Discriminant Capacity of Acoustic, Perceptual, and Vocal Self: The Effects of Vocal Demands

*Ana Cristina Côrtes Gama, †Zuleica Camargo, ‡Marco Aurélio Rocha Santos, and §Luiz Carlos Rusilo,
*‡*Belo Horizonte, Minas Gerais and* †§*São Paulo, Brazil*

Summary: Objectives. To analyze the discrimination ability of acoustic, auditory parameters, and perception of vocal effort during professional and social voice use, and the correlations of these parameters with the vocal demands.

Study Design. Longitudinal study.

Methods. Seventy-three subjects participated in the study: 31 females aged from 28 to 65 years (G1; professional voice users) and 42 females aged from 31 to 59 years (G2, social voice users;). All the subjects were subjected to acoustic voice analysis including F0 median, semiamplitude interquartile, quantile 99.5%, and skewness; first F0 derivate mean, standard deviation (SD), and skewness; intensity skewness; spectral slope mean, SD, and skewness; long-term average spectrum–frequency SD, perceptual parameters (GRBASI scale), and self-perception of vocal effort, before and after 2 hours and 30 minutes of voice use. Statistical analyses were completed via multivariate discriminant analysis and canonical correlation analysis.

Results. Discriminant analysis of acoustic, perceptual, and self-rating variables and analysis of the grouped parameters did not differentiate the samples before and after vocal use. Higher levels of canonical correlation were found for the professional voice group after voice use, with a correlation between perceptual analysis and acoustic measures.

Conclusions. The current measures could not discriminate the differences of the type of vocal demands, professional or social.

Key Words: Voice–Speech acoustics–Dysphonia–Faculty–Speech–Language and hearing sciences.

INTRODUCTION

Teachers are often referred to as vocal athletes because of the significant vocal demands placed on them in the workplace.[1,2] The American College of Sports Medicine guidelines established clinical protocols for physical training and conditioning.[3] The guidelines outline the fundamental tenet of training to include progressive overload, or a gradual increase in the physical stresses associated with physical exercise.[3] The correlation between vocal loading and vocal fatigue is not well established. Extended periods of vocal fold vibration and collision of vocal fold tissues during continuous speech and more so during increased sound pressure levels in teachers,[4] may lead to vocal fatigue. Titze suggests that intense voice use leads to vocal fatigue and subsequent changes in blood circulation, tension, viscosity, and the composition of fluids, with deleterious effect on phonatory efficiency.[5]

The literature defines vocal fatigue as a negative adaptation as a consequence of prolonged vocal use,[6] and this negative adaptation can be observed in auditory-perceptual, acoustic, aerodynamic, or physiological evaluation, indicating an undesirable effect on functional phonatory physiology.[7] Recent literature suggests that after vocal loading, the most sensitive parameters indicative of hyperfunctional voice production are acoustic signals related to fundamental frequency (F0),[8–10] laryngeal tremor,[11,12] glottic chink on videolaryngoscopic evaluation,[13,14] fatigue signals in electromyography (EMG),[12] and changes in aerodynamic parameters.[9,15] Most previous research focused on the acoustic characteristics (F0, vocal tremor) and aerodynamic voice analysis of sustained vowels[8–10] in self-perception of the voice[16] or in physiological and laryngeal characteristics.[12–14] Less attention was directed to the study of long-term vocal properties (perceptive and acoustic) of prolonged speech. These measures, when extracted from continuous speech, showed results more compatible with the habitual individual vocalizations compared with perceptual and acoustic measures extracted from sustained vowels.[17] Moreover, few studies analyzed vocal and speech modifications in specific groups of professionals,[8] such as teachers in ergonomic work situations of professional vocal use.[18] The effects of professional and social vocal use in women submitted to differentiate levels of vocal demand evaluated by acoustic and perceptual analysis extracted from continuous speech and self-perception of vocal effort are still less known. The aim of this study was to analyze the discrimination ability of acoustic, auditory parameters and perception of vocal effort of the effects of professional and social use of the voice during an extended period of voice use, and the correlations of these parameters with the vocal demands.

METHODS

Voice samples

This longitudinal study was approved by the committee of ethics in research of the institution (ETIC 0531/2011). Seventy-three subjects participated in the study: 31 females aged from 28 to 65 years (mode = 59); all were teachers of the municipal educational system at Belo Horizonte and were classified as professional voice users (G1). Forty-two females aged from 31 to 59 years (mode = 42) were included as

Accepted for publication June 16, 2014.
From the *Department of Speech-language Pathology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; †Integrated Laboratory on Acoustic Analysis and Cognition, Postgraduate Studies Program in Applied Linguistics and Language Studies of the Pontifical Catholic University of São Paulo, São Paulo, Brazil; ‡Otorhinolaryngology Service of Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; and the §Department of Quantitative Methods, Faculty of Economics and Administration of the Pontifical Catholic University of São Paulo, São Paulo, Brazil.
Address correspondence and reprint requests to Ana Cristina Côrtes Gama, Federal University of Minas Gerais, Av. Alfredo Balena, 190/249, Belo Horizon, CEP 30130-100, Brazil. E-mail: anacgama@medicina.ufmg.br
Journal of Voice, Vol. 29, No. 2, pp. 260.e45-260.e50
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2014.06.012

nonprofessional voice users (G2). The professional activities of G2 included government employees (11), housewives (6), traders (7), managers (6), nurses (3), hairdressers (3), engineers (2), a nutritionist (1), student (1), nun (1), and pedagogue (1).

Exclusion criterion in G1 was speech therapy during the research; the onset of vocal complaints in G2 was considered exclusion criteria. Professional voice use in G1 involved continuous voice use for 2 hours and 30 minutes. In G2, social vocal activity or voice use at home was determined during the same time interval.

The number of subjects was defined by a sample size calculation using the one-way ANOVA test with the power of 90%, considering the higher standard deviation (SD) and a maximal difference between the groups to be one SD of the acoustic measures. *Minitab Release 14.1* (Minitab, Inc, State College, PA) was used for these calculations.

Subjects were accrued from four schools within Belo Horizonte; a convenient sample of female teachers who agreed to take part in the study was obtained. Subjects in G2 were recruited from different places according to the inclusion and exclusion criteria.

Voice recordings of G1 subjects were performed in a quiet room at the schools where the teachers worked in the morning to ensure relative voice rest before data acquisition. Speech samples were obtained twice; before and after 2.5 hours of classroom teaching. G2 subjects were recorded in a quiet room within their home in the morning before and after 2.5 hours of social vocal use. The same equipment was used in both groups, and the interval between the recordings was defined based on the period of voice use during classroom teaching.

The stimulus for recordings included sentences and a semi-spontaneous speech protocol used to evaluate phonetic vocal quality[19] based on the principle of susceptibility of the vowel and consonantal segments to the effects of the long-term voice quality settings. Subjects were asked to respond to "Tell about the city where you live" and read six sentences in random order, totaling three repetitions per sentence. Subjects could read the sentences before the recording and ask any question.

Recordings were performed with the subjects seated using an omnidiretional condensor microphone (CO1 Samson Technologies Corp., Hauppauge, NY) positioned 20 cm from the subject at a 45° angle, linked to a sound card (Quad Capture Interface; Roland DG Corporation, Japan) and a notebook (Intel® Pentium® Processor P6200, USA). Data were stored and analyzed via the software *SONAR LE* (Roland DG Corporation, Japan) in wave files, edited and named based on acoustic and perceptual analysis.

Perceptual analysis

Perceptual analysis was performed by two speech language pathologists with 10 years of experience in work in a voice clinic and the samples were randomized, and therefore, the reviewers were blinded to the group (G1 vs G2) and also the recording condition (pre vs post loading).

The GRBASI scale was used for perceptual analysis; G-degree of dysphonia, R-roughness, B-breathiness, A-asthenia, S-stiffness, and the last parameter I-instability inserted by Dejonckere et al (1996).[20] With this scale, each speech pathologist evaluated the parameters according on a 0 to 3 scale with 0 reflecting absence of disturbance, 1 for slight, 2 for moderate, and 3 for intense. The speech language pathologists performed the perceptual analysis with *Coby* CV-3000 (Coby Electronics, Lake Success, NY) earphones and were allowed to hear the voice recordings as many times as needed.

Acoustic analysis

Acoustic analysis of the vocal samples (read sentences and semispontaneous speech) was done by the script *Expression Evaluator*[21] applied to the free software *Praat*[22] for automatic extraction of acoustic measures. The script generated the acoustic parameters of F0 (median, semiamplitude interquartile, quantile 99.5%, and skewness), first F0 derivative (mean, SD, and skewness); intensity (skewness), spectral slope (mean, SD, and skewness) and long-term average spectrum (LTAS) frequency (SD) measures.

Vocal self-perception

The subjects in both G1 and G2 were asked to evaluate their effort during sentence reading before and after 2.5 hours of voice use via a 10 point visual analogue scale (VAS). Zero represented the absence of effort and 10 maximum effort. A visual image was provided to compliment the scale (Figure 1).

Statistical analysis

Statistical analysis of acoustic parameters, perceptual analysis, and vocal self-perception was done by multivariate discriminant analysis and a canonical correlation analysis to evaluate the strength of the association and to discriminate the parameters related to vocal loading in both G1 and G2.

Discriminant analysis of the acoustic parameters, perceptual analysis, and vocal self-perception was done both separately and grouped to evaluate the possibility of prediction and discrimination of the samples in each group and at each moment of vocal use. Canonical correlation analysis

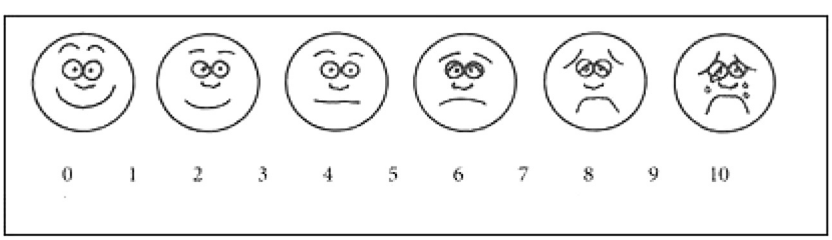

FIGURE 1. Visual analogue scale (VAS). See http://www.jkns.or.kr/fulltext/htm/0042011125f1.htm.