Auditory-perceptual Evaluation of Normal and Dysphonic Voices Using the Voice Deviation Scale

***Rosiane Yamasaki, †Glaucya Madazio, ‡Sylvia H.S. Leão, §Marina Padovani, ||Renata Azevedo, and ¶Mara Behlau,** *†||¶*São Paulo and* ‡*Auckland, Brazil, and* §*Brazil*

Summary: Objective. This study aims to (1) determine the cutoff values of the overall severity (OS) of vocal deviation on the visual analog scale (VAS) based on the ratings of the numerical scale (NS); and (2) verify the power of discrimination of these cutoff values according to different degrees of vocal deviation.

Study Design. This is a prospective study.

Methods. The auditory-perceptual evaluation was performed by four speech-language pathologists who used two protocols with different scalar properties: the VAS and the 4-point NS. Vocal samples from142 women and 69 men, plus 10% of repetition, with and without vocal complaints, ranging from 19 to 60 years were included. The analyzed speaking task was the counting from 1 to 10. For both protocols, the judges rated the OS.

Results. Based on the correspondence between the two scales, the cutoff values of the OS on the VAS obtained were 35.5, 50.5, and 90.5 points. The 35.5 value corresponds to the cutoff point between normal variability and mild/moderate vocal deviations; the 50.5 value corresponds to the cutoff point between mild/moderate and moderate vocal deviations; the 90.5 value corresponds to the cutoff point between moderate and severe deviations. Areas under the receiver operating characteristic curve for the three cutoff values were 0.918, 0.948, and 0.970, respectively.

Conclusions. The cutoff values of the OS on the VAS were obtained, and the areas under the ROC curve showed that all the three cutoff points had excellent accuracy that represents a higher power of discrimination of different degrees of vocal deviations.

Key Words: Auditory-perceptual evaluation–Voice assessment–Rating scales–Vocal screening–Dysphonia.

INTRODUCTION

The auditory-perceptual evaluation is the traditional evaluation method in voice clinic, and it is still considered as the "gold standard" for documenting the severity of voice impairment.[1–3] This analysis allows the characterization of voice quality and the quantification of voice deviation.[3] Being essentially a perceptual task,[4,5] the reliability of this evaluation depends on the training, type of rating scale, task design, type of stimulus, and the listener's attention and experience.[1,2,4–6] Thus, controlling for these main interference factors decreases subjectivity and increases the internal validity of the evaluation.

Currently, auditory-perceptual evaluations have been performed with structured scales and protocols. Rating scales have been used to measure a variety of psychometric phenomena, including the perception of voice quality.[4] Voice quality is multidimensional, and the differences on how listeners focus their attention on different aspects of multidimensional perceptual qualities are apparently a significant source of inter-rater unreliability.[7] Therefore, it is important to use vocal parameters that present high intra-rater and inter-rater reliability, such as the overall severity (OS) of voice deviation.[8–10] This parameter has been rated by the use of different perceptual scales such as the 4-point numerical scale (NS) (0 = absence of disorder; 1 = mild; 2 = moderate; 3 = severe) and a visual analog scale (VAS) of 100 mm. Both scales have been widely used in voice assessment. However, the VAS seems more sensitive to small differences of voice quality deviations than the NS,[11] and therefore, it is an appropriate instrument for screening purposes. It is also critical to know the boundaries between normal and disordered voices. Simberg et al[12] used the score of 34 mm on the G variable (GRADE), from GRBAS scale, as a limit to separate normal variation of voice quality (NVVQ) and abnormal voice quality. Due to the interesting clinical applications of this study, we are conducting it in Brazil.[13] The auditory-perceptual evaluation was performed by three judges with high reliability, and the cutoff value of 34.5 mm was indicated as differentiating normal and abnormal vocal qualities. The result was very close to the Finish study,[12] demonstrating that this form of analysis is robust, and it was not influenced by cultural aspects. We used this cutoff rating in a new study to determine other degrees of vocal deviation, as well as their correspondent distribution range on the VAS.

Thus, the purposes of this study were (1) to determine the cutoff values of the OS of vocal deviation on the VAS based on the ratings of the NS; and (2) to verify the power of discrimination of these cutoff values according to different degrees of vocal deviation.

METHOD

Voice samples

For this research, 250 voice samples were selected from a voice center database. However, 39 voice samples were excluded

Accepted for publication January 5, 2016.

From the *Centro de Estudos da Voz (CEV), Department of Otolaryngology, School of Medicine, Universidade de São Paulo (USP), São Paulo, Brazil; †Centro de Estudos da Voz (CEV), São Paulo, Brazil; ‡Speech Science, School of Psychology and Centre for Brain Research, The University of Auckland, Auckland; §Centro de Estudos da Voz (CEV), Instructor Professor of Speech Language Pathology and Audiology Course of Faculdade de Ciências Médicas da Santa Casa de São Paulo, Assistant Professor of Speech Language Pathology and Audiology Course of UniFMU Rede Laureate; ||Department of Speech-language Pathology and Audiology, Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil; and the ¶Centro de Estudos da Voz (CEV), Department of Speech-language Pathology and Audiology, Universidade Federal de São Paulo (UNIFESF), São Paulo, Brazil

Address correspondence and reprint requests to Rosiane Yamasaki, Centro de Estudos da Voz (CEV), Rua Machado Bittencourt, 361, 10th floor, CEP 04044-905, São Paulo, Brazil. E-mail: r.yamasaki@uol.com.br

Journal of Voice, Vol. 31, No. 1, pp. 67–71
0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.01.004

because of poor recording quality. A total of 211 voice recordings were analyzed. Vocal samples from 142 women and 69 men, with and without vocal complaints and ranging in age from 19 to 60 years, were included in the study. Voice samples of individuals without vocal complaint were sourced from vocal screening for telemarketing services. All voice samples were recorded in similar conditions: quite environment, voices recorded directly on a desktop computer with an off board Creative Sound Blaster Audigy 7.1 soundboards, *Sonic Foundry Sound Forge 4.5* program, unidirectional microphone with flat frequency response, located at 45° and 5-cm distance from the speaker's mouth. The sound entry on *Sound Forge* was approximately −6 dB, and it was regulated on the Windows (Windows XP) volume control.

The analyzed speaking task was the counting from 1 to 10, in an auto-selected frequency and intensity, comfortable for the speaker.

Protocols and perceptual evaluation

For the auditory-perceptual evaluation, the following protocols were used: VAS, formed by an undifferentiated 100-mm horizontal line, in which the extreme left side represented the absence of vocal deviation and the extreme right side represented the maximum level of disorder; and a 4-point NS, 0 (zero) being the absence of vocal deviation, 1 being mild deviation, 2 being moderate, and 3 being severe deviation. For both protocols, the judges rated the OS, representing the global impression of voice deviation.

The perceptual evaluation was performed by four speech-language pathologists who had undergone two previous trainings sessions 1 month before the main session. The training stimuli consisted of dysphonic and normal voices and four nominal anchor voice samples—absence of any, mild, moderate, and severe vocal deviation. During training, judges were asked to rate the OS of vocal deviation of 100 voice samples. In the first training session, the judges used the NS, and in the second one, the VAS was used. Each training session lasted 1 hour.

The four speech-language pathologists, who had more than 15 years of clinical experience in voice assessment, performed the auditory-perceptual evaluation in two different sessions. The VAS was used during the first assessment, and during the second assessment, the same judges applied the NS. The scale used in each session was randomly chosen.

The list of sonorous stimuli was the same. However, two different versions were presented because of the two sessions, with random samples and 10% of the samples were played twice to verify intra-rater reliability. The perceptual analysis was accomplish individually with bilateral earphone, Plantronics model Audio 90, on the same computer that recorded the voices. To be sure of their answers, the judges could repeat the stimulus whenever necessary.

Statistical analysis

Calculation of the gradual values on VAS

The cutoff values on VAS were obtained according to the correspondence between the two scales and through the application of ROC curve analysis. The estimation of each cutoff point based on the values of sensitivity, specificity, and efficiency.

The accuracy of the cutoff values when separating different degrees of vocal deviations was determined by the area under the ROC curve. Accuracy level can range from excellent to poor, being that: 0.90–1 represent an excellent accuracy, 0.80–0.90 a good accuracy, 0.70–0.80 a fair, and 0.60–0.70 represent a poor accuracy. A value of poor and a 0.50–0.60 represented a failure.

Reliability

The level of inter-rater agreement was verified using the application of the Coefficient Analysis of the Intraclass Correlation. The intra-rater reliability was tested through 10% of stimuli repetition.

RESULTS

Based from the judges' results, the Coefficient Analysis of the Intra-class Correlation showed that the four judges presented a high level of agreement for both scales with statistical significance. However, the variable VAS presented a coefficient rate slightly higher (0.849) than the variable NS (0.821), indicating that the agreement was even higher on the VAS (Table 1). The intra-rater reliability was superior to 75% for all judges.

Estimation of cutoff values on the VAS

The VAS cutoff values were obtained by correlation between two scales, VAS and NS, and by ROC curve analysis. Based on sensitivity, specificity and efficiency values, the cutoff values obtained were 35.5, 50.5 and 90.5 points (Table 2). The 35.5 value corresponds to the cutoff point between normal variability and mild/moderate vocal deviation; the 50.5 value corresponds to the cutoff point between mild/moderate and moderate vocal deviation; the 90.5 value corresponds to the cutoff point between moderate and severe deviations (Figure 1).

The three cutoff values generated four distribution bands on VAS. The correspondences between VAS values and NS values are presented in Table 3.

TABLE 1.
Values of the Correlation Coefficient and of the Confidence Interval of the VAS and NS

		Confidence Interval		
Variables	Coefficient	Inferior Limit	Superior Limit	Significance (*P*)
VAS	0.849	0.818	0.877	<0.001*
NS	0.821	0.785	0.853	<0.001*

Abbreviation: VAS: Visual analog scale; NS: Numerical scale.