Auditory-perceptual Parameters as Predictors of Voice Acoustic Measures

Duy Duong Nguyen, and Catherine Madill, *Sydney, Australia*

Summary: Background. Much research has examined the relationship between perceptual and acoustic measures. However, little is known about the prediction values of perceptual measures on an acoustic parameter.

Aims. This study utilized simulated and disordered voice samples to investigate the prediction values of breathiness, roughness, and strain ratings on the selection of some time-based and spectral-based measures of voice quality.

Method. This study retrospectively analysed two sets of precollected data. The experimental data had been collected from nine trained speakers manipulating false vocal fold activity, true vocal fold mass, and larynx height. The voice-disordered data had been extracted from a clinical database for 68 patients with muscle tension voice disorders (MTVD). Both data sets had been perceptually rated for breathiness, roughness, and strain. Voice samples (prolonged vowel /ɑ/ and Rainbow Passage readings) had undergone acoustic analysis using Praat for harmonics-to-noise ratio (HNR) and the program "Analysis of Dysphonia in Speech and Voice" (ADSV) for cepstral peak prominence (CPP), Cepstral/Spectral Index of Dysphonia (CSID), and Low/High spectral ratio (L/H ratio). Perceptual parameters were regressed against these acoustic measures to test their prediction values.

Results. Reliability data showed satisfactory intra- and inter-reliability of perceptual ratings for both data sets. Breathiness significantly predicted CPP (both vocal tasks) and CSID (Rainbow Passage) in experimental data and predicted all the acoustic measures in MTVD data. Roughness significantly predicted HNR, CPP, and CSID in experimental data, and CPP (Rainbow Passage) and CSID (both vocal tasks) in MTVD data. Strain (both vocal tasks) significantly predicted L/H ratio in both data sets.

Conclusions. Breathiness ratings predicted selection of HNR, CPP and CSID; roughness ratings predicted selection of CPP and CSID, and strain ratings predicted L/H ratio.

Key words: Vocal fold−Perceptual analysis−Acoustic analysis−Voice quality−Laryngeal manipulation −Muscle tension voice disorder.

INTRODUCTION

The rationale for selection of acoustic voice measures

Auditory-perceptual judgment has been standardized to form an inevitable part in clinical voice assessment.[1] However, there are limitations affecting its clinical applicability[2] eg, it relies upon the variable internal standards/criteria of a rater and other factors eg, rating scale validity/reliability and poorly defined perceptual dimensions.[3] Therefore, there is a need to use a combination of perceptual assessment and an objective evaluation method[4] such as acoustic analysis.[5] There has been an ongoing search for reliable and valid acoustic voice measures.[6] Several protocol recommendations have been made regarding the choice of relevant/robust acoustic measures for clinical voice assessment. For example, cepstral-based measurement has been recommended based on the basis that it allows analysing a range of dysphonic severity.[7] The relevance of an acoustic measure has also been determined by its correlation with a perceptual dimension eg, the use of cepstral peak prominence (CPP) is supported by its correlation with breathiness ratings.[8,9] However, from a statistical point of view, correlation is not the same as prediction. Using linear regression analysis, it is possible to clarify for a given unit of change in a predictor, how much the outcome would change.[10] Such an analysis could be applied to acoustic-perceptual relation to allow better determining whether an acoustic measure would be reasonably selected using a specific auditory-perceptual quality as predictor. To date no study has determined how a perceptual dimension predicts an acoustic voice outcome.

Time-based acoustic measurements include jitter, shimmer, and the harmonics-to-noise ratio (HNR).[11] Perturbation analyses have been less preferred in voice assessment as they are only reliably applicable to prolonged vowels with nearly periodic signals.[12] Errors in pitch tracking due to aperiodicity result in poor reliability of perturbation measurements.[9] Harmonics-to-noise ratio still remains an important analysis and its algorithms have been revised and improved for more accurate estimation.[13,14] This measure has been considered an acoustic measure reflecting voice clarity[15] and is an acoustic index of hoarseness.[11] Many studies have used this to evaluate voice treatment outcomes.[16,17]

Accepted for publication February 23, 2023.

Supported by the Dr Liang Voice Program, The University of Sydney. The authors have no other funding or financial relationships to disclose.

From the Voice Research Laboratory, Sydney School of Health Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia.

Address correspondence and reprint requests to Duy Duong Nguyen, Voice Research Laboratory, Sydney School of Health Sciences, Faculty of Medicine and Health, Susan Wakil Health Building, The University of Sydney, D18 Western Avenue, Sydney, New South Wales 2050, Australia E-mail addresses: duong.nguyen@sydney.edu.au, duongtmh@gmail.com

Journal of Voice, Vol. ■■, No. ■■, pp. ■■−■■

0892-1997

https://doi.org/10.1016/j.jvoice.2023.02.030

Recently, cepstral and spectral analyses have been used more frequently. Current cepstral/spectral measures include the cepstral peak prominence (CPP), Low/high energy ratio, and cepstral/spectral index of dysphonia. Cepstral peak prominence and its smoothed variant (CPPS) are computed based on cepstral analysis by fitting a linear regression line relating quefrency to cepstral magnitude.[8] The difference in amplitude between the cepstral peak and the corresponding value on the regression line directly below the peak represents CPP.[8] Advantages of CPP/CPPS are that they can be used in both prolonged vowels and connected speech without depending upon accurate frequency-tracking.[18,19] Limitations of CPP include being affected by vocal tasks and vocal intensity[20,21] and gender.[21] Low/high spectral ratio (L/H ratio) is a ratio of spectral energy between the spectral region below 4000 Hz and the region above 4000 Hz.[19] The Cepstral/spectral index of dysphonia (CSID) is a combined measures calculated using multiple linear regression from CPP, L/H ratio, and standard deviation (SD) of L/H ratio.[22] Previous studies have shown the role of these cepstral/spectral measures in differentiating pathological from healthy voices and in estimating dysphonic severity.[22,23] It has been demonstrated that cepstral/spectral measures can predict voice disorder status,[24] quantify dysphonic severity,[19] classify/differentiate dysphonic from vocally healthy voices,[22,25] and screen voice disorders.[22] These capabilities have been used as rationales to select cepstral/spectral measures in voice analysis.

The major auditory-perceptual qualities and their relationship with HNR and cepstral-based measures

There are three major auditory-perceptual qualities in clinical voice assessment including breathiness, roughness, and strain. Breathy voice contains the audible sound of breathing during phonation resulting from air turbulence at the glottis.[12,26] Roughness is defined as having uneven, bumpy quality that is unsteady in the short-term but stationary in the long-term, resulting from aperiodic waveform.[12,26] Strained voice is the perception of increased vocal effort (hyperfunction) during phonation.[1] Previous studies have shown correlations of these parameters with a range of acoustic outcomes.

Breathiness

A very limited number of studies have examined the correlation between breathiness and HNR although this auditory-perceptual parameter is the perceptual attribute of glottal turbulent noise and HNR quantifies harmonics relative to noise levels in the voice signals.[27] de Krom[28] found that HNR was the best single predictor for both rated breathiness and roughness. A previous study has investigated the correlation between the Grade, Roughness, Breathiness, Aesthenia, Strain of the GRBAS scale and the multi-dimensional voice program (MDVP) scheme.[29] They found significant correlation between dysphonic severity ("Grade") and noise harmonic ratio (NHR) and significant correlation between roughness and NHR.[29] It is also associated with voice clarity.[15] The HNR has been shown to correlate with the level of hoarseness.[30] HNR has also been shown to be the most useful time-based measure for predicting severity of voice quality.[31] Vaz Freitas et al[32] observed that HNR significantly predicted perceptual ratings of the following perceptual qualities in GRBAS scale: grade (dysphonic severity level), breathiness, and anaesthesia. No study has investigated whether breathiness significantly predicted HNR.

Breathiness has also been found to correlate with CPP/CPPS. Hillenbrand and Houde[9] found that breathiness ratings of sustained vowels were best predicted by CPP and its smoothed measure (CPPS). For Rainbow Passage reading, they found that both CPP and CPPS significantly predicted breathiness ratings. However, it has not been examined whether breathiness ratings predict cepstral/spectral measures

Roughness

Heman-Ackah et al[33] found that CPPS of connected speech and sustained vowel /ɑ/ correlated with perception of roughness although they only explained for a limited amount of the variance (25% for CPPS of connected speech and 18% for CPPS of /ɑ/). Cannito et al[34] found that CPPS of connected speech moderately correlated with the perception of roughness of connected speech. Lopes et al[35] found that CPPS correlated moderately and negatively with vocal roughness of the sustained /ɛ/ vowel. Antonetti et al[36] found that CPPS showed moderate to strong negative correlations with breathiness, roughness, and strain. No study has clarified whether roughness predicted HNR and any cepstral measures.

Strain

Lopes et al[35] found that CPPS correlated negatively and weakly with strain ratings of sustained /ɛ/ vowel. Lowell et al[37] demonstrated significant correlations between CAPE-V strain severity ratings and CPP and CSID of the Rainbow Passage. They also found significant correlations between L/H ratio (cut-off at 4000 Hz) and L/H ratio SD with strain ratings of the Rainbow Passage. According to their findings, higher levels of strain correlated with lower CPP, lower L/H ratio, and lower L/H ratio SD. Anand et al[38] found that perceived strain was strongly correlated with CPP. McKenna and Stepp[39] found that L/H ratio and HNR were significant predictors of vocal effort ratings.

The above-mentioned studies have examined acoustic-perceptual relationship from the lens of using acoustic measures as correlates or validation tools for perceptual parameters. To our knowledge, no study has investigated the prediction values of perceptual measures on selecting an acoustic measure. In addition, the relationship between acoustic and perceptual measures has been mostly studied in voice-disordered speakers. There is limited research on acoustic-perceptual relationship and the prediction of perceptual measures on acoustic outcome using well-controlled voice quality simulation data. Rosenthal et al[40] examined how different levels of vocal effort produced by vocally healthy people affected physiologic and acoustic parameters of the voice. Two previous studies have shown that laryngeal manipulation by trained vocally healthy speakers resulted in simulated voice qualities that reflect laryngeal abnormal muscle tension.[41,42] Using simulated perceptual voice quality measures to predict acoustic measures would give more insight into how a given acoustic measure could be selected based on perceptual judgment.

The present study tested the predictive values of breathiness, roughness, and strain ratings on HNR and spectral-based acoustic voice measures using two sets of voice data. The first was voice data in an instructed laryngeal manipulation experiment, which have been published.[43,44] The second data set was extracted from a clinical database of patients seeking voice therapy treatment. These three perceptual dimensions were selected as they were the major voice qualities heard in common voice disorders in clinics and have been selected as the most relevant features for auditory-perceptual measurement.[1] It was hypothesized that: 1) Breathiness ratings would significantly predict HNR, CPP, CSID, and L/H ratio; 2) Roughness would significantly predict HNR and CPP; and 3) Strain ratings would significantly predict L/H ratio.

METHODS

Ethical approval

This was a retrospective study using previously collected data from two projects approved by The University of Sydney Human Research Ethics Committee (Protocol numbers: 2019/281 and 2019/529). The study was implemented in accordance with relevant ethical regulations.

Participants

Experimental group

The data set for the experimental part was extracted from nine participants (five females and four males) aged 19-36 years whose characteristics were described in previous two publications.[43,44] These participants had been trained to manipulate three laryngeal parameters, including false vocal folds (constriction and release of constriction), true vocal fold mass (thick and thin), and larynx height (low and normal positions) using the Voicecraft®[41] vocal training system. Details of voice training program specific to this study have been described in previous studies.[43,44] A habitual voice production condition was also included as baseline.

Clinical group

Voice data were randomly extracted from a database for 68 patients (60 females and 8 males) with a mean age of 35.6 years (standard deviation, SD = 13.6, range = 20-84). All had a confirmed diagnosis of muscle tension voice disorders (MTVD) who sought voice therapy treatment at a speech pathology clinic. All had comprehensive voice assessment including speech pathology and laryngoscopy or strobolaryngoscopy. In this group, 34 had primary MTVD (without visible vocal fold mucosal lesions) and 34 had MTVD with vocal fold benign mucosal lesions such as pre-nodular mucosal swellings, nodules, or mucosal thickening. Mean voice handicap index (VHI-10) score for the whole MTVD group was 17.8 (SD = 9.4). Detailed characteristics of this group have been described elsewhere.[45]

Voice data

Experimental voice data

Details of voice recording for the experimental group have also been described in detail in previous publications.[43,44] In summary, voice samples were recorded using an AKG C420 series II condenser cardioid ear-mounted microphone positioned 5 cm and 45° off the mouth axis. The microphone was connected to a Tascam DA20MkII digital audiotape recorder via a Behringer Eurotrack MX802A−ULN 8 Channel 2-Bus Mixing Amplifier at 48 kHz/16-bit. The signals were then transferred to a personal computer in the same sampling rate and resolution and saved in wav file format.

Clinical data

Voice recordings were extracted from a clinical voice database where all patients underwent standardized voice assessment tasks, including reading the Rainbow Passage,[46] the consensus auditory perceptual evaluation -voice (CAPE-V) phrases,[1] and prolonged vowel (/ɑ/). Voice samples were recorded using an AKG C520 cardioid head-mounted microphone[47] placed at 6 cm, 45° off the mouth axis via a professional external sound card (Roland Quadcapture[48]) at 44.1kHz/16-bit. The signals were saved to a laptop computer using the Audacity sound editing software[49] in *.wav format.

Acoustic analyses

Voice recordings were edited in Audacity 3.1.3[49] to extract middle three seconds of the prolonged vowel /ɑ/ and the 2nd and 3rd sentences of the Rainbow Passage.

Harmonics-to-noise ratio (HNR) in decibels (dB)

This time-based measure was extracted using Praat 6.2.14[50] using the command "Voice Report". As this measure depends upon reliable fundamental frequency tracking, signals classified as types 3 and 4 were excluded from HNR measurement. The protocols for signal typing followed Praat settings and signal classification documented in the literature.[51] Signal typing was performed by the first author with consensus from the last author.

Cepstral peak prominence (CPP) in dB

The acoustic analysis program *Analysis of Dysphonia in Speech and Voice* (ADSV) Model 5109 Version 3.4.2[52] was used to measure CPP for both vowel (CPPv) and Rainbow Passage (CPPrp) samples. Details of settings for CPP settings in ADSV have been mentioned in previous studies.[44]

Cepstral/spectral index of dysphonia (CSID)

This was also measured using ADSV.[52] CSID data were obtained automatically in ADSV for vowel (CSIDv) and manually calculated for the Rainbow Passage task (CSIDrp) based on CPP, L/H ratio, and L/H ratio standard deviation (SDL/H) values in ADSV using the following formula[22]:

$$CSID\ of\ Rainbow\ Passage = 154.59 - 10.39 \times CPP - 1.08 \times L/H - 3.71 \times SDL/H$$

L/H ratio in dB

This was measured automatically in the ADSV program. This is a ratio between low-frequency and high-frequency spectral energy using a cut-off value of 4000 Hz and measured in decibel (dB).[19]

Auditory-perceptual data

The present study was interested in the prediction values of auditory perceptual ratings of three parameters (roughness, breathiness, and strain) on acoustic measures. Auditory-perceptual rating scores of these parameters for the experiment dataset were taken from the database of a previously published study[43] using median rating scores calculated from two experienced speech language pathologists (SLPs) who demonstrated high intrarater reliability. Ratings had been made using a Visual Analogue Scale (VAS) of 100-mm line with 1 and 100 representing minimum and maximum degree of the perceptual dimension being rated. Intrarater reliability using intraclass correlation coefficients (ICC) for these two raters were 0.852 and 0.903.[43] Inter-rater ICC values of the two raters for the perceptual dimensions were as follows: breathiness: 0.622; roughness: 0.345; and strain: 0.597.[43]

Rating scores of the same perceptual measures for the clinical dataset were extracted from baseline data of a database that had been analysed in a previous publication[45] where three raters (two SLPs and one ENT surgeon) judged the degree of breathiness, roughness, and strain using an online VAS rating tool of 100-point with 1 for “minimum” and 100 for “maximum”. Mean rating score across the three raters was used for analysis. Intrarater ICC values (single measures) for the three raters were as follows[45]: breathiness: 0.738, 0.948, 0.810; roughness: 0.869, 0.889, 0.812; and strain: 0.862, 0.896, 0.829. Inter-rater ICCs (single measures) were as follows: breathiness: 0.659, roughness: 0.696, and strain: 0.691.[45]

Statistical analyses

Data were analysed using IBM SPSS 25.0[53] and GraphPad Prism 8.1.2.[54] Descriptive statistics were calculated on acoustic measures. Prior to analyses normal distribution of the data was examined using Kolmogorov-Smirnov tests.[55] As perceptual and acoustic variables were measured using different type of unit of measurement, all data were normalized in the same class/category so that the max and the min values were 1 and 0, respectively. Linear regression was used to examine the prediction values of each perceptual parameter on acoustic measures. The assumption of multicollinearity was checked and if a predictor violated this assumption, it was excluded from the multiple regression model. Significance level of 0.05 was used.

RESULTS

Experimental group characteristics

Figure 1 shows perceptual rating scores for all laryngeal conditions in the experimental group. The median of rating

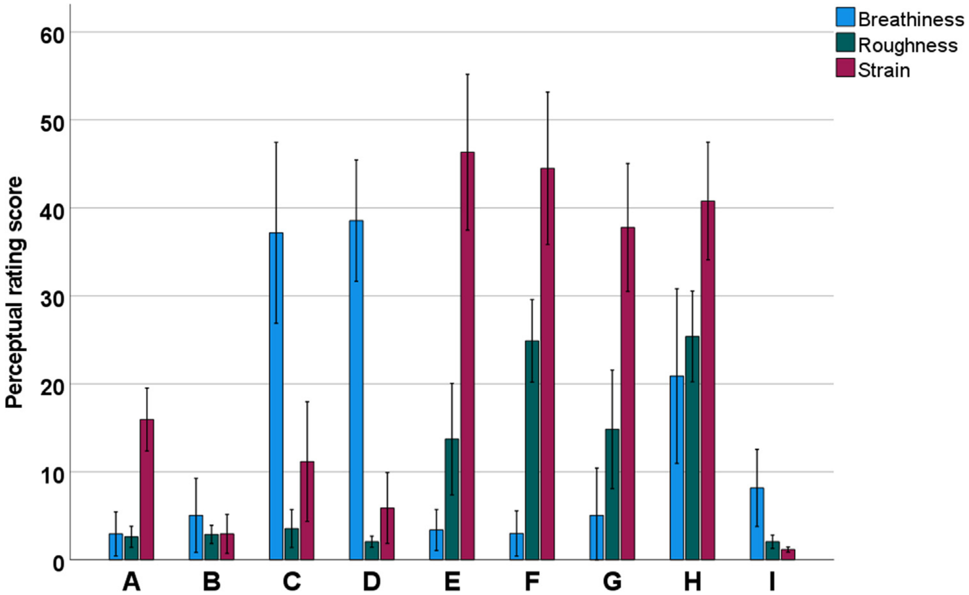

FIGURE 1. Perceptual rating scores for all experimental conditions. In order of conditions of the false vocal folds, true vocal folds, and larynx height: A. released, thick, low; B. released, thick, normal; C. released, thin, low; D. released, thin, normal; E. constricted, thick, low; F. constricted, thick, normal; G. constricted, thin, low; H. constricted, thin, normal; I. habitual voice. Error bars represent standard errors.

TABLE 1.
Acoustic Data of Experimental Group

Measures	Mean (SD)	95% CI for Mean	Min - Max
HNR (dB)	21.2 (4.9)	19.9–22.5	10.5–29.6
CPPv (dB)	10.5 (4.5)	9.5–11.5	1.1–18.7
CPPrp (dB)	5.6 (1.7)	5.2–5.9	1.0–8.6
CSIDv	33.2 (38.8)	24.6–41.8	−15.1 to 111.2
CSIDrp	21.9 (21.4)	17.1–26.6	−15.0 to 70.6
L/Hv (dB)	30.2 (6.1)	28.8–31.5	12.7–42.4
L/Hrp (dB)	25.2 (3.4)	24.5–26.0	17.5–34.5

Abbreviations: v, vowel; rp, Rainbow Passage.

score of the two raters was calculated and was used for all data analyses. Figure 1 shows that rating scores for strain and roughness were more prominent in constricted false vocal fold conditions and breathiness were relatively high in two released false vocal fold conditions with thin true vocal fold mass. These data represented mild to moderate levels for these perceptual dimensions.

Table 1 shows acoustic data for the experimental group calculated from all conditions combined (n = 9 × 9 conditions = 81). Figure 2 show cepstral and spectral measures for each condition in the experimental group. For statistical comparison across laryngeal parameters (false vocal fold, true vocal fold mass, and larynx heigh), refer to a previous study.[44]

MTVD group characteristics

Figure 3 shows perceptual rating scores for the MTVD groups. For both primary MTVD and MTVD with mucosal lesions, rating scores for all perceptual parameters were within mild to moderate range.

Table 2 shows the acoustic data for MTVD group in comparison with normative cut-off from the literature. The HNR and CSID data showed a wide range extending from pathological range to normative range, and the means of these measures were within the normal values for both primary and secondary MTVD groups. Only the mean of CPP of both vocal tasks was within the pathological range. In combination with auditory-perceptual data, these showed that the MTVD groups demonstrated mild to moderate levels of vocal dysfunction.

Prediction value of the perceptual parameters on acoustic measures

Linear regression was calculated to estimate the value of auditory-perceptual parameters in predicting each of the acoustic measure. All acoustic and perceptual measures were normalized so that the minimum and maximum were 0 and 1, respectively. This normalization eliminated variabilities due to the involvement of different types of unit of measurements in each category (ie, the perceptual rating scale and unit of CPP and CSID). The normalization was expected to result in more robust regression models where standardized regression coefficients (β) were used to determine which perceptual parameters were a better predictor of the acoustic outcome. The predictors were input into the regression equation using a stepwise method. Eligibility of each predictor was tested using collinearity diagnostics in SPSS via the variance inflation factor (VIF) and correlation coefficients (r).[10] A predictor was excluded from the multiple regression model if VIF was ≥ 2.5 and $r \geq 0.8$. As strain rating was associated with a high VIF and highly correlated with both breathiness and roughness, it was excluded from the multiple linear regression. Normalized breathiness and roughness scores were used in the same multiple regression equation whilst normalized strain was used in a separate simple linear model.

Hypothesis 1: Breathiness ratings would significantly predict HNR, CPP, CSID, and L/H ratio

Table 3 shows standardized regression coefficient β and P values for breathiness in the multiple regression equation involving normalized rating scores of breathiness and roughness as independent variables. In the experimental data set, breathiness did not significantly predict HNR. In MTVD data, this perceptual parameter significantly predicted HNR.

Table 3 also shows that breathiness was a significant predictor of CSID of Rainbow Passage in the experimental data and CSID of both vocal tasks in MTVD data. Breathiness did not significantly predict CSID of vowel in experimental dataset.

The prediction value of breathiness on L/H ratio was also hypothesized. Table 3 shows inconsistency in prediction results between experimental and MTVD data: breathiness did not significantly predict L/H of both tasks in experimental data, but significantly predicted L/H of both tasks in MTVD data.

Hypothesis 2: Roughness would significantly predict HNR and CPP

Standardized regression coefficients for roughness are presented in Table 3. This perceptual dimension predicted HNR in the experimental data (the higher roughness level, the lower HNR) but did not significantly predict HNR in clinical data.

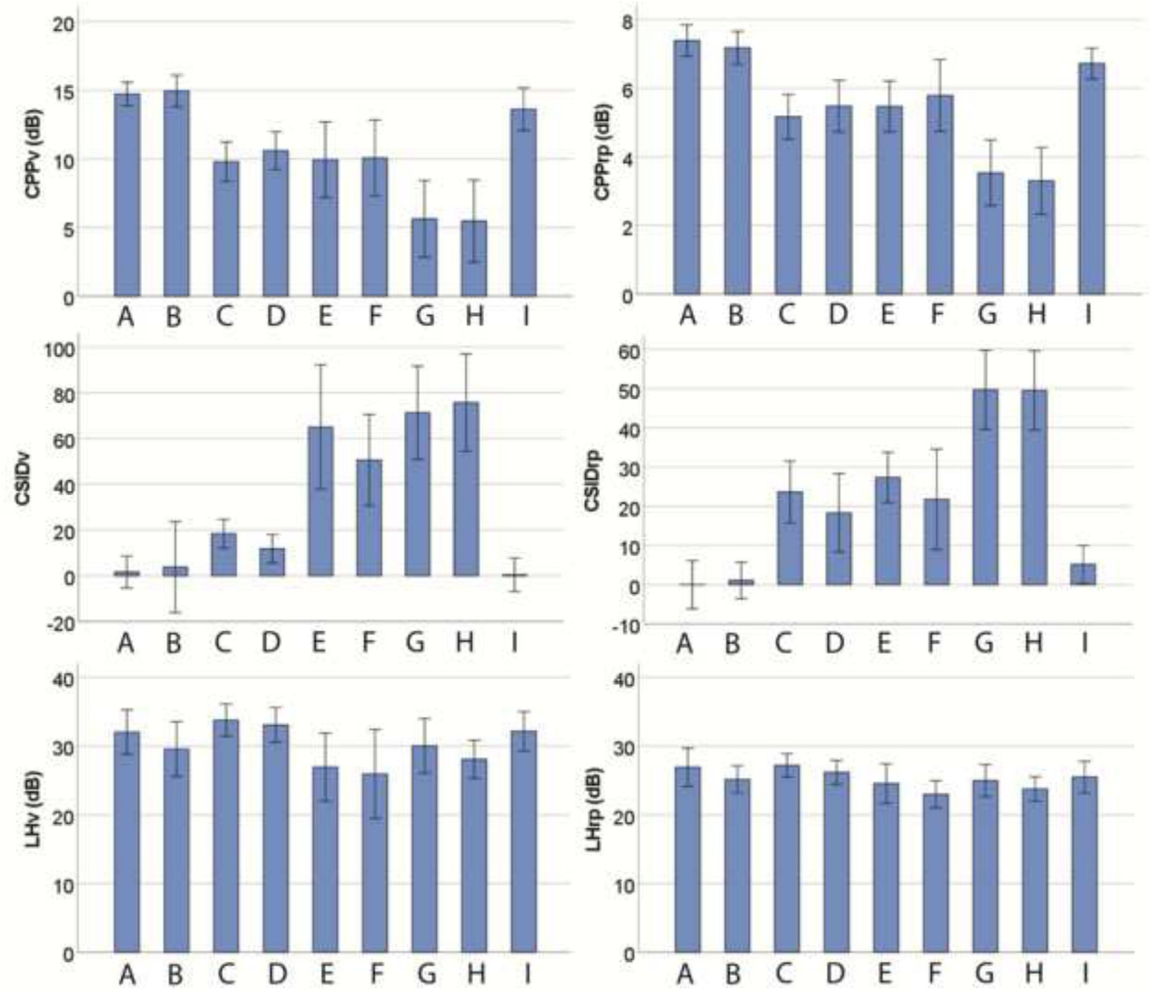

FIGURE 2. Cepstral/spectral measures of experimental data. In order of conditions of the false vocal folds, true vocal folds, and larynx height: A. released, thick, low; B. released, thick, normal; C. released, thin, low; D. released, thin, normal; E. constricted, thick, low; F. constricted, thick, normal; G. constricted, thin, low; H. constricted, thin, normal; I. habitual voice. Error bars represent standard errors.

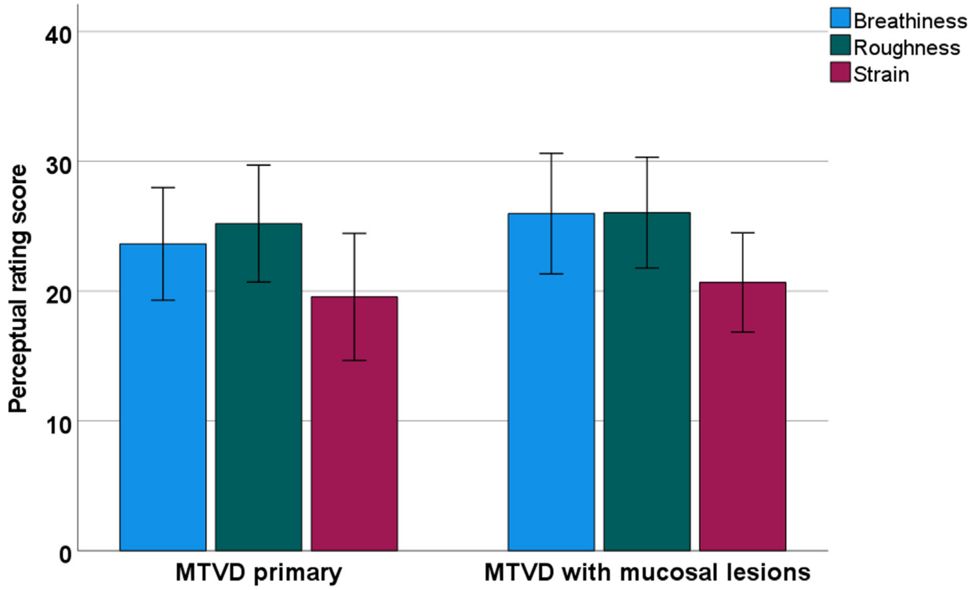

FIGURE 3. Perceptual rating scores for two muscle tension voice disorder (MTVD) sub-groups. Error bars represent standard errors.

TABLE 2.
Acoustic Data of MTVD Groups (n = 34 for Primary MTVD; n = 34 for MTVD With Mucosal Lesions)

Measures	MTVD Groups	Mean (SD)	95% CI for Mean	Min-Max	Normative Cut-offs
HNR (dB)	Primary	23.3 (4.2)	21.9–24.8	13.2–30.6	> 20dB[50]
	With lesions	23.9 (4.5)	22.3–25.6	12.6–32.5	
CPPv (dB)	Primary	10.7 (2.6)	9.8–11.6	4.2–15.8	> 11.46 dB[23]
	With lesions	10.7 (2.6)	9.8–11.6	5.5–17.0	
CPPrp (dB)	Primary	5.3 (1.4)	4.8–5.8	1.8–7.3	> 6.11dB[23] – 6.6dB[56]
	With lesions	5.4 (1.0)	5.0–5.7	2.8–7.6	
CSIDv	Primary	6.9 (19.2)	0.2–13.6	−17.9 to 64.4	< 11.8[57]
	With lesions	9.5 (18.2)	3.1–15.8	−22.6 to 60.6	
CSIDrp	Primary	14.3 (17.5)	8.2–20.4	−9.3 to 59.7	< 24.27[22]
	With lesions	13.9 (12.6)	9.5–18.3	−11.9 to 39.8	
L/Hv (dB)	Primary	40.4 (6.0)	38.3–42.5	29.6–51.2	> 25.26dB[25]
	With lesions	38.1 (4.3)	36.6–39.6	28.5–50.8	
L/Hrp (dB)	Primary	30.6 (3.5)	29.4–31.8	25.1–37.2	> 28.08dB[25]
	With lesions	30.6 (3.1)	29.5–31.7	23.7–36.6	

Abbreviations: v, vowel; rp, Rainbow Passage.

CPP of both vowel and Rainbow Passage was statistically significantly predicted by roughness rating in the experimental data: CPP would drop if roughness increased. In MTVD data, only CPP of Rainbow Passage was significantly predicted by roughness.

We did not hypothesize CSID to be predicted by roughness given the lack of literature evidence supporting the use of this acoustic measure in rough voices. However, results of regression analysis showed that this acoustic measure was statistically significantly predicted by roughness in both data sets (Table 3).

Hypothesis 3: Strain ratings would significantly predict L/H ratio

For both experimental data and MTVD data, the prediction of the acoustic measures using strain ratings was calculated using simple linear regression given the multicollinearity when adding strain to the multiple regression model.

Results from experimental data showed that strain significantly predicted L/H ratio of vowel and Rainbow Passage (Table 4). The β values of strain in predicting L/H ratio of vowel (β = -0.349, t = -3.311, P = 0.001) and L/H ratio of Rainbow Passage (β = -0.260, t = -

TABLE 3.
Prediction Values of Breathiness and Roughness Ratings on Acoustic Measures Based on Normalized Data

Conditions	Dependent Variables	Vocal Tasks	Breathiness		Roughness	
			β	P	β	P
Laryngeal manipulation	HNR	Vowel	0.016	0.890	−.538	<0.001*
	CPP	Vowel	−0.281	0.004*	−.525	<0.001*
		RP	−0.350	<0.001*	−.518	<0.001*
	CSID	Vowel	0.003	0.976	.649	<0.001*
		RP	0.262	0.006*	.565	<0.001*
	L/H ratio	Vowel	0.148	0.172	N/A	N/A
		RP	0.068	0.540	N/A	N/A
Muscle tension voice disorders	HNR	Vowel	−0.327	0.028*	−0.180	0.220
	CPP	Vowel	−0.615	<0.001*	−0.050	0.688
		RP	−0.287	0.012*	−0.509	<0.001*
	CSID	Vowel	0.441	<0.001*	0.341	0.004*
		RP	0.398	<0.001*	0.414	<0.001*
	L/H ratio	Vowel	−0.732	<0.001*	N/A	N/A
		RP	−0.530	<0.001*	N/A	N/A

Notes: Breathiness rating significantly predicted CPP of both vowel and Rainbow Passage in both experimental data and MTVD data. Based on the β and P value, in experimental data, the prediction value of breathiness was similar for CPP of both vocal tasks; meanwhile, in the MTVD data, breathiness contributed more to predicting CPP of vowel (β = -0.615) than CPP of Rainbow Passage (β = -0.287). The sign of β indicated that as breathiness rating increased, CPP decreased.

* Statistically SignificantN/A, Calculation was not Performed.

β, Standardized Regression Coefficient.

TABLE 4.

Analysis of Variance Results Testing Model Fit for Normalized Strain Ratings in Predicting L/H Ratio of Vowel (L/Hv) and Rainbow Passage Reading (L/Hrp)

Groups	Dependent Variables	df	F	*P*
Experimental data	L/Hv	1, 79	10.963	0.001
	L/Hrp	1, 79	5.748	0.019
MTVD data	L/Hv	1, 66	14.892	<0.001
	L/Hrp	1, 66	7.992	0.006

In MTVD group, strain also significantly predicted L/H of vowel and Rainbow Passage (Table 4). The β values of both equations were significant in predicting L/H of vowel (β = -0.429, t = -3.859, P <0.001) and Rainbow Passage (β = -0.329, t = -2.827, P = 0.006) in which higher strain ratings indicated lower L/H ratio values. Abbreviations: df, degrees of freedom.

2.398, P = 0.019) indicated that as strained rating increased, L/H ratio decreased.

DISCUSSION

Although HNR and cepstral and spectral measures (CPP, CSID, and L/H ratio) have proven useful in clinical application and voice research, their selection based on auditory-perceptual consideration has not been well justified. In the present study, we showed the prediction values of three commonly used perceptual dimensions on these acoustic parameters. This provides further evidence on which acoustic measures should be selected to objectively evaluate a particular auditory-perceptual dimension. We examined these relationships using both experimental and clinical data. The laryngeal manipulation data set was collected in well-controlled conditions where trained speakers intentionally differentially changed the three laryngeal parameters to simulate laryngeal configurations seen in voice disorders with abnormal muscle tension patterns. The MTVD data represented real-world changes in laryngeal function associated with the commonly observed laryngeal muscle tension patterns, which were modelled in the experimental data, eg, supraglottic constriction.[58] Overall, the findings accepted our hypotheses despite some inconsistencies between the two datasets.

Prediction values of perceptual measures on cepstral/spectral parameters

Breathiness

We found that breathiness rating significantly predicted HNR in the MTVD data but not in the experimental data. This acoustic feature reflects the amount of noise relative to harmonic components in the vocal signals and has been used to quantify pathological voices eg, hoarseness.[11] Breathy voice results from incomplete glottal closure, leading to unmodulated glottal flow presented perceptually as noise present in the perceived signals.[59] Therefore, in general, it is understandable that perceived breathiness is stipulated by noise in the voice signals, which is reflected via the signal-to-noise ratio.[60] Previous studies have found significant correlations between breathiness and HNR.[28] However, the perception of noise levels in the voice is also determined by the complicated interaction between the spectral shape and the relative levels of harmonics and noise components in the signals.[61] It is likely that the noise characteristics (noise spectrum) in the experimental data may not be the same as those in the clinical data due to differences in how each condition affected the phonation mechanism. These differences may explain for the discrepancy in the prediction of breathiness on HNR between the two datasets. In addition, in the MTVD data, HNR was averaged from a relatively homogeneous population with reasonable sample sizes. Meanwhile, the laryngeal manipulation represented changes in HNR in three combined laryngeal parameters manipulated by a small number of speakers (n = 9) and HNR was averaged across all laryngeal conditions (9), some of which resulted in great variability in the data eg, constricted false vocal fold conditions. In the present study, breathiness was rated with significantly higher scores in two release of constriction conditions, including released false vocal folds/thin true vocal folds/low larynx height, released false vocal folds/thin true vocal folds/normal larynx height, and constricted false vocal folds/thin true vocal folds/normal larynx height, all of which had in common the “thin” true vocal folds (Figure 1). These were the most likely reasons leading to the discrepancy in prediction value of breathiness on HNR between the two study groups.

This study showed that breathiness was a significant predictor of CPP of both vocal tasks in both experimental and MTVD data, with CPP dropping if breathiness ratings increased. The findings agreed with previous research that had found significant relationship between CPP of both vowel and connected speech (Rainbow Passage) and breathiness ratings.[8,9] Eadie and Baylor[62] also found that both CPP and CPPS significantly predicted breathiness rating of vowel, and CPP was significant predictor of breathiness rating of connected speech. Other authors[34] also found significant correlation between CPPS and breathiness ratings. A previous study[44] showed that thin TVFM caused CPP to drop significantly, adding objective evidence to the voice characteristics of this vocal fold condition.

As the calculation of CSID included CPP,[22] this explains the similar trend in the regression models between CPP and CSID with one exception; CSID of the vowel was not significantly predicted by breathiness in the laryngeal manipulation data (Table 3). These findings supported the use of CPP (both vocal tasks) and CSID (Rainbow Passage) as an objective measure for breathy voice quality.

We also found that breathiness only significantly predicted L/H ratio in the MTVD group and not in the experimental group. Previous research has shown that the spectral slope reflects and is affected by phonation types and voice quality, eg, breathiness,[59] possibly resulting from the addition of high-frequency energy level in breathy voices,[63] amongst other mechanisms.[59] As such, it is more reasonable for breathiness to be able to predict L/H ratio than not. The finding that breathiness did not significantly predict L/H ratio in the laryngeal manipulation data may be caused by the variabilities associated with different laryngeal conditions in a small sample size. Future studies should replicate laryngeal manipulation in a larger study group to clarify this.

Roughness

Our finding in the experimental data showed that roughness significantly predicted HNR. This agreed with a study by de Krom[28] who found that HNR was the best single acoustic correlate of roughness and Eskenazi et al[31] who showed that HNR was significantly correlated with roughness ratings. However, a discrepancy was found between the experimental and MTVD datasets in the prediction value of roughness on HNR. Whilst roughness rating significantly predicted HNR of laryngeal manipulation data, it did not predict HNR in MTVD data (Table 3). Data from Table 1 and Table 2 show that mean HNR was lower in the experimental group than in MTVD group. It is possible that this difference in HNR between the two groups may contribute to this prediction discrepancy. Differences in auditory-perceptual rating protocols between the two datasets may also be a factor explaining for this discrepancy. Additionally, it is necessary to note that the mechanism for roughness may be different across the two datasets. In the experimental data, all participants were vocally healthy. Therefore, the irregularities in vocal fold vibration were intentionally generated, which may be different from the patterns of irregularities in the MTVD groups where vocal roughness may have been caused by several factors at the same time. These factors may include consistent abnormal muscle tension patterns, vocal fold mucosal lesions, and possibly abnormal glottal closure.

Vocal roughness was a significant predictor for CPP and CSID (both vocal tasks) in the experimental group and CPP (Rainbow Passage) and CSID (both tasks) in MTVD group (Table 3), implying that these two acoustic measures can be used to characterize rough voices. These findings are explainable given that roughness is thought to result from irregular vocal fold vibration[60,64] and cepstral analysis is a measure that reflects vibration periodicity.[8] This finding confirmed the significant relationship between cepstral/spectral measures and vocal roughness and agreed with study by Eadie and Baylor[62] who found that CPP accounted for 58.1% of the variance in prediction of vocal roughness of connected speech. CPPS and fundamental frequency (F0) combined predicted 51.6% of vocal roughness of connected speech.[62] Eadie and Baylor[62] did not find CPP and CPPS to be statistically significant predictors of vocal roughness of vowels.

Strain

Vocal strain is believed to relate to increased laryngeal muscle tension[65] eg, in muscle tension voice disorders.[58,66] The perceptual data showed that strain ratings were higher than those of breathiness and roughness in laryngeal manipulated conditions with constricted false vocal folds (Figure 1). However, strain was rated as lower than the other voice dimensions in MTVD data (Figure 3). Although these data gave experimental and clinical implications on the possible involvement of increased muscle tension patterns, perceived strain levels may not be equivalent between the two data sets. The effects of perceived breathiness and roughness on strain ratings was likely to be different between simulated voices and real disordered voices. It is possible that the higher levels of both breathiness and roughness in the pathological voices in the MTVD group than in the experimental group may result in the discrepancy in strain ratings between the two groups. This is in line with the literature which has shown that strain ratings are influenced more for the vocal signals with added noise than for the signals without added noise.[67]

Previous studies have tried to find objective ways to quantify strain including cepstral and spectral measurements.[38] However, acoustic correlates of perceptions of strained voice have been notoriously difficult to establish.[38,67] Given the lack of supporting data in the literature on the relationship between vocal strain and a range of acoustic measures, we only calculated the prediction of strain on L/H ratio. The findings observed in both experimental and clinical data supported our hypothesis that strain ratings significantly predicted the L/H ratio. Our findings agreed with Lowell et al[37] who found significant correlations between L/H ratio and strain ratings of the Rainbow Passage. A previous study[37] has also showed significant correlations between strain ratings and CSID of the Rainbow Passage, a measure that includes L/H ratio in the calculation formula.[22]

To date, it seems that the most consistent acoustic correlate of strain is a decreased spectral tilt due to added energy at high frequency regions[37,59] and/or the lower fundamental component relative to the high frequency.[59] Decreased relative fundamental frequency (RFF) has also been observed in people with vocal hyperfunction.[68] The present study added more evidence supporting the use of spectral slope in examining vocal strain. It is necessary to note that, although the increased higher spectral components can lead to the lower spectral slope in strained voices, the high-frequency regions contain vocal tract information in addition to the voice source variables. For example, the high-frequency region above 4-5 kHz can reflect speech segmental information, speaker identity, and gender identity.[69] There may be great between-speaker variability in spectral slope[70] that can preclude reliable application of this acoustic index in strain assessment. The coexistence of other voice qualities eg, breathiness and roughness in voice-disordered populations may add extra acoustic features affecting the source spectrum, which also have effects on spectral slope or other acoustic measures intended to quantify vocal strain. For example, Shoji et al[63] showed significant increase in high-frequency energy levels in breathy voices. As such, if breathiness coexists with strain, which is common in pathological voices, the spectral slope (L/H ratio) may reflect not only strain but also breathiness.

Limitations

The present study had several limitations that could be improved in future studies. First, we only examined a limited number of perceptual and acoustic measures (HNR and cepstral/spectral measures). We did not examine perturbation measures (jitter and shimmer) due to their reliability problems in severely aperiodic signals. Future studies should investigate the prediction values of a wider range of perceptual measures on acoustic outcomes. Second, in laryngeal manipulation, the effects of each laryngeal parameter were not isolated. Therefore, it would be difficult to attribute the particular action of a laryngeal parameter to the perceptual voice features. The findings should be interpreted as “instructed laryngeal manipulation” until a more objective method can be used to validate the implementation of laryngeal manipulation. Third, the MTVD group demonstrated mild to moderate degrees of dysphonia. It is not clear whether the findings would be replicable with higher dysphonic levels. Last, the clinical group only consisted of a functional voice disorder. The findings may not be generalizable to the wider spectrum of voice disorders until more representative study groups are investigated.

CONCLUSION

The following conclusions can be made from this study:

5.1. Clinicians can select relevant acoustic measures for objective evaluation of breathiness, roughness, and strain using the prediction models presented in this study.

5.2. The simulation of pathological voice qualities using laryngeal manipulation is useful to investigate acoustic-perceptual interactions without confounding effects related to laryngeal dysfunction in pathological conditions. The combination of both simulated and clinical voice data provided a more comprehensive picture of the acoustic-perceptual relationship.

5.3. Breathiness significantly predicted CPP (both vocal tasks) and CSID (connected speech) in experimental data; breathiness predicted HNR and all cepstral/spectral measures in MTVD data.

5.4. Roughness significantly predicted HNR, CPP, and CSID in experimental data, and CPP (connected speech) and CSID (both vocal tasks) in MTVD data.

5.5. Strain ratings of vowel and connected speech significantly predicted L/H ratio in both laryngeal manipulation data and MTVD data.

5.6. The inconsistencies in prediction values of the perceptual parameters between the experimental and clinical data implied complicated involvement of different laryngeal factors in voice production, which would need larger experimental studies to clarify.

DECLARATION OF COMPETING INTERESTS

None.

REFERENCES

1. Kempster GB, Gerratt BR, Verdolini Abbott K, et al. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol*. 2009;18:124–132.
2. Awan SN, Solomon NP, Helou LB, et al. Spectral-cepstral estimation of dysphonia severity: external validation. *Ann Otol Rhinol Laryngol*. 2013;122:40–48.
3. 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.
4. Brinca LF, Batista AP, 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.
5. 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.
6. Peterson EA, Roy N, Awan SN, et al. Toward validation of the cepstral spectral index of dysphonia (CSID) as an objective treatment outcomes measure. *J Voice*. 2013;27:401–410.
7. 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.
8. Hillenbrand J, Cleveland RA, Erickson RL. Acoustic correlates of breathy vocal quality. *J Speech Hear Res*. 1994;37:769–778.
9. Hillenbrand J, Houde RA. Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *J Speech Hear Res*. 1996;39:311–321.
10. Field A. *Discovering Statistics Using IBM SPSS Statistics*. 4th ed. London: SAGE Publications Ltd.; 2013.
11. Yumoto E, Gould WJ, Baer T. Harmonics-to-noise ratio as an index of the degree of hoarseness. *J Acoust Soc Am*. 1982;71:1544–1549.
12. Titze IR. *Workshop on Acoustic Voice Analysis: Summary Statement*. Iowa City, IA: National Center for Voice and Speech; 1995.
13. Awan SN, Frenkel ML. Improvements in estimating the harmonics-to-noise ratio of the voice. *J Voice*. 1994;8:255–262.
14. de Krom G. A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. *J Speech Hear Res*. 1993;36:254–266.
15. Warhurst S, Madill C, McCabe P, et al. The vocal clarity of female speech-language pathology students: an exploratory study. *J Voice*. 2012;26:63–68.
16. Tuomi L, Bjorkner E, Finizia C. Voice outcome in patients treated for laryngeal cancer: efficacy of voice rehabilitation. *J Voice*. 2014;28:62–68.
17. Wu CH, Chan RW. Effects of a 6-week straw phonation in water exercise program on the aging voice. *J Speech Lang Hear Res*. 2020;63:1018–1032.
18. Fraile R, Godino-Llorente JI. Cepstral peak prominence: a comprehensive analysis. *Biomed Signal Process Control*. 2014;14:42–54.
19. Awan SN, Roy N, Jette ME, et al. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: comparisons with auditory-perceptual judgements from the CAPE-V. *Clin Linguist Phon*. 2010;24:742–758.

20. Phadke KV, Laukkanen AM, Ilomaki I, et al. Cepstral and perceptual investigations in female teachers with functionally healthy voice. *J Voice*. 2020;34:485.e33–485.e43.
21. Awan SN, Giovinco A, Owens J. Effects of vocal intensity and vowel type on cepstral analysis of voice. *J Voice*. 2012;26:670.e15–670.e20.
22. Awan SN, Roy N, Zhang D, et al. Validation of the cepstral spectral index of dysphonia (CSID) as a screening tool for voice disorders: development of clinical cutoff scores. *J Voice*. 2016;30:130–144.
23. Murton O, Hillman R, Mehta D. Cepstral peak prominence values for clinical voice evaluation. *Am J Speech Lang Pathol*. 2020;29:1596–1607.
24. Sauder C, Bretl M, Eadie T. Predicting voice disorder status from smoothed measures of cepstral peak prominence using praat and Analysis of Dysphonia in Speech and Voice (ADSV). *J Voice*. 2017;31:557–566.
25. Watts CR, Awan SN. Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. *J Speech Lang Hear Res*. 2011;54:1525–1537.
26. Titze IR. *Principles of Voice Production*. Englewood Cliffs, NJ: Prentice-Hall, Inc.; 1994.
27. Baken RJ, Orlikoff RF. *Clinical Measurement of Speech and Voice*. 2nd ed. San Diego: Singular Thomson Learning; 2000.
28. de Krom G. Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *J Speech Hear Res*. 1995;38:794–811.
29. Bhuta T, Patrick L, Garnett JD. Perceptual evaluation of voice quality and its correlation with acoustic measurements. *J Voice*. 2004;18:299–304.
30. Yumoto E, Sasaki Y, Okamura H. Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. *J Speech Hear Res*. 1984;27:2–6.
31. Eskenazi L, Childers DG, Hicks DM. Acoustic correlates of vocal quality. *J Speech Hear Res*. 1990;33:298–306.
32. Vaz Freitas S, Melo Pestana P, Almeida V, et al. Integrating voice evaluation: correlation between acoustic and audio-perceptual measures. *J Voice*. 2015;29:390.e1–390.e7.
33. Heman-Ackah YD, Michael DD, Goding Jr. GS. The relationship between cepstral peak prominence and selected parameters of dysphonia. *J Voice*. 2002;16:20–27.
34. Cannito MP, Doiuchi M, Murry T, et al. Perceptual structure of adductor spasmodic dysphonia and its acoustic correlates. *J Voice*. 2012;26:818.e5–818.e13.
35. Lopes LW, Sousa E, Silva A, et al. Cepstral measures in the assessment of severity of voice disorders. *Codas*. 2019;31: e20180175.
36. Antonetti AE, Siqueira LT, Gobbo MP, et al. Relationship of cepstral peak prominence-smoothed and long-term average spectrum with auditory−perceptual analysis. *Appl Sci*. 2020;10:8598.
37. Lowell SY, Kelley RT, Awan SN, et al. Spectral- and cepstral-based acoustic features of dysphonic, strained voice quality. *Ann Otol Rhinol Laryngol*. 2012;121:539–548.
38. Anand S, Kopf LM, Shrivastav R, et al. Objective indices of perceived vocal strain. *J Voice*. 2019;33:838–845.
39. McKenna VS, Stepp CE. The relationship between acoustical and perceptual measures of vocal effort. *J Acoust Soc Am*. 2018;144:1643–1658.
40. Rosenthal AL, Lowell SY, Colton RH. Aerodynamic and acoustic features of vocal effort. *J Voice*. 2014;28:144–153.
41. Bagnall AD. *Voicecraft Workshop Manual*. Adelaide, SA: Voicecraft International; 1997.
42. Estill J. *Primer of Compulsory Figures: Level One*. Santa Rosa, CA: Estill Voice Training Systems; 1996.
43. Madill CJ, Sheard C, Heard R. Are instructions to manipulate specific parameters of laryngeal function associated with auditory-perceptual ratings of voice quality in nondisordered speakers? *J Voice*. 2017;31: 504.e21–504.e33.
44. Madill C, Nguyen DD. Impact of instructed laryngeal manipulation on acoustic measures of voice-preliminary results. *J Voice*. 2023;37: 143.e1–143.e11.
45. Madill C, Chacon A, Kirby E, et al. Active ingredients of voice therapy for muscle tension voice disorders: a retrospective data audit. *J Clin Med*. 2021;10:4135.
46. Fairbanks G. *Voice and Articulation Drillbook*. 2nd ed. New York: Harper & Row; 1960.
47. AKG Acoustics. C520 Available from: https://www.akg.com/Microphones/Headset%20Microphones/C520.html. Access date: 5/6/2018.
48. Roland Corp. Quad-capture - USB 2.0 Audio Interface 2019 Available from: https://www.roland.com/au/products/quad-capture/. Access date: 2/8/2021.
49. Audacity Team. Audacity(R): Free Audio Editor and Recorder [Computer application] Available from: https://www.audacityteam.org/. Access date: 16/5/2022.
50. Boersma P, Weenink D. Praat: doing phonetics by computer Available from: http://www.fon.hum.uva.nl/praat/. Access date: 26/1/2018.
51. Sprecher A, Olszewski A, Jiang JJ, et al. Updating signal typing in voice: addition of type 4 signals. *J Acoust Soc Am*. 2010;127:3710–3716.
52. PentaxMedical. Analysis of dysphonia in speech and voice— ADSV [Computer application]Available from: https://www.pentaxmedical.com/pentax/en/99/1/Analysis-of-Dysphonia-in-Speech-and-Voice-ADSV. Access date: 30/3/2018.
53. IBM Corp. IBM SPSS Software 2018 Available from: https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software. Access date: 8/2/2018.
54. GraphPad Software. Prism 8 2018 Available from: https://www.graphpad.com/scientific-software/prism/. Access date: 20/4/2018.
55. Massey FJ. The Kolmogorov-Smirnov test for goodness of fit. *J Am Stat Assoc*. 1951;46:68–78.
56. Garrett R. *Cepstral- and spectral-based acoustic measures of normal voices [doctoral thesis]*. Milwaukee, USA: University of Wisconsin; 2013.
57. Shim H-J, Jung H, Koul R, et al. Spectral and cepstral based acoustic features of voices with muscle tension dysphonia. *Clin Arch Commun Disord*. 2016;1:42–47.
58. Morrison MD, Rammage LA. Muscle misuse voice disorders: description and classification. *Acta Otolaryngol*. 1993;113:428–434.
59. Klatt D, Klatt L. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am*. 1990;87: 820–857.
60. Hillenbrand J. Perception of aperiodicities in synthetically generated voices. *J Acoust Soc Am*. 1988;83:2361–2371.
61. Kreiman J, Gerratt BR. Perceptual interaction of the harmonic source and noise in voice. *J Acoust Soc Am*. 2012;131:492–500.
62. Eadie TL, Baylor CR. The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *J Voice*. 2006;20:527–544.
63. Shoji K, Regenbogen E, Yu JD, et al. High-frequency power ratio of breathy voice. *Laryngoscope*. 1992;102:267–271.
64. Emanuel FW, Whitehead RL. Harmonic levels and vowel roughness. *J Speech Hear Res*. 1979;22:829–840.
65. Oates J, Russell A. Learning voice analysis using an interactive multimedia package: development and preliminary evaluation. *J Voice*. 1998;12:500–512.
66. Morrison MD, Nichol H, Rammage LA. Diagnostic criteria in functional dysphonia. *Laryngoscope*. 1986;96:1–8.
67. Park Y, Cadiz MD, Nagle KF, et al. Perceptual and acoustic assessment of strain using synthetically modified voice samples. *J Speech Lang Hear Res*. 2020;63:3897–3908.
68. Stepp CE, Hillman RE, Heaton JT. The impact of vocal hyperfunction on relative fundamental frequency during voicing offset and onset. *J Speech Lang Hear Res*. 2010;53:1220–1226.
69. Boyd-Pratt HA, Donai JJ. The perception and use of high-frequency speech energy: clinical and research implications. *Perspectives of the ASHA Special Interest Groups*. 2020;5:1–9.
70. Kitzing P. LTAS criteria pertinent to the measurement of voice quality. *J Phon*. 1986;14:477–482.