Frequency, Intensity, and Target Matching Effects on Photoglottographic Measures of Open Quotient and Speed Quotient

Article in Journal of speech and hearing research · April 1990
DOI: 10.1044/jshr.3301.45 · Source: PubMed

3 authors, including:

Bruce Gerratt
University of California, Los Angeles
157 PUBLICATIONS 3,647 CITATIONS

Some of the authors of this publication are also working on these related projects:

Towards standardizing voice quality measures View project
Historically, most investigations of the vibratory motions of the vocal folds in phonation have relied on visual observation, via high-speed photography or stroboscopy. More recently, electronic techniques, such as photoglottography (optical glottography) have been used to reflect vocal fold movements because of the relative ease of recording and subsequent analysis of these electronic signals (Baer, Lofqvist, & McGarr, 1983; Baer, Titze, & Yoshioka, 1983; Lofqvist & Yoshioka, 1980; McGarr & Lofqvist, 1988). To interpret the applications of photoglottography (PGG) to phonatory pathology (Hanson, Gerratt, Karin, & Berke, 1988; Hanson, Gerratt, & Ward 1984; Kitzing, 1977, 1986) more data on the normal range of PGG data is desirable. One goal of this study was to provide normative information from PGG measures in male speakers. A second goal addressed a question of comparison between spontaneous phonation and frequency/intensity matched phonation samples.

MEASURES FROM VISUAL STUDIES OF THE GLOTTAL CYCLE

Musehold (1913), using stroboscopy, observed that the glottis may be closed during more than half of the glottal vibratory cycle. The measurement of “Open Quotient” (OQ), has subsequently been used to express the ratio of the open period to the entire glottal cycle’s duration (Tarnoczy, 1951). This physiologically significant measure represents the proportion of the glottal cycle during which airflow can occur. It was not until the development of high-speed cine-photography of vocal fold vibration that the durations of these periods of the glottal cycle could be measured with reasonable accuracy.

Although there have been relatively few high-speed cine-photography studies that have measured the durational relationships of the events in a normal glottal cycle, data indicate that for normal modal phonation, the duration of the closed period tends to comprise less of the glottal cycle as fundamental frequency (F₀) increases, and in the lower- and mid-frequency ranges the closed period tends to be of longer relative duration as vocal intensity is increased (Bell Telephone Laboratories, 1937; Brackett 1948; Farnsworth, 1940; Luchsinger 1954; Luchsinger & Pfister, 1959; Moore & von Leden, 1958; Smith, 1954; Tarnoczy, 1951; Timcke, von Leden, & Moore, 1958).

In addition to the demarcation of open and closed periods, the glottal cycle can be further divided by a period during which the folds are moving apart and a period during which they are returning to the midline (Brackett, 1948). The term Speed Quotient (SQ) was used by Timcke, von Leden, and Moore (1958) to describe the ratio of duration of the opening phase to duration of the closing phase. Timcke, von Leden, and Moore (1958), and Luchsinger and Pfister (1959) found SQ to be independent of fundamental frequency.

Hildebrand (1976) used ultra-high-speed cine-photography to study the effects of frequency and intensity on both OQ and SQ in a systematic analysis of 10 female voices. She saw no significant interactions between OQ and SQ or among the variables of F₀, intensity, and voice training. She found that the OQ increased with rising fundamental frequency in the low- to mid-frequency range and decreased with rising intensity in mid to high intensities. From low to medium F₀, SQ decreased significantly. SQ variations from mid to high F₀ and with change in intensity were not significant for female voices.

MEASURES FROM PHOTOGLOTTOGRAPHY

PGG has been used to obtain measures of the glottal cycle in several studies. Sonesson (1960) examined 10 subjects with PGG during phonation of two different intensities (differing 5 to 6 dB) at five frequencies from...
125 to 325 Hz. The relative duration of the open period (and, therefore, also the OQ) tended to increase as \( F_o \) was elevated, and to decrease with greater intensity. These observations were similar to those made from high-speed cine-studies. Sonesson also noted a significant interindividual variation. Kitzing and Sonesson (1974) used PGG to study 20 female subjects in a systematic study of PGG patterns. Their data tended to confirm the prior observations of Sonesson (1960), and were consistent with the previous high-speed cine-photography studies. For more than 60% of the female subjects, there was no closed period at \( F_o \)s higher than 225 Hz.

A few studies have applied time related PGG measures to vocal pathology. Kitzing (1977, 1986) observed that the perception of vocal strain was associated with elevation of SQ in PGG signals when samples of "strained" and normal phonation were compared. Photoglottography OQ and SQ values measured for patients with abnormal neuromuscular function may vary greatly from values reported for normal voices (Gerratt, Hanson, & Berke, 1986; Hanson, Gerratt, Karin, & Berke, 1988; Hanson, Gerratt, & Ward, 1984). Recent evidence indicates that SQ values derived from PGG signals differentiate among patients with recurrent paralysis, idiopathic vagal paralysis, and normal male speakers (Hanson, Gerratt, Karin, & Berke, 1988). Thus there is some evidence that measures of relative durations of periods of vocal fold movement derived from PGG signals might be clinically useful for objective measurement of vocal physiology.

**Rationale for This Study**

In order to make valid judgments about characteristics that constitute pathological phonatory patterns, more information is needed about the range of PGG patterns in normal voices. In particular, there are very few data on PGG derived measures for normal male speakers. Consequently, one goal of this study was to examine the effects on OQ and SQ of change in vocal intensity and frequency for specific frequency and intensity targets for male voices.

Investigators usually ask subjects to match \( F_o \) and intensity levels in the hope of sampling a range of phonation relevant to a particular research question, as well as preventing these two variables from exerting a contaminating effect on the results. An implicit assumption is that these results can be generalized to normal laryngeal behavior. However, no study has examined how phonation matched to specific frequency and intensity targets compares to phonation produced in a more natural setting in which the subject is given no special phonatory instructions. Therefore, in this study phonation sampled at comfortable spontaneous \( F_o \) and intensity was compared to a phonation produced by the same subject at similar \( F_o \) and intensity while matching a particular frequency/intensity target.

**METHODS**

### Subjects

Twelve adult male volunteers 25 to 45 years of age (\( M = 32 \) years) who were in excellent physical condition were studied. The subjects had no history of voice abnormality and demonstrated no pathology on physical examination by a laryngologist and a speech pathologist.

### Data Acquisition

PGG and electroglottography (EGG) signals were recorded simultaneously to compare the measures. EGG signals were obtained using a Synchrovoice Research Electroglossograph. The device provided a signal proportional to the dynamic impedance between two electrodes located lateral to the thyroid cartilage with a ground electrode on the neck. Both the raw EGG signal and the first derivative of the signal were monitored.

PGG signals were obtained after illuminating the larynx with light projected from a xenon arc lamp via a fiberoptic telescope passed through the nose and situated so that the glottis filled 50% of the viewing field. Light transilluminated through the larynx and neck was transduced by a Centronic single element photo-voltaic detector with an active area of 50 mm², followed by a preamplifier with a bandwidth of approximately 5 kHz. The photodetector and pre-amplifier were encapsulated in plastic and shielded electromagnetically. The light sensor was placed on the skin of the neck over the cricothyroid membrane. The studies were carried out in a light shielded double wall construction sound booth. The voice was transduced by a microphone, placed 10 cm from the lips (Bruel & Kjaer, Model 4144).

### Visual and Auditory Feedback

Each subject was aided in phonating at nine specified frequency/intensity targets by visual and auditory feedback. The acoustic signal and first derivative of the EGG were digitized simultaneously at 20,000 samples per second with 14-bit resolution. The computer calculated \( F_o \) from the prominent negative peaks in the first derivative of the EGG. Sound pressure level in dB was calculated from the calibrated DC output of a microphone measuring amplifier (Bruel & Kjaer, Model 2609). The calculated \( F_o \) and dB values were sent to a computer terminal in the sound booth every 250 ms to inform the experimenter of each subject's matching performance.

At the same time, a voltage proportional to the \( F_o \) was sent from the computer to an XY oscilloscope for visual feedback to the subject. If the subject's \( F_o \) matched the target \( F_o \), the voltage output resulted in a horizontal line positioned in the center of the display. If the subject's \( F_o \) was higher or lower than the target \( F_o \), the horizontal line...
was positioned proportionally above or below the center of the display. For further assistance in the subject’s attempt to match pitch, a sine wave of the target frequency was played through a speaker in the sound field.

Feedback of vocal intensity was provided by the amplitude display on the measuring amplifier. The target intensity was marked on the amplifier display by a narrow line. The subject was instructed to alter vocal intensity until the needle in the amplifier display closely approximated the position of this line target.

**Experimental Task**

Samples during spontaneous phonation, as well as for phonation matching particular frequency and intensity targets, were obtained for the vowel /i/. This vowel was chosen because it is produced with elevation of the epiglottis, allowing unobstructed (monitored via the fiberscope) transmission of supraglottic light for PGG. Each subject was asked to produce the vowel /i/ at comfortable pitch and normal conversational loudness level. This was designated as a spontaneous choice of pitch and loudness. Next, the subject glided from the lowest to highest comfortable pitch within modal register while maintaining a comfortable loudness level. The 25th, 50th, and 75th percentiles of this pitch range were calculated and designated as “low,” “mid,” and “high.” The subject was then instructed to match each of these three $F_o$ levels at his softest and loudest comfortable vocal intensity. The 25th, 50th, and 75th percentiles of the comfortable intensity range at each $F_o$ level were calculated and designated as “soft,” “medium,” and “loud.”

The subject was then asked to produce the vowel /i/ at the three levels of intensity for each of three target frequencies. The experimenter monitored the subject’s performance via a computer terminal display. When both the $F_o$ and intensity were acceptably close to the target, the experimenter instructed the subject to take a short rest before attempting the next target phonation. After a brief trial using the visual feedback clues, the subjects were able to hold the specified target frequency within a range of $\pm 4$ Hz. Intensity targets were held within a range of $\pm 2$ dB.

**Analysis**

One-second portions from the middle portion of stable representative vowel productions were selected for analysis. The recorded PGG, EGG, and acoustic signals were low-pass filtered at 3 kHz, and synchronously digitized at 20,000 samples per second.

Points of glottal opening, peak aperture, and closing were selected on the PGG waveforms interactively, using a method described in detail by Gerratt, Hanson, and Berke (1986). Briefly, the moments of opening and closing were marked with reference to changes in acceleration of the PGG signal. Glottal opening was marked at the positive peak in the third derivative of the PGG occurring in close proximity to the positive peak in the first derivative of the EGG (see Figure 1). Glottal closing was marked at the negative peak of deceleration in the third derivative of the PGG occurring close to the prominent negative peak in the velocity of the EGG. Peak glottal aperture was marked at the peak of the PGG signal. Timing measures were then calculated from the marked photoglottographic signals.

OQ was calculated by dividing the duration of each marked open period by the duration of the corresponding entire glottal period. SQ was determined for each cycle by dividing the duration from the point of opening to the peak of the PGG by the duration from the peak amplitude of the PGG to the point of closing. Mean values and standard deviation of OQ and SQ were calculated for a minimum of 50 consecutive cycles for each sample.

**RESULTS**

**Open Quotient**

OQ provided a measure of the proportion of the glottal cycle during which the edges of the vibratory portions of the vocal folds were separated. Figure 2 displays means and standard deviations of OQ for the targeted phonatory samples at the 25th, 50th and 75th percentile levels of comfortable $F_o$ and vocal intensity for all subjects. Analysis of variance performed with the two within-subjects factors of $F_o$ and intensity indicated a significant difference in OQ with changes in $F_o$ [$F(2,18) = 6.152, p < .05$]. Significant differences in OQ values with variation in intensity, and significant interaction effects were absent. Post hoc repeated contrasts of the three levels of $F_o$. 

![Figure 1. The photoglottographic signal (PGG), the third derivative of the PGG signal (3DPGG) and the first derivative of a simultaneously recorded electroglossiograph signal (1DEGG) from normal phonation of an adult male. The vertical, dotted lines separate the open and closed glottal periods.](image)
revealed significant differences in OQ ($p < .025$) between low and high $F_o$, and between mid and high $F_o$. No significant difference was found for OQ between low and mid $F_o$.

**Speed Quotient**

The durational relationship of opening and closing movements of the vocal folds was expressed as SQ. Means and standard deviations of SQ for targeted phonatory samples at nine combinations of $F_o$ and intensity are shown in Figure 3. Analysis of variance of these data revealed no significant differences ($p > .05$) with variation of $F_o$ or intensity, and no interaction effects. SQ did not change significantly across the range of $F_o$ and intensity targets.

**Spontaneous versus Targeted Phonation Samples**

OQ and SQ data from spontaneous phonations were compared to samples of similar intensity and $F_o$ that were produced during the target matching protocol. Table 1 presents the SQ, OQ, $F_o$, and intensity values of targeted and spontaneous phonation samples. Means and standard deviations for OQ are shown in Figure 4, and data for SQ are shown in Figure 5. Paired samples $t$ tests of these OQ and SQ data revealed significant differences ($p < .05$ for both measures) between the targeted and spontaneous phonations.

**DISCUSSION**

OQ values for individual subjects indicated that the closed period of the glottal cycle comprised a smaller percentage of the period with increasing $F_o$. This is in agreement with previous reports from both cine-photography and transillumination studies. Although the PGG OQ data for male voices indicated that the closed period tended to comprise a larger portion of glottal duration with increase in intensity, there was sufficient variation among subjects that the increase in mean OQ values with intensity was not significant.

In contrast to Hildebrand's (1976) conclusion that SQ changed significantly over the lower range of $F_o$ in female voices, these SQ data for male voices measured from PGG signals did not show a statistically significant trend. Although mean values suggested that SQ increased with intensity, variation was such that the changes were not significant. Several reports have noted that variation in the duration of periods of the vibratory pattern occurred
even though pitch and intensity of the sampled phonations were of the same magnitude (Farnsworth, 1940; Smith, 1954, 1957). These observations suggest that factors other than frequency/intensity are of equal or greater importance in determining the relative durations of phonatory vibratory cycle. In contrast to the lack of significant variation in SQ across $F_0$/intensity change, both OQ and SQ were significantly increased when targeted phonation samples were compared to spontaneous samples.

These results suggest that data from phonation obtained for subjects while matching frequency and intensity targets may not be directly comparable to normal spontaneous phonation. The effects of attempting to match a phonatory target may be a more important variable than factors related to frequency or intensity. The observation is not entirely surprising. The subjects in this study reported that matching of frequency and intensity targets required much greater effort than did spontaneous phonation. Thus, it is possible that the phonatory mode used during the targeted phonation was different from that for spontaneous phonation. These results, however, suggest that generalization of data from studies employing phonatory targets to natural speech should be made with caution.

The relatively small range of change in OQ and SQ for normal phonation seen in targeted phonatory samples contrasts with the range of values reported for pathological phonation. Hanson, Gerratt, Karin, and Berke (1988) found OQ and SQ values derived from PGG significantly different from normal for patients with laryngeal paralysis. PGG samples demonstrated mean SQ values as low as 0.46 for signals from patients with recurrent laryngeal nerve paralysis. SQ values over 2.0 were obtained for patients with vagal nerve paralysis. For such studies of subjects with pathologic phonation, the influence of vocal $F_0$ and intensity on OQ and SQ may be relatively unimportant compared to the large effects associated with particular laryngeal pathologies. Because frequency/intensity effects appear to be relatively small for OQ and insignificant for SQ, it appears that matching of specific $F_0$ and intensity targets in clinical studies may be unnecessary, and even undesirable, for time-related measures of vibratory physiology.

Examination of simultaneous records of PGG and EGG with stroboscopy indicated that PGG signals appeared to provide a more accurate and reliable reflection of opening and closing of the glottic aperture. It should be understood that the “closed period” identified in the PGG signal actually represents a period of approximation of the upper margins of the vocal fold cover. A glottic “chink” posterior to the vocal processes is not uncommon in normal phonation. Thus, closed period should not be assumed to indicate complete absence of air flow. Normally, there is an increase in the transillumination of light through the vocal folds as they thin from below. When there is a persisting glottic chink, this change in light transmission is even greater, and the transition between vertical and horizontal opening movements is less clear. We attempted to reduce the error of marking the point of opening by examining the third derivative of the PGG for a rapid change in the acceleration of the PGG signal. Our studies of simultaneous PGG and stroboscopy (Gerratt, Hanson, & Berke, 1988) indicated that a peak in jerk (change in acceleration) of the PGG signal most closely approximated the point at which the upper vocal margins began to unzip.

In higher $F_0$ range, and in some anatomical configurations, EGG may provide a relatively low signal to noise ratio. Opening velocity of the EGG may for some normal voices not clearly indicate the event of upper fold separation. The EGG signal impedance may also be significantly affected by the mucus cover breaking in strings across the opening glottis. Stroboscopic studies (Gerratt, Hanson, & Berke, 1988) indicate that when this occurs, the opening impedance shoulder of the EGG waveform is rounded and velocity derived estimates of opening from the EGG will be delayed in relation to actual unzipping. Change in acceleration of the PGG signal appeared more...
reliable than the velocity of the EGG signal as an indication of horizontal unzipping of the upper margins of the vocal folds.

It should be emphasized that glottic transillumination can be affected by change in position of the light source, the photosensor, or by interposition of the epiglottis. This is particularly true when the light source has a fairly narrow angle of illumination, as in this study. Change in the level of supraglottic illumination is readily evident when monitoring the PGG signal, and measurements should be made from stable signal configurations. Video monitoring of the glottic image through the fiberoptic telescope used for the light source in this study can assure that artifacts are not introduced by these factors. Movements of the transnasal telescope, however, occur with soft palate contraction and are a problem with the transnasal technique. We have, subsequent to this study, modified the technique of PGG to eliminate the need for transnasal introduction of the light source. We now use a halogen pen flashlight, which is introduced about a centimeter into the oral cavity and focused on the soft palate and posterior pharyngeal wall. This provides a bright diffuse level of supraglottic illumination that is free of AC artifact. The technique appears to be less sensitive to slight changes in positioning of the supraglottic structures. PGG obtained in this way is noninvasive, can be performed by nonmedical personnel, provides reasonably consistent data for normal voices, and appears to offer significant potential for objective phonatory measures in clinical patient populations.

ACKNOWLEDGMENTS

We wish to thank Thomas Chasse for his important contributions to this work, and his efforts in the acquisition and analysis of this data. We also wish to acknowledge the assistance of Rom Karin, Kristin Precoda, and Mark Torres. This work was supported by Grant NS 20707 from the National Institutes of Health and by Veterans Administration Medical Research funds.

REFERENCES


Received December 15, 1988
Accepted July 18, 1989

Requests for reprints should be sent to David G. Hanson, Department of Otolaryngology—Head and Neck Surgery, 303 East Chicago Avenue, Chicago, IL 60611-3008.