

## **Acoustic Voice Quality Index as a potential tool for voice screening**

**<sup>1&2)</sup>Maryam Faham, <sup>3)</sup>Anne-Maria Laukkanen, <sup>3)</sup>Tero Ikävalko, <sup>4)</sup>Leena Rantala, <sup>5)</sup>Ahmed Geneid, <sup>6)</sup>Sofia Holmqvist-Jämsén, <sup>5)</sup>Kaarina Ruusuvirta, <sup>7)</sup>Sirpa Pirilä**

<sup>1)</sup> Speech and Language Pathology Department, Rehabilitation Faculty, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>2)</sup> Rehabilitation Sciences Research Center, Shiraz University of Medical Sciences, Shiraz, Iran.

<sup>3)</sup> Speech and Voice Research Laboratory, Faculty of Social Sciences, Tampere University, Tampere, Finland

<sup>4)</sup> Unit of Logopedics, Faculty of Social Sciences, Tampere University, Tampere, Finland

<sup>5)</sup> Department of Otorhinolaryngology and Phoniatics – Head and Neck Surgery, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

<sup>6)</sup> Åbo Akademi University, Faculty of Arts, Psychology and Theology, Helsinki, Finland

<sup>7)</sup> Faculty of Humanities/Logopedics, Child Language Research, University of Oulu, Oulu, Finland

*Corresponding Author:*

*Maryam Faham*; Speech and Language Pathology Department, Rehabilitation Faculty, Shiraz University of Medical Sciences, Shiraz, Iran.

Tel: +989126709404

E-mail: [mfaham@sums.ac.ir](mailto:mfaham@sums.ac.ir); [Maryam.faham@yahoo.com](mailto:Maryam.faham@yahoo.com)

Address: First Abiverdi, Chamran street; Rehabilitation Faculty-

71945-33699

## Abstract

**Introduction:** To diminish the risk of voice disorders in people who are highly dependent on their voices, such as teachers, vocal screening is important already at the beginning of such individuals' professional studies. A reliable, specified screening tool is needed. The Acoustic Voice Quality Index (AVQI) has been found to differentiate normal voices from abnormal voices and to serve as a treatment outcome measure. This study investigated whether AVQI could be a screening tool in combination with auditory- and self-perception of the voice to discriminate normal from slightly poor voices.

**Type of Study:** Experimental

**Methods:** Some 128 female teaching students (mean age 26.39 years, SD 9.80 years) with no diagnosed voice disorders participated in this study. They read aloud a text in Finnish, sustained the vowel /a:/, and filled the Voice Handicap Index (VHI) questionnaire. Voice samples were recorded with an AKG C544L headset microphone, iFocusrite soundcard, and Praat software using a 44100 sample rate and 16-bit amplitude quantization. Five expert voice therapists evaluated the samples to determine the grade of dysphonia (G) using a scale of 0–0.5 (=normal), 0.5–1 (=mild), 1–2 (=moderate), and 2–3 (=severe). Three medial seconds of [a:] and the first 31 syllables of the text were analyzed using AVQI script version 03.01 in Praat (5.3.55). The analysis gives one AVQI score per participant (scale 0–10).

The AVQI threshold of normal and disordered voices for Finnish speakers is 1.83; a  $G_{\text{mean}}=0.0-0.5$  and VHI score <19 were considered normal. Statistical analysis was done using the receiver operating characteristic (ROC) curve, Spearman's correlation coefficient, and the independent samples *t*-test.

**Results:** According to the AVQI results, the area under the curve ( $A_{\text{ROC}}$ ) was 0.554, which is fair. The Youden index gave a cutoff value of 0.30 with a sensitivity of 85% and a specificity of 81.1%. There were weak but significant correlations between  $G_{\text{mean}}$  and AVQI and two AVQI parameters, smoothed cepstral peak prominence (CPPS) and harmonic-to-noise ratio (HNR;  $r=0.27$ ;  $-0.24$ ;  $-0.20$ , respectively;  $p < 0.05$ ); and between total VHI and AVQI score and CPPS ( $r=0.21$ ;  $0.20$ , respectively;  $p < 0.05$ ). Furthermore, the AVQI scores differed significantly between the groups with a VHI total score <19 and  $\geq 19$ .

**Conclusions:** AVQI did not differentiate between voices that had been perceptually judged as normal or slightly abnormal, but a combination of perceptual assessment in the form of AVQI and VHI could better screen slightly deviant voices.

**Key Words:** voice disorder; AVQI; acoustic analysis; auditory-perceptual assessment; Voice Handicap Index (VHI)

## INTRODUCTION

In clinical practice, it is important not only to treat voice disorders, but also to prevent them. The need is especially crucial in high-risk groups, such as teachers. The incidence of voice disorders among teachers is higher than in the general population<sup>(1)</sup> due to high vocal loading in a noisy environment, poor knowledge of the voice, and a lack of vocal training that would prepare teaching students for the demands of the profession. A preventive voice care program could help to improve the voice knowledge and vocal skills of teachers and teaching students and thus reduce the incidence of voice disorders.<sup>(2-4)</sup> On the other hand, a robust screening tool is needed to increase the cost-effectiveness of such preventive care programs.<sup>(1)</sup> Proper vocal screening of teaching students would help to uncover voice problems at an early stage and direct those who are most urgently in need to voice training and therapy to protect them from different severe voice disorders. However, there is a lack of a proper screening tool both as far as objective and subjective screening methods are concerned.

Among the existing voice quality evaluation methods, acoustic analysis is attractive in clinical voice assessments, since it is easy to use and may offer reliable objective results for vocal assessment. In traditional acoustic analysis, sustained vowels – mostly /a:/– have been used for evaluation, which has various advantages and disadvantages. As the most simple voice samples, vowels are easy to produce and thus quick to record. They do not have variation in time domain, non-voiced segments, or rapid onsets and offsets. They are also free from the effects of language, prosody, phonetic context, and speech rate. While sustained vowels may be well suited to the analysis of vocal perturbations, they do not offer a comprehensive representation of voice quality. Various fluctuations related to voice onsets and offsets and voice breaks may be indicative of voice disorders, and they may not be represented in sustained vowels but require connected speech for evaluation.<sup>(5, 6)</sup>

To yield an “ecologically valid” acoustic analysis, two multi-parametric models have been introduced, the Cepstral Spectral Index of Dysphonia (CSID) by Awan et al. (2009),<sup>(7)</sup> and the Acoustic Voice Quality Index (AVQI) by Maryn et al. (2010).<sup>(5, 8)</sup> Both models were developed to meet the limitations of traditional acoustic analysis, and there are some differences between them. First, CSID provides separate dysphonia severity indexes for the sustained vowel and connected speech (range 0–100), while AVQI gives one combined dysphonia severity index for the sustained vowel and connected speech (range 0–10). The parameters included in these models are

also somewhat different. For connected speech, CSID parameters include smoothed cepstral peak prominence (CPPS), a low/high spectral ratio for sound energy below and above 4 kHz, and a standard deviation (SD) of this ratio. For the vowel segment, the above-mentioned parameters are calculated; furthermore, the SD of the CPPS is included and the gender is taken into account.<sup>(9-11)</sup> AVQI parameters include CPPS, harmonic-to-noise ratio (HNR), two values for shimmer, and two values that describe the slope of the long-term average spectrum of the sample (a high/low spectral ratio for sound energy below and above 1 kHz, and the tilt of the regression line through the spectrum).<sup>(5)</sup> CSID is part of the Analysis of Dysphonia in Speech and Voice program (ADSV, model 5109; KayPENTAX, Montvale, NJ) that operates using the Computerized Speech Laboratory (CSL; KayPENTAX). AVQI runs in Praat software.<sup>(9)</sup>

Both CSID and AVQI have been found to distinguish normal and disordered voices reliably.<sup>(12)</sup> CSID has been found to yield a high reliability in measuring dysphonia severity in various voice disorders.<sup>(9, 13)</sup> AVQI has been found to correlate well with perceptual assessment results<sup>(8)</sup> and to differentiate dysphonic from normophonic voices with 79% accuracy.<sup>(14)</sup> Barsties et al. have also shown that age and gender do not significantly affect AVQI. <sup>(15)</sup> As connected speech is included in AVQI, it needs to be validated for different languages. Up to now, AVQI has been validated for many languages – including Lithuanian,<sup>(16)</sup> Korean,<sup>(17)</sup> Finnish,<sup>(18)</sup> Dutch,<sup>(5, 8, 19)</sup> German, French, and English<sup>(19)</sup> – showing high diagnostic accuracy, sensitivity, and specificity. The results thus suggest that AVQI is a robust clinical tool. There is also evidence that AVQI is able to track differences before and after voice therapy, which emphasizes its clinical usability.<sup>(20)</sup>

Alongside acoustic analysis, auditory perceptual assessment is often the first step in clinical and research practices to differentiate normal from abnormal voice quality. Various perceptual assessment tools, including GRBAS, RHB, and CAPE-V, have been shown to be reliable.<sup>(6)</sup> Awan et al. found a strong correlation between CSID and CAPE-V results.<sup>(21)</sup> Both AVQI and CSID have yielded a high correlation with the perceptual assessment of dysphonia severity (similar to CAPE-V), although the correlation for CSID was somewhat higher.<sup>(9)</sup>

Voice clients' self-assessment of the effects of a probable voice disorder on their daily life is also very important, especially for professional voice users when taking into account their high level of voice dependency. Questionnaires like the Voice Handicap Index (VHI),<sup>(22)</sup> Voice-Related Quality of Life (V-RQOL),<sup>(23)</sup> and Voice Activity and Participation Profile (VAPP)<sup>(24)</sup> are commonly used and found to be reliable objective methods to measure treatment outcomes<sup>(25, 26)</sup> or evaluate the

present quality of life of patients with a voice disorder.<sup>(27-30)</sup> Thomas et al. found that 17.2% of teaching students had voice complaints and higher VHI scores than those without voice complaints and the non-teacher population; long, intensive hours of voice use were perceived as risk factors that may lead to a voice problem.<sup>(29)</sup>

In total, it seems that both AVQI and CSID are highly reliable tools for the assessment of a clearly disordered voice. This raises the question of whether they are also suited for use as screening tools. The present study focused on AVQI, since it is widely used, functions free of charge in Praat, and has been validated in the Finnish language. The aim of the present study was to investigate whether AVQI can be used as a quick, noninvasive screening tool to discriminate normal voices from slightly poor, potentially dysphonic ones. As a reference in such a large-scale screening, where the laryngeal investigation is not possible in practice due to time and financial constraints, we used perceptual voice analysis with GRBAS and the self-evaluation of the participants applying VHI.

## **MATERIAL AND METHODS**

### *Participants*

All 128 participants were female teaching students with a mean age of 26.39 (SD 9.80) years. They were informed about the purpose of the study and the procedure, and they volunteered to participate as part of their orientation to university studies. The participants filled a questionnaire concerning health issues that can affect the voice, like asthma, allergies, reflux, and the use of chemicals (smoking, medication, etc.). Based on the background information, the participants had no known disorders of speech, voice, or hearing. There were only seven smokers. No one was excluded from the study, as the aim was to see whether AVQI and auditory analysis are able to distinguish participants with signs of mild dysphonia. None of the participants was recorded while suffering from an acute respiratory infection or other acute deterioration of the voice, since the voice sample would then not represent the ordinary state of the participant's voice.

### *Voice samples*

Participants read aloud a text ("North wind and the sun"; in Finnish "Pohjantuuli ja aurinko") and sustained the vowel [a:] three times. Both tasks were performed at a comfortable habitual speaking pitch and loudness. All voice samples were recorded with an AKG C544L head-mounted

condenser microphone and digitized at 44100 samples per second and a 16-bit depth to a PC using a Focusrite iTrack Solo soundcard. The recordings were made in a well-damped studio of the Speech and Voice Research Laboratory at Tampere University. All voice samples were analyzed using AVQI script version 03.01<sup>(31)</sup> in Praat (5.3.55). Three medial seconds of the second [a:] and the first 31 syllables of the text were used for the analyses. The AVQI script automatically gives one index value per participant (on a 0–10 scale) and the analysis results for the six parameters included in the AVQI script (jitter, shimmer in % and dB, HNR, CPPS, and two parameters quantifying spectral slope). The threshold for dysphonia in Finnish speakers has been found to be 3.09 for AVQI version 2<sup>(18)</sup>, and 1.83 for AVQI 03.01.<sup>(32)</sup>

### *Auditory-perceptual evaluation*

For auditory-perceptual evaluation, concatenated speech samples containing 31 syllables followed by three seconds of vowel prolongation were examined. Five expert voice therapists with at least three years of voice therapy experience evaluated the voice quality perceptually while wearing well-sealed headphones in a quiet room using an ordinal scale for the grade (G) of dysphonia as follows: 0–0.5=normal, 0.5–1=mild, 1–2=moderate, and 2–3=severe. The listeners were free to listen to each sample as many times as needed. To measure intra-rater reliability, 10% of the voice samples were repeated in random order in the total set of samples. The mean rating for G was used in further analyses.  $G_{\text{mean}} \leq 0.5$  was considered an indicator of a quite normal voice.<sup>(33)</sup>

### *Self-assessment*

For the self-assessment part, the participants were asked to fill in the VHI questionnaire. VHI is a 30-item questionnaire rating the effect of a voice disorder on quality of life. VHI has three subgroups of questions rating voice-related physical, functional, and emotional handicaps, and each subgroup contains 10 items. Participants should score each item from 0 (never) to 4 (always); so the total score ranges from 0 to 120, where 0 means no negative effect of any voice problem on quality of life. Behlau et al. defined the threshold as a score of 19.<sup>(34)</sup> Following this, the participants were divided into two groups: Participants in Group A had a total score <19 and those with total scores  $\geq 19$  were placed in Group B.

## *Statistical analysis*

The Statistical Package for the Social Sciences, version 22.0 (IBM Corp, Armonk, New York, USA) was used for the analysis of descriptive and analytic data. The inter-rater and intra-rater reliability of the auditory-perceptual assessment were calculated with Spearman's correlation coefficient and intraclass correlation coefficient (ICC), respectively.

To evaluate the perceptual diagnostic accuracy of the AVQI based on the perceptual evaluation, the receiver operating characteristic (ROC) curve analysis was performed.  $G_{\text{mean}}=0.0-0.5$  was defined as a normal voice without any sign of dysphonia. In the ROC curve analysis, the best threshold level for the AVQI was determined using the Youden index, which produces the best threshold provided by the maximum of the sensitivity + (specificity - 1).

Relations between auditory-perceptual evaluation, AVQI total score and its sub-parameter values, and VHI score were studied with Spearman's correlation coefficient. The  $r$  was interpreted based on the guidelines of Frey et al.<sup>(35)</sup> An  $r < 0.20$  indicates a slight correlation,  $r = 0.20-0.40$  a low correlation,  $r = 0.41-0.70$  a moderate correlation;  $r = 0.71-0.90$  a high correlation, and  $r > 0.90$  a very high correlation.

Finally, the independent samples  $t$ -test was applied to study whether participants with total VHI scores  $< 19$  and  $\geq 19$  differed significantly from each other in terms of the AVQI total score. This was also done for all six AVQI parameters in both groups. All results were considered statistically significant at  $p \leq 0.05$ .

## **RESULTS**

### *Reliability of perceptual assessment*

The inter- and intra-rater agreements were analyzed using Spearman's correlation and ICC. The results showed low inter-rater agreement (0.2-0.3) and low-to-moderate intra-rater agreement (0.3-0.5).

### *Audio-perceptual results*

Figure 1 illustrates the  $G_{\text{mean}}$  distribution over each severity rank. There were weak but statistically significant correlations between  $G_{\text{mean}}$  and AVQI, CPPs, and HNR ( $r = 0.27$ ;  $-0.24$ ; and  $-0.20$ , respectively,  $p < 0.05$ ).

### *Diagnostic accuracy of AVQI*

The mean of the AVQI score was 1.07 (SD 0.77). The descriptive data of AVQI and the parameters it consists of are presented in Table 1. A ROC analysis (Figure 2) was performed to see if AVQI 03.01 is able to distinguish normal voices from voices with subtle deviations. The  $G_{\text{mean}}$  was calculated and then grouped as 0–0.5, 0.5–1, 1–2, and 2–3. A  $G_{\text{mean}} \leq 0.5$  was considered an indicator of a quite normal voice. The area under the curve ( $A_{\text{ROC}}$ ) was equal to 0.554, which means that the accuracy of AVQI 03.01 for screening is fair based on the perceptual-auditory judgment of the overall voice quality. The Youden index gave a cutoff value of 0.30 with a sensitivity of 85% and a specificity of 81.1%.

### *Results of self-assessment*

The VHI scores showed a mean of 15.95 (SD 10.01). There were weak significant correlations between total VHI and AVQI score and CPPS ( $r=0.21$  and  $0.20$ , respectively,  $p < 0.05$ ). VHI and HNR did not correlate with each other ( $r=0.03$ ,  $p=0.77$ ).

A VHI score of 19 was assigned as the diagnostic threshold. (36) Thus, the 88 participants whose VHI score was  $<19$  were assigned to Group A, and the 40 participants whose VHI score was  $\geq 19$  were assigned to Group B. The independent samples  $t$ -test showed that the AVQI and CPPS differed significantly between the groups, as shown in Table 2.

## **DISCUSSION**

Earlier studies have shown that AVQI is able to differentiate dysphonic from normophonic voices in different languages<sup>(16, 17, 31, 36-38)</sup> and also in different age groups of both genders<sup>(15)</sup> with a high accuracy of around 79%. (14) This is of importance in clinical practice. However, the great prevalence of voice disorders in vocally loading voice occupations like teaching calls for the establishment of effective preventive procedures. There is a need for the reliable screening of the early signs of voice disorders as early as possible, preferably already during professional education. The aim of this study was to investigate whether AVQI could distinguish subtle voice deviations from the normal voice, and thus whether it would be applicable for screening the early stages of voice problems in people studying for a voice profession with a high risk of developing voice disorders, such as teaching.

Based on our results, it seems that with some precautions AVQI could be usable to distinguish between normal voices and voices with subtle characteristics of potential early-stage dysphonia. The  $A_{\text{ROC}}$  was 0.54, which is not good, but fair. The participants were grouped based on the voice



experts' auditory-perceptual assessment.  $G_{\text{mean}} < 0.5$  was considered indicative of a normal voice (without any deviation) and a higher  $G_{\text{mean}}$  was indicative of voices with subtle changes. The AVQI ranged from -1 to 3 in our participants, and an index of 0.3 was defined as a threshold for subtle voice changes. However, the  $A_{\text{ROC}}$  was not strong, so the result should be interpreted with caution. The AVQI 03.01 threshold for normophonic and dysphonic voices in Finnish speakers is 1.83 (Kankare et al. in press). Based on this, 30.2% of the voice samples evaluated in the present study received an AVQI that suggests dysphonia. It is important to note that the way of speaking affects the values obtained in the acoustic analysis. For instance, speaking softly increases jitter and shimmer and decreases HNR, while speaking loudly has the opposite effect.<sup>(39-41)</sup> This suggests that it is important to instruct the participants to use a normal habitual conversational loudness and pitch level in any recordings of the voice samples to be used in voice screening. To control the influence of vocal loudness, we instructed the participants to read the text and sustain the vowel at their habitual conversational loudness.

The correlations between AVQI parameters and  $G_{\text{mean}}$  were weak but significant. This is in line with earlier studies.<sup>(5, 33)</sup> In the first version of AVQI introduced by Maryn et al. (2010),<sup>(5)</sup> the highest correlations were obtained for CPPS ( $r_s = 0.71$ ) followed by HNR ( $r_s = 0.68$ ).

The inter- and intra-rater reliability of the auditory-perceptual analysis was low in the present study. This seems to suggest difficulties in judging voices without clear signs of dysphonia. On the one hand, this may be because no anchor samples were presented to the judges of the present study. On the other hand, there seems to be a need for further special training in perceptual analysis for screening purposes in the curriculum of logopedics students. For instance, speaking softly and with a hypofunctional type of voice production increases breathiness and turbulence noise in the voice. The use of vocal fry and creaky voice may be misjudged as roughness. These risks for misjudgment apply both to auditory-perceptual and acoustic analysis. Low inter- and intra-rater reliability is recognized as one of the limitations of our study. The short length of the connected speech sample used for the AVQI analysis may be one of the factors that impair perceptual analysis, especially in those voice samples that are normal or near-normal with only subtle deviations from the normal voice. It would be worth studying the effect of voice sample length on the reliability of perceptual assessment and its correlation with AVQI scores in normal or near-normal

The VHI results of the present study also showed a slight but significant correlation with AVQI and CPPS, but not with HNR, unlike what we saw in the perceptual results. Awan et al.<sup>(10)</sup> also found that CPPS had a weak-to-moderate correlation with VHI in a group of patients with various voice pathologies. Pommee et al.<sup>(33)</sup> have also reported that there were moderate correlations between VHI and VHI sub-scores and AVQI. They stated that there was a strong association between the diagnostic accuracy of these two tools. This connection is also clear in the present study, where we could see significant differences in AVQI scores between participants with total VHI scores of <19 and ≥19.

## **CONCLUSIONS**

Our results suggest that in the absence of a reliable laryngeal investigation, a combination of auditory-perceptual evaluation, AVQI, and VHI may be applicable to screen the subtle voice changes that may be early signs of dysphonia. This kind of screening would be necessary especially for students studying for vocally loading occupations. However, the auditory-perceptual results obtained in the study were not strong enough to draw a firm conclusion. Special auditory-perceptual training for judging slight vocal deviations might increase the reliability of the perceptual analysis.

## **DECLARATION OF INTERESTS**

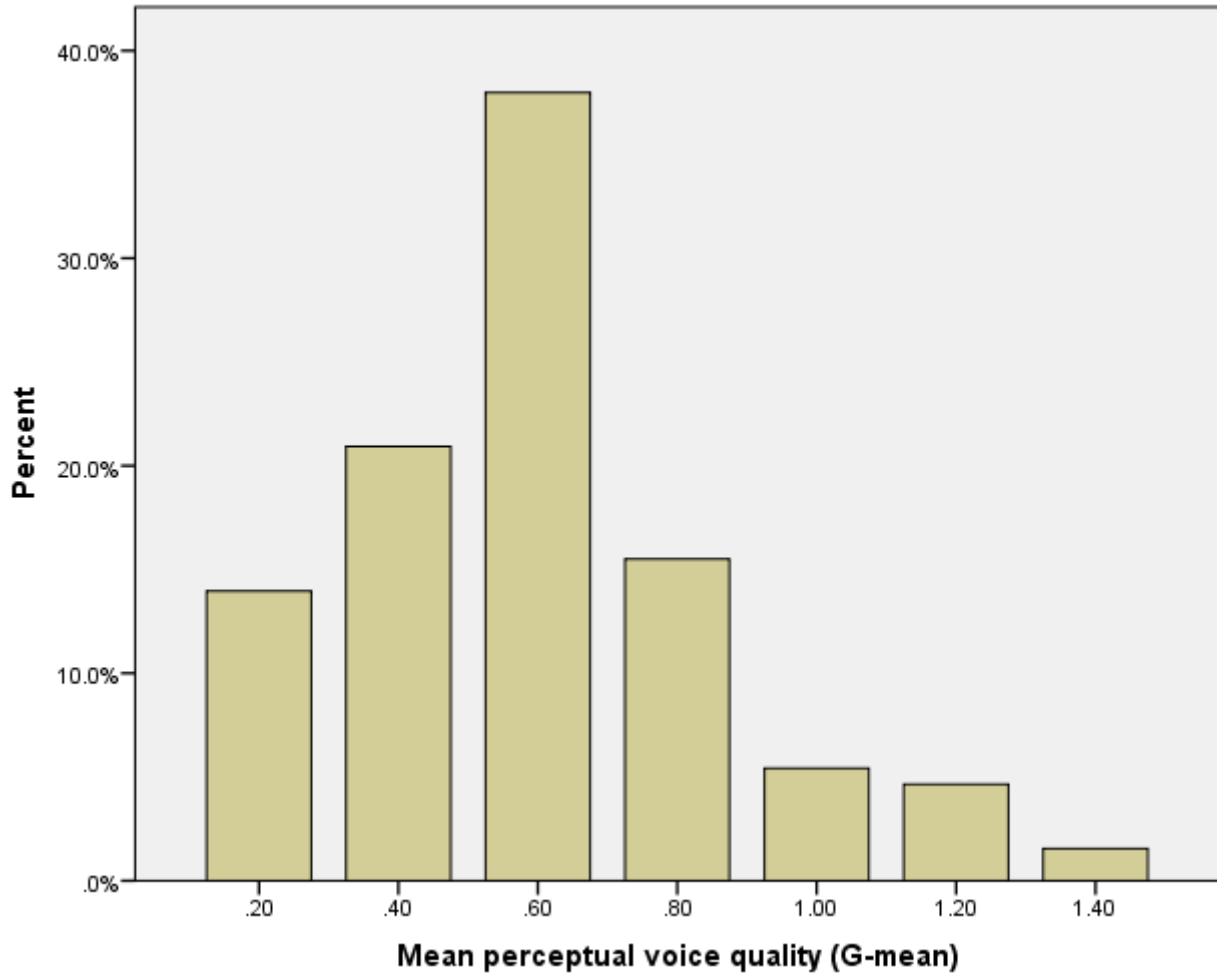
The authors have no interests to report.

## **REFERENCES**

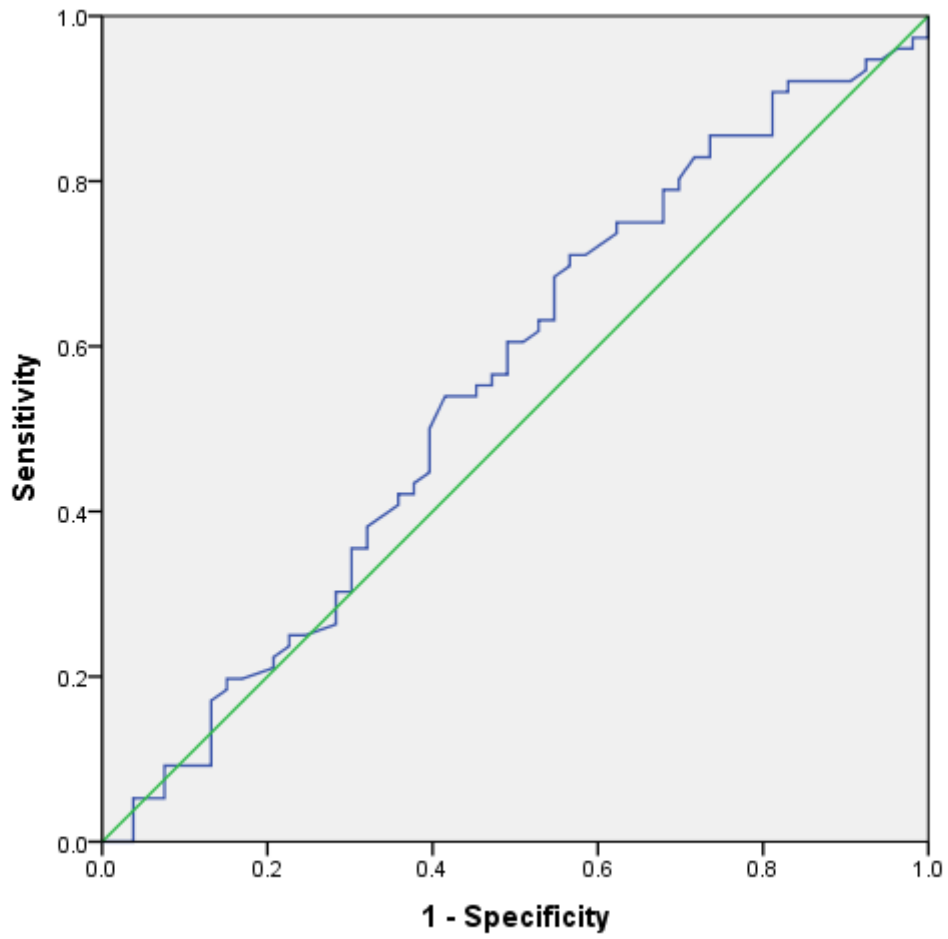
1. Hazlett DE, Duffy OM, Moorhead SA. Review of the Impact of Voice Training on the Vocal Quality of Professional Voice Users: Implications for Vocal Health and Recommendations for Further Research. *Journal of Voice*. 2011;25(2):181-91.
2. Bovo R, Galceran M, Petruccelli J, Hatzopoulos S. Vocal problems among teachers: Evaluation of a preventive voice program. *Journal of Voice*. 2007;21(6):705-22.
3. Nanjundeswaran C, Li NY, Chan KM, Wong RK, Yiu EM, Verdolini-Abbott K. Preliminary data on prevention and treatment of voice problems in student teachers. *Journal of voice : official journal of the Voice Foundation*. 2012;26(6):816 e1-12.
4. Timmermans B, Coveliers Y, Meeus W, Vandenabeele F, Van Looy L, Wuyts F. The Effect of a Short Voice Training Program in Future Teachers. *Journal of Voice*. 2011;25(4):E191-E8.
5. Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *Journal of voice : official journal of the Voice Foundation*. 2010;24(5):540-55.
6. Barsties B, De Bodt M. Assessment of voice quality: current state-of-the-art. *Auris Nasus Larynx*. 2015;42(3):183-8.
7. Awan SN, Roy N, Dromey C. Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model. *Clin Linguist Phon*. 2009;23(11):825-41.

8. Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. *J Commun Disord*. 2010;43(3):161-74.
9. Lee JM, Roy N, Peterson E, Merrill RM. Comparison of two multiparameter acoustic indices of dysphonia severity: the Acoustic Voice Quality Index and Cepstral Spectral Index of dysphonia. *Journal of Voice*. 2018;32(4):515. e1-. e13.
10. Awan SN, Roy N, Cohen SM. Exploring the Relationship Between Spectral and Cepstral Measures of Voice and the Voice Handicap Index (VHI). *Journal of Voice*. 2014;28(4):430-9.
11. Watts CR. The Effect of CAPE-V Sentences on Cepstral/Spectral Acoustic Measures in Dysphonic Speakers. *Folia Phoniatr Logop*. 2015;67(1):15-20.
12. Barsties B, Maryn Y. The improvement of internal consistency of the Acoustic Voice Quality Index. *Am J Otolaryngol*. 2015;36(5):647-56.
13. Peterson EA, Roy N, Awan SN, Merrill RM, Banks R, Tanner K. Toward Validation of the Cepstral Spectral Index of Dysphonia (CSID) as an Objective Treatment Outcomes Measure. *Journal of Voice*. 2013;27(4):401-10.
14. Barsties VLB, Ulozaite-Staniene N, Petrauskas T, Uloza V, Maryn Y. Diagnostic Accuracy of Dysphonia Classification of DSI and AVQI. *Laryngoscope*. 2019;129(3):692-8.
15. Barsties V, Latoszek B, Ulozaite-Staniene N, Maryn Y, Petrauskas T, Uloza V. The Influence of Gender and Age on the Acoustic Voice Quality Index and Dysphonia Severity Index: A Normative Study. *Journal of voice : official journal of the Voice Foundation*. 2019;33(3):340-5.
16. Uloza V, Petrauskas T, Padervinskis E, Ulozaitė N, Barsties B, Maryn Y. Validation of the acoustic voice quality index in the Lithuanian language. *Journal of Voice*. 2017;31(2):257. e1-. e11.
17. Maryn Y, Kim HT, Kim J. Auditory-Perceptual and Acoustic Methods in Measuring Dysphonia Severity of Korean Speech. *Journal of voice : official journal of the Voice Foundation*. 2016;30(5):587-94.
18. Kankare E, Barsties V, Latoszek B, Maryn Y, Asikainen M, Rorarius E, Vilpas S, et al. The acoustic voice quality index version 02.02 in the Finnish-speaking population. *Logopedics Phoniatics Vocology*. 2019:1-8.
19. Maryn Y, De Bodt M, Barsties B, Roy N. The value of the Acoustic Voice Quality Index as a measure of dysphonia severity in subjects speaking different languages. *European Archives of Oto-Rhino-Laryngology*. 2014;271(6):1609-19.
20. Hosokawa K, Barsties B, Iwahashi T, Iwahashi M, Kato C, Iwaki S, et al. Validation of the Acoustic Voice Quality Index in the Japanese Language. *Journal of voice : official journal of the Voice Foundation*. 2017;31(2):260 e1- e9.
21. Awan SN, Roy N, Jette ME, Meltzner GS, Hillman RE. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clin Linguist Phon*. 2010;24(9):742-58.
22. Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, et al. The voice handicap index (VHI) development and validation. *American Journal of Speech-Language Pathology*. 1997;6(3):66-70.
23. Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of voice : official journal of the Voice Foundation*. 1999;13(4):557-69.
24. Ma EP, Yiu EM. Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *J Speech Lang Hear Res*. 2001;44(3):511-24.
25. Faham M, Ahmadi A, Drinnan M, Saadatmand N, Fatahi E, Jalalipour M. The Effects of a Voice Education Program on VHI Scores of Elementary School Teachers. *Journal of voice : official journal of the Voice Foundation*. 2016;30(6):755 e1- e11.
26. Roy N, Weinrich B, Gray SD, Tanner K, Stemple JC, Sapienza CM. Three treatments for teachers with voice disorders: a randomized clinical trial. *J Speech Lang Hear Res*. 2003;46(3):670-88.
27. Niebudek-Bogusz E, Woznicka E, Zamyslowska-Szmytke E, Sliwinska-Kowalska M. Correlation between acoustic parameters and Voice Handicap Index in dysphonic teachers. *Folia Phoniatr Logop*. 2010;62(1-2):55-60.

28. Lu D, Wen B, Yang H, Chen F, Liu J, Xu Y, et al. A Comparative Study of the VHI-10 and the V-RQOL for Quality of Life Among Chinese Teachers With and Without Voice Disorders. *Journal of voice : official journal of the Voice Foundation*. 2017;31(4):509 e1- e6.
29. Thomas G, Kooijman PGC, Donders ART, Cremers CWRJ, de Jong FICRS. The voice handicap of student-teachers and risk factors perceived to have a negative influence on the voice. *Journal of Voice*. 2007;21(3):325-36.
30. Vanhoudt I, Thomas G, Wellens WA, Vertommen H, de Jong FI. The background biopsychosocial status of teachers with voice problems. *J Psychosom Res*. 2008;65(4):371-80.
31. Barsties B, Maryn Y. External Validation of the Acoustic Voice Quality Index Version 03.01 With Extended Representativity. *Ann Otol Rhinol Laryngol*. 2016;125(7):571-83.
32. Kankare E RL, Ikävalko T, Barsties v. Latoszek B, Laukkanen A-M. Akustisen äänenlaatuindeksin (AVQI) version 03.01 validointi suomenkielisille puhujille [Validation of AVQI version 03.01 for Finnish speakers, in Finnish. *Puhe ja kieli*. 2019;In press.
33. Pommée T, Maryn Y, Finck C, Morsomme D. Validation of the Acoustic Voice Quality Index, Version 03.01, in French. *Journal of Voice*. 2018.
34. Behlau M, Madazio G, Moreti F, Oliveira G, Dos Santos Lde M, Paulinelli BR, et al. Efficiency and Cutoff Values of Self-Assessment Instruments on the Impact of a Voice Problem. *Journal of voice : official journal of the Voice Foundation*. 2016;30(4):506 e9- e18.
35. Frey LR BC, Friedman PGKG. *Investigating Communication: An Introduction to Research Methods* Englewood Cliffs, NJ: Prentice-Hall; 1991.
36. Hosokawa K, Barsties VLB, Iwahashi T, Iwahashi M, Iwaki S, Kato C, et al. The Acoustic Voice Quality Index Version 03.01 for the Japanese-speaking Population. *Journal of voice : official journal of the Voice Foundation*. 2019;33(1):125 e1- e12.
37. Delgado Hernández J, León Gómez NM, Jiménez A, Izquierdo LM, Barsties v. Latoszek B. Validation of the Acoustic Voice Quality Index Version 03.01 and the Acoustic Breathiness Index in the Spanish language. *Annals of Otology, Rhinology & Laryngology*. 2018;127(5):317-26.
38. Kankare E, Barsties B, Maryn Y, Ilomäki I, Laukkanen A, Tyrmi J, et al., editors. A preliminary study of the Acoustic Voice Quality Index in Finnish speaking population. 11th Pan European Voice Conference; 2015 August 31th—September 4th; Florence; 2015: Firenze University Press, Italy.
39. Orlikoff RF, Kahane JC. Influence of Mean Sound Pressure Level on Jitter and Shimmer Measures. *Journal of Voice*. 1991;5(2):113-9.
40. Brockmann-Bauser M, Bohlender J, Mehta D. Acoustic perturbation measures improve with increasing vocal intensity in individuals with and without voice disorders. *Journal of Voice*. 2018;32(2):162-8.
41. Brockmann M, Storck C, Carding PN, Drinnan MJ. Voice loudness and gender effects on jitter and shimmer in healthy adults. *J Speech Lang Hear Res*. 2008;51(5):1152-60.



**Figure1.** Distribution of the voice quality ratings ( $G_{\text{mean}}$ ) based on auditory –perceptual evaluation



**Figure2.** The ROC curve illustrating the validity of AVQI03-01 for voice screening.

**Table 1: Descriptive data of AVQI and its parameters for all participants**

AVQI parameters	Minimum	Mazimum	Mean (SD)
CPPs	10	18	14.30 (1.29)
HNR	14	26	20.88(2.38)
Shimmer (%)	2	8	4.20(1.29)
Shimmer-dB	0	1	0.45 (0.1)
Slope	-29	-15	-22.09(2.95)
Tilt	-14	-11	-12.63(0.82)
AVQI_03_01	-1	3	1.07(0.77)

**Table 2.** Differences in AVQI scores between participants with low and high VHI scores.

AVQI parameters	<i>Mean(SD)</i>		<i>P-Value</i>
	<i>VHI&lt;19</i>	<i>VHI≥19</i>	
AVQI	0.95 (0.76)	1.33 (0.75)	<b>0.011</b>
CPPs	14.50 (1.26)	13.88 (1.29)	<b>0.014</b>
HNR	20.92 (2.34)	20.77 (2.53)	0.742
Shimmer (%)	4.18 (1.23)	4.27 (1.43)	0.737
Shimmer-dB	0.45 (0.10)	0.46 (0.11)	0.642
Slope	-22.05 (2.96)	-22.19 (2.99)	0.803
Tilt	-12.70 (0.57)	-12.49 (0.70)	0.115