The Exploration of an Objective Model for Roughness With Several Acoustic Markers

*,†Ben Barsties v. Latoszek, *,‡,§Marc De Bodt, ¶,**,††Ellen Gerrits, and *,‡‡,§§Youri Maryn, *‡‡‡Antwerp, §§§Ghent, Belgium, †Nijmegen, ¶**††Utrecht, The Netherlands

Summary: Objective. In voice assessment, the evaluation of voice quality is a major component in which roughness has received wide acceptance as a major subtype of abnormal voice quality. The aim of the present study was to develop a new multivariate acoustic model for the evaluation of roughness.

Method. In total, 970 participants with dysphonia and 88 participants with normal voice were included. Concatenated voice samples of continuous speech and sustained vowel [a:] were perceptually judged on roughness severity. Acoustic analyses were conducted on the voiced segments of the continuous speech sample plus sustained vowel as well. A stepwise multiple linear regression analysis was applied to construct an acoustic model of the best acoustic predictors. Concurrent validity, diagnostic accuracy, and cross-validation were verified on the basis of Spearman correlation coefficient (r_s), several estimates of the receiver operating characteristics plus the likelihood ratio, and iterated internal cross-correlations.

Results. Six experts were included for perceptual analysis based on acceptable rater reliability. Stepwise multiple regression analysis yielded a 12-variable acoustic model. A marked correlation was identified between the model and the perceptual judgment ($r_s = 0.731$, $P = 0.000$). The cross-correlations confirmed a high comparable degree of association. However, the receiver operating characteristics and likelihood ratio results showed the best diagnostic outcome at a threshold of 2.92, with a sensitivity of 51.9% and a specificity of 94.9%.

Conclusions. Currently, the newly developed roughness model is not recommended for clinical practice. Further research is needed to detect the acoustic complexity of roughness (eg, multiplophonia, irregularity, chaotic structure, glottal fry, etc).

Key Words: Voice assessment–Voice quality–Roughness–Acoustic measurement–Auditory-perceptual judgment.

INTRODUCTION

The evaluation of voice quality in the acoustic voice signal is a main component of clinical assessment and research. Roughness is one of the major subtypes that refer to abnormal overall voice quality. Some vocal pathologies are dominantly characterized in roughness like laryngeal polyp,[1] polypoid degeneration,[1] Reinke edema,[2] and ventricular dysphonia.[2] The voice sound of rough voices is produced by irregular vocal fold vibrations caused by various laryngeal pathologies characterized by frequent, rapid, and random changes of the regular movement patterns of the vocal fold.[3] A rough voice sound is a low-frequency aperiodic noise with a chaotic waveform and clearly defined subharmonics. These subharmonics have the intensities nearing the strength of fundamental frequency.[3] The concept of roughness is auditory-perceptually based. This is a behavioral response to a stimulus or stimulus acoustic features in the voice sound. Voice clinicians generally use standardized and quantified judgment scales like the grade, roughness, breathiness, asthenia, strain (GRBAS) scale[4] or the Consensus Auditory-Perceptual Evaluation of Voice protocol.[5] The evaluation of the voice sound is subjectively based, and it induces notable intra-rater and inter-rater variability by the listeners. There are many factors that influence the reliability and accuracy of the perceptual evaluation.[6] Despite all these limitations, perceptual evaluation remains the candidate for gold-standard assessment.[6,7] First, voice quality is a perceptual phenomenon that is related to the response of the brain to specific acoustic features in the voice sound. These acoustic features are presumably related to periodicity prominence.[8] Second, the judgment scales are a simple and efficient method in daily clinical practice to document the presence, degree, and progression of any type of abnormal voice quality.[8]

In the past, several attempts were undertaken to develop various kinds of instrumental methods to improve the validity and reliability in abnormal voice quality judgment. For several reasons, acoustic measurements have received attention as a useful tool in the voice quality assessment. Research indicates that acoustic measurements are the most frequently used diagnostic instruments in the evaluation of voice disorders.[9] The technology used in acoustic measurement is noninvasive.[10] Finally, acoustic measurements are easy to use[10] and have relatively low costs.[10,11] All in all, acoustic measurements might have the potential to offer an objective adjunct to existing perceptual assessments. On the one hand, acoustic measurements are traditionally applied on sustained vowels and less in continuous speech[6]; however, continuous speech approaches more everyday conversation. The consideration of only one speech task reduces the ecological validity in the judgment of voice quality.[12] Furthermore, several studies showed poor reliability or poor documentation of improvement in voice quality using single acoustic

Accepted for publication April 19, 2017.

Conflict of interest: This article has no current or potential conflict of interest.

From the *Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium; †Institute of Health Studies, HAN University of Applied Sciences, Nijmegen, The Netherlands; ‡Department of Otorhinolaryngology and Head & Neck Surgery, University Hospital, Antwerp, Belgium; §Faculty of Medicine & Health Sciences, University of Ghent, Ghent, Belgium; ¶Faculty of Health Care, HU University of Applied Sciences Utrecht, The Netherlands; **Faculty of Humanities, University of Utrecht, Utrecht, The Netherlands; ††Department of Otolaryngology, University Medical Center Utrecht, Utrecht, The Netherlands; ‡‡European Institute for ORL, Sint-Augustinus Hospital, Antwerp, Belgium; and the §§Faculty of Education, Health & Social Work, University College Ghent, Ghent, Belgium.

Address correspondence and reprint requests to Ben Barsties v. Latoszek, Universiteitsplein 1, 2610 WILRIJK, Antwerp, Belgium. E-mail: ben.barsties@t-online.de

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

0892-1997

http://dx.doi.org/10.1016/j.jvoice.2017.04.017

measures.[13,14] Additionally, many acoustic single parameters revealed low or poor concurrent validity.[3,15] A recent meta-analysis showed only a moderate to strong correlation between auditory-perceptual judgment of roughness for a few acoustic markers (ie, 13 out of a possible 87 acoustic measures).[3] The most promising evidence of moderate to strong correlation coefficients between perceptual ratings of roughness and acoustic measures was found in different frequency bands of the spectral noise level (SNL),[16,17] the second amplitude of a spectrum (H2A),[18] jitter factor,[19] glottal-to-noise excitation ratio (GNE),[20] amplitude variability index,[21] and smoothed cepstral peak prominence (CPPs).[22,23] Generally, these six measures revealed the highest effect size to objectively evaluate roughness for a larger group of voice samples.[3] A greater validity can be expected using multivariate models with several acoustic markers.[6] For the evaluation of roughness, Table 1 shows several attempts to create a multivariate model as well.[24–26] Relevant methodological features, concurrent validity, and classification accuracy between the acoustic measurement and the auditory-perceptual judgment of roughness are listed there. The validity outcomes demonstrated that these multivariate acoustic models reveal moderate results in accuracy (59%–64%) and concurrent validity (r = 0.77–0.80). However, all studies include small numbers of subjects and considered only one speech task. A newly developed acoustic model for roughness must include both sustained vowels and continuous speech to be considered ecologically valid. A recent breakthrough in acoustic measurements was achieved in two multivariate acoustic models for overall voice quality, which included both speech tasks. They are called the Acoustic Voice Quality Index (AVQI)[27] and the Cepstral Spectral Index of Dysphonia.[28,29] Both models showed sufficient results in accuracy and reliability in detecting abnormal voice quality. A literature review showed that these two acoustic models are valid and robust tools in the evaluation of overall voice quality.[6]

The aim of the present study was to explore an acoustic multivariate model for the quantification of roughness, with the auditory-perceptual judgment of roughness as criterion. This study considered the following criteria of concatenated voice samples (ie, by following the example of the AVQI): a large number of voice samples, a rater panel with homogeneous and representative reliability, and various validity investigations.

METHOD

Participants

All participants were selected on an informed consent basis at the otolaryngology caseload of the Sint-Jan General Hospital in Bruges, Belgium. They were seen for an interdisciplinary otolaryngology and speech-language pathology assessment in the period from October 2002 to February 2014. In total, 970 patients with dysphonia and 88 vocally healthy participants were included in the study. The dysphonia group presented various organic and nonorganic etiologies and various degrees in dysphonia severity. An Olympus ENF-V flexible transnasal chip-on-tip laryngostroboscope (Olympus Medical Systems Europa GmbH, Hamburg, Germany) was used to assess the dysphonia diagnosis. Further details of the dysphonia group are presented

TABLE 1.
Multivariate Acoustic Models in the Evaluation of Roughness Summarized in Their Methodology Features and Outcome

Source	Number of Subjects	Speech Task	Multivariate Statistical Method	Included Objective Measures in the Model	Scale of the Auditory-Perceptual Judgment of Roughness	Outcome: Concurrent Validity	Outcome: Classification Accuracy
Hammarberg et al[24]	17	A short story of 92 words	Stepwise multiple regression analysis	Spectral level from the frequency band 2–5 kHz to the frequency band 5–8 kHz, mean fundamental frequency	EAI: 5 points	Multiple correlation: r = 0.65	42%
Hammarberg et al[25]	16	A short story of 92 words	Stepwise multiple regression analysis	Waveform perturbation, fundamental frequency	EAI: 5 points	Multiple correlation: r = 0.80	64%
Eskenazi et al[26]	73	Sustained vowel /i/	Prediction sum of squares owing to regression analysis	Spectral flatness of the residue signal, harmonics-to-noise ratio	EAI: 7 points	Correlation: r = 0.77	59%

Abbreviation: EAI, equal-appearing interval.