal group was CSID × Roughness, suggesting that CSID may be robust to the confounding effect of hypernasality.

Conclusions

The results of this study show that both perturbation/noise-based and cepstral/spectral acoustic measures derived from a sustained /a/ demonstrate adequate criterion validity and diagnostic accuracy for identifying dysphonia among individuals with ALS. However, likely due to the complexity of multisubsystem involvement—including the potentially confounding influence of articulatory and resonatory impairments common in ALS—the acoustic assessment of dysphonia during continuous speech is more complicated in ALS.

Limitations and Future Directions

Several limitations associated with this project should be considered in future studies. While some subjects were rated as highly dysphonic, our sample was skewed relatively low in dysphonia severity, and thus, conclusions drawn from this study should be interpreted cautiously in the context of speakers with more severe dysphonia. The Bamboo Passage used in this study had not been validated for use with the CSID, and results for this analysis should be interpreted with caution. Additionally, cepstral measures tend to be calculated only over particular segments of continuous speech, such as an all-voiced sentence from the CAPE-V, and such an approach would make findings from future projects more comparable to prior work. We limited our cepstral/spectral analyses to measures with substantial representation in prior literature; other cepstral measures such as CPP distribution across an utterance should be evaluated for use in speakers with ALS. Future research may also further investigate the significant correlations that we found among jitter/shimmer and breathiness and strain, as well as the significant correlation between HNR and strain, which are not commonly reported in other studies. Perhaps the most important future research direction suggested by the results of this study is to further investigate and quantify the unique impacts of complex neurodegenerative diseases such as ALS on the acoustic assessment of dysphonia. The relative influence of UMN and LMN damage resulting in different types of voice impairments may be investigated by looking at speakers with primarily spastic speech features and those with primarily flaccid speech features separately. Additionally, the potentially confounding influence of articulatory and resonatory impairments should be investigated in more detail.

Data Availability Statement

The data sets generated and/or analyzed during this study are not publicly available as the participants did not consent to data sharing.

Acknowledgments

References

Akif Kiliç, M., Öğüt, F., Dursun, G., Okur, E., Yildirim, I., & Midilli, R. (2004). The effects of vowels on voice perturbation measures. *Journal of Voice, 18*(3), 318–324. https://doi.org/10.1016/j.jvoice.2003.09.007

Anand, S., Kopf, L. M., Shrivastav, R., & Eddins, D. A. (2019). Objective indices of perceived vocal strain. *Journal of Voice, 33*(6), 838–845. https://doi.org/10.1016/j.jvoice.2018.06.005

Arends, N., Povel, D.-J., Os, E. V., & Speth, L. (1990). Predicting voice quality of deaf speakers on the basis of glottal characteristics. *Journal of Speech and Hearing Research, 33*(1), 116–122. https://doi.org/10.1044/jshr.3301.116

Askenfelt, A. G., & Hammarberg, B. (1986). Speech waveform perturbation analysis: A perceptual-acoustical comparison of seven measures. *Journal of Speech and Hearing Research, 29*(1), 50–64. https://doi.org/10.1044/jshr.2901.50

Audacity Team. (2021). *Audacity: Free audio editor and recorder* (Version 3.0.0) [Computer software]. https://audacityteam.org

Awan, S. N., Giovinco, A., & Owens, J. (2012). Effects of vocal intensity and vowel type on cepstral analysis of voice. *Journal of Voice, 26*(5), 670.e15–670.e20. https://doi.org/10.1016/j.jvoice.2011.12.001

Awan, S. N., & Roy, N. (2005). Acoustic prediction of voice type in women with functional dysphonia. *Journal of Voice, 19*(2), 268–282. https://doi.org/10.1016/j.jvoice.2004.03.005

Awan, S. N., & Roy, N. (2006). Toward the development of an objective index of dysphonia severity: A four-factor acoustic model. *Clinical Linguistics & Phonetics, 20*(1), 35–49. https://doi.org/10.1080/02699200400008353

Awan, S. N., Roy, N., & Cohen, S. M. (2014). Exploring the relationship between spectral and cepstral measures of voice and the Voice Handicap Index (VHI). *Journal of Voice, 28*(4), 430–439. https://doi.org/10.1016/j.jvoice.2013.12.008

Awan, S. N., Roy, N., Jetté, M. E., Meltzner, G. S., & Hillman, R. E. (2010). Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics, 24*(9), 742–758. https://doi.org/10.3109/02699206.2010.492446

Awan, S. N., Roy, N., Zhang, D., & Cohen, S. M. (2016). Validation of the Cepstral Spectral Index of Dysphonia (CSID) as a screening tool for voice disorders: Development of clinical cutoff scores. *Journal of Voice, 30*(2), 130–144. https://doi.org/10.1016/j.jvoice.2015.04.009

Baylis, A., Chapman, K., & Whitehill, T. L. (2015). Validity and reliability of visual analog scaling for assessment of hypernasality and audible nasal emission in children with repaired cleft palate. *The Cleft Palate–Craniofacial Journal, 52*(6), 660–670. https://doi.org/10.1597/14-040

Behrman, A., Cody, J., Elandary, S., Flom, P., & Chitnis, S. (2020). The effect of SPEAK OUT! and The LOUD Crowd on dysarthria due to Parkinson's disease. *American Journal of Speech-Language Pathology, 29*(3), 1448–1465. https://doi.org/10.1044/2020_AJSLP-19-00024

Benba, A., Jilbab, A., & Hammouch, A. (2017). Using human factor cepstral coefficient on multiple types of voice recordings for detecting patients with Parkinson's disease. *IRBM, 38*(6), 346–351. https://doi.org/10.1016/j.irbm.2017.10.002

Bhuta, T., Patrick, L., & Garnett, J. D. (2004). Perceptual evaluation of voice quality and its correlation with acoustic measurements. *Journal of Voice, 18*(3), 299–304. https://doi.org/10.1016/j.jvoice.2003.12.004

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (Version 6.2.20) [Computer software]. http://www.praat.org/

Brinca, L. F., Batista, A. P. F., Tavares, A. I., Gonçalves, I. C., & Moreno, M. L. (2014). Use of cepstral analyses for differentiating normal from dysphonic voices: A comparative study of connected speech versus sustained vowel in European Portuguese female speakers. *Journal of Voice, 28*(3), 282–286. https://doi.org/10.1016/j.jvoice.2013.10.001

Brockmann-Bauser, M., & Drinnan, M. J. (2011). Routine acoustic voice analysis: Time to think again? *Current Opinion in Otolaryngology & Head and Neck Surgery, 19*(3), 165–170. https://doi.org/10.1097/MOO.0b013e32834575fe

Brockmann, M., Drinnan, M. J., Storck, C., & Carding, P. N. (2011). Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *Journal of Voice, 25*(1), 44–53. https://doi.org/10.1016/j.jvoice.2009.07.002

Brooks, B. R., Miller, R. G., Swash, M., & Munsat, T. L. (2000). El Escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders, 1*(5), 293–299. https://doi.org/10.1080/146608200300079536

Byeon, H. (2021). Comparing ensemble-based machine learning classifiers developed for distinguishing hypokinetic dysarthria from presbyphonia. *Applied Sciences, 11*(5), 2235. https://doi.org/10.3390/app11052235

Byeon, H., Yu, S., & Cho, S. (2016). Characteristics of amyotrophic lateral sclerosis speakers drawn out through spectral and cepstral analysis. *Information, 19*(11(B)), 5491–5496.

Carding, P. N., Steen, I. N., Webb, A., MacKenzie, K., Deary, I. J., & Wilson, J. A. (2004). The reliability and sensitivity to change of acoustic measures of voice quality. *Clinical Otolaryngology & Allied Sciences, 29*(5), 538–544. https://doi.org/10.1111/j.1365-2273.2004.00846.x

Carrow, E., Rivera, V., Mauldin, M., & Shamblin, L. (1974). Deviant speech characteristics in motor neuron disease. *Archives of Otolaryngology—Head & Neck Surgery, 100*(3), 212–218. https://doi.org/10.1001/archotol.1974.00780040220014

Carson, C., Ryalls, J., Hardin-Hollingsworth, K., Le Normand, M.-T., & Ruddy, B. (2016). Acoustic analyses of prolonged vowels in young adults with Friedreich ataxia. *Journal of Voice, 30*(3), 272–280. https://doi.org/10.1016/j.jvoice.2015.05.008

Chen, A., & Garrett, C. G. (2005). Otolaryngologic presentations of amyotrophic lateral sclerosis. *Otolaryngology–Head and*

Neck Surgery, 132(3), 500–504. https://doi.org/10.1016/j.otohns.2004.09.092

Chiaramonte, R., & Bonfiglio, M. (2020). Acoustic analysis of voice in bulbar amyotrophic lateral sclerosis: A systematic review and meta-analysis of studies. *Logopedics Phoniatrics Vocology, 45*(4), 151–163. https://doi.org/10.1080/14015439.2019.1687748

Colton, R. H., Casper, J. K., & Leonard, R. (2011). *Understanding voice problems: A physiological perspective for diagnosis and treatment* (4th ed.). Lippincott Williams & Wilkins.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research, 12*(3), 462–496. https://doi.org/10.1044/jshr.1203.462

da Silva Antonetti, A. E., Siqueira, L. T. D., de Almeida Gobbo, M. P., Brasolotto, A. G., & Silverio, K. C. A. (2020). Relationship of cepstral peak prominence-smoothed and long-term average spectrum with auditory–perceptual analysis. *Applied Sciences, 10*(23), 8598. https://doi.org/10.3390/app10238598

de Krom, G. (1994). Consistency and reliability of voice quality ratings for different types of speech fragments. *Journal of Speech and Hearing Research, 37*(5), 985–1000. https://doi.org/10.1044/jshr.3705.985

de Krom, G. (1995). Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *Journal of Speech and Hearing Research, 38*(4), 794–811. https://doi.org/10.1044/jshr.3804.794

Eadie, T. L., & Baylor, C. R. (2006). The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *Journal of Voice, 20*(4), 527–544. https://doi.org/10.1016/j.jvoice.2005.08.007

Eshghi, M., Connaghan, K. P., Gutz, S. E., Berry, J. D., Yunusova, Y., & Green, J. R. (2021). Co-occurrence of hypernasality and voice impairment in amyotrophic lateral sclerosis: Acoustic quantification. *Journal of Speech, Language, and Hearing Research, 64*(12), 4772–4783. https://doi.org/10.1044/2021_JSLHR-21-00123

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). *irr: Various coefficients of interrater reliability and agreement* (Version 0.84.1). https://cran.r-project.org/web/packages/irr/index.html

Garrett, R. K. M. (2013). *Cepstral- and spectral-based acoustic measures of normal voices* [Master's thesis, The University of Wisconsin–Milwaukee]. The University of Wisconsin–Milwaukee ProQuest Dissertations Publishing.

Green, J. R., Beukelman, D. R., & Ball, L. J. (2004). Algorithmic estimation of pauses in extended speech samples of dysarthric and typical speech. *Journal of Medical Speech-Language Pathology, 12*(4), 149–154.

Harris, P., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. (2009). Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Haverkamp, L. J., Appel, V., & Appel, S. H. (1995). Natural history of amyotrophic lateral sclerosis in a database population validation of a scoring system and a model for survival prediction. *Brain, 118*(3), 707–719. https://doi.org/10.1093/brain/118.3.707

Heman-Ackah, Y. D., Michael, D. D., Baroody, M. M., Ostrowski, R., Hillenbrand, J., Heuer, R. J., Horman, M., & Sataloff, R. T. (2003). Cepstral peak prominence: A more reliable measure of dysphonia. *Annals of Otology, Rhinology & Laryngology, 112*(4), 324–333. https://doi.org/10.1177/000348940311200406

Heman-Ackah, Y. D., Michael, D. D., & Goding, G. S. J. (2002). The relationship between cepstral peak prominence and selected parameters of dysphonia. *Journal of Voice, 16*(1), 20–27. https://doi.org/10.1016/S0892-1997(02)00067-X

Heman-Ackah, Y. D., Sataloff, R., Laureyns, G., Lurie, D., Michael, D., Heuer, R., Rubin, A., Eller, R., Chandran, S., Abaza, M., Lyons, K., Divi, V., Lott, J., Johnson, J., & Hillenbrand, J. (2014). Quantifying the cepstral peak prominence, a measure of dysphonia. *Journal of Voice, 28*(6), 783–788. https://doi.org/10.1016/j.jvoice.2014.05.005

Hillenbrand, J. (1988). Perception of aperiodicities in synthetically generated voices. *The Journal of the Acoustical Society of America, 83*(6), 2361–2371. https://doi.org/10.1121/1.396367

Hillenbrand, J., & Houde, R. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research, 39*(2), 311–321. https://doi.org/10.1044/jshr.3902.311

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

Imatomi, S., Arai, T., & Kato, M. (2003). Effects of hoarseness on ratings of hypernasality: Source–filter theory approach. *The Japan Journal of Logopedics and Phoniatrics, 44*(4), 304–314. https://doi.org/10.5112/jjlp.44.304

Jannetts, S., & Lowit, A. (2014). Cepstral analysis of hypokinetic and ataxic voices: Correlations with perceptual and other acoustic measures. *Journal of Voice, 28*(6), 673–680. https://doi.org/10.1016/j.jvoice.2014.01.013

Kapsner-Smith, M. R., Díaz-Cádiz, M. E., Vojtech, J. M., Buckley, D. P., Mehta, D. D., Hillman, R. E., Tracy, L. F., Noordzij, J. P., Eadie, T. L., & Stepp, C. E. (2022). Clinical cutoff scores for acoustic indices of vocal hyperfunction that combine relative fundamental frequency and cepstral peak prominence. *Journal of Speech, Language, and Hearing Research, 65*(4), 1349–1369. https://doi.org/10.1044/2021_JSLHR-21-00466

Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., & Hoffman, H. T. (2007). Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. *Journal of Voice, 21*(5), 576–590. https://doi.org/10.1016/j.jvoice.2006.05.001

Kataoka, R., Warren, D., Zajac, D. J., Mayo, R., & Lutz, R. W. (2001). The relationship between spectral characteristics and perceived hypernasality in children. *The Journal of the Acoustical Society of America, 109*(5), 2181–2189. https://doi.org/10.1121/1.1360717

Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus Auditory-Perceptual Evaluation of Voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology, 18*(2), 124–132. https://doi.org/10.1044/1058-0360(2008/08-0017)

Kendall, K. (2007). Presbyphonia: A review. *Current Opinion in Otolaryngology & Head and Neck Surgery, 15*(3), 137–140. https://doi.org/10.1097/MOO.0b013e328166794f

Kent, J., Kent, R., Rosenbek, J., Weismer, G., Martin, R., Sufit, R., & Brooks, B. (1992). Quantitative description of the dysarthria in women with amyotrophic lateral sclerosis. *Journal of Speech and Hearing Research, 35*(4), 723–733. https://doi.org/10.1044/jshr.3504.723

Kent, R. D., Sufit, R. L., Rosenbek, J. C., Kent, J. F., Weismer, G., Martin, R. E., & Brooks, B. R. (1991). Speech deterioration in amyotrophic lateral sclerosis: A case study. *Journal of Speech and Hearing Research, 34*(6), 1269–1275. https://doi.org/10.1044/jshr.3406.1269

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kreiman, J., & Gerratt, B. R. (2010). Perceptual assessment of voice quality: Past, present, and future. *SIG 3 Perspectives on Voice and Voice Disorders, 20*(2), 62–67. https://doi.org/10.1044/vvd20.2.62

Kreiman, J., Gerratt, B. R., & Gabelman, B. (2002). Jitter, shimmer, and noise in pathological voice quality perception. *The Journal of the Acoustical Society of America, 112*(5), 2446–2446. https://doi.org/10.1121/1.4780067

Lee, A. S.-Y., Ciocca, V., & Whitehill, T. L. (2003). Acoustic correlates of hypernasality. *Clinical Linguistics & Phonetics, 17*(4–5), 259–264. https://doi.org/10.1080/0269920031000080091

Lopes, L. W., Barbosa Lima, I. L., Alves Almeida, L. N., Cavalcante, D. P., & de Almeida, A. A. F. (2012). Severity of voice disorders in children: Correlations between perceptual and acoustic data. *Journal of Voice, 26*(6), 819.e7–819.e12. https://doi.org/10.1016/j.jvoice.2012.05.008

Lowell, S. Y., Kelley, R. T., Awan, S. N., Colton, R. H., & Chan, N. H. (2012). Spectral- and cepstral-based acoustic features of dysphonic, strained voice quality. *Annals of Otology, Rhinology & Laryngology, 121*(8), 539–548. https://doi.org/10.1177/000348941212100808

Lu, F.-L., & Matteson, S. (2014). Speech tasks and interrater reliability in perceptual voice evaluation. *Journal of Voice, 28*(6), 725–732. https://doi.org/10.1016/j.jvoice.2014.01.018

Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology, 5*(9), 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d

Martin, D., Fitch, J., & Wolfe, V. (1995). Pathologic voice type and the acoustic prediction of severity. *Journal of Speech and Hearing Research, 38*(4), 765–771. https://doi.org/10.1044/jshr.3804.765

Murton, O., Hillman, R., & Mehta, D. (2020). Cepstral peak prominence values for clinical voice evaluation. *American Journal of Speech-Language Pathology, 29*(3), 1596–1607. https://doi.org/10.1044/2020_AJSLP-20-00001

Orozco-Arroyave, J. R., Hönig, F., Arias-Londoño, J. D., Vargas-Bonilla, J. F., & Nöth, E. (2015). Spectral and cepstral analyses for Parkinson's disease detection in Spanish vowels and words. *Expert Systems, 32*(6), 688–697. https://doi.org/10.1111/exsy.12106

Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology, 27*(3), 887–905. https://doi.org/10.1044/2018_AJSLP-17-0009

PENTAX Medical. (2011). *Analysis of Dysphonia in Speech and Voice (ADSV)* (Model 3950c, Version 4.0.0). https://www.pentaxmedical.com/pentax/en/99/1/Analysis-of-Dysphonia-in-Speech-and-Voice-ADSV

Peterson, E. A., Roy, N., Awan, S. N., Merrill, R. M., Banks, R., & Tanner, K. (2013). Toward validation of the Cepstral Spectral Index of Dysphonia (CSID) as an objective treatment outcomes measure. *Journal of Voice, 27*(4), 401–410. https://doi.org/10.1016/j.jvoice.2013.04.002

Phadke, K. V., Laukkanen, A.-M., Ilomäki, I., Kankare, E., Geneid, A., & Švec, J. G. (2020). Cepstral and perceptual investigations in female teachers with functionally healthy voice. *Journal of Voice, 34*(3), 485.e33–485.e43. https://doi.org/10.1016/j.jvoice.2018.09.010

Qi, Y., & Hillman, R. E. (1997). Temporal and spectral estimations of harmonics-to-noise ratio in human voice signals. *The Journal of the Acoustical Society of America, 102*(1), 537–543. https://doi.org/10.1121/1.419726

Rabinov, C. R., Kreiman, J., Gerratt, B. R., & Bielamowicz, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measure of jitter. *Journal of Speech and Hearing Research, 38*(1), 26–32. https://doi.org/10.1044/jshr.3801.26

Ramig, L. O., Scherer, R. C., Klasner, E. R., Titze, I. R., & Horii, Y. (1990). Acoustic analysis of voice in amyotrophic lateral sclerosis: A longitudinal case study. *Journal of Speech and Hearing Disorders, 55*(1), 2–14. https://doi.org/10.1044/jshd.5501.02

Ravits, J. M., & La Spada, A. R. (2009). ALS motor phenotype heterogeneity, focality, and spread: Deconstructing motor neuron degeneration. *Neurology, 73*(10), 805–811. https://doi.org/10.1212/WNL.0b013e3181b6bbbd

Robert, D., Pouget, J., Giovanni, A., Azulay, J.-P., & Triglia, J.-M. (1999). Quantitative voice analysis in the assessment of bulbar involvement in amyotrophic lateral sclerosis. *Acta Oto-Laryngologica, 119*(6), 724–731. https://doi.org/10.1080/00016489950180702

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12*(1), Article 77. https://doi.org/10.1186/1471-2105-12-77

Sampaio, M., Vaz Masson, M. L., de Paula Soares, M. F., Bohlender, J. E., & Brockmann-Bauser, M. (2020). Effects of fundamental frequency, vocal intensity, sample duration, and vowel context in cepstral and spectral measures of dysphonic voices. *Journal of Speech, Language, and Hearing Research, 63*(5), 1326–1339. https://doi.org/10.1044/2020_JSLHR-19-00049

Sauder, C., Bretl, M., & Eadie, T. (2017). Predicting voice disorder status from smoothed measures of cepstral peak prominence using Praat and Analysis of Dysphonia in Speech and Voice (ADSV). *Journal of Voice, 31*(5), 557–566. https://doi.org/10.1016/j.jvoice.2017.01.006

Schultz, B. G., Rojas, S., St John, M., Kefalianos, E., & Vogel, A. P. (2021). A cross-sectional study of perceptual and acoustic voice characteristics in healthy aging. *Journal of Voice.* Advance online publication. https://doi.org/10.1016/j.jvoice.2021.06.007

Silbergleit, A. K., Johnson, A. F., & Jacobson, B. H. (1997). Acoustic analysis of voice in individuals with amyotrophic lateral sclerosis and perceptually normal vocal quality. *Journal of Voice, 11*(2), 222–231. https://doi.org/10.1016/S0892-1997(97)80081-1

Sing, T., Sander, O., Beerenwinkel, N., & Lengauer, T. (2005). ROCR: Visualizing classifier performance in R. *Bioinformatics, 21*(20), 3940–3941. https://doi.org/10.1093/bioinformatics/bti623

Stipancic, K. L., Palmer, K. M., Rowe, H. P., Yunusova, Y., Berry, J. D., & Green, J. R. (2021). "You say severe, I say mild": Toward an empirical classification of dysarthria severity. *Journal of Speech, Language, and Hearing Research, 64*(12), 4718–4735. https://doi.org/10.1044/2021_JSLHR-21-00197

Strand, E. A., Buder, E. H., Yorkston, K. M., & Ramig, L. O. (1994). Differential phonatory characteristics of four women with amyotrophic lateral sclerosis. *Journal of Voice, 8*(4), 327–339. https://doi.org/10.1016/S0892-1997(05)80281-4

Swinnen, B., & Robberecht, W. (2014). The phenotypic variability of amyotrophic lateral sclerosis. *Nature Reviews Neurology, 10*(11), 661–670. https://doi.org/10.1038/nrneurol.2014.184

Tomik, B., & Guiloff, R. J. (2010). Dysarthria in amyotrophic lateral sclerosis: A review. *Amyotrophic Lateral Sclerosis, 11*(1–2), 4–15. https://doi.org/10.3109/17482960802379004

Tomik, J., Tomik, B., Wiatr, M., Składzień, J., Stręk, P., & Szczudlik, A. (2015). The evaluation of abnormal voice qualities in patients with amyotrophic lateral sclerosis. *Neurodegenerative Diseases, 15*(4), 225–232. https://doi.org/10.1159/000381956

Watts, C. R. (2015). The effect of CAPE-V sentences on cepstral/spectral acoustic measures in dysphonic speakers. *Folia Phoniatrica et Logopaedica, 67*(1), 15–20. https://doi.org/10.1159/000371656

Watts, C. R., & Awan, S. N. (2011). Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. *Journal of Speech, Language, and Hearing Research, 54*(6), 1525–1537. https://doi.org/10.1044/1092-4388(2011/10-0209)

Webb, A. L., Carding, P., Deary, I. J., MacKenzie, K., Steen, N., & Wilson, J. A. (2004). The reliability of three perceptual evaluation scales for dysphonia. *European Archives of Oto-Rhino-Laryngology and Head & Neck, 261*(8), 429–434. https://doi.org/10.1007/s00405-003-0707-7

Wolfe, V., Cornell, R., & Fitch, J. (1995). Sentence/vowel correlation in the evaluation of dysphonia. *Journal of Voice, 9*(3), 297–303. https://doi.org/10.1016/s0892-1997(05)80237-1

Wolfe, V., & Martin, D. (1997). Acoustic correlates of dysphonia: Type and severity. *Journal of Communication Disorders, 30*(5), 403–416. https://doi.org/10.1016/S0021-9924(96)00112-8

Yiu, E. M.-L. (1999). Limitations of perturbation measures in clinical acoustic voice analysis. *Asia Pacific Journal of Speech, Language and Hearing, 4*(3), 155–166. https://doi.org/10.1179/136132899807557475

Yorkston, K., Beukelman, D. R., & Hakel, D. M. (2007). *Speech Intelligibility Test (SIT) for Windows* [Computer software]. https://www.madonna.org/institute/software

Yu, M., Choi, S. H., Choi, C.-H., & Choi, B. (2018). Predicting normal and pathological voice using a cepstral-based acoustic index in sustained vowels versus connected speech. *Communication Sciences & Disorders, 23*(4), 1055–1064. https://doi.org/10.12963/csd.18550

Zraick, R. I., Wendel, K., & Smith-Olinde, L. (2005). The effect of speaking task on perceptual judgment of the severity of dysphonic voice. *Journal of Voice, 19*(4), 574–581. https://doi.org/10.1016/j.jvoice.2004.08.009