Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation

Tomoyuki Haji
Vocal Dynamics Laboratory, Lenox Hill Hospital, 100 East 77th Street, New York, New York 10021

Satoshi Horiguchi and Thomas Baer
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

Wilbur J. Gould
Vocal Dynamics Laboratory, Lenox Hill Hospital, 100 East 77th Street, New York, New York 10021

(Received 8 October 1985; accepted for publication 2 April 1986)

Electroglottography (EGG) was used to monitor vocal fold vibration patterns in normal subjects and patients with various laryngeal disorders. In order to evaluate the regularity of vocal fold vibration, frequency and amplitude perturbation of EGG waves during sustained phonation were measured with a laboratory computer. The data were compared to the degree of hoarseness evaluated by auditory perception and by sound spectrographic analysis. Frequency and amplitude perturbation measures showed some overlap between normal and pathological groups. However, there was a close relation between perturbation analysis of EGG waves and degree of hoarseness (Spearman's rank correlation coefficient $rs = 0.73$, $p < 0.0005$). Amplitude perturbation was found to be a more sensitive measure of the irregularity of vocal fold vibration than frequency perturbation.

PACS numbers: 43.70.Aj, 43.70.Dn, 43.70.Jt, 43.72.Ar

INTRODUCTION

Electroglottography (EGG) is a technique for the indirect examination of vocal fold contact during vibration through measurement of electrical impedance changes (Fant *et al.*, 1966; Fourcin, 1974). The EGG does not interfere with phonation and it is still somewhat controversial (Baer *et al.*, 1983; Childers *et al.*, 1984; Fourcin, 1981; Kelman, 1981; Lecluse *et al.*, 1975; Pederson, 1977; Smith, 1981). Whatever details it represents, EGG certainly reflects the vibratory cycle of the vocal folds with fairly high fidelity. Irregularities in the EGG thus correspond to irregularities in the vibratory pattern of the folds.

Perturbation analyses of frequency and amplitude in a voice signal have been developed in an attempt to provide objective parameters for early detection of laryngeal disorders. Since the waveform obtained is much less complex than that of the voice signal, and is unaffected by the acoustic resonances of the vocal tract, it was considered that EGG was more suitable than the voice signal for perturbation analyses.

Frequency and amplitude perturbations of the EGG reflect different aspects of irregularity of vocal fold vibration. To assess the clinical significance of frequency and amplitude perturbations of the EGG signal, indices were calculated and compared to the degree of hoarseness evaluated by auditorily based judgments of the recorded voice and visually based analysis of spectrograms. The utility of these indices in providing objective measures of hoarseness was then examined.

I. SUBJECTS

Subjects were 30 normal speakers and 33 patients with vocal pathologies. The normal speaking group ranged in age from 24–59 years, and was comprised of 18 males and 12 females. They had either an indirect laryngoscopy or a laryngofiberscopy, and proved to have no laryngeal or pulmonary disorders. Their voices were perceived auditorily, by two otolaryngologists, as normal with no significant noise component. The patients, 19 males and 14 females, ranged in age from 14–77 years, and had various pathologies as can be seen in Table I.

II. METHOD

The EGG signals were recorded with a Fourcin laryngograph (Fourcin, 1974) and electrodes were placed symmetrically over the thyroid alae. The subjects produced the sustained vowel /a/ in a natural speaking voice. The EGG signal was recorded, using one channel of an audio tape recorder (OTARI MX5050), while the voice signal was recorded simultaneously on another channel using a microphone (Electrovoice 666) placed about 20 cm from the lips. A steady portion of the signals was selected for analysis on a laboratory computer (DEC VAX 11/780). The signals were sampled at 20 000 samples/s and digitized with 12-bit resolution (Fig. 1).

Fifty consecutive periods of the EGG signals were measured for each sample, using base-line crossing to define each period. In most of the samples, a measurement of 50 cycles

TABLE I. Pathologies of the patients. RLN: recurrent laryngeal nerve.

Pathology	No.
polyp	12
nodule(s)	10
glottic reconstruction after hemilaryngectomy	3
laryngitis	2
RLN palsy	2
laryngeal ca.	2
laryngeal tbc.	1
vocal atrophy	1
total	33

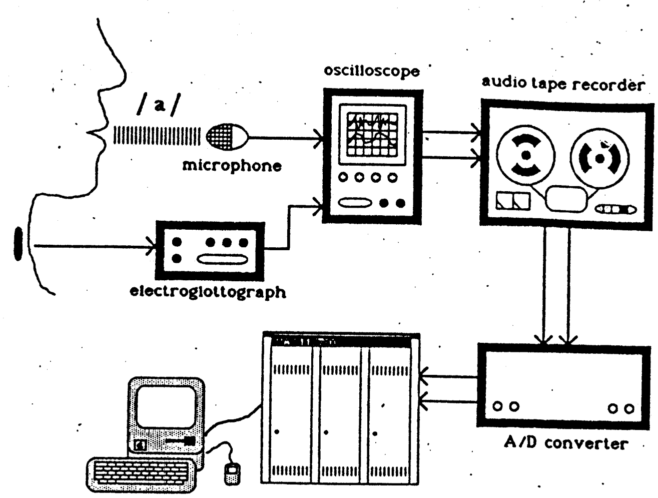

FIG. 1. Recording and processing system.

was started at a point 250 ms from the beginning of the sample. When this particular portion was too irregular to identify periods, a steadier portion was selected. The first base-line crossing was determined semi-automatically, using a cursor. Then the other 49 periods were determined automatically. The frequency of each cycle was defined as the inverse of the period. The amplitude was defined as the peak-to-peak distance for each cycle (Fig. 2).

Frequency perturbation was expressed as the mean difference in frequency between consecutive cycles measured in semitones. "Semitone" was used to reduce the influence of fundamental frequency. Amplitude perturbation was expressed as the mean difference in amplitude between consecutive cycles in dB.

In addition to these measurements from the EGG, the pathological voice samples were rated auditorily from the recordings and visually from spectrograms. The auditory ratings were on a four-point scale; 0 = none, 1 = slight, 2 = moderate, and 3 = extreme (Issiki *et al.*, 1969). Ratings were performed by two otolaryngologists and one speech pathologist. All judges were well accustomed to rating on this scale, advocated by the Japan Society of Logopedics and Phoniatrics. As mentioned above, all of the normal controls were rated 0 by two otolaryngologists. Pathological voice samples were recorded again on the audio tape with a random addition of some normal and pathological voices (These voices were not the ones used in this study.) without any particular order. Each judge worked alone and was blind to the subjects' identities. The individual scores for each subject were then averaged. The spectrographic ratings were made on a five-point scale, using Yanagihara's categories, which are based on the appearance of the harmonic structure and noise (Yanagihara, 1967). (Category 0, for normal voices and those with no significant noise component, was added.) The spectrographic judgments were done separately from the auditory judgments. Again, the judges were blind to the subjects' identities when ratings were performed. Individual scores were also averaged.

III. RESULTS

Sine waves of 50, 100, and 200 Hz, recorded using the audio tape recorder from a signal generator (Wavetek model 184), were used to test the accuracy and limitations of the processing system. The frequency perturbation obtained for sine waves of 50, 100, and 200 Hz were all less than 0.05 semitones. Amplitude perturbations for the same signals were less than 0.04 dB.

To test the interjudge reliability of auditory rating, the Spearman's rank correlation coefficients between the judges were calculated. The coefficients were 0.670, 0.666, and 0.640, and all were statistically significant ($p<0.01$). To test the intrajudge reliability of auditory rating, one of the judges performed the rating again, several days after the first rating. The Spearman's rank correlation coefficient between two judgments was 0.703, with a significant level of 0.01 and higher. These results justified an averaging of scores in order to represent each subjects' rating.

To test the interjudge reliability of spectrographic rating, the Spearman's rank correlation coefficients between the judges were calculated. The coefficients were 0.819, 0.883, and 0.813, and again, all were statistically significant ($p<0.01$). To test the intrajudge reliability of spectrographic rating, one of the judges performed the rating again, several days after the first rating. The Spearman's rank-correlation coefficient between two judgments was 0.924, with a significant level of 0.01 and higher. These results justified an averaging of scores in order to represent each subjects' rating.

A scatter plot of amplitude and frequency perturbations, for normal subjects, can be seen in Fig. 3. Frequency

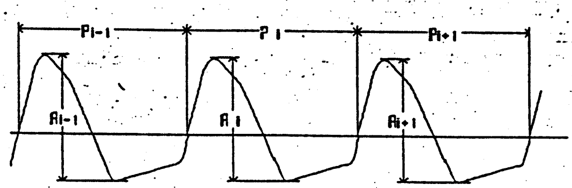

FIG. 2. Waveform measurements. $\Delta F(\text{semitone}) = (\Sigma_{i=1}^{N-1}|F_i - F_{i+1}|)/(N-1)$, where $F_i = 39.86\times\log(f_i/C_0)$, $f_i = 1/P_i$, where N = the number of periods measured, F_i = the frequency of the ith cycle (in semitone), f_i = the frequency of the ith cycle (in Hz), P_i = the period of the ith cycle (in s), $C_0 = 16.352$ (Hz). $\Delta A(\text{dB}) = [\Sigma_{i=1}^{N-1}|20\log(A_i + 1/Ai)|]/(N-1)$, where N = the number of periods measured and A_i = the amplitude of the ith cycle.

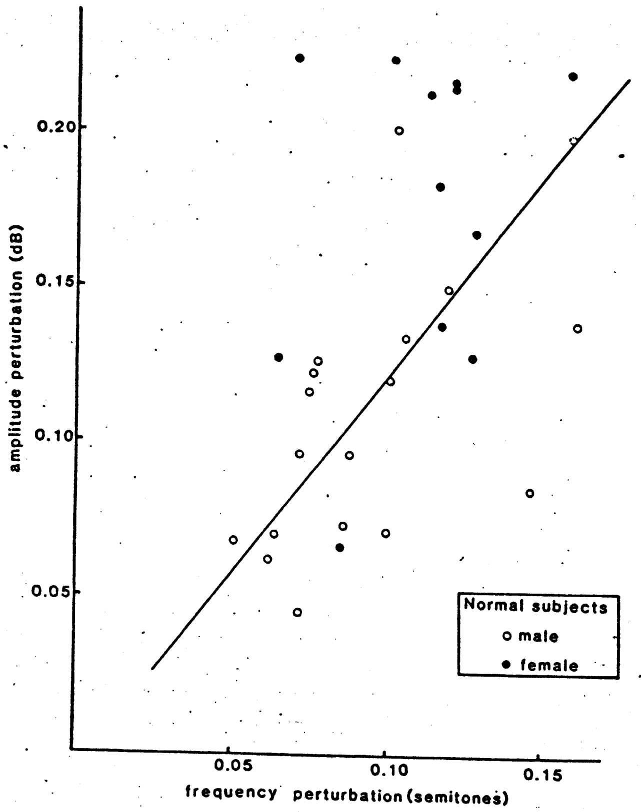

FIG. 3. Scatter plot of frequency and amplitude perturbation of normal subjects (correlation coefficient $r = 0.476$). The line is the least-squares regression line-fit of all the data points ($Y = 1.13X - 0.053$).

perturbations ranged from 0.05–0.16 semitones, with a mean of 0.1 semitones. The amplitude perturbations in normals ranged from 0.05–0.22 dB, with a mean of 0.14 dB. These two variables were significantly correlated at the level of 0.01, with a Pearson coefficient of 0.476. Male–female differences were analyzed in these data. Males and females did not differ significantly in frequency perturbation, but they did differ significantly in amplitude perturbation at the level of 0.01 when using a t test. Larger perturbations were exhibited for females. Table II shows correlations between perturbation analyses of EGG and auditory-perceptual and sound-spectrographic evaluations of hoarseness. Spearman's rank correlations were used to assess the association between variables. Each of the four pairs showed significant correlation. The highest correlation was noted between amplitude perturbation of EGG and the auditory-perceptual rating of hoarseness (rank correlation coefficient $rs = 0.73$, $p < 0.0005$). The lowest correlation was noted between frequency perturbation and the spectrographic rating (rank correlation coefficient $rs = 0.48$, $p < 0.001$). This pattern suggested that amplitude perturbation was a better indicator of hoarseness than frequency perturbation of EGG. It also suggested that perturbation analyses were more closely related to auditory perception of hoarseness than to sound-spectrographic evaluation.

Figure 4 shows the relationship between amplitude perturbation of EGG and auditory-perceptual rating ($rs = 0.73$, $p < 0.0005$). To further examine these data, the patients were classified into three groups, "slight," "moderate," and "extreme," according to the averaged perceptual rating (0–1.5; 1.5–2.5; and 2.5–3.0, respectively). Amplitude perturbations for the slightly, moderately, and extremely hoarse groups were significantly different from each other, but no significant difference was found between the normal and slightly hoarse groups. In frequency perturbation of EGG, however, the only significant difference was between the moderately and extremely hoarse groups; there was no significant difference among the normal, slightly, and moderately hoarse groups. The t tests were used for differences between groups. Significant levels were all at 0.01 and higher.

TABLE II. Correlations between perturbation analyses of EGG and perceptual and sound spectrographic evaluation of hoarseness. Spearman's rank correlation was used to assess the association. The table shows Spearman's rank correlation coefficients (rs) between variables.

	Perceptual rating	Spectrographical rating
Amplitude perturbation	0.73[a]	0.54[b]
Frequency perturbation	0.50[b]	0.48[b]
	$n = 33$	

[a] $p < 0.0005$.
[b] $p < 0.001$.

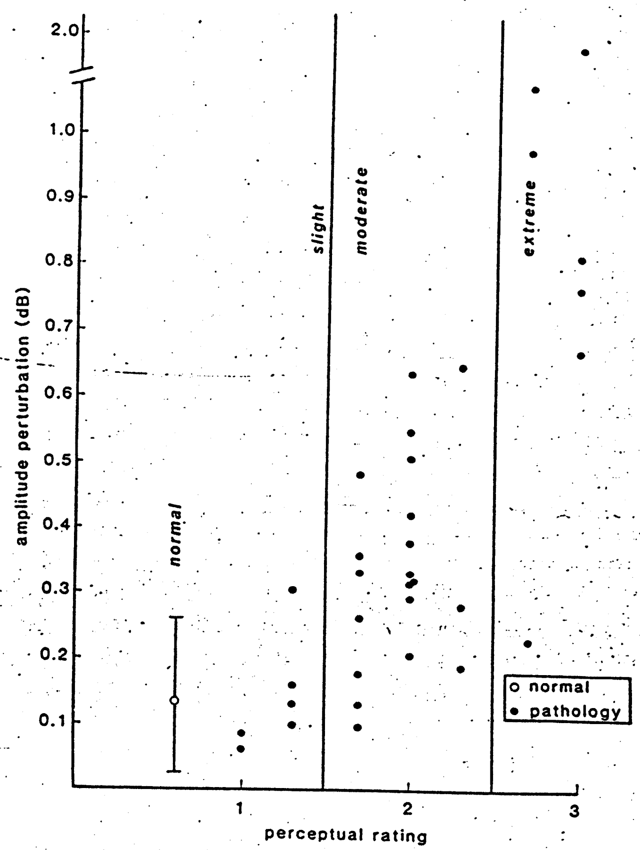

FIG. 4. Relationship between amplitude perturbation of EGG and perceptual rating of hoarseness.

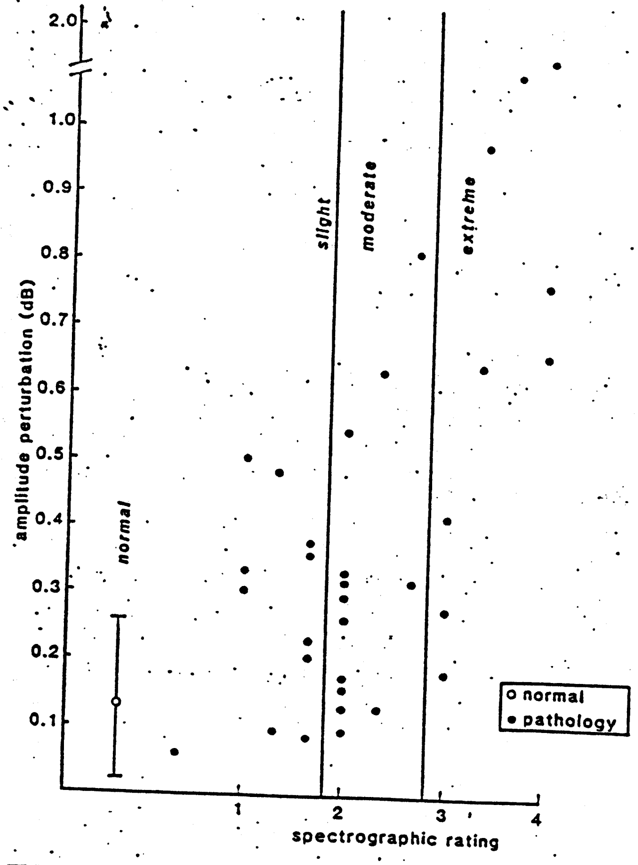

FIG. 5. Relationship between amplitude perturbation of EGG and sound spectrographic evaluation of hoarseness.

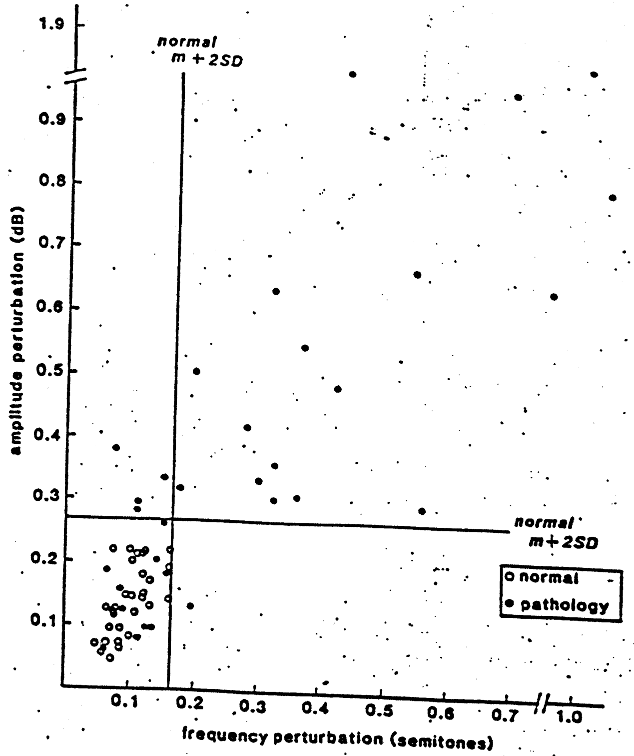

FIG. 6. Correlation between frequency and amplitude perturbation of EGG in the normal voiced subjects and the patients (correlation coefficient $r = 0.82$). *m*: mean. SD: standard deviation.

Figure 5 shows the relationship between amplitude perturbation of EGG and the sound spectrographic rating ($rs = 0.54$ $p < 0.001$). The correlation was considerably lower than that between amplitude perturbation and perceptual rating, as pictured in Fig. 4.

Figure 6 shows the correlation between frequency and amplitude perturbation of EGG in the normal-voiced subjects and in the patients. The two indices were significantly correlated at the 0.01 level, with a Pearson coefficient of 0.82. However, amplitude-perturbation differentiated pathologies from normal subjects were better than frequency perturbation, when the critical values were defined as the means plus standard deviations. Considering frequency and amplitude perturbation together would identify only one more case than amplitude perturbation alone.

IV. DISCUSSION

Many researchers have attempted to differentiate laryngeal pathologies from normal-voiced subjects by using various methods of acoustic analysis in voice signals. In doing so, attention must be paid to the degree of hoarseness or the severity of disorders of the samples in the pathological group. If only patients with extremely hoarse voices are included in the pathological group, there is no question that differentiation between normals and pathologies is remarkably good. On the other hand, if the pathological group includes mainly patients with only slightly hoarse voices, the differentiation rate will be poor. Some of the differences in evaluations of the utility of acoustic analysis seem to be attributable to differences between pathological samples. To evaluate the samples properly, auditory-perceptual rating and sound spectrographic rating were carried out for each voice sample. These procedures, though not completely objective, proved to be reliable enough to evaluate the degree of hoarseness.

Significant correlations were found between perturbation analyses of EGG waves and the degree of hoarseness evaluated by auditory perception and from sound spectrograms. Additionally, objectiveness of the EGG perturbation measures could give them more advantage than auditory or spectrographic ratings of the degree of hoarseness to both scientific and clinical use. Comparing both EGG measures, amplitude perturbation correlated more highly with hoarseness, evaluated both perceptually and sound spectrographically than frequency perturbation did, and there was less overlap between normals and abnormals.

To examine the detail of the correlation between these two perturbation measures of EGG and auditory-perceptual ratings, patients were divided into three groups; slightly, moderately, and extremely hoarse, according to the perceptual rating.

Both frequency perturbation and amplitude perturbation could significantly differentiate between extremely hoarse voices and moderately hoarse voices. However, only amplitude perturbation could significantly differentiate between moderately hoarse voices and slightly hoarse voices.

Neither EGG measurements could significantly differentiate between slightly hoarse voices and normal voices. These measurements showed some overlap between the normally and abnormally voiced. There may have been several reasons for the overlap. The length of EGG samples analyzed in this study was usually less than 0.5 s. This is so short that it might not reflect the portions that could differentiate between the normally and abnormally voiced. Another reason may be that perturbation in frequency or amplitude basically reflects the irregularity of vocal fold vibration or "roughness," while hoarseness consists of "breathiness" or noise, caused by turbulent airflow through the glottis. Breathiness would not be reflected in our EGG measures, because EGG signals rely on vocal fold contact, whereas turbulent airflow or "breathy voice," is usually caused by insufficient glottic closure. Technical problems may provide another reason. We often failed to obtain clear EGG signals from subjects with thick necks. In general, clear EGG signals are more difficult to obtain from women than from men, probably due to the anatomical difference between females and males: The shape of the thyroid cartilage, the length of their vocal folds, the amplitude of vocal fold vibration, and the amount of cervical fatty tissue. These factors may have contributed to the difference in amplitude perturbation values between normal subject males and females. The degree of hoarseness, as judged by the presence of harmonics and noise on the spectrogram according to Yanagihara's criteria, seems to highlight the noise-component characteristic of breathiness rather than perturbation in frequency or amplitude. The fact that this breathiness noise is not represented in our measurements of the EGG signal may account for the spectrographic ratings having lower correlations than perceptual ratings with the perturbation measurements.

Many researchers have evaluated perturbations analyses of frequency and amplitude in the voice signal. Though there is some disagreement over the efficacy of these perturbation analyses for early detection of laryngeal disorders (Ludlow *et al.*, 1984), frequency and amplitude perturbation of the voice signal are generally considered as the important parameters to represent the irregularity of vocal fold vibration. Frequency perturbation of EGG is thought to be basically the same as that of the voice signal. The normal values we obtained from EGG are close to those obtained from voice signals by other researchers (Kitajima, 1983; Horii, 1980; Yumoto, 1983). Also, mean fundamental frequencies of EGG signals showed very close values to those of voice signals in our data. On the other hand, the amplitude perturbation of EGG reflects different aspects of the irregularity of vocal fold vibration from either frequency or amplitude perturbation of the voice measured at the lips. Although the meaning of amplitude perturbation of the EGG is still not completely clear, we do know that it somehow reflects the irregularity of the mode of contact of the vocal folds. From this study, EGG amplitude perturbation seems to be a more sensitive indicator of hoarseness than frequency perturbation. Thus EGG shows the possibility of bringing us useful information which cannot be obtained from the voice signal alone. In addition, the simple waveform of EGG facilitates computer analysis and does not require filtering or other processing. The fact that EGG is not invasive and does not interfere with speaking is clearly advantageous for clinical applications.

There remains some questions about the relationship between EGG and voice signal perturbation measures. Direct comparison of EGG perturbation measures with voice signal perturbation measures is the next step. These results will be presented elsewhere.

In conclusion, the frequency and amplitude perturbation of EGG, especially the amplitude perturbation, can be a useful clinical adjunct for evaluating the irregularity of vocal fold vibration. Further basic research on EGG waveforms may add more promise to the analysis of the EGG signal.

ACKNOWLEDGMENTS

This work was supported in part by NIH Grant NS-13870 to Haskins Laboratories.

Send reprint requests to Tomoyuki Haji, Department of Otolaryngology, Faculty of Medicine, Kyoto University, Shogoin, Sakyo-ku, Kyoto 606, Japan.

Baer, T., Lofqvist, A., and McGarr, N. S. (1983). "Laryngeal vibrations: A comparison between high-speed filming and glottographic techniques," J. Acoust. Soc. Am. 73, 1304–1308.

Childers, D. G., Smith, A. M., and Moore, G. P. (1984). "Relationships between electroglottograph, speech, and vocal cord contact," Folia Phoniatr. 36, 105–118.

Fant, G., Ondrackova, J., Lindqvist, J., and Sonesson, B. (1966). Electrical glottography. Quarterly Status and Progress Report (Speech Transmission Lab., Royal Institute of Technology, Stockholm) 4, 5–21.

Fourcin, A. (1974). "Laryngograph examination of vocal fold vibration," in *Ventilatory and Phonatory Control Mechanism*, edited by B. Wyke (Oxford U. P., London), pp. 315–333.

Fourcin, A. (1981). "Laryngographic assessment of phonatory function," ASHA Reports 11, 116–127.

Horii, Y. (1980). "Vocal shimmer in sustained phonation," J. Speech Hear. Res. 23, 202–209.

Isshiki, N., Okamura, H., Tanabe, M., and Morimoto, M. (1969). "Differential diagnosis of hoarseness," Folia Phoniatr. 21, 9–19.

Kelman, A.W. (1981). "Vibratory pattern of the vocal folds," Phonia Phoniatr. 33, 73–99.

Kitajima, K. (1973). "An analysis of pitch perturbation in normal and pathologic voices" (in Japanese). Pract. Otol. 66, 1195–1213.

Lecluse, F. L. E., Brocaar, M. P., and Vershuure, J. (1975). "The electroglottography and its relation to glottal activity," Folia Phoniatr. 27, 215–224.

Ludlow, C., Bassich, C. J., Lee, Y. J., Connor, N. P., and Coulter, D. C. (1984). "The validity of using phonatory jitter to detect laryngeal pathology," J. Acoust. Soc. Am. Suppl. 1 75, S8.

Pederson, M. F. (1977). "Electroglottography compared with synchronized stroboscopy in normal persons," Folia Phoniatr. 29, 191–199.

Smith, S. (1981). "Research on the principle of electroglottography," Folia Foniatr. 33, 105–114.

Yanagihara, N. (1967). "Significance of harmonic changes and noise components in hoarseness," J. Acoust. Hear. Res. 10, 531–541.

Yumoto, E. (1983). "Quantitative measurement of the degree of hoarseness" (in Japanese). Pract. Otol. 76, 2151–2168.