Does the Acoustic Voice Quality Index (AVQI) correlate with perceived creak and strain in normophonic young adult Finnish females?

Anne-Maria Laukkanen, Leena Rantala

Speech and Voice Research Laboratory, Faculty of Social Sciences,

Tampere University, Tampere, Finland

Short title: AVQI vs. creak and strain

Number of tables: 5

Number of figures: 2 (Fig. 1 a and b)

Number of words: 2955

Corresponding author

Anne-Maria Laukkanen, Ph.D.

Speech and Voice Research Laboratory

Tampere University, Tampere, Finland

Address: Åkerlundinkatu 5, 33100 Tampere, Finland

E-Mail: Anne-Maria.Laukkanen@tuni.fi

tel. +358503635152

Key Words: perturbation, CPPS, HNR, spectrum slope

ABSTRACT

Background: The Acoustic Voice Quality Index (AVQI) is a correlate of dysphonia. It has been found to differentiate between dysphonic and normophonic speakers and to indicate the effects of voice therapy. This study investigates how the AVQI reacts towards creak and strain, which are common in normophonic speakers.

Methods: The material was obtained from an earlier study, where Finnish female university students (N 104, mean age 24.3 years, SD 6.3 years) with no known pathology of voice or hearing and a perceptually normal voice (G = 0 in GRBAS) were recorded while reading aloud a standard text and sustaining the vowel [a:]. Perceptual analysis for the amount of creak and strain was carried out by two expert listeners. In this study, the AVQI (03.01) was analyzed and correlated with perceptual evaluations. Samples with low and high amounts of creak and strain were compared with t-tests.

Results: On average, the AVQI was below the threshold value of dysphonia in the Finnish population. The AVQI (rho 0.35, p = 0.000) and its sub-parameters CPPS (rho -0.35, p = 0.000) and HNR (rho -0.30, p = 0.002) showed low but significant correlations with creak. Strain had low but significant correlations with spectral Slope (rho 0.38, p = 0.000) and Tilt (rho -0.40, p = 0.009). The AVQI was lower (better) in samples that were evaluated as having a high amount of strain, but the difference was not significant. Only CPPS differentiated significantly between low and high amounts of creak.

Conclusion: The AVQI does not seem to differentiate between high and low amounts of creak and strain in normophonic speakers.

INTRODUCTION

The Acoustic Voice Quality Index (AVQI) has been developed for indicating acoustic correlates of perceived hoarseness, and for differentiating dysphonia from normophonia [1]. The index consists of six parameters, the smoothed cepstral peak prominence (CPPS, dB), harmonics-to-noise ratio (HNR, dB), dB, shimmer local (SL%), shimmer local dB (SL dB), general slope of the spectrum (Slope, dB; voice energy difference between the levels of 0–1000 Hz and 1000–10000 Hz), and tilt of the regression line through the spectrum (Tilt, dB). To improve ecological validity, the AVQI is calculated for a combination of a short text sample and sustained vowel. Since it includes running speech, it requires validation in each language. Up to now it has been validated in several types of language including Germanic languages (Dutch, German, English) [2–6], Romance languages (French and Brazilian Portuguese) [7, 8], a Baltic language (Lithuanian) [9], a Japonic language (Japanese) [10], and other languages like Korean [11], Finnish [12, 13] and Turkish [14]. In addition to distinguishing between dysphonia and normophonia, the AVQI has also been reported to detect voice changes after voice therapy [2, 10]. It has also been tested for the screening of mild dysphonia [15], and while it was not found to differentiate perceptually mild dysphonia from perceptually normal voices, it was found to improve the strength of screening when combined with perceptual evaluation and self-evaluation (VHI) [15].

Roughness is an important characteristic of dysphonia. On the other hand, a rough sounding voice quality may also be a variant of voice use in healthy speakers. Creaky voice, vocal fry, and laryngealization have been used more or less synonymously in voice science [16] to refer to a rough sounding, often (but not always) low-pitched voice quality. Keating et al. suggest the term 'creaky voice' (or creak) as a general term referring to voices with different acoustic types of creak [17]. The present study follows this terminology. Nowadays creak has become very common in speakers without voice disorders in different countries and cultures [see 18–21]. Women seem to use it more frequently than men [21]. Creak has been detected in the voices of vocally healthy speakers of different ages, and it may occur in all sentence positions [22].

In Finland not only ordinary speakers but also many professional speakers in media use almost continuous creaky voice. Finnish teenagers of both genders have been reported to use a lot of creak [20, 21]. According to [23] some Finnish female teachers had vocal fry in as much as 16−54% of their speech. A recent study investigated Finnish female university students without any diagnosed voice disorders [24]. It was found that 73.2 % of the participants (N=104) used a slight or moderate amount of creaky voice. Furthermore, perceived creak correlated with perceived strain. Strain in turn refers to the perception of excessive effort related to hyperfunctional voice production.[25] The prevalence of slight or moderate strain was 88.4 %. Since creak is characterized by irregularities in the voice signal [17] and an irregular, rough-sounding voice, and strain are characteristics that may be related to dysphonia, the question arises of how the AVQI reacts to these characteristics in the voices of otherwise non-dysphonic speakers. Based on the material of [24], this study aims to answer this question.

MATERIALS AND METHODS

Participants and tasks

The material of our previous study consisted of voices from 104 Finnish female university students (mean age 24.3 years, SD 6.3 years) who had no known pathology of the voice and hearing and who had been pre-evaluated to have a perceptually non-dysphonic voice (G=0 in GRBAS scale) [24]. The participants used more or less creak in their speech and text reading. The judgement of voice normality followed the clinical experience-based interpretation that non-dysphonic, speaking habit-related creak mainly occurs in the phrase and sentence endings, has a more regular quality than dysphonic roughness, and typically does not include turbulence noise. Some research evidence also supports this [26]. Furthermore, the abnormally high amount of audio-perceived effort and noise typically heard in dysphonic strain, see e.g. [27] differentiate it from habitual strain in speech.

The participants had recorded a standard text (the Finnish version of *The Northwind and the Sun*) and sustained a prolonged [a:] in a sound-treated studio using a head-set microphone (AKG C44L) placed at a 6 cm distance from the middle of the speaker's lips at a 45-degree angle, and a PC with an iFocusrite soundcard. A sampling rate of 44.1 kHz and 16-bit depth was used.

Analyses

Acoustic analyses of the text reading and vowel samples were carried out using Praat software (5.3.57) [28] and the AVQI-script (version 03.01; 6. 22). The AVQI was calculated from the first 31 syllables of the text and 3 seconds from the middle, most stable part of the sustained vowel [1]. According to the validation of the AVQI 03.01 in Finnish, the threshold value between dysphonic and normophonic voices is 1.83 [13]. The mean values of the AVQI were as follows: 1.13 (range

-1.45 – 3.66) in normophonic speakers and 2.89 (range -0.98 – 8.86) in dysphonic speakers. The graphical output of the Praat script for the AVQI illustrates the index on a scale from 0 to 10 but as it can be seen from the results of the validation in Finnish [13], the AVQI scores can be even lower than 0.

Perceptual analysis results were obtained from the previous study [24]. There two very experienced voice experts (a speech therapist and a speech trainer) evaluated the text reading and vowel samples together for the amount of (1) creak and (2) strain using a four-point scale (0 = not at all, 1 = small amount, 2 = moderate amount, 3 = a lot). The samples were listened to in a free field from a PC, via a Genelec loudspeaker. Each sample could be listened to as many times as deemed necessary. Creak was used as a general term, like in [17], i.e. including both a low-pitched, tapping type of vocal fry, and a rough, dry rasping quality that can be produced also at higher pitches. Strain was defined as perception of excessive vocal effort (hyperfunction), following the terminology of the Cape -V protocol [25]. The inter-rater reliability of the two listeners was acceptable (Pearson Chi-Square 100.159, p = 0.000 for creak, 69.199, p = 0.000 for strain; Spearman's rho 0.63, p = 0.000 and rho 0.68, p 0.000, respectively). Therefore, the mean values for both creak and strain were calculated to be used for correlation analyses.

Further, the participants were divided into two groups for the analysis: those with the score of 0 or 1 were classified as the *Low group* and those with 2 or 3 points as the *High group*.

Statistical analyses were made using SPSS (version 25). Firstly, the normality of the distribution of the perceptual and acoustic parameters was studied using the One sample Kolmogorov-Smirnov test. Samples evaluated to have a high amount of creak or strain were compared to those with a low amount of these characteristics using the independent samples Student's *t*-test (for normally distributed parameters) or the non-parametric Mann-Whitney *U* test (for parameters that were not normally distributed). Bonferroni correction was used (p-value = 0.05 was divided by the number of comparisons made between the acoustic and perceptual parameters, i.e. 0.05/14 = 0.003; thus differences between parameters were regarded as statistically significant when p = ≤ 0.003). The associations between the perceptual and acoustic parameters were studied with the Spearman's *rho*.

RESULTS

Of the participants, 30.8 % used a moderate or high amount of creak in their voice samples (High group; Table 1). Almost as many participants (29.8 %) had a moderate or high amount of strain (High group; Table 2). Tables 1 and 2 show the mean values for the acoustic parameters in the groups of samples with a Low or High amount of creak or strain. On average, the participants' AVQI values (version 03.01) were below the threshold for dysphonia in the Finnish population (1.83). It should be noted, however, that there is always some overlap in the individual AVQI values between normophonic and dysphonic speakers. This can be seen in [13], where the AVQI ranged from -1.45 to 3.66 in normophonic speakers and from -0.98 to 8.86 in dysphonic speakers. In the present study, the mean AVQI was 1.04 (range from -0.94 to 3.33, SD 0.79). Seventeen out of 104 (16 %) received values higher than 1.83 (range from 1.94 to 3.33), and 13 (12.5 %) received values below 0.

Table 1. Mean values for the acoustic parameters in samples representing a low or high amount of perceived creak. Low = evaluation of 0-1, High = evaluation of 2-3 on a scale 0 = not at all, 1 = small amount, 2 = moderate amount, 3 = a lot.

Table 2. Mean values for the acoustic parameters in samples representing a low or high amount of perceived strain. Low = evaluation of 0-1, High = evaluation of 2-3 on a scale 0 = not at all, 1 = small amount, 2 = moderate amount, 3 = a lot.

Creak seemed to be indicated in CPPS, HNR, and in the AVQI. Table 1 shows that the CPPS and HNR were lower and the AVQI higher in the samples from the High creak group. Table 3 in turn shows that the *t*-testing differentiated the groups with p-values below 0.05 for those parameters. In this material, however, only CPPS remained statistically significant after Bonferroni correction. Strain seemed to be indicated in a larger value for Slope and lower values for Tilt and HNR (Table 4; mean values for the groups in Table 2), but the differences were not significant after Bonferroni correction.

Table 3. Differences between the samples with Low and High amount of perceived creak; independent samples *t*-test (top) and Mann-Whitney *U* test (bottom).

Table 4. Differences between the samples with Low and High amount of perceived strain; independent samples *t*-test (top) and Mann-Whitney *U* test (bottom).

Table 5. Correlations (Spearman’s rho) between perceptual evaluations and acoustic parameters.

Table 5 shows that creak correlated negatively with CPPS (rho = -0.35, p = 0.000) and HNR (rho = -0.30, p = 0.002) and positively with the AVQI (rho = 0.35, p = 0.000). Strain correlated positively with Slope (rho = 0.38, p = 0.000) and negatively with Tilt (rho = -0.40, p = 0.000). These correlations were rather weak but nevertheless significant. Creak and strain were also associated with each other (rho = 0.40, p = 0.000). Relations between the AVQI and the perception of creak and strain are illustrated in the scatter plots of Figure 1.

Figure 1. Scatter plots illustrating the relations between the AVQI and (a) the mean evaluation of creak, and (b) the mean evaluation of strain.

DISCUSSION

The aim of this study was to assess how creak and strain are associated with the Acoustic Voice Quality Index (AVQI) and the parameters constructing it. The participants were perceptually evaluated as normophonic (G = 0), and on average their AVQI was below the threshold of dysphonia in Finnish speakers [13]. There was a low but significant positive correlation between the AVQI and the amount of creak. CPPS and HNR decreased with creak which is to be expected, since creak includes irregularities of the signal [17]. Some participants (12.5 %) had a negative AVQI, possibly related to a higher sound pressure level (SPL) or strain which, in turn, may decrease noise and irregularities in the vocal fold vibration. These hypotheses are supported by the results of [29, 30, 31] showing that higher SPL decreases the perturbation measures of jitter, shimmer, and HNR. Furthermore, it is known that increased psychophysiological stress may decrease perturbation, most likely by increasing laryngeal muscle control [32]. In the present study, strain was related to Slope and Tilt. This is understandable, since hyperfunctional voice production is known to even the spectrum [33-36], thus Slope increases (stronger overtones). Tilt decreases, which in turn reflects the decrease in noise energy in the higher frequency range of the spectrum in strain. The same phenomenon (i.e. the opposite connections between strained voice quality and the two parameters measuring the energy level differences in the spectrum) was found in our study, where we assessed how voices produced with different loudness levels affect the AVQI values [37].

Creak and strain correlated with each other (rho = 0.40, p = 0.000). This may sound surprising if one relates creak only with low-pitched vocal fry and low subglottic pressure [38] and strain with high subglottic pressure and *f0* [39]. However, we have used creak as a general term comprising different types of rough sounding voice qualities since these different types do occur in normophonic speakers [17, 22]. Furthermore, a small level difference between the first and second harmonic in the voice has been seen as the most common measure of creak [40]. This measure, in turn, is a common sign of glottal constriction [41]. Clinicians often relate creaky voice to excessive laryngeal tension [42], and lesions, such as contact granulomas, are typical in speakers with a creaky voice [43]. All these studies support the finding of a relation between creak and strain.

The results are thus logical and in line with earlier studies. As the participants were normophonic, it is understandable that the AVQI and most of its' sub-parameters did not distinguish a low and high amount of creak or strain significantly (after Bonferroni correction). Only CPPS distinguished a high and low amount of creak. This is in line with earlier studies showing that CPPS is a highly robust parameter for detecting noise and irregularity in the signal [44-47], i.e. those characteristics that cause the perception of hoarseness. The finding that the AVQI did not distinguish a low and high amount of creak and strain in this material may be interpreted to support the earlier results that the AVQI is a robust and reliable tool for distinguishing dysphonia from normophonia; thus, it is not on average very vulnerable to false positives [e.g. 1, 3]. However, the results of the present study also support the earlier findings that it is not strictly possible to use the AVQI as the sole diagnostic tool in individual cases either, since there is always some overlap [e.g. 1, 3, 13]. As many as 16 % of the participants in the present study had an AVQI above the threshold. Since one of the limitations in the present study is the fact that no laryngoscopic evaluation was made, it may be speculated that perhaps these participants were not healthy. On the other hand, a high AVQI value may just reflect the human vocal variation in some normophonic people. For instance, speaking softly may give a larger than 'normal' AVQI [37], or relatively large AVQI values may be related to other voice use patterns, such as the excessive habitual use of creaky voice as in this study. Therefore, when using the AVQI for clinical voice analysis, it is important that the sound samples have been recorded at the participant's normal conversational loudness and such characteristics as the frequent voluntary use of creak are avoided.

Defining normophonia is of course an interesting and important question. According to clinical tradition, it is possible to suspect a voice disorder if a person's voice quality, pitch, or loudness largely differs from the voices of individuals of a similar age, gender, cultural background, and geographic location [41], and particularly if the person him/herself suffers from voice-related problems in life. Thus, we can hardly classify creak as a sign of dysphonia, if it occurs in most individuals in a culture and the person him/herself does not suffer as a result of it, particularly if he/she is capable of voluntarily changing from creaky to non-creaky phonation. The sound samples for the present study were obtained from a previous study by the authors [24]. As the participants' voices had received the result G = 0 in GRBAS in pre-evaluation, and as the previous study had failed to find a relation between creak and self-reported vocal symptoms, we regard the participants of this study to be normophonic.

One further limitation in the present study is that there were only two listeners performing the evaluation. However, the listeners had a lot of experience in voice analysis, they represented two different fields of expertise (clinical and non-clinical), and the agreement between their evaluations was reasonably high - either well within the scope or even better than those reported earlier [48].Therefore, the evaluation seems to be reliable. The evaluation concerned each participant's text sample and vowel sample together, as the aim was to relate perception to the results of the AVQI, which measures both text and vowel phonation together. However, based on the listener's remarks, it is possible to say that creak was mainly present in the text, and quite rarely also in the sustained vowel phonation. This is also a fact that seems to support the normophonia of the participants, since dysphonic roughness occurs also in sustained vowels [31, 49].

Because the participants of this study were young females, the results cannot be generalized to the voices of men and speakers of different ages. This should be studied further. As the results of the present study suggest that the AVQI may not be very sensitive to strain (at least in normophonic speakers), a further study comparing the AVQI in normophonic and dysphonic speakers with strained voice quality is warranted.

CONCLUSIONS

The AVQI (03.01) did not differentiate a low and high amount of creak and strain in normophonic speakers. CPPS was the only sub-parameter of the AVQI that distinguished a low and high amount of creak.

ACKNOWLEDGEMENTS

We thank special laboratory technician Tero Ikävalko, MA, for the assistance in the acoustic analyses.

STATEMENT OF ETHICS

Material was obtained from our earlier study [24]. The material was collected as a part of the normal teaching procedure, which does not involve any physical or other risks for the participants. The participants have been volunteers and given their written informed consent for the collected material to be used for research. Material was collected and saved, and the results of the analyses were published respecting the Finnish Data Protection Act (Chapter 5, section 31).

A permission for the study was not applied as the research did not risk the physical integrity of the participants. The Ethical Committee of Research in Humanities in Tampere University does not require a permission to apply for research that does not concern the physical integrity of the participants. This follows the directives of the Finnish National Board on Research Integrity (TENK) concerning the ethical principles of research with human participants (Publications of TENK 3/2019, ISSN 2490-161X; https://tenk.fi/en/advice-and-materials/guidelines-ethical-review-human-sciences).

DISCLOSURE STATEMENT

The authors have no commercial or other interests to declare.

AUTHOR CONTRIBUTIONS

Both authors have contributed in the planning of the study, in organizing the listening tests, running the statistical analyses and writing the manuscript.

FUNDING

Neither of the authors received any funding for the present study.

REFERENCES

1. 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. J Voice 2010; 24: 540–555.

2. 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: 161–174.

3. 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. Eur Arch Otorhinolaryngol 2014; 271: 1609–1619.

4. Barsties B, Maryn Y: The acoustic voice quality index. toward expanded measurement of dysphonia severity in German subjects. HNO 2012; 60: 715–720.

5. Barsties B, Maryn Y: The improvement of internal consistency of the acoustic voice quality index. Am J Otolaryngol 2015; 36: 647–656.

6. Barsties B, Maryn Y: External validation of the acoustic voice quality index version 03.01 with extended representativity. Ann Otol Rhinol Laryngol 2016; 125: 571–583.

7. Pommee T, Maryn Y, Finck C, Morsomme D: The Acoustic Voice Quality Index, version 03.01, in French and the Voice Handicap Index. J Voice 2020; 34(4) pp 646.e1-646.e10, in press.

8. Englert M, Barsties v Latoszek B, Maryn Y, Behlau M: Valdation of the Acoustic Voice Quality Index, version 03.01, to the Brazilian Portuguese Language. J Voice 2019; in press.

9. Uloza V, Petrauskas T, Padervinskis E, Ulozaitė N, Barsties B, Maryn Y: Validation of the acoustic voice quality index in the Lithuanian language. J Voice 2017; 31: 257.e11.

10. 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. J Voice 2017; 31: 260.e9.

11. Maryn Y, Kim H, Kim J: Auditory-perceptual and acoustic methods in measuring dysphonia severity of korean speech. J Voice 2016; 30: 587–594.

12. Kankare E, Barsties v. Latoszek B, Maryn Y, Asikainen M, Rorarius E, et al.: The Acoustic Voice Quality Index Version 02.02 in the Finnish-speaking Population. Logoped Phoniatr Vocol 2019; 30: 1–8.

13. Kankare E, Rantala L, Ikävalko T, Barsties v. Latoszek B, Laukkanen A–M: Validation of AVQI version 03.01 for Finnish speakers [In Finnish; Akustisen äänenlaatuindeksin (AVQI) version 03.01 validointi suomenkielisille puhujille]. Puhe ja kieli 2020; 40(3):165-182.

14. Yesilli-Puzelli G, Tadihan-Özkan E, Maryn YJ: Validation and Test-Retest reliability of Acoustic Voice Quality Index, version 02.06, in the Turkish language. J Voice 2020 ; in Press.

15. Faham M, Laukkanen A–M, Ikävalko T, Rantala L, Geneid A, Holmqvist–Jämsén S, Ruusuvirta K, Pirilä S: Acoustic voice quality index as a potential tool for voice screening. J Voice 2019; in press.

16. Hollien H: On vocal registers. J Phon 1974;2:125-144.

17. Keating P, Garellek M, Kreiman J: Acoustic properties of different kinds of creaky voice. Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland, 2011.

18. Yuasa IP: Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? American Speech 2010; 85:315–337. DOI: 10.1215/00031283-2010-018.

19. Hornibrook J, Ormond T, Maclagan M: Creaky voice or extreme vocal fry in young women. The New Zealand Medical Journal (Online) 2018; 30:131:36–40.

20. Härkönen R: Acoustical and perceptual analysis of the situational effect on 14-year-old's speech. [In Finnish]. MA thesis, Tampere, Finland: University of Tampere, 2016.

21. Ketolainen I, Laakso M, Simberg S: Speaking pitch in 16–17-year-old Finnish teenagers [in Finnish; 16-17-vuotiaiden suomalaisnuorten puheäänen korkeus]. Puhe ja kieli 2017; 37:259–277.

22. Oliveira G, Davidson A, Holczer R, Kaplan S, Paretzky A: A comparison of the use of glottal fry in the spontaneous speech of young and middle-aged American women. J Voice 2016; 30: 684–687

23. Pirilä S, Pirilä P, Ansamaa T, Yliherva A, Sonning S, Rantala L: Relationship between activity noise, voice parameters, and voice symptoms among female teachers. Folia Phoniatr Logop 2017; 69: 94–102.

24. Laukkanen A–M, Rantala L: Relations between creaky voice and vocal symptoms of fatigue. Folia Phoniatr Logop 2020; in press, https://doi.org/10.1159/000506901

25. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier-Kramer J, Hillman RE: Consensus Auditory-Perceptual Evaluation of Voice: Development of a standardized clinical protocol. American Journal of Speech Language Pathology 2009; 18: 124–132.

26. Wolk L, Abdelli-Beruh NB, Slavin D: Habitual use of vocal fry in young adult female speakers. J Voice 2012;26(3): e111-e116.

27. Yamaguchi H, Shrivastav R, Andwers ML: A Comparison of Voice Quality Ratings Made by Japanese and American Listeners Using the GRBAS Scale. Folia Phoniatr Logop 2003;55:147–157. DOI: 10.1159/000070726

28. Boersma P, Weenink D: Praat [computer program]. Doing Phonetics by Computer (Version 5.3.57). Amsterdam, The Netherlands: Institute of Phonetic Sciences, 2013.

29. Orlikoff RB, Kahane JC: Influence of mean sound pressure level on jitter and shimmer measures. J Voice 1981; 5(2): 113-119.

30. 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: 1152–1160.

31. Teixeira JP, Fernandes PO: Jitter, shimmer and HNR classification within gender, tones and vowels in healthy voices. Procedia Technology 2014; 16:228-1237.

32. Van Puyvelde M, Neyt X, McGlone F, Pattyn N: Voice Stress Analysis: A New Framework for Voice and Effort in Human Performance. Front Psychol 2018;20. https://doi.org/10.3389/fpsyg.2018.01994.

33. Frokjaer-Jensen B, Prytz S: Registration of voice quality. Bruel Kjaer Technol Review 1976; 3: 3–17.

34. Hammarberg B, Fritzell B, Gaufin J, Sundberg J, Wedin L: Perceptual and acoustic correlates of abnormal voice qualities. Acta Oto-Laryngol 1980; 90: 441–451.

35. Kitzing P: LTAS criteria pertinent to the measurement of voice quality. J Phonetics 1986; 14: 477–482.

36. Narasimhan SV, Vishal K: Spectral measures of hoarseness in persons with hyperfunctional voice disorder. J Voice 2017; 31: 57–61. http://dx.doi.org/10.1016/j.jvoice.2016.03.005

37. Murry T, Brown WS (Jr): Subglottal air pressure during two types of vocal activity: vocal fry and modal phonation. Folia Phoniatr 1971;23:440-449.

38. Hirano M, Hibi S, Terasawa R, Fujiu M: Relationship between aerodynamic, vibratory acoustic and psychoacoustic correlates in dysphonia. J Phonetics 1986;14:445–456.

39. Gordon M, Ladefoged P: Phonation types: A cross-linguistic overview. J Phonetics 2001;29: 383-406.

40. Kreiman J, Shue Y-L, Chen G, Iseli M, Gerratt BR, Neubauer J, Alwan A: Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. J Acoust Soc Am 2012; 132:2625-2632.

41. Stemple JC, Glaze KB, & Klaben BG: Pathologies of the laryngeal mechanism. In: JC Stemple, N Roy, and BK Klaben (eds.) *Clinical Voice pathology: Theory and Management*, San Diego: Plural publishing, 2018, 129-30.Electronic version.

42. Ylitalo Y, Hammarberg B: Voice characteristics, effects of voice therapy and long-term follow-up of contact granuloma patients. J Voice 2000;14(4):557-566.

43. Hillenbrand J, Houde RA: Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. J Speech Lang Hear Res 1996; 39: 311–321.

44. Heman-Ackah YD, Michael DD, Goding GS: The relationship between cepstral peak prominence and selected parameters of dysphonia. J Voice 2002; 16: 20–27. https://doi.org/10.1016/S0892-1997(02)00067-X

45. 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 phonet 2010; 24: 742–758. doi:10.3109/02699206.2010.492446

46. Maryn Y, Weenink D: Objective dysphonia measures in the program praat: Smoothed cepstral peak prominence and acoustic voice quality index. J Voice 2015; 29: 35–43.

47. Laukkanen A–M, Ikävalko T, Rantala L, Kankare E: The Acoustic Voice Quality Index (AVQI) in the assessment of voice quality: A preliminary multiple-case study on the effects of recording level, room noise and phonation type [In Finnish; Akustinen äänenlaatuindeksi (AVQI) äänen arvioinnissa: alustava monitapaustutkimus äänitystason, -tilan ja äänentuottotavan vaikutuksista]. Puhe ja kieli 2020; 40(3):143-160.

48. Zraick RI, Kempster GB, Connor NP etc.: Establishing validity of the concensus auditory-perceptual evaluation of voice (CAPE-V). American Journal of Speech-Language Pathology 2011; 20:14–22.

49. Schoentgen J: Jitter in sustained vowels and isolated sentences produced by dysphonic speakers. Speech Communication 1989;8(1):61-79.

LEGEND FOR THE FIGURE

Figure 1. Scatter plots illustrating the relations between the AVQI and (a) the mean evaluation of creak, and (b) the mean evaluation of strain.

Table 1. Mean values for the acoustic parameters in samples representing low or high amount of perceived creakiness. Low = evaluation from 0-1, High = evaluation from 2-3 on a scale 0 = not at all, 1 = small amount, 2 = moderate amount, 3 = a lot.

	Creakiness	N	Mean	SD	SE
CPPS (dB)	Low	72	14.63	1.22	0.14
	High	32	13.8	1.4	0.25
HNR (dB)	Low	72	21.29	2.48	0.29
	High	32	20.22	1.76	0.31
SHIM (%)	Low	72	4.16	1.33	0.16
	High	32	4.29	1.12	0.2
SHIM (dB)	Low	72	0.45	0.12	0.01
	High	32	0.45	0.1	0.02
SLOPE (dB)	Low	72	-21.99	2.89	0.34
	High	32	-22.1	3.07	0.54
TILT (dB)	Low	72	-12.62	0.68	0.08
	High	32	-12.64	0.46	0.08
AVQI	Low	72	0.91	0.76	0.09
	High	32	1.33	0.79	0.14

Table 2. Mean values for the acoustic parameters in samples representing low or high amount of perceived strain. Low = evaluation from 0-1, High = evaluation from 2-3 on a scale 0 = not at all, 1 = small amount, 2 = moderate amount, 3 = a lot.

	Strain	N	Mean	SD	SE
CPPS (dB)	Low	73	14.31	1.34	0.16
	High	31	14.52	1.31	0.24
HNR (dB)	Low	73	21.14	2.52	0.3
	High	31	20.54	1.75	0.32
SHIM (%)	Low	73	4.18	1.4	0.16
	High	31	4.24	0.87	0.16
SHIM (dB)	Low	73	0.45	0.12	0.01
	High	31	0.45	0.07	0.01
SLOPE (dB)	Low	73	-22.46	2.99	0.35
	High	31	-21	2.56	0.46
TILT (dB)	Low	73	-12.51	0.61	0.07
	High	31	-12.9	0.55	0.1
AVQI	Low	73	1.09	0.83	0.1
	High	31	0.91	0.69	0.12

Table 3. Differences between the samples with Low and High amount of perceived creakiness; independent samples *t*-test (top) and Mann-Whitney *U* test (bottom)

Parameter	Levene's Test for Equality of Variances		*t*-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	SE Difference	Lower	Upper
CPPS	0.143	0.706	3.074	102	0.003	0.83410	0.27132	0.29593	1.37226
SLOPE	0.122	0.728	0.173	102	0.863	0.10816	0.62576	-1.13302	1.34934
AVQI	0.034	0.853	-2.593	102	0.011	-0.42289	0.16307	-0.74634	-0.09944

Parameter	*U*	Sig.
HNR	749.000	0.005
SHIM (%)	1017.500	0.343
SHIM (dB)	1009.500	0.315
TILT	1117.500	0.808

Table 4. Differences between the samples with Low and High amount of perceived strain; independent samples *t*-test (top) and Mann-Whitney *U* test (bottom)

Parameter	Levene's Test for Equality of Variances		*t*-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	SE Difference	Lower	Upper
CPPS	0,019	0,889	-0,727	102	0,469	-0,20753	0,28543	-0,77369	0,35862
SLOPE	0,506	0,478	-2,371	102	0,020	-1,45759	0,61478	-2,67700	-0,23818
AVQI	0,986	0,323	1,093	102	0,277	0,18453	0,16890	-0,15047	0,51954

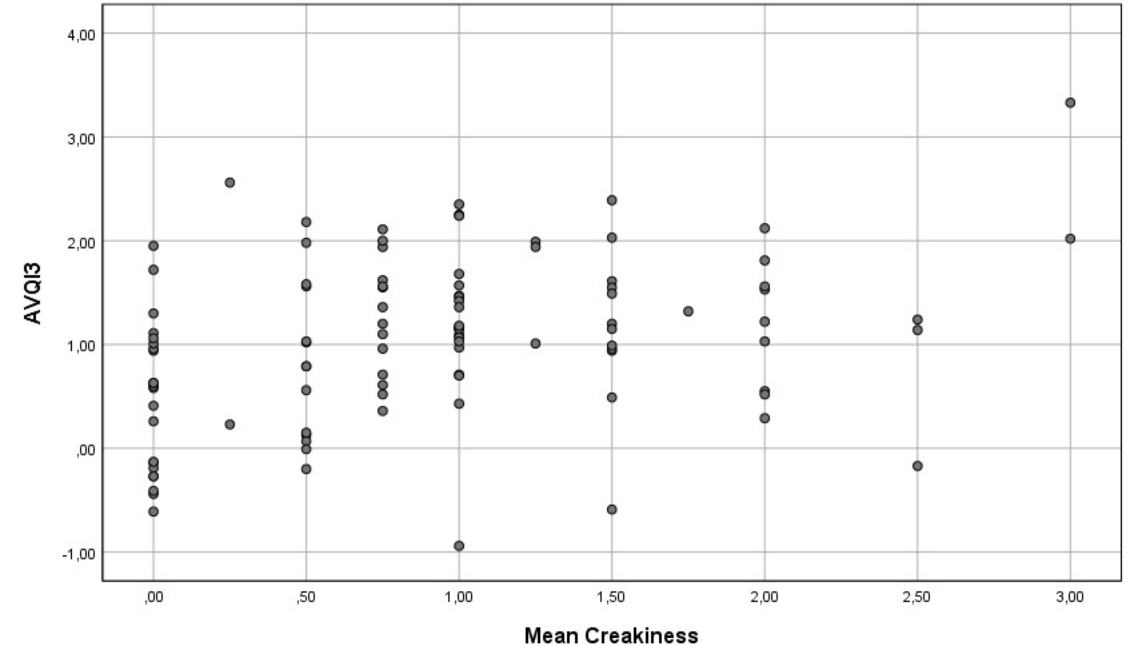
Parameter	*U*	Sig.
HNR	856.000	0.05
SHIM (%)	970.000	0.251
SHIM (dB)	949,500	0.195
TILT	731.500	0.004

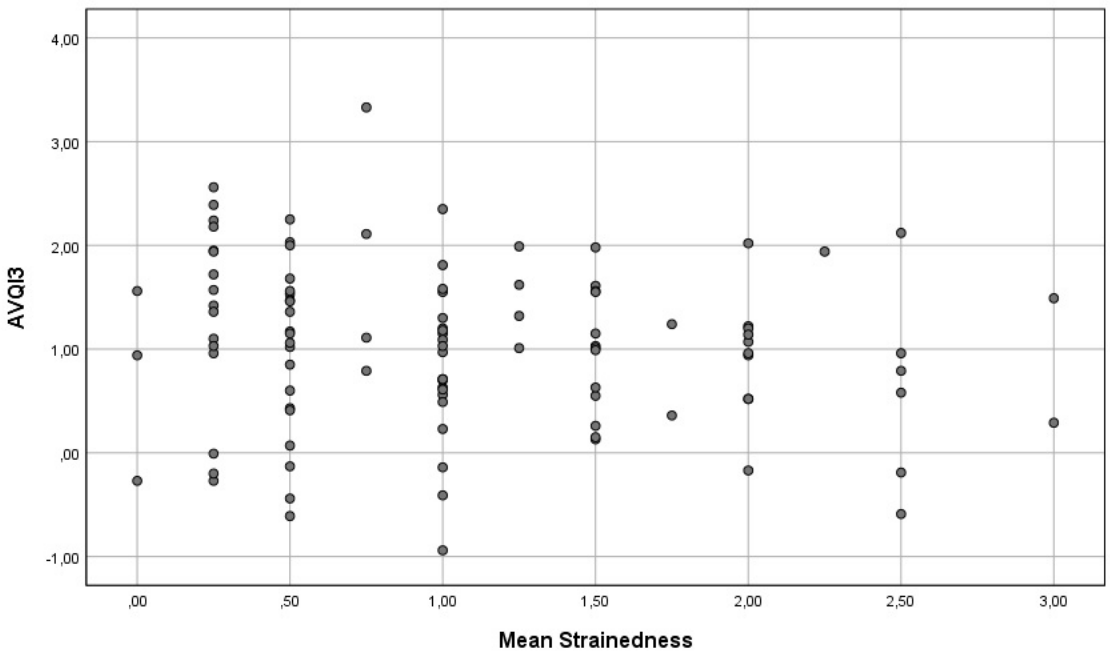
Table 5. Correlations (Spearman's rho) between perceptual evaluations and acoustic parameters.

		Creakiness	Strain	CPPS	HNR	SHIM %	SHIM dB	SLOPE	TILT	AVQI
Creakiness	Correl. Coeff.	1.000	.404**	-.348**	-.305**	0.131	0.144	-0.111	0.133	.351**
	Sig. (2-tailed)		0.000	0.000	0.002	0.186	0.146	0.264	0.179	0.000
	N	104	104	104	104	104	104	104	104	104
Strain	Correl. Coeff.	.404**	1.000	0.151	-0.143	0.046	0.090	.380**	-.398**	-0.162
	Sig. (2-tailed)	0.000		0.125	0.147	0.646	0.363	0.000	0.000	0.101
	N	104	104	104	104	104	104	104	104	104

** Significance at 0.01 level (2-tailed)

*Significance at 0.05 level (2-tailed)


4,00
3,00
2,00
1,00
,00
-1,00
AVQI3
,00
,50
1,00
1,50
2,00
2,50
3,00
Mean Creakiness


AVQI3
4,00
3,00
2,00
1,00
,00
-1,00
,00
,50
1,00
1,50
2,00
2,50
3,00
Mean Strainedness