Research Article

Acoustic Voice Quality Index for Discriminating Across Normal and Different Vocal Pathological Conditions

JAIISH(2019)
Vol 38 pp. 16-25

Srushti Shabnam[a] & Pushpavathi Mariswamy[b]

Affiliations

[a]All India Institute of Speech and Hearing, Department of Speech Language Pathology, University of Mysore, Manasagnagothri, Mysuru, Karnataka, India
[b]All India Institute of Speech and Hearing, Manasagnagothri, Mysuru, Karnataka, India

Corresponding Author

Srushti Shabnam
All India Institute of Speech and Hearing, University of Mysore
Mysuru
Karnataka
India.
simpleshabnam@gmail.com

Key Words

Acoustic voice quality index
vocal pathology
dysphonia

Abstract

The Acoustic Voice Quality Index (AVQI) is a multiparametric measure to assess the overall voice quality using both sustained and continuous speech.The study aimed to compare the AVQI values across normophonic voice and dysphonic voice secondary to different pathological conditions and to compare the values obtained from constituent parameters of AVQI across normophonic and different vocal pathological conditions. Seventy-four participants in the dysphonic group and twenty-eight in the normophonic group were considered. The auditory-perceptual analysis was done by three Speech-Language Pathologists using the Grade, Roughness, Breathiness, Asthenia, and Strain (GRBAS) scale to classify the participants into normophonic and dysphonic groups. Phonation samples of /a/ and reading samples were recorded using the Praat program, and AVQI was calculated using PraatAVQI script v.2.03. The results revealed that among the dysphonic group, higher AVQI values were seen in unilateral vocal fold (VF) palsy, followed by bilateral mass lesion, unilateral mass lesion, MTD-I, MTD-II & III, and acute laryngitis. CPPS and HNR values were lowest in the unilateral VF palsy group indicating high breathiness and noise component respectively. Shimmer local and shimmer local dB values were high in the unilateral VF palsy group and mass lesion group suggesting maximum aperiodic vibration of vocal folds in these groups. To conclude,AVQI and constituent parameters might help in discriminating vocal pathological conditions acoustically. As the present study is a preliminary attempt, future studies can be carried out with larger sample size, restricted age range, and due consideration of the auditory-perceptual dysphonia severity.

INTRODUCTION

An individual is diagnosed with voice disorder when his/her voice quality, loudness, and pitch vary from people of similar age, gender, geographical location, and cultural background (Aronson, 1980; Boone, 1977). The Diagnostic Classification System of Voice Disorders (DCSVC)grossly divides the voice disorders into two groups, i.e., Organic voice disorder (OVD) and Functional voice disorder (FVD). Further, FVD consists of two groups, i.e., psychogenic voice disorders (PVD) and muscle tension voice disorders (MTVD) (Baker, Ben-Tovim, Butcher, Esterman, & McLaughlin, 2007). The prevalence of voice disorders among communication disorders is around 4-7% in the Indian context (Sinha, Shivaswamy, Barman, Seth, Seshadri & Savithri, 2017; Konadath, Chatni, Lakshmi, & Saini, 2017).The prevalence rate of voice disorders among professional voice users is also relatively high; 86% of the politicians, 74% of vendors, 59% of singers, and 49% of teachers exhibited voice problems (Boominathan, Rajendran, Nagarajan, Seethapathy, & Gnanasekar, 2008).These studies suggest that voice disorder is a prevalent condition in the Indian context and warrants attention to precise diagnosis and effective intervention.

The auditory-perceptual and acoustic analyses of voice are vital components of the voice evaluation carried out by Speech-Language Pathologists (SLPs), as they provide excellent measures of intervention outcome (Stemple, Roy, & Klaben, 2018). Auditory-perceptual analysis of voice is a process of listening to and describing the abnormalities of a voice, specifically the deviations in terms of pitch, loudness, and quality. Acoustic analysis of voice provides quantitative data on vocal fold (VF) vibration in terms of pitch and amplitude, perturbation measures, harmonics to noise ratio, spectral, and cepstral measures, which in turn provide a better understanding of the pattern of VF vibrations (Maryn, Roy, De Bodt, Van Cauwenberge, & Corthals, 2009).

Maryn, Corthals, Van Cauwenberge, Roy, and De Bodt (2010) developed AVQI, a multi-parametric acoustic model to assess voice quality, which used both sustained vowel and continuous speech to improve the ecological validity, auditory-perceptual, and instrumental assessment of dysphonia. For this purpose, sustained and continuous speech (reading phonetically balanced text) samples were collected from 251 participants (229 with dysphonia and 22 without dysphonia) and were combined. The samples were given to five experienced speech language pathologists for the auditory-perceptual rating of overall voice quality. The non-voiced segments within the continuous speech were removed using a custom voicing detection algorithm. Concatenated samples were analyzed using 13 acoustic parameters based on spectral and cepstral analyses, amplitude perturbation and fundamental frequency perturbation. The AVQI equation consists of six acoustic parameters (smoothened cepstral peak prominence, shimmer local, harmonics-to-noise ratio, shimmer local dB, general slope of the spectrum, the tilt of the regression line through the spectrum). The definitions of these parameters are provided in Appendix I.

AVQI= [3.295 – (0.111*CPPS) - (0.073*HNR) – (0.213*shimmer local) + (2.789*shimmer local dB) – (0.032*slope) + (0.077*tilt)]*2.571.

The diagnostic efficacy of combining both continuous speech and sustained vowel samples in the acoustic and auditory-perceptual assessment of dysphonic voice has been discussed by several researchers (Heman-Ackah, Michael, & Goding, 2002; Maryn et al., 2010). Studies have also reported that AVQI has diagnostic accuracy,concurrent validity, and reflects changes following intervention (Heman-Ackah, et al., 2002; Maryn, et al., 2010; Maryn, De Bodt, & Roy, 2010). AVQI has been validated across different languages (Dutch, Lithuanian,Japanese, Korean, German, Spanish, and Kannada) and found to be reliable (Maryn, et al., 2010; Uloza, et al., 2017; Hosokawa, et al., 2017; Kim, Barsties, & Lee, 2019; Barsties, Lehnert, & Janotte, 2020; Delgado, et al., 2018; Benoy, 2017; Pebbili, et al., 2019). AVQI is also found to be useful in discriminating normophonic and dysphonic voices.

Englert, et al. (2020) evaluated the precision of AVQI and its isolated acoustic measures (CPPS, HNR, shimmer local, shimmer local dB, slope, and tilt) in discriminating voices with different degrees of deviation (normal, mild, moderate, and severe). The results suggested that AVQI was a reliable tool in differentiating across the degrees of dysphonia. It was also observed to have more accuracy in differentiating between moderate and severe dysphonia. Also, the isolated acoustic measures showed good precision at a higher degree of dysphonia.

There have been no studies that investigated whether AVQI and its isolated acoustic measures can be a useful tool to discriminate across the different vocal pathological conditions. Laryngeal/ vocal imaging is currently considered a standard tool for understanding VF physiology, its pathologies, and the differential diagnosis. Visual examination using endoscopy or stroboscopy are commonly used instruments for vocal imaging. Vocal pathologies can be treated through voice therapy, and the prognosis can be assessed through acoustic evaluation and vocal imaging. However, frequent evaluations are required to monitor the progress of voice therapy. In this scenario, the feasibility of endoscopy/stroboscopy reduces because of high operational time and cost factors. Further, younger children with voice problems often do not cooperate for endoscopic procedures. Meanwhile, acoustic analysis has a significant advantage over vocal imaging, such as ease of administration, cost-effectiveness, time saving, and can be administered multiple times. The isolated acoustic measures of AVQI provide information about different vocal aspects like phonatory gap, irregularity in VF vibration, and hyper- and hypo-adduction of VFs. CPPS values provide evidence regarding the presence of phonatory gap and breathiness component in voice; HNR reveals information about the phonatory gap as well the noise component in voice; shimmer local and shimmer local dB provides information about aperiodicity in VF vibration, and Spectral slope and spectral tilt gives evidence regarding the presence of hyper- and hypo-adduction of VFs (Lieberman,1963; Hillenbrand & Houde, 1996; Hartl, Hans, Vaissiere, Riquet, & Brasnu, 2001; Ludlow, Kent, &Gray, 2018). Hence, it is worth considering all the individual measures of AVQI and the overall AVQI value, as it provides a holistic understanding of an individual's voice. Also, compared to other software programs for voice analysis such as VAGHMI (Speech and Voice Sytems, Bangalore, India), Dr.Speech (Tiger Electronics, Seattle, WA, USA), and Multi Dimentional Voice Program (MDVP; Computerised Speech Lab, Kay Elemetrics Corporation, Lincoln Park, NJ, USA). AVQI has advantages such as (i) it provides overall severity of dysphonia, (ii) it analyses both sustained vowel and continuous speech sample in a single analysis, which adds to the ecological validity of the tool, (iii) it gives values for all the isolated acoustic measures which comprise of noise, perturbation, cepstral and spectral measures, and (iv) developers have made it freely available to the public.

Considering the advantages of AVQI, and to address the concerns related to vocal imaging, a preliminary attempt was made to investigate if AVQI and its isolated acoustic measures can help discriminate normophonic voice and different vocal pathological conditions. It is hypothesized that AVQI and its isolated acoustic measures might differentiate various vocal pathological conditions. The present study aimed to investigate the utility of AVQI and its isolated acoustic measures in discriminating across normophonic voice and different vocal pathological conditions. The specific objectives were to compare the AVQI values across normophonic voice and dysphonic voice due to different pathological conditions and to compare values obtained from isolated acoustic measures of AVQI across normophonic and different vocal pathological conditions.

METHODS

Participants

There were 74 participants in the dysphonic group (51 males and 23 females, age range=11 to 82 years, mean age= 39.4±15.5 years). The dysphonic group had 13 individuals with bilateral mass lesion; 16 with unilateral mass lesion; 10 with unilateral palsy; 9 with muscle tension dysphonia type II and III (MTD-II & III); 18 with muscle tension dysphonia type I (MTD-I); and 8 with acute laryngitis. The bilateral and unilateral mass lesion group majorly consisted of individuals with vocal nodules and vocal polyps. There were 28 participants in the normophonic group with 13 males and 15 females (Age range= 19 to 39 years; Mean age= 24.7±4.2 years).

All the participants considered in the study were native Kannada speakers and had normal hearing and cognitive abilities. The individuals under the dysphonic group had to undergo the routine clinical examination, which involved detailed case history, auditory-perceptual evaluation, acoustic evaluation, aerodynamic evaluation, and vocal imaging (stroboscopy). The underlying vocal pathology was diagnosed by the team involving a SLP and an Otolaryngologist using the videostroboscopy Xion Endostrob E with a 70-degree rigid scope and the Xenon R-180 LED light source for illumination. Individuals with organic and functional voice disorders were included in the dysphonic group. However, individuals with neurological problems with a total laryngectomy, and resonance disorder were excluded from the dysphonic group. The vocal usage of the dysphonic group ranged from Level I (Elite Vocal Performer) to Level IV (Non-Vocal Professional) and the degree of dysphonia ranged from slight to severe.

For the normophonic group, participants were selected through convenience sampling from among the staff and students of the Institute. The individuals had to undergo an informal screening and an auditory-perceptual examination of voice by an experienced SLP. Individuals with perceptually normal voice (G=0 on GRBAS scale; Hirano, 1981) and no complaints of voice problems or upper respiratory tract infections, asthma, or allergic disease on the day of recording were included in the normophonic group. Individuals with neurological problems, hormonal disturbances, resonance disorder, and history of laryngeal surgeries or related laryngeal pathologies were excluded from the normophonic group. A written consent was obtained from each participant, where the information regarding the aim, objectives, research method, and approximate duration of the procedure were mentioned.

Procedure for voice recording

For the acoustic analysis, the voice recording was done in a sound-treated room where the average ambient noise level was 25 dB. The participants were made to sit comfortably and a table-mounted dynamic microphone (Mipro MM-107; Supercardioid vocal microphone, Mipro Elcetronics, Co. Ltd., Chiayi, Taiwan) was placed at a distance of 4-5 cm and at 30^{o} angle from the participant's mouth. All the recordings were done at a sampling frequency of 44.1 kHz, 16-bit resolution, in the mono channel, using the program Praat v. 6.0.40, and were saved in .wav format. To estimate AVQI, both phonation and continuous speech samples were required. Hence, the participants were instructed to phonate vowel /a/ for more than three seconds, and three trials of phonation were taken at their comfortable pitch and loudness. The most stable recording was considered for further analysis. An interval of 2 minutes was given between each recording, to avoid vocal fatigue on subsequent trials. Next, they were asked to read the first paragraph of the standardized Voiced Kannada passage (Shasidhar, 1984) at their comfortable pitch and loudness. The second sentence of the Voiced Kannada passage (/i: u:rannu namma ra:dzjada bomba:i ennuvaru/) was considered for the continuous speech sample. The obtained samples were opened in the Praat program and were truncated, renamed, and were saved as .wav format accordingly.

Acoustic analysis of voice samples

For the calculation of AVQI, both phonation and contionuous speech samples were opened in the Praat program, the Praat script of AVQI version 02.03 (Maryn, 2013) was run, and then AVQI value and values of constituent parameters were obtained on output window (Figure 1).

Auditory-perceptual analysis of voice samples

The voice samples of all the participants were subjected to auditory-perceptual analysis using the GRBAS scale, in order to group the voice samples into normophonic and dysphonic categories. In the GRBAS scale, the overall grade (severity of dysphonia) was rated on a 0-3 scale (0, 1, 2, and 3 representing normal, slight, moderate, and severe, respectively). To categorize the voice samples into normophonic and dysphonic groups, the auditory-perceptual analysis was carried out by three experienced SLPs (raters). The raters had a minimum of 5 years of clinical experience in dealing with the diagnosis and management of voice disorder. All the voice samples (from individuals with normophonic and dysphonic voice quality) were randomized and double-blinded for auditory-perceptual analysis. The recorded samples were provided to each judge individually in a quiet room with an ambient noise level lower than 40 dB. The samples were presented using Sennheiser HD 180 headphones (Sennheiser

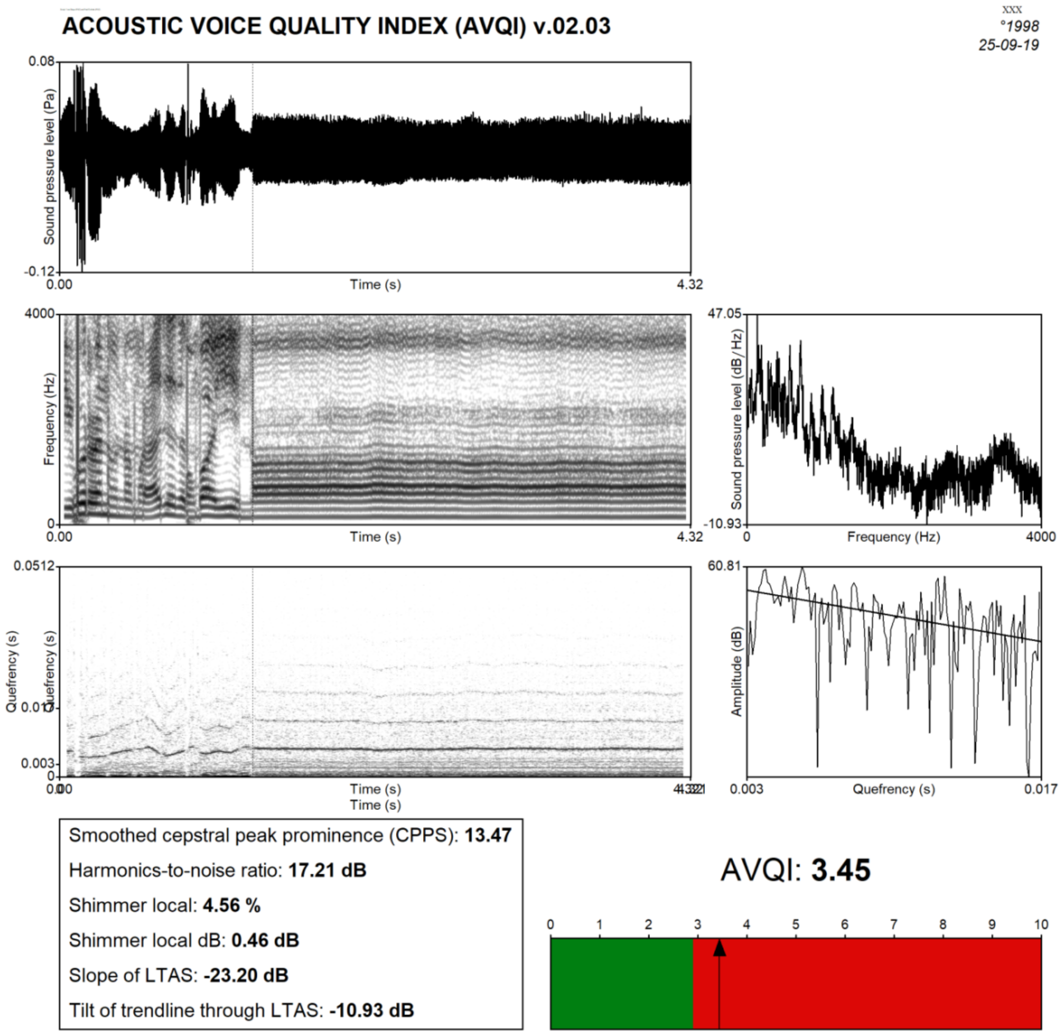

Figure 1: An image of the AVQI 02.03 output

electronic GmbH & Co. KG, Germany) at a comfortable listening level and were instructed to write their responses in the response sheet. The judges were asked to listen to the samples carefully before making a final decision and were allowed to take a break in between to avoid fatigue.

In the present study, voice samples which were rated as '0' were categorized as normophonic and those rated as '1', '2', or '3' were categorized as dysphonic. Based on the consensus across at least two of the three raters, a particular grade was assigned to each sample. The voice samples were categorized based on the overall Grade of dysphonia (G): 28 samples were considered normophonic, 40 slight, 25 moderate, and 9 were rated as severe (Normophonic group= 28 and Dysphonic group= 74). The consistency among the raters, were confirmed through the assessment of inter-rater agreement for each pair using Cohen's Kappa.

Statistical Analysis

The software program IBM Statistical Package for Social Sciences (SPSS) version 20 was used for statistical analysis. The Shapiro -Wilk test was carried out to test the assumption of normality for AVQI and its constituent parameters across the normophonic and vocal pathological groups. Descriptive statistics were done to obtain the mean and standard deviation (SD) values for AVQI and its constituent parameters across all the groups. A one-way analysis of variance (ANOVA) with Tukey post hoc analysis was carried out to determine the group mean difference for AVQI. MANOVA and Tukey post hoc test was carried out to investigate the main effect of type of pathology on the acoustic measures. Mann Whitney U test was used to observe the effect of type of pathology on CPPS, HNR, shimmer local and shimmer local dB.

RESULTS

Inter-rater agreement

Cohen's Kappa coefficient was used to assess the inter-rater agreement for the overall grade (G) of dysphonia severity. The coefficient for Rater 1 vs. Rater 2 was 0.66; Rater 1 vs. Rater 3 was 0.64, and for Rater 2 vs. Rater 3 was 0.76, signifying good agreement among the raters.

Test of Normality

The results of the Shapiro-Wilks' test revealed that the overall AVQI and slope followed normal distribution for the overall dysphonia group ($p > 0.05$). All the measures except shimmer local and shimmer local dB followed a normal distribution ($p > 0.05$) for the normophonic group and individual vocal pathological groups. Shimmer local and shimmer local dB values followed a normal distribution ($p > 0.05$) in all vocal pathological groups except for unilateral mass lesion group (p< 0.05).

Comparison of AVQI values across normophonic and dysphonic

The results of descriptive statistics showed that the dysphonic group had higher AVQI values (3.80±1.58) compared to the normophonic group (1.94±0.83). The results of MANOVA showed an overall significant main effect of dysphonia on AVQI and slope (Wilks' Lambda = 0.73, F (2, 99) = 17.49, $p < 0.001$). The results of subsequent ANOVA indicated the difference to be statistically significant for AVQI (F (1, 100) =35.00, $p < 0.001$) between the two groups. No significant difference was obtained for slope (F (1, 100) = 2.04, $p > 0.001$). The results of the Mann Whitney U test indicated that CPPS (|z|= 5.18, $p < 0.001$), HNR (|z|= 3.92, $p < 0.001$), shimmer local (|z|= 4.61, $p < 0.001$), and shimmer local dB (|z|= 4.92, $p < 0.001$) values differed significantly between the dysphonic and normophonic groups. However, no significant difference was seen for tilt (|z|= 1.75, $p > 0.001$) between the two groups.

Comparison of AVQI values across different pathological groups

Among the dysphonic group, higher values were obtained for unilateral VF palsy, followed by bilateral mass lesion, unilateral mass lesion, MTD-I, MTD-II & III, and laryngitis (Table 1). A higher AVQI value indicates poorer overall voice quality. The groups with unilateral mass lesion, MTD-I, MTD-II & III, and laryngitis were found to have similar AVQI values. The results of ANOVA showed that there was a significant effect of pathologies on AVQI values (F(6, 95) = 11.77, $p < 0.001$). Tukey post hoc test results uggested that the normophonic group had significantly lower AVQI values compared to unilateral VF palsy, bilateral mass lesion, and unilateral mass lesion groups ($p < 0.05$). Moreover, unilateral VF palsy had significantly higher AVQI values than other pathological groups ($p < 0.05$).

Comparison of CPPS, HNR, slope, and tilt values across different groups

The group with unilateral VF palsy were found to have the least values for CPPS and the value increased in the bilateral mass lesion, unilateral mass lesion, MTD-I, acute laryngitis, and MTD-II & III. Similarly, HNR values were found to be least in the group with unilateral VF palsy, followed by bilateral mass lesion, unilateral mass lesion, MTD-I, MTD-II & III, and acute laryngitis. The CPPS and HNR values were found to be highest in the normophonic group (Table 2). The values obtained for slope and tilt did not vary much across the groups. The MANOVA showed an overall significant main effect of type of pathology on the acoustic measures [Wilks' Lambda = 0.38, F (24, 322.16) = 4.20, $p < 0.001$]. The subsequent ANOVA result for each parameter has been summarized in Table 3. The Partial Eta Squared values ranged between 0.3 to 0.4 for CPPS and HNR, suggesting a medium effect of pathology on CPPS and HNR. For tilt, Partial Eta Squared value obtained was 0.196, suggesting a small effect of pathology on tilt.

The Tukey post hoc test results indicated that CPPS values were significantly high for the normophonic group compared to unilateral VF palsy, unilateral mass lesion, and bilateral mass lesion ($p < 0.05$). The group with unilateral VF palsy was found to have significantly lesser CPPS values compared to other pathological groups ($p < 0.05$). The CPPS values obtained for MTD-I, MTD-II& III, acute laryngitis, and normophonic groups did not differ significantly. HNR values were significantly lesser in the unilateral VF palsy group compared to normophonic and other pathological groups ($p < 0.05$). The values obtained for slope were found to have no significant difference across the groups. tilt values varied significantly across the groups ($p < 0.05$), but no definite pattern could be discerned.

Comparison of shimmer local and shimmer local dB values across different groups

The shimmer local and shimmer local dB values were the highest for the unilateral VF palsy group, followed by bilateral mass lesion, unilateral mass lesion, MTD-I, MTD-II & III, acute laryngitis, and it was the lowest for the normophonic group (Table 4, in page no. 52). The MANOVA showed a significant main effect of type of pathology on shimmer local and shimmer local dB [Wilks' Lambda = 0.67, F (8, 104) = 2.84, P = 0.007]. The subsequent ANOVA

Table 1: Mean (SD), Minimum, and Maximum AVQI values across different pathological conditions and normophonic group

Groups	Mean	Minimum	Maximum
Unilateral VF palsy	5.62 (1.56)	2.81	8.26
Bilateral mass lesion	4.07 (1.41)	2.27	6.69
Unilateral mass lesion	3.60 (1.68)	1.49	6.50
MTD-I	3.48 (1.32)	1.06	5.72
MTD-II & III	3.08 (1.14)	1.06	4.91
Laryngitis	3.06 (1.12)	1.25	4.60
Normophonic	1.94 (0.83)	0.26	3.50

Table 2: Mean (SD) of CPPS, HNR, slope values across different pathological conditions and normophonic group

Groups	CPPS	HNR	Slope	Tilt
Unilateral VF palsy	8.39 (3.35)	12.15 (5.73)	-23.24 (4.09)	-11.49 (2.34)
Bilateral mass lesion	10.74 (2.86)	18.35 (5.23)	-24.12 (7.02)	-11.44 (1.62)
Unilateral mass lesion	11.85 (2.73)	18.85 (6.03)	-25.35 (4.19)	-12.43 (0.72)
MTD-I	11.96 (2.89)	19.58 (4.17)	-25.79 (4.72)	-12.84 (0.64)
MTD-II & III	12.90 (1.57)	21.95 (3.64)	-27.67 (6.10)	-11.45 (1.39)
Laryngitis	13.74 (1.04)	20.63 (2.82)	-24.01 (4.06)	-10.59 (2.51)
Normophonic	14.82 (1.71)	22.87 (3.38)	-23.42 (5.61)	-12.53 (1.25)

Table 3: ANOVA results of CPPS, HNR, slope, and tilt across the groups

Parameters	F (6, 95)	Sig.	Partial Eta Squared
CPPS	10.86	0.000	0.407
HNR	7.67	0.000	0.326
slope	1.08	0.375	0.064
tilt	3.85	0.002	0.196

result for shimmer local and shimmer local dB is summarized in Table 5 (in page no. 52). The results of the Tukey post hoc test suggested that shimmer local was found to be significantly high for unilateral VF palsy compared to acute laryngitis and MTD-II & III ($p < 0.05$). However, Partial Eta squared values suggest the lesser effect of pathology. There was no significant difference observed across the bilateral mass lesion, MTD-I, MTD-II & III, and acute laryngitis groups for shimmer local. Shimmer local dB was found to be significantly high for unilateral VF palsy compared to acute laryngitis, MTD-I, and MTD-II & III. Also, no significant difference was found between unilateral palsy and bilateral mass lesion groups.

The results of the Mann Whitney U test indicated that the group with unilateral mass lesion had a significantly higher value compared to the normophonic group for shimmer local (|z|= 2.34, p= 0.019) and shimmer local dB (|z|= 2.30, p= 0.021). The effect size was calculated manually for shimmer local (r= 0.35) and shimmer local dB (r= 0.34) using the formula $r = |z| / \sqrt{N}$, and the results revealed a medium effect of pathology on them. There was no significant difference between unilateral mass lesion and other pathological conditions for both shimmer local and shimmer local dB. The normophonic group was also found to have significantly lower shimmer local and shimmer local dB values compared to all pathological conditions ($p < 0.05$).

DISCUSSION

Comparison of AVQI values across different groups

The first objective was to compare the AVQI values across normophonic voice and dysphonic voice due to different pathological conditions. The results suggest that the normophonic group had significantly lower AVQI values compared to the dysphonic group. These results are in agreement with previous studies (Pebbili et al., 2019; Benoy, 2017)where they had obtained significantly lower AVQI values for the normophonic group compared to the dysphonic group.

The unilateral VF palsy group had significantly higher AVQI values compared to other pathological groups. This could be due to the presence of larger glottic chink and asynchronous VF vibration compared to other pathological conditions. This result supports the finding of Dedo (1992), who reported wide phonatory gap in VF palsy which results in extremely breathy voice quality. The unilateral mass lesion, MTD-I, MTD-II & III, and acute laryngitis were found to have similar AVQI values.AVQI values in these groups were found to be significantly lower than the unilateral VF palsy group and significantly higher than the normophonic group.

The lower AVQI values in the other pathological groups in comparison to the group with unilateral VF palsy can be attributed to the lesser extent

Table 4: Mean (SD) and Median values of shimmer local and shimmer local dB for different pathological conditions and normophonic group

Groups	Shimmer local		Shimmer local dB	
	Mean (SD)	Median	Mean (SD)	Median
Unilateral VF palsy	10.51 (8.03)	10.13	0.98 (0.52)	0.91
Bilateral mass lesion	6.94 (3.52)	6.27	0.66 (0.29)	0.55
Unilateral mass lesion	6.47 (4.69)	4.80	0.66 (0.43)	0.51
MTD-I	5.87 (2.80)	5.32	0.58 (0.21)	0.53
MTD-II & III	4.89 (2.07)	4.75	0.48 (0.15)	0.46
Laryngitis	4.39 (1.36)	4.07	0.46 (0.14)	0.44
Normophonic	3.22 (1.18)	3.13	0.35 (0.46)	0.32

Table 5: ANOVA results of shimmer local and shimmer local dB differentiating across the groups

Parameter	F (4, 53)	Sig.	Partial Eta Squared
Shimmer local	3.328	0.017	0.201
Shimmer local dB	4.863	0.002	0.268

of phonatory gap and irregularity in VF adduction. For example, nodules and polyps are reported to have increased mass and stiffness of the VFs, and hourglass-shaped glottic closure with reduced vibratory amplitude and mucosal wave (Hirano & Bless, 1993). Acute laryngitis is reported to have generalized edema, decreased or absent mucosal wave, and slightly decreased vibratory amplitude (Sapienza & Hoffman-Rudy, 2009). Excessive glottic and supraglottic medial contraction, anterior-posterior contraction of the supraglottic musculature, decreased vibratory amplitude, or psychogenic bowing of VFs is reported in MTD (Altman, Atkinson, & Lazarus, 2005; Lee, & Son, 2005), while unilateral VF palsy is characterized by weakened or bowed VF and the presence of passive vibration around the paralyzed VF. Also, slower initiation of the mucosal wave on the affected side along with a slower period and reduced amplitude of vibration is reported (Sercarz, Berke, Gerratt, Ming, &Natividad, 1992). Hence, the extent of pathology seems to be greater in unilateral VF palsy, resulting in higher AVQI value compared to other pathological conditions.

Comparison of CPPS, HNR, shimmer local, shimmer local dB, slope, and tilt values across different groups

The next objective of the study was to compare values obtained from constituent parameters of AVQI across normophonic and different vocal pathological conditions. The CPPS values were significantly high for the normophonic group compared to unilateral VF palsy, unilateral mass lesion, and bilateral mass lesion. Literature reports high CPPS value for normophonic individuals due to well-defined harmonic structure, and low in severe dysphonic voices as the harmonic formation is restricted by irregular adduction of VFs (Heman-Ackah et al., 2002). The unilateral VF palsy group was found to have significantly lesser CPPS values compared to other pathological groups. Lesser CPPS values could be because of the fact that CPPS have a high correlation with breathiness (Hillenbrand & Houde, 1996) and individuals with unilateral VF palsy generally tend to have greater breathiness component due to a large phonatory gap. The CPPS values obtained for MTD-I, MTD-II &III, acute laryngitis, and normophonic group did not differ significantly, which could be due to lesser severity of dysphonia and lesser extent of pathology in MTD and acute laryngitis compared to palsy and mass lesion conditions.

HNR values were significantly lesser in the unilateral VF palsy group than the normophonic and other pathological groups, due to the presence of high noise components in palsy conditions owing to the wide phonatory gap. In addition, good mobility of VFs would result in better glottic closure in other pathological conditions. The high noise component results from incomplete glottal closure that creates excess air during phonation, which increases the noise amplitude, and in turn, lowers the HNR (Hartlet al., 2001; Oguz, Demirci, Safak, Arslan, Islam, & Kargin, 2007).

Significantly high shimmer local and shimmer local dB values were obtained for the group with unilateral VF palsy compared to other groups, suggesting maximum aperiodic vibration of VFs. Patel and Parsram (2005) had reported significantly higher shimmer values in individuals with VF paralysis compared to normophonic individuals, as a result of asynchronous vibration of VFs. The study also reported higher shimmer values in the mass lesion group compared to normophonics. This can be attributed to the inflammation or small masses on VFs leading to inconsistent glottal closure and poorer VF median edge contact (Oguz, Tarhan, Korkmaz, Yilmaz, Safak, Demirci, &Ozluoglu, 2007). The result of the study is in agreement with Davis (1979), who reported higher values of amplitude perturbation quotient (APQ) in unilateral paralysis followed by nodules and then laryngitis. Lieberman (1963) reported that inflammation and very small growth on

VFs only minimally affected the perturbation measures, while larger masses produced increased perturbation.

Spectral slope and spectral tilt are measures obtained from Long-term Average Spectrum (LTAS) analysis. The signal attained through LTAS represents the vocal function taking place at the larynx as sound and transfer through the vocal tract (Lofqvist & Mandersson, 1987). The spectral slope has been identified as a correlate of hoarseness in the voice. The smaller values of the spectral slope values indicate a slower decline of energy with frequency, which is associated with VF hyperfunction. In comparison, larger spectral slope values indicate a faster decline of energy with frequency, associated with vocal hypofunction (Ludlow et al., 2018). Similarly, spectral tilt was found to be associated with glottal closure during phonation. A reduction in spectral tilt value is associated with hyperadduction and high values are associated with hypoadduction (Ludlow et al., 2018). In the current study, the values obtained for slope and tilt did not vary much across the groups, indicating that slope and tilt might not help discriminate between the pathological conditions when considered in isolation.

CONCLUSIONS

To summarize, the overall AVQI value was found to be useful in discriminating between normophonic and dysphonic voice. However, this cannot be used to differentiate across different pathological conditions. The results obtained from isolated acoustic measures such as CPPS, HNR, shimmer local, and shimmer local dB were able to significantly discriminate between VF palsy, mass lesion, and muscle tension dysphonia. However, the present study is a preliminary attempt and studies are warranted to establish cut-off values for different vocal pathologies. Future studies can be conducted on higher and an equal number of participants in each group restricting the age range, as age affects the acoustic measures. The overall auditory-perceptual dysphonia severity should be taken into consideration for categorizing the voice samples. The size of mass lesions, if considered also might provide us with some remarkable and supporting results. Further such studies can assist SLPs in screening and diagnosis of voice disorders and monitoring the prognosis during the voice therapy effectively. AVQI 02.03 is a non-commercial tool that runs in the Praat program making it cost-effective; it is also less time consuming, and non-invasive.

REFERENCES

Altman, K. W., Atkinson, C., & Lazarus, C. (2005). Current and emerging concepts in muscle tension dysphonia: A 30-month review. *Journal of Voice, 19*(2), 261-267.

Aronson, A. E. (1980). *Clinical Voice Disorders. An Interdisciplinary Approach.* New York, NY: Thieme-Stratton, Inc.

Baker, J., Ben-Tovim, D. I., Butcher, A., Esterman, A., & McLaughlin, K. (2007). Development of a modified diagnostic classification system for voice disorders with inter-rater reliability study. *Logopedics Phoniatrics Vocology, 32*(3), 99-112.

Barsties, B., Lehnert, B., & Janotte, B. (2020). Validation of the acoustic voice quality index version 03.01 and acoustic breathiness index in German. *Journal of Voice, 34*(1), 157.e17-157.e25.

Benoy, J. J. (2017). *Acoustic Voice Quality Index (AVQI) and perceptual measures in the Indian population.* Unpublished Master's Dissertation submitted to the University of Mysore, Mysore, India.

Boominathan, P., Rajendran, A., Nagarajan, R., Seethapathy, J., & Gnanasekar, M. (2008). Vocal abuse and vocal hygiene practices among different level professional voice users in India: A survey. *Asia Pacific Journal of Speech, Language and Hearing, 11*(1), 47-53.

Boone, D. R. (1977). *The Voice and Voice Therapy.* Englewood Cliffs, NJ. Prentice-Hall.

Davis, S.B. (1979). Acoustic characteristics of normal and pathological voices. *Speech and Language, 1*, 271-335.

Dedo, H. H. (1992). Injection and removal of Teflon for unilateral vocal cord paralysis. *Annals of Otology, Rhinology & Laryngology, 101*(1), 81-86.

Delgado H, J., Leon G, N. M., Jiménez, A., Izquierdo, L. M., & Barsties, V. Latoszek, B. (2018). Validation of the acoustic voice quality index version 03.01 and the acoustic breathiness index in the Spanish language. *Annals of Otology, Rhinology & Laryngology, 127*(5), 317-326.

Englert, M., Lopes, L., Vieira, V., & Behlau, M. (2020). Accuracy of acoustic voice quality index and its isolated acoustic measures to discriminate the severity of voice Disorders. *Journal of Voice.* S0892-1997(20)30293-9. Online ahead of print. https://doi.org/10.1016/j.jvoice.2020.08.010

Hartl, D. M., Hans, S., Vaissière, J., Riquet, M., & Brasnu, D. F. (2001). Objective voice quality analysis before and after onset of unilateral vocal fold paralysis. *Journal of Voice, 15*(3), 351-361.

Heman-Ackah, Y. D., Michael, D. D., & GodingJr, G. S. (2002). The relationship between cepstral peak prominence and selected parameters of dysphonia. *Journal of Voice, 16*(1), 20-27.

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech, Language, and Hearing Research, 39*(2), 311-321.

Hirano, M. (1981). *"GRBAS" scale for evaluating the hoarse voice & frequency range of phonation.* In

M. Hirano (Ed.), Clinical Examination of Voice, (pp. 83-84). New York, NY: Springer-Verlag/Wein.

Hirano, M., & Bless, D. M (1993). *Videostroboscopic Examination of the Larynx.* San Diego, CA: Singular.

Hosokawa, K., Barsties, B., Iwahashi, T., Iwahashi, M., Kato, C., Iwaki, S., ... & Ogawa, M. (2017). Validation of the acoustic voice quality index in the Japanese language. *Journal of Voice, 31*(2), 260-e1.

Kim, G. H., Barsties, B., & Lee, Y. W. (2019). Validation of acoustic voice quality index version 3.01 and Acoustic Breathiness Index in Korean Population. *Journal of Voice.* S0892-1997(19)30342-X. Online ahead of print. doi.org/10.1016/j.jvoice.2019.10.005.

Konadath, S., Chatni, S., Lakshmi, M. S., & Saini, J. K. (2017). Prevalence of communication disorders in a group of islands in India. *Clinical Epidemiology and Global Health, 5*(2), 79-86.

Lee, E. K., & Son, Y. I. (2005). Muscle tension dysphonia in children: Voice characteristics and outcome of voice therapy. *International Journal of Pediatric Otorhinolaryngology, 69*(7), 911-917.

Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *The Journal of the Acoustical Society of America, 35*(3), 344-353.

Löfqvist, A., & Mandersson, B. (1987). Long-time average spectrum of speech and voice analysis. *Folia Phoniatricaet Logopaedica, 39*(5), 221-229.

Ludlow, C. L., Kent, R. D., & Gray, L. C. (2018). *Measuring Voice, Speech, and Swallowing in the Clinic and Laboratory.* San Diego, CA: Plural Publishing.

Maryn, Y. (2013). The Acoustic Voice Quality Index in the Praat program: A practical guide. *Belsele:Flemish Association for Speech Language Therapists.*

Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. *Journal of Voice, 24*(5), 540-555.

Maryn, Y., De Bodt, M., & Roy, N. (2010). The acoustic voice quality index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders, 43*(3), 161-174.

Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. *The Journal of the Acoustical Society of America, 126*(5), 2619-2634.

Oguz, H., Demirci, M., Safak, M. A., Arslan, N., Islam, A., & Kargin, S. (2007). Effects of unilateral vocal cord paralysis on objective voice measures obtained by Praat. *European Archives of Oto-Rhino-Laryngology, 264*(3), 257-261.

Oguz, H., Tarhan, E., Korkmaz, M., Yilmaz, U., Safak, M. A., Demirci, M., & Ozluoglu, L. N. (2007). Acoustic analysis findings in objective laryngopharyngeal reflux patients. *Journal of Voice, 21*(2), 203-210.

Patel, R., & Parsram, K. S. (2005). Acoustic analysis of subjects with vocal cord paralysis. *Indian Journal of Otolaryngology and Head and Neck Surgery, 57*(1), 48-51.

Pebbili, G. K., Shabnam, S., Pushpavathi, M., Rashmi, J., Sankar, R. G., Nethra, R., ... & Shashish, G. (2019). Diagnostic Accuracy of Acoustic Voice Quality Index Version 02.03 in Discriminating across the Perceptual Degrees of Dysphonia Severity in Kannada Language. *Journal of Voice.* https://doi.org/10.1016/j.jvoice.2019.07.010

Sapienza, C., & Hoffman-Rudy, B. (2009). *Voice Disorders.* San Diego, CA: Plural Publishing.

Sercarz, J. A., Berke, G. S., Gerratt, B. R., Ming, Y., & Natividad, M. (1992). Videostroboscopy of human vocal fold paralysis. *Annals of Otology, Rhinology & Laryngology, 101*(7), 567-577.

Shasidhar, K.N. (1984). Analysis of Speech of Stutterers. *An unpublished Master's Dissertation submitted to University of Mysore*, Mysore, India.

Sinha, S. K., Shivaswamy, J., Barman, A., Seth, D., Seshadri, D., & Savithri, S. R. (2017). Prevalence of communication disorders in a rural population at taluq level of Gujarat, India. *Clinical Epidemiology and Global Health, 5*(2), 73-78.

Stemple, J. C., Roy, N., & Klaben, B. K. (2018). *Clinical Voice Pathology: Theory and Management.* San Diego, CA: Plural Publishing.

Uloza, V., Petrauskas, T., Padervinskis, E., Ulozaitė, N., Barsties, B., & Maryn, Y. (2017). Validation of the acoustic voice quality index in the Lithuanian language. *Journal of Voice, 31*(2), 257.e1-257.e11.

APPENDIX I

1. Smoothened cepstral peak prominence (CPPS) is defined as the distance between the first harmonic's peak and the point with equal quefrency on the regression line through the smoothed cepstrum.

2. Harmonics-to-noise ratio (HNR) is defined as the base-10-logarithm of the ratio between the periodic energy and the noise energy, multiplied by 10.

3. Shimmer Local is the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude.

4. Shimmer Local dB is the average absolute base-10 logarithm of the difference between the amplitudes of consecutive periods, multiplied by 20.

5. The general slope of the spectrum (slope)is defined as the difference between the energy in 0-1000Hz and the energy in 1000- 10,000 Hz of the long-term average spectrum.

6. The tilt of the regression line through the spectrum (tilt) is defined as the difference between the energy in 0-1000 Hz and the energy in 1000- 10, 000Hz of the trendline through the long-term average spectrum.