Statistical Analysis of the Reliability of Acoustic and Electroglottographic Perturbation Parameters for the Detection of Vocal Roughness

***Kiyohito Hosokawa, *Makoto Ogawa, †Michiko Hashimoto, and *Hidenori Inohara,** **Osaka and †Hyogo, Japan*

Summary: Objectives. The aim of this study was to verify whether the electroglottography (EGG) perturbation parameters could be superior indicators to traditional acoustic (Ac) measures, particularly for the detection of mild vocal roughness.

Methods. Thirty-nine participants with muscle tension dysphonia (the MTD group), 48 dysphonic participants with vocal fold lesions (the organic group), and 40 nondysphonic participants (the control group) were enrolled in the study. Based on the severity of vocal roughness, each of the two dysphonic groups was divided into mildly and severely dysphonic subgroups. The Ac and EGG signals during sustained /e:/ phonation were recorded simultaneously. The period and amplitude perturbation quotients of both signals (Ac-PPQ/-APQ and EGG-PPQ/-APQ) were calculated. The receiver operating characteristic (ROC) analyses were applied to evaluate the discriminative capabilities.

Results. In the analyses between the control and each of the two dysphonic groups, the values of the areas under the curve (AUC) for EGG-PPQ were significantly higher than those for Ac-PPQ. Next, the ROC analyses between the control and mildly dysphonic MTD subgroup demonstrated that the AUC values for EGG-PPQ/EGG-APQ were significantly higher than those for Ac-PPQ/Ac-APQ. In the analyses of the mildly dysphonic organic group, the AUC value for EGG-PPQ was significantly higher than that for Ac-PPQ.

Conclusions. The present study demonstrated that both the period and the amplitude perturbation parameters of the EGG signals showed higher diagnostic accuracies than those of Ac signals, especially for the detection of mild vocal roughness. These results suggest that the EGG perturbation parameters could provide better information than the traditional Ac perturbations.

Key Words: Voice disorders–Electroglottography–Perturbation analysis–ROC analysis.

INTRODUCTION

The severity of pathologic voice disorders has generally been evaluated using two major methodologies: auditory perceptual ratings and computed analyses of acoustic (Ac) signals. The former are based on the degrees of psychologically abnormal characteristics of perceived voices, whereas the latter are techniques that can quantify the level of cycle-to-cycle irregularities or the ratio of contaminated spectral noise components in Ac signals.

As one of the representative ratings in the former, the GRBAS scale consists of the scores indicating the overall grade of hoarseness (G: grade) and four fundamental components (R: roughness, B: breathiness, A: asthenicity, and S: strain).[1–3] In particular, vocal roughness and breathiness have been considered to be the two critical components of the perceptual characteristics of hoarseness.[4] This scale has been used worldwide because there are sufficient coincidences between experienced examiners and no requirements for specific equipment.[5,6] However, the major disadvantage of this method is the poor objectivity associated with human prejudice, wishful thinking, or judgment error.

Concerning the latter techniques, it has been reported that computer-synthesized voice signals with a rapid period or amplitude modulation result in perceived vocal roughness.[7–9] From this viewpoint, the period and amplitude perturbation quotients (PPQ and APQ) were developed as representatives of computed Ac measures to quantitate the severity of vocal roughness.[10,11] Indeed, these parameters were demonstrated to show higher values in rough voices in dysphonic patients with vocal pathologies,[12,13] as well as a simulated rough voice in normal speakers[14] compared with normal voices, and to be positively correlated with the severities of vocal roughness.[3,15–18] In particular, Zyski et al[15] reported that a single discriminant function analysis demonstrated that these parameters could separate normal and pathologic groups. Consequently, numerous studies have used the period and amplitude perturbations of Ac signals as an established objective methodology to estimate the severity of vocal disturbances associated with a variety of voice disorders and to assess the outcomes of different therapies.[19–21]

However, several recent studies have cast doubt on the reliability of these Ac measures.[22–25] For example, Carding et al[22] reported that the jitter and shimmer had arguable validity but poor reliability because these parameters could insufficiently discriminate between normal and dysphonic speakers and between voices before and after therapies. Ma and Yiu[23] reported that the relative amplitude perturbation and shimmer of Ac signals could differentiate between dysphonic and nondysphonic voices only in female, but not in male speakers. Brockmann-Bauser and Drinnan,[25] and Werth et al[24] also described that the majority of mildly or moderately pathologic

Accepted for publication July 15, 2013.

From the *Department of Otorhinolaryngology, Head and Neck Surgery, Osaka University Graduate School of Medicine, Osaka, Japan; and the †Department of Otorhinolaryngology, Kinki Central Hospital of Mutual Aid Association of Public School Teachers, Hyogo, Japan.

Address correspondence and reprint requests to Makoto Ogawa, Department of Otorhinolaryngology, Head and Neck Surgery, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan. E-mail: mogawa@ent.med.osaka-u.ac.jp

Journal of Voice, Vol. 28, No. 2, pp. 263.e9-263.e16
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2013.07.005

voices showed values for the Ac jitter and shimmer within the range of those in normal voices. It, therefore, remains controversial whether the Ac perturbation measures are able to discriminate normal and pathologic voices, especially normal and mildly disturbed voices.

Electroglottography (EGG) is a technique used to analyze the vibratory characteristics of the vocal folds by detecting impedance changes across the larynx.[26] In comparison with the waveforms of Ac signals, those of EGG signals consist of a continuum of a simpler waveform reflected by repeated contacts and dissociations between the bilateral vocal folds. A number of studies have already calculated the perturbation parameters of EGG signals in dysphonic patients[27–33] and compared the changes in these measures during therapies or task loading.[30–33] In particular, several studies implied that these EGG measures could be finer indicators to detect vocal abnormalities than those of Ac signals.[27–29,33] However, the conclusions in studies were based only on the existence of significant differences between the normal and dysphonic groups or the degrees of correlation with the perceptual vocal qualities, and no studies have compared the reliability of Ac and EGG measures in the detection of vocal pathologies using appropriate statistical methodologies. Recently, the receiver operating characteristic (ROC) analysis has become one of the most common means to quantify the discrimination performance of a diagnostic procedure.

We therefore undertook a statistical analysis of the reliability of the Ac and EGG parameters, and the aims of this study were (1) to compare the values for the Ac and EGG perturbation parameters between normal and abnormal rough voices in the organic and nonorganic voice disorders, separately and (2) to verify whether the EGG perturbation parameters could be superior indicators for the detection of mild vocal roughness to traditional Ac measures using an ROC analysis.

MATERIALS AND METHODS

Participants

The Osaka University Graduate School of Medicine gave approval for this study. Informed consent from all the participants was obtained before data collection. The dysphonic participants with vocal roughness were recruited from the Department of Otorhinolaryngology-Head and Neck Surgery of Osaka University from June 2010 to February 2012. The participants presented with a variety of etiologies and were referred for voice assessment. After otolaryngologic assessments, they were examined by rigid or flexible laryngoscopy and diagnosed with various kinds of dysphonia. The inclusion criteria for the dysphonic group were as follows: a diagnosis of organic or nonorganic voice disorders defined by the findings of laryngoscopy and the presence of a disturbed voice. In particular, the diagnosis of muscle tension dysphonia (MTD) was defined as the presence of a rough/strained voice and supraglottic contraction during phonation, without any organic abnormalities in the laryngeal region. The exclusion criteria for the dysphonic group included any participants diagnosed as having (1) voice disorders dominantly showing breathiness, rather than roughness, such as vocal fold paralysis, vocal fold atrophy, sulcus vocalis, and hypofunctional dysphonia and (2) neurologic diseases such as spasmodic dysphonia or Parkinson disease. Consequently, 39 MTD participants (the MTD group: 27 males and 12 females, mean age 64.6 years, range 35–82) and 48 dysphonic participants with vocal fold lesions (the organic dysphonia [OD] group: 20 males and 28 females, mean age 57.0 years, range 23–82) were enrolled in the study. Table 1 summarizes the distribution of the participants by groups and diagnosis. In addition, 40 nondysphonic participants without any laryngeal abnormal findings were also enrolled in the study (the control group; 20 males and 20 females, mean age 42.3 years, range 28–75).

Recording of the Ac and EGG signals

A head-worn microphone (SE50; Samson Technologies Corp., Hauppauge, NY) was laterally positioned 2 cm from the lips of a participant, and the EGG electrodes (model 6103; KayPENTAX, Lincoln Park, NJ) were fitted on the neck over the bilateral lamina of the thyroid cartilage by a Velcro strap. Simultaneous Ac and EGG recordings were made in a sound-treated room. Each participant was directed to phonate a sustained vowel /e:/ at a natural and comfortable pitch/loudness for approximately 3 seconds. The lines of the two signals were connected to a linear PCM recorder (H4n; Zoom Corp., Tokyo, Japan) for digitization at a sampling rate of 44.1 kHz and 16 bits per sample quantization.

Calculation of the perturbation measures of the Ac and EGG signals

A single speech therapist who was not offering clinical services to the participants extracted 1000 milliseconds simultaneous intervals from the Ac and EGG signals, avoiding the areas of voice onset, offset, and the aperiodic part of the waveform in which there is a failure to detect the voice pulse, as much as possible. The Ac signals were low-pass-filtered at 10 kHz for a more accurate estimation of the perturbation parameters.[34] In addition, the EGG signals were band-pass-filtered from

TABLE 1.
Distribution of the Dysphonic Participants According to the Diagnosis and Group

Group	MTD Group		OD Group		
Subgroup	M-MTD	S-MTD	M-OD	S-OD	Total
Diagnosis					
MTD	19	20			39
Polypoid mucosa			3	8	11
Nodule(s)			8	1	9
Polyp(s)			5	3	8
Acute/chronic laryngitis			8	0	8
Cyst			3	3	6
Tumor			2	2	4
Postradiotherapy			1	1	2
Total	19	20	30	18	87