

# Laryngeal vibrations: A comparison between high-speed filming and glottographic techniques<sup>a)</sup>

Thomas Baer, Anders Löfqvist,<sup>b)</sup> and Nancy S. McGarr<sup>c)</sup>

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06510

(Received 28 September 1982; accepted for publication 3 January 1983)

This study was designed to compare information on laryngeal vibrations obtained by high-speed filming, photoglottography (PGG), and electroglottography (EGG). Simultaneous glottographic signals and high-speed films were obtained from two subjects producing steady phonation. Measurements of glottal width were made at three points along the glottis in the anterior-posterior dimension and aligned with the other records. Results indicate that PGG and film measurements give essentially the same information for peak glottal opening and glottal closure. The EGG signal appears to reliably indicate vocal-fold contact. Together, PGG and EGG may provide much of the information obtained from high-speed filming as well as potentially detect horizontal phase differences during opening and closing.

PACS numbers: 43.70.Bk, 43.70.Ny

## INTRODUCTION

High-speed films are most commonly used to monitor details of the glottal cycle. However, this technique is difficult and expensive and cannot be performed during natural conditions because of the need to use a laryngeal mirror. It is therefore desirable to use glottographic techniques such as photoglottography and electroglottography in place of this more difficult and more invasive technique.

Photoglottography (PGG) or transillumination is a semi-invasive technique for monitoring laryngeal behavior. Briefly, transillumination involves directing a light source toward the glottis from above or below and monitoring its intensity on the other side, to measure glottal width (Soneson, 1960). This technique has proven extremely useful for studying the coordination of glottal movements with those of the supralaryngeal articulators (Löfqvist and Yoshioka, 1981; McGarr and Löfqvist, 1982). For studies of phonation it may supply measures of opening and closing time during the glottal cycle which may be clinically or pedagogically useful. It may also be useful for monitoring glottal activity preparatory to phonation or at its initiation. In comparison with filming—especially high-speed filming—transillumination can be performed more easily and under more natural conditions, including natural speech. Perhaps more importantly, the transillumination signal is more easily analyzed in parallel with other instrumental measures of vocal-fold activity. In combination with these other measures, such as electroglottography and EMG, we believe transillumination can be useful for examining the relationship between vibratory performance and acoustic output on one hand, and between glottographic signals and those such as EMG which can be obtained more invasively on the other hand.

Although photoglottography has been in practical use for several years, there is some question about its reliability

and validity. Notably, many authors seem to agree that it can reliably indicate timing of peak glottal opening and closure, although there may be some uncertainty about the moment of glottal opening (Kitzing and Sonesson, 1974; Hutter, 1976). In studies comparing glottal area variations measured by transillumination and from high-speed films, Harden (1975) found good correspondence during most of the glottal cycle. However, Coleman and Wendahl (1968), in a similar study, challenged the reliability of the technique. The different results obtained in these two studies may be due to different apparatus and techniques employed in the two investigations. For example, differences in the size of the sensor and its placement may be significant. A comparison between glottal width measures obtained by transillumination and from simultaneous fiber-optic filming during voiceless obstruent production showed that temporal information supplied by the two methods was virtually identical (Löfqvist and Yoshioka, 1980; Yoshioka *et al.*, 1981). However, to compare smaller faster movements during phonation, a high-speed filming system is required in place of the fiberoptic endoscope.

While photoglottography, or transillumination, carries information about the pattern of glottal opening, electroglottography (EGG) is thought to convey information about the patterns of vocal-fold contact. Briefly, the technique involves the transmission of an electrical field between electrodes placed bilaterally on the neck of the subject so that the electrical impedance is expected to vary as a function of the degree of vocal-fold contact. That is, impedance should decrease as the area of vocal-fold contact increases, other factors remaining the same. While it is clear that the pattern of electroglottographic signals is related to the patterns of laryngeal vibrations, there has been some disagreement whether the EGG signal accurately represents vocal-fold contact area. Most studies indicate good agreement between apparent vocal-fold contact and deflections of the EGG signal, with either normal (Fant *et al.*, 1966; Fourcin, 1974; Kitzing, 1977; Baer *et al.*, in press; Childers *et al.*, in press) or excised (Lecluse *et al.*, 1975) larynges. On the other hand, Smith (1981) argues that the EGG registers acoustic and me-

<sup>a)</sup> Parts of this paper were presented at the 102nd Meeting of the Acoustical Society of America, Miami Beach, 30 November–4 December 1981.

<sup>b)</sup> Also Department of Phonetics, Lund University, Sweden.

<sup>c)</sup> Also Center for Research in Speech and Hearing Sciences, Graduate School and University Center, The City University of New York.

chanical effects and that the conventional interpretation of the EGG signal is untenable. This evidence is, however, unconvincing and not very well documented. We thus believe that the conventional interpretation is still valid until disproven in a more convincing way.

In general, the EGG and PGG signals provide information about complementary parts of the glottal cycle—PGG about the open period and EGG about the closed period. As noted by Rothenberg (1981), however, the glottis rarely either opens or closes abruptly over its entire length. Rather, for part of the cycle, the folds are likely to be in contact or separated over only part of their length. Thus, EGG and PGG signals are likely to overlap. Baer *et al.* (in press) argued that by obtaining both glottographic signals in parallel, and observing the overlap, the usefulness of each is increased because horizontal phase differences can be detected. A comparison between high-speed film and these measures is still needed to validate this assertion, however.

It therefore seemed appropriate to perform a validation study using our own equipment and techniques for transillumination and electroglottography in collaboration with the high-speed filming system provided by colleagues at the National Technical Institute for the Deaf. Specifically, the validity of glottographic techniques, namely photoglottography (PGG) and electroglottography (EGG), are examined to assess comparable information available in high-speed films.

## I. METHOD

The subjects were one female and one male with no evidence of laryngeal pathology. Because of the requirements for effective glottal illumination, each of the subjects was asked to produce steady phonation of the vowel /i/.

During these productions, high-speed laryngeal films at 4000 frames/s were taken using procedures described by Metz *et al.* (1980). Briefly, this system provides a xenon arc light source coupled with an optical system to project a high intensity light beam on the vocal folds. Reduction of infrared and ultraviolet radiation in the light source is accomplished by filtering. The cold light is then projected paraxial to the camera lens to intersect on a laryngeal mirror positioned in the oropharynx of the subject. During the positioning and filming, the subject was able to view the vocal folds by means of extrinsic mirrors mounted on the equipment housing. Similarly, the view of the vocal folds could be monitored throughout the filming by means of a reflex viewfinder installed on the camera lens.

High quality acoustic recordings were obtained at the time of the filming. The microphone was positioned on the shaft supporting the laryngeal mirror so that the subject maintained a lip to microphone distance of about 7 cm. The acoustic propagation delay between the glottis and the microphone was thus expected to be about 0.7 ms. Noise from the camera and optical-filming system was virtually eliminated since the subject was isolated in a sound treated room separate from the equipment.

Glottographic signals—transillumination and electroglottography—were obtained simultaneously with the high-speed films. Light from the filming system passing through

the glottis was sensed by a phototransistor placed on the surface of the neck just below the cricoid cartilage and coupled to the skin by a light tight enclosure. Electroglottographic signals were obtained from the one subject (BW) using the FJ Electroglottograph, and from the other subject (KH) using the Fourcin Laryngograph. According to Lecluse *et al.* (1975) there is no substantial difference between the signals recorded with those two instruments. The electrodes were placed on the neck at the level of the thyroid prominence. All glottographic signals were recorded on FM channels of an instrumentation tape recorder with a bandwidth of 2.5 kHz. Audio and timing codes were recorded on parallel direct channels. The timing codes were also recorded photographically on the film and were subsequently used for synchronization.

Using a computer-assisted measuring system, frame by frame measurements were made from the films during those portions where the film speed was constant at about 4000 frames/s. Measures of glottal width (WID) were made at the widest point along the anterior-posterior dimension of the glottis for each frame for purposes of comparison with the other glottographic records. Three additional measures of glottal opening were made along the anterior-posterior dimension as follows. The first (ANT) was made as close to the anterior commissure as possible. Since the view of the anterior commissure was sometimes blocked, the exact location of the point used for measurements differed slightly between films. The second measurement (MID) was made in the middle of the membranous glottis, and the third (POS), close to the vocal processes.

Audio and glottographic signals as well as timing codes were sampled and digitized at 10-K samples/s. Records from each of these were aligned with the film measurements.

## II. RESULTS

Figure 1 shows data for about three cycles of steady phonation at 145 Hz for the male speaker (KH). Records are, from top to bottom, the film measurements, photoglottography (PGG), electroglottography (EGG), and the audio signal, respectively. First, measures of glottal width from the

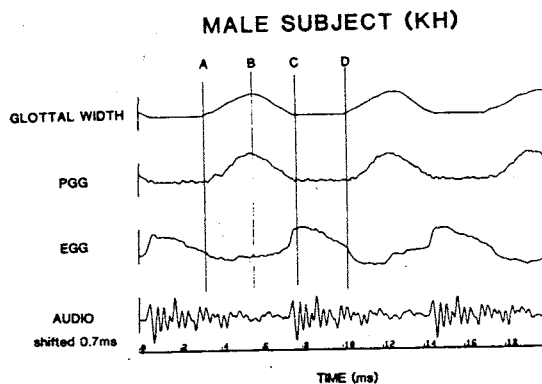


FIG. 1. Results for subject KH. The curves represent from top to bottom, glottal width measured from film, photoglottogram, electroglottogram, and audio signal.

films and transillumination (PGG) are shown to be practically identical. Both signals produce the same measures of onset (line A), peak glottal opening (line B), and glottal closure (line C). The EGG signal is plotted with increasing transconductance upwards. As expected, the EGG signal is complementary to the other records. Deflections in the EGG signal correspond roughly with glottal closure indicated by the other two methods. Due to technical problems, the EGG signal, for this subject, is somewhat noisy. Simultaneous audio has been sampled with pre-emphasis and has been shifted by 0.7 ms to compensate for the delay due to acoustic propagation from the glottis to the microphone. It can be noted that acoustic excitation appears to correspond with the end of the open period.

Looking in more detail, deflection in the EGG signal occurs slightly before the glottis is completely closed—as evidenced in the film records and the PGG signals. Peak deflection, corresponding to maximum area of contact, appears to occur about the moment of glottal closure. In examination of the films, the period of overlap in the three records corresponds to the interval when the region of contact between the folds moves from the anterior–posterior ends towards the center for this speaker. As indicated by line D, the descent of the EGG becomes rapid at the point of glottal opening, producing a knee in the curve. Examination of the film shows that glottal opening propagates from the center to the anterior–posterior ends during the interval between the knee in the EGG curve and its return to baseline (cf. also Fig. 3 below).

Figure 2 shows about four cycles of steady phonation at 250 Hz for the female speaker (BW). The moments of opening (lines A and D), peak opening (line B), and glottal closure (line C) as indicated by the film measurements are marked. As with the other speaker, the moments of peak glottal opening (line B) and glottal closure (line C) indicated by the film records and PGG are similar. However, the correspondence between the film records and PGG for this speaker is more subtle. That is, the relative slope of glottal opening in the interval D–E, is greater when measured by glottal width of the films than when indicated by PGG. In the PGG signal, the onset is so gradual that it is difficult to identify a single

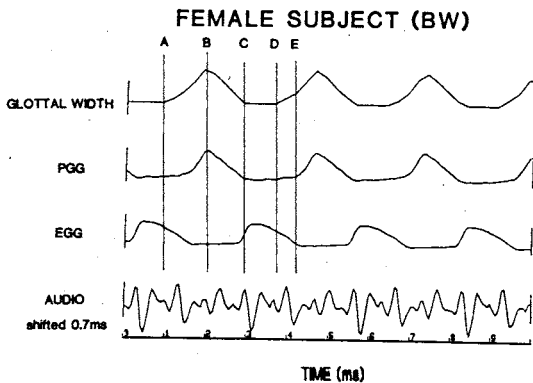


FIG. 2. Results for subject BW. Curves as in Fig. 1.

MALE SUBJECT (KH)

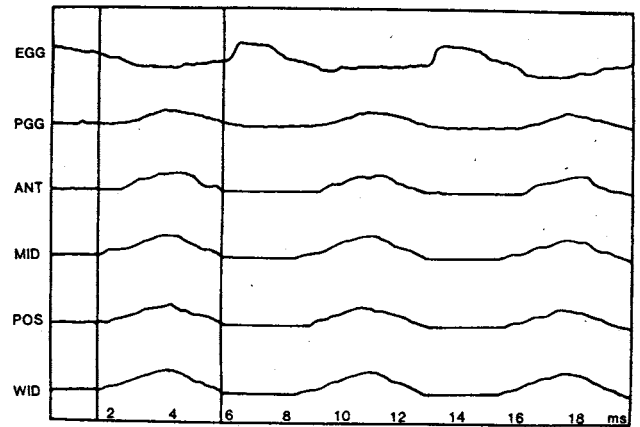


FIG. 3. Subject KH, comparison between glottograms and glottal opening measured at different points along the glottis. The curves represent from top to bottom, EGG, PGG, ANT, MID, POS, and WID.

point as the moment of opening. Further, the EGG signal does not show a knee as in Fig. 1. Thus, correspondence between the film and PGG, as well as PGG and EGG at opening, indicates that the glottal opening was gradual and showed large horizontal phase differences in these records (cf. also Fig. 4 below). This gradual opening could explain the absence of the knee in the EGG. Again there is acoustic excitation at the end of the open period.

Figure 3 shows the glottograms and the three measurements from the film (ANT, MID, and POS, respectively), as well as the measures of glottal width (WID) for speaker KH. From the film measures, two observations are apparent. First, the glottis does not open simultaneously along its entire length. Opening occurs slightly earlier in the medial region, and then propagates to the anterior and posterior ends. Glottal closure, on the other hand, occurs almost simultaneously along the entire length of the glottis. Second, the relative duration of the closed phase of the glottal cycle is

FEMALE SUBJECT (BW)

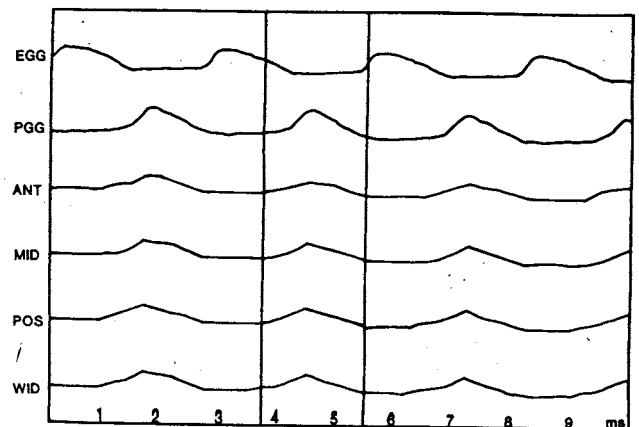


FIG. 4. Subject BW, comparison between glottograms and glottal opening measured at different points along the glottis. Curves as in Fig. 3.

longer anteriorly than posteriorly. The transillumination signal reflects the longer closed phase, and corresponds fairly well to the rise measured in the ANT portion of the film measures.

This correspondence is again illustrated in Fig. 4 for speaker BW. In the photoglottographic signal, the lower portion of the trace begins to rise at about the same time as the trace in the ANT film record. Unlike speaker KH, opening of the glottis occurs slightly earlier in the anterior and posterior portions than in the medial. For this speaker, the anterior part of the glottis was not visible. The film image suggested that the opening was occurring earlier in the anterior portion than was reflected in the film measures. However, both speakers are alike in that glottal closure again occurs almost simultaneously along the entire length of the folds as shown across all of these measures.

### III. DISCUSSION

The results concerning the reliability of transillumination confirm that the PGG and film measures give essentially the same information about peak glottal opening and glottal closure in normal phonation. We also confirm the observations of other investigators, in that there is more uncertainty about the moment of glottal opening, and this uncertainty appears to arise from the fact that glottal opening is more gradual than glottal closure. It is well known that the depth of glottal closure is quite small just prior to opening, while it becomes quite large immediately after closure. There also tend to be greater horizontal phase differences during opening than closing. "Opening" therefore occurs at different times along the anterior-posterior extent of the glottis.

Concerning the relationship between photoglottography and high-speed film measurements, it appears that the PGG signal can be thought of as representing a weighted sum of the widths along the length of the glottis. The weighting function depends on the location of both the light source and sensor with respect to the glottis. When the weights are high near the portion of the glottis that opens first, the agreement is better than when the weights are relatively low. We believe that the weighting functions in our experiment differed for the two subjects. Thus for subject KH, the PGG signal was in agreement with the opening measured at the anterior portion of the glottis. For subject BW, on the other hand, the PGG appeared to be relatively insensitive to the opening movement at the anterior and posterior ends, and the slope of the PGG signal thus increases after the midportion of the glottis opens.

Considering the EGG signal, its correspondence with other measures of glottal activity appears to confirm its validity as an indicator of vocal-fold contact. Although it is not possible to obtain independent measures of vocal-fold contact area, it is plausible that the EGG represents a measure of this quantity. The EGG signal reaches peak amplitude at about the moment of glottal closure indicated by the other measures, suggesting that the depth of glottal contact is maximum at this time. The rate of deflection of the EGG signal just prior to this maximum is very sharp, and it occurs over an interval that is comparable to the interval between film frames (cf. Childers *et al.*, in press). This aspect of the EGG

signal agrees with the interpretation that glottal closure is quite abrupt and demonstrates small horizontal phase differences. The EGG signal is also consistent with the notion that glottal opening is more gradual in both the vertical and horizontal dimensions. For the female subject, glottal opening cannot be clearly identified in the EGG waveform; for the male subject, it corresponds only to a mild increase in the rate-of-fall of the curve. A more gradual opening of the glottis for female subjects has also been reported by Kitzing and Sonesson (1974).

In conclusion, glottographic signals appear to be capable of supplying much of the significant information available in high-speed films. In comparison, films give not only measures of glottal area, but also the distribution of width along the glottis. However, filming procedures are prohibitively difficult and the introduction of the laryngeal mirror for this procedure may have some effect on the phonations that are produced. While the glottographic techniques we have employed cannot detect the distribution of width along the glottis, they can be used to detect the presence of horizontal phase differences during opening and closing and can be used under nearly natural speaking conditions. It appears, therefore, that simultaneous photo- and electroglottographic signals can be used to great advantage in studies of voice production for monitoring the patterns of laryngeal vibrations.

### ACKNOWLEDGMENTS

We are indebted to our colleagues at the National Technical Institute for the Deaf for generously sharing their time and laboratories: Robert L. Whitehead, Chairperson of Communication Research Department, and Dale E. Metz whose special expertise as cameraman made the study possible. We are also grateful to Jo Estill and Rosemary Szczesiul for assisting in film measurements. This work was supported by NINCDS Grants Nos. NS-13617 and NS-13870, and NIH Biomedical Research Support Grant No. RR-05596 to Haskins Laboratories.

- Baer, T., Titze, I. R., and Yoshioka, H. (in press). "Multiple simultaneous measures of vocal fold activity," Proceedings of the Vocal Fold Physiology Conference, Madison, Wisconsin.
- Childers, D. G., Smith, A. M., and Moore, G. P. (in press). "Relationships between electroglottograph, speech and vocal cord contact," *Folia Phoniatrica*.
- Coleman, R., and Wendahl, R. (1968). "On the validity of laryngeal photosensor monitoring," *J. Acoust. Soc. Am.* **44**, 1733-1735.
- Fant, G., Ondrackova, J., Lindqvist, J., and Sonesson, B. (1966). "Electrical glottography," *Q. Stat. Prog. Rep.*, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm 4/1966, 15-21.
- Fourcin, A. (1974). "Laryngographic examination of vocal fold vibration," in *Ventilatory and Phonatory Control Systems*, edited by B. Wyke (Oxford U. P., London), pp. 315-333.
- Harden, R. J. (1975). "Comparison of glottal area changes as measured from ultrahigh-speed photographs and photoelectric glottographs," *J. Speech Hear. Res.* **18**, 728-738.
- Hutters, B. (1976). "Problems in the use of the photoelectric glottograph," *A. Rep. Inst. Phonet., Univ. of Copenhagen* **10**, 275-312.

- Kitzing, P. (1977). "Methode zur kombinierten photo- und elektroglottographischen Registrierung von Stimmlippenschwingungen," *Folia Phoniatr.* **29**, 249-260.
- Kitzing, P., and Sonesson, B. (1974). "A photoglottographic study of the female vocal folds during phonation," *Folia Phoniatr.* **26**, 138-149.
- Lecluse, F. L. E., Brocaar, M. P., and Verschurre, J. (1975). "The electroglottography and its relation to glottal activity," *Folia Phoniatr.* **27**, 215-224.
- Löfqvist, A., and Yoshioka, H. (1980). "Laryngeal activity in Swedish obstruent clusters," *J. Acoust. Soc. Am.* **68**, 792-801.
- Löfqvist, A., and Yoshioka, H. (1981). "Interarticulator programming in obstruent production," *Phonetica* **38**, 21-34.
- McGarr, N. S., and Löfqvist, A. (1982). "Obstruent production by hearing-impaired speakers," *J. Acoust. Soc. Am.* **42**, 34-72.
- Metz, D. E., Whitehead, R., and Peterson, D. (1980). "An optical-illumination system for high-speed laryngeal cinematography," *J. Acoust. Soc. Am.* **67**, 719-720.
- Rothenberg, M. (1981). "Some relations between glottal air flow and vocal contact area," in *ASHA Reports: Proceedings of the Conference on the Assessment of Vocal Pathology*, Vol. 11, edited by C. L. Ludlow and M. O. Hart, pp. 88-94.
- Smith, S. (1981). "Research on the principle of electroglottography," *Folia Phoniatr.* **33**, 105-114.
- Sonesson, B. (1960). "On the anatomy and vibratory pattern of the human vocal folds," *Acta Oto-laryngol. Suppl.* **156**, 1-80.
- Yoshioka, H., Löfqvist, A., and Hirose, H. (1981). "Laryngeal adjustments in the production of consonant clusters and geminates in American English," *J. Acoust. Soc. Am.* **70**, 1615-1623.