Electroglottographic wavegrams: A technique for visualizing vocal fold dynamics noninvasively

Christian T. Herbst
Laboratory of Biophysics, Department of Experimental Physics, Faculty of Science, Palacký University Olomouc, tř. 17. listopadu 12, 771 46 Olomouc, Czech Republic

W. Tecumseh S. Fitch
Laboratory of Bio-Acoustics, Department of Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Wien, Austria

Jan G. Švec
Laboratory of Biophysics, Department of Experimental Physics, Faculty of Science, Palacký University Olomouc, tř. 17. listopadu 12, 771 46 Olomouc, Czech Republic

(Received 4 June 2010; revised 5 August 2010; accepted 25 August 2010)

A method for analyzing and displaying electroglottographic (EGG) signals (and their first derivative, DEGG) is introduced: the electroglottographic wavegram (“wavegram” hereafter). To construct a wavegram, the time-varying fundamental frequency is measured and consecutive individual glottal cycles are identified. Each cycle is locally normalized in duration and amplitude, the signal values are encoded by color intensity and the cycles are concatenated to display the entire voice sample in a single image, similar as in sound spectrography. The wavegram provides an intuitive means for quickly assessing vocal fold contact phenomena and their variation over time. Variations in vocal fold contact appear here as a sequence of events rather than single phenomena, taking place over a certain period of time, and changing with pitch, loudness and register. Multiple DEGG peaks are revealed in wavegrams to behave systematically, indicating subtle changes of vocal fold oscillatory regime. As such, EGG wavegrams promise to reveal more information on vocal fold contacting and de-contacting events than previous methods.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3493423]

PACS number(s): 43.70.Jt, 43.70.Gr, 43.75.Rs [AL]

Pages: 3070–3078

I. INTRODUCTION

Electroglottography was invented by Fabre (1957) to monitor the vibration of vocal folds in vivo. A low intensity, high-frequency current is passed between two electrodes that are placed externally on the neck, on each side of the thyroid cartilage at vocal fold level. The contacting and de-contacting of the vocal folds causes variations in the electrical impedance across the larynx, resulting in a variation in current between the two electrodes (Baken, 1992; Fourcin and Abberton, 1971). This variation in the current flow has been found to be related to changes in vocal fold contact area (Scherer et al., 1988).

The electroglottographic signal is a time-varying one-dimensional representation of the complex three-dimensional motion of the vocal folds. Experimental research (Baer et al., 1983; Childers et al., 1983) has confirmed a close relation between peaks in the derivative of the EGG signal and the contacting and de-contacting events of the vocal folds. It has been shown that landmarks in the EGG waveform are related to the movement and position of the vocal folds during phonation (Hess and Ludwigs, 2000; Rothenberg, 1979). The physiologic relevance of the EGG signal has been examined theoretically by Titze (1989, 1990), who discussed the effects of (a) increased glottal adduction; (b) glottal convergence and vertical phasing; (c) medial vocal fold surface bulging; and (d) increased vertical phasing in vocal fold vibration.

The first mathematical derivative of the EGG waveform (DEGG) reflects the rate of change of the EGG with time (Childers and Krishnamurthy, 1985; Teeny and Fourcin, 1980). The timing of glottal closure can often be derived from a single maximum in the DEGG signal (Childers et al., 1983). However, in some subjects, multiple peaks for both the contacting and the de-contacting phase can be seen. It has been suggested that the multiple peaks in the DEGG signal might be related to zipper-like opening and closing of the vocal folds (either anterior-to-posterior or posterior-to-anterior) (Henrich et al., 2004; Hess and Ludwigs, 2000). It is also conceivable that they result from artifacts, such as mucus strands (Colton and Conture, 1990), or abnormalities in vocal fold tissue structure distorting the regularity of the EGG waveform (Baken and Orlikoff, 2000; Kitzing, 1990). Such DEGG multiple peaks phenomena have, however, not yet been investigated in detail.

Quantitative analysis of the EGG waveform has been achieved by measuring the relative proportion of glottal...
 closure within a glottal vibratory period (Rothenberg and Mahshie, 1988), known as the “larynx closed quotient” (Howard, 1995) or “contact quotient” (Orlikoff, 1991) (CQ\textsubscript{EGG}). This quotient has been found useful in clinical as well as basic voice research (Schutte and Miller, 2001; Švec et al., 2008). A sudden decrease of CQ\textsubscript{EGG} has been reported, e.g., during the transition from chest/modal to falsettoregister in singing (Henrich et al., 2005; Miller et al., 2002). Research has shown, however, that the CQ\textsubscript{EGG} is dependent on the choice of algorithm used to determine the contacting and de-contacting events, and must therefore be used with caution (Herbst and Ternström, 2006; Higgins and Schulte, 2002; Kania et al., 2004; Sapienza et al., 1998).

In order to study longer sequences of speech and singing, hundreds of glottal cycles need to be examined, especially when investigating phonations with changing physiologic parameters. In such cases, entire EGG signals are often represented by time-varying analysis parameters (e.g., EGG contact quotient or EGG signal amplitude). Since a single analysis parameter does not reflect all nuances seen in the EGG waveform, valuable information might be disregarded in the signal processing stage.

The need of extracting meaningful information from large amounts of data has been addressed by several authors. For better visualization of signals conveying information concerning phonation, methods have been conceived which segment the signal into single glottal cycles and show how these cycles change over time. Among those techniques are: the “dreedimensionale Periodizitättsanalyse” (3D-PAN), for analyzing acoustic signals (Sedlacek and Michek, 1988); the voice cascademethod for analysis of acoustic and EGG signals (Berg and Gäll, 1999a, 1999b); videokymography (VKG) for documenting vibratory pattern of the vocal folds (Švec and Schutte, 1996; Švec et al., 2007); and phonovibrography for reducing the amount of data gathered in high-speed imaging (Lohscheller et al., 2008). The advantage of these approaches is that ongoing developments and gradual changes of laryngeal behavior can be seen in one single image.

In the present study, a similar approach for data reduction of the electrolaryngographic signal will be presented. A new method for analyzing and displaying EGG signals and their first derivative (DEGG) is introduced, which (a) allows monitoring the EGG (or DEGG) signal over time within a single image; and (b) provides an intuitive means for quickly assessing the vocal fold contact phase and its variation over time.

II. ASSEMBLING THE WAVEGRAM: METHOD

An algorithm has been developed to show how the EGG waveform changes over time, from cycle to cycle, in one single image (called a “wavegram”). An EGG signal (or a portion thereof) is taken and converted into a graph where time is displayed on the x-axis from left to right; the progress of consecutive glottal cycles is displayed on the y-axis from bottom to top; and the locally normalized vocal contact area is encoded in the z-axis as a color intensity value. The algorithm (implemented with a C++ signal processing library written by author CH) consists of the following stages: (1) detection and separation of EGG cycles, (2) normalization of the amplitudes of the EGG cycles, (3) coding the normalized EGG values with color intensity, (4) concatenation of the color-intensity-coded EGG cycles, and (5) normalization of the period length to obtain a uniform graph display. These steps are described below in detail.

A. Detection of glottal cycles

To detect glottal cycles in quasi-periodic phonation, the period duration of the input signal is initially determined by performing an auto-correlation analysis as described by Boersma (1993). The part of the EGG (or DEGG) signal that represents single glottal cycles is defined and determined by repeatedly cross-correlating an ideal EGG waveform “template” [taken from Titze (1990)] with the actual EGG (or DEGG) signal (see Fig. 1). During this process, the templates are first stretched to the approximate cycle duration (as determined by the initial auto-correlation analysis), and then each template is correlated with a region of the EGG signal having twice the size of the computed glottal cycle duration. The actual beginning of each glottal cycle is defined as the beginning of the ideal fit.

For non-periodic phonation and EGG waveforms deviating from the ideal EGG waveform template (e.g., samples obtained by dysphonic speakers), the method as described above might fail. In such a case, the period should be rather determined on a cycle-to-cycle basis from direct inspection of the electrolaryngographic signal (and its first derivative) in the time domain (Fourcin and Abberton, 2008). Such an alternative algorithm for glottal cycle detection is described in the Appendix.

B. Normalization and color-coding

Every single extracted glottal cycle is locally normalized in amplitude (ordinate). The amplitude values ($y_j$) are then coded into monochrome color information:

\[
\text{col}_j = 255(1 - y_j), \quad j \in [1 \ldots n],
\]  

FIG. 1. (Color online) Extraction of one glottal cycle. Ideal EGG waveform templates are (a) stretched to the estimated glottal cycle duration, and (b) correlated with a portion of the EGG signal having twice the size of the estimated glottal cycle (the “region of interest”). The beginning of the glottal cycle is defined as the beginning of the ideal waveform fit.
where \( n \) is the size of the glottal cycle, \( 0 \leq y_j \leq 1 \) is the locally normalized EGG amplitude (which is proportional to the relative vocal fold contact area); and \( col_j \) is the resulting monochrome color information. Here, high values correspond to dark colors. Apart from the normalization, the color coding process is, in principle, similar to the one used for converting the consecutive sound spectra to sound spectrograms (Koenig et al., 1946). The process and the resulting color-coded EGG cycle is displayed in Fig. 2.

C. Cycle-concatenation and normalization of the final display

The color-coded strips, corresponding to individual glottal cycles are rotated by 90° counter-clockwise [Fig. 3(c)]. The height of the individual cycle plots corresponds now to the period duration, and a specific point along the \( y \)-axis represents a particular phase of the glottal cycle. The width of the individual strips can be proportionally squeezed to display as many waveforms as possible, allowing us to represent the whole phonation in a single image. As a final step, the heights of the individual cycle plots (i.e., period durations) are normalized by means of interpolation to form the final graph. As shown later, this height normalization is useful for investigating modifications of the events occurring within the glottal cycle. In the resulting wavegram, time is displayed on the \( x \)-axis; normalized progress of consecutive glottal cycles is displayed on the \( y \)-axis; and normalized vo-
calfold contact area within a cycle is shown on the z-axis by means of (monochrome) color intensity [see Fig. 3(d)].

As an alternative display option, the first derivative of the EGG signal (DEGG) can be used as an input (rather than the EGG signal itself) to the algorithm of generating the wavegram—see Fig. 4. DEGG-based wavegrams provide a clearer view of the moments of most rapid change in the vocal fold contact area and allow us to track multiple peaks in both the contacting and de-contacting phase of the DEGG-signal (see later in Fig. 7).

In order to reduce the resulting graph width, an optional data reduction task can be performed by only considering glottal cycles at user-defined time intervals. This method of data reduction was used for the examples shown in Figs. 6–10 in Sec. III and the Appendix of this manuscript.

III. RESULTS

Wavegrams for a sustained stable phonation (sung by a male singer) are shown in Fig. 5 on the right. On the left side of the figure, there are the corresponding EGG (dark color) and DEGG (light color) waveforms. The following landmarks (Childers and Krishnamurthy, 1985; Rothenberg, 1979) for the typical EGG waveform are identified in the figure:

1. Start of the glottal cycle (x=0); this landmark has been chosen arbitrarily around the moment of minimum vocal fold contact area in order to be able to display the entire contact phase of the glottal cycle in the EGG wavegram.

2. Initiation of glottal closure (x=0.13), identified by a minor, but pronounced change in EGG signal steepness and a secondary positive peak in the DEGG signal. The EGG wavegram shows a minute change in color intensity (light to dark in upward vertical dimension). In the DEGG wavegram this phenomenon is documented as a fine line of gray color.

3. Maximum increase of vocal fold contact (x=0.18), identified by the positive peak of the DEGG signal. The EGG wavegram exhibits an abrupt change in color intensity (light to dark in upward vertical dimension). In the DEGG wavegram the positive peak of the DEGG signal appears as a line of dark gray color.

4. Maximum decrease of vocal fold contact (x=0.57), identified by the negative peak of the DEGG signal. This event in the glottal cycle appears in the DEGG wavegram as a bright line. In the EGG wavegram, this event is not clearly visible, since the de-contacting phase of a glottal cycle is distributed over a longer period of time.

5. End of glottal cycle.

The temporal locations of the strongest positive and negative peaks of the DEGG signal can be used to calculate the EGG contact quotient (Henrich et al., 2004). Other methods of calculating the EGG contact quotient, e.g., based on a criterion level, can also be used. One has to be aware, however, that these different methods are likely to produce different results, based on algorithm settings (Herbst and Ternström, 2006). When applying the DEGG criterion to the
The presence of multiple glottal cycles displayed in Fig. 5, a contact quotient of ca. 0.39 is calculated. The DEGG wavegram shows this clearly, since the positive and negative peak of the DEGG signal are displayed as distinct lines of darker and lighter color, respectively.

In Fig. 6, phonation sustained at a stable fundamental frequency of approximately 233 Hz with increasing vocal intensity is shown (female classical singer, 4 years of academic vocal training). The DEGG wavegram [Fig. 6(c)] reveals a steady change from phonation with a shorter duration of glottal closure to phonation with a longer duration of glottal closure (see Fig. 6(a)): EGG waveforms extracted at $t=1.5$ s and $t=5$ s, respectively). The gradual change in both the contact phase and the amplitude of the EGG waveform [Fig. 6(c)] suggests that no abrupt change in vocal fold vibratory pattern has occurred (Roubeau et al., 2009).

Phonation involving increasing and decreasing vocal intensity (a so-called messa-di-voce) produced at a stable pitch (F#3, ca. 185 Hz) by an untrained male amateur singer is shown in Fig. 7. The most significant difference from the phonation displayed in Fig. 6 is the presence of multiple converging DEGG peaks in the contacting phase of vocal fold vibration at lower intensity levels [Fig. 7(e)]. At $t \approx 4.5$ s, after the converging DEGG peaks have fully merged, a sudden increase of vocal fold contact phase occurred. This is reflected by the fact that the strongest negative peak of the DEGG signal [Fig. 7(a)] extracted at $t=5$ s occurred at a later stage of the glottal cycle, as compared to the strongest negative DEGG peak of a glottal cycle extracted at $t=4.3$ s. A reversed phenomenon was seen around $t=7.5$ s, where an abrupt shortening of glottal contact phase is accompanied by the separation of multiple DEGG peaks in the contacting phase (see ellipses in Fig. 7).

This relationship between multiple peaks and other aspects of vocal fold behavior, as often observed in wavegrams, suggests that such peaks reveal subtle aspects of vocal fold dynamics.

In Fig. 8, phonation with increasing fundamental frequency (ca. 208–415 Hz) is illustrated in a wavegram (female classical singer, 2 years of academic vocal training). Analysis data revealed a sudden change of EGG contact quotient around $t=0.95$ s. This coincided with (i) a sudden increase of vocal fold contact phase.
change of vocal fold contact phase [EGG waveforms and wavegrams, Figs. 8(a)–8(e)]; (ii) a sudden change of overall EGG signal amplitude [Fig. 8(e)] (Roubeau et al., 2009); (iii) an abrupt pitch change [Fig. 8(f)], that was accompanied by an audible change of vocal timbre; and (iv) a sudden decrease of sound pressure level by 6 dB (not shown in figure). The change of EGG waveform happened over a period of ca. 0.028 s (7 glottal cycles).

Phonation with increasing fundamental frequency (ca. 208–415 Hz) with a continuous register transition is shown in Fig. 9 (female classical singer, 4 years of academic vocal training). The wavegram [Figs. 9(b) and 9(c)] exhibited no abrupt change of glottal contact phase. This was corroborated by a stable overall EGG signal amplitude [Fig. 9(e)] and smooth changes of fundamental frequency [Fig. 9(f)]. Careful examination of the acoustic signal revealed a barely audible timbral change around $t=2$ s, which did, however, not coincide with a significant change of sound pressure level. The EGG waveform [Fig. 9(a)] gradually changed from a chest phonation ($t=0.5$ s) with a strong “knee,” i.e., a “bump” in the de-contacting phase of the EGG waveform where the signal amplitude is beginning to decrease more rapidly (Hess and Ludwigs, 2000), to a falsetto-like waveform ($t=3.5$ s, see Fig. 9). In particular, the knee was gradually dissipating. From $t=0.8$ s to $t=2.8$ s, two negative DEGG peaks were seen. The evidence suggests that—unlike the phonation shown in Fig. 8—no abrupt change of vocal fold oscillatory regime took place, i.e., the transition from chest to falsetto register was accomplished gradually, as intended.

IV. DISCUSSION AND CONCLUSIONS

The wavegram technique provides a new and potentially powerful method for displaying entire electroglottographic signals, or parts thereof. Much like in the spectrogram, information on vibratory behavior developing in time is compacted into one single graph providing insight into changes of vocal fold dynamics. Nevertheless, waveform details of individual glottal cycles (and their gradual development over time) are preserved, thus providing a useful tool to quickly gain physiologic insights into longer bouts of phonation than are visible in a simple amplitude graph.

The wavegram reveals changes of vocal fold contact phase, as well as of phenomena that cannot easily be seen with other methods for displaying electrolaryngographic signals/waveforms. Those phenomena, which develop over time, are expected to be related to physiologic behavior of vocal fold vibration, the nature of which has not yet been fully explored.

When analyzing electrolaryngographic signals, derivative analysis parameters such as the contact (CQ) or the contact index (CI) are calculated. Those consist of only one time-varying variable, the calculation of which is dependent on
user inputs which are partially arbitrary, e.g., the specification of a certain threshold value (Herbst and Ternström, 2006). In the waveform, on the other hand, the EGG waveform is treated “as is.” Apart from period and alignment of glottal cycles, no absolute values are calculated. Thus, the waveform constitutes an alternative to the contact quotient monitoring, since it provides additional information, and does not depend on any arbitrary threshold criterion. Our observations suggest that this can provide novel insights into details of vocal fold behavior that are easily overlooked in other analysis techniques.

As a prominent example, converging multiple DEGG peaks in the contacting phase appear to constitute a systematic and consistent phenomenon seen in a considerable proportion of subjects (an example is shown in Fig. 7). The same is true for multiple DEGG peaks in the de-contacting phase, which may be related to gradual changes in the vocal fold oscillatory regime. The consistent behavior of these peaks over time provides evidence that they reveal physiologic phenomena, and should not be considered artifacts in the strict sense. In particular, multiple DEGG peaks are induced by vocal fold vibratory phenomena during which vocal fold contact abruptly increases (in the contacting phase) or decreases (in the de-contacting phase). The exact universal interpretation of these phenomena has not yet been available. It can be speculated that they are related to phase differences (in the superior-inferior as well as the anterior-posterior vocal folds dimension), commonly termed “zipper-like” vocal fold opening or closure (Childers et al., 1986). Further research investigating the physiological relevance of these phenomena is warranted.

The waveform reveals that vocal fold contacting and de-contacting “events” are more complex than commonly assumed. Rather than a single incident, vocal fold contacting and de-contacting should be considered a sequence of events, taking place over a certain period of time. The analyses presented here indicate that this sequence of events can change with fundamental frequency, loudness and register. The EGG signal appears to convey more physiological information on vocal fold contacting and de-contacting events than what is offered by more traditional representations of the EGG signal. The waveform technique promises to provide a method to visualize, further explore and understand this “hidden” information.

The waveform method focuses on the contacting and de-contacting phenomena occurring within individual glottal cycles and monitors their changes as the phonation progresses. As such, the method is not primarily intended for detecting pathologies, but rather for better understanding the physiological phenomena occurring during vocal fold vibration. The potential limitations of the waveform technique in its current state are twofold: (a) It is dependent on proper period detection, which is problematic in non-periodic/ pathologic phonations (Titze, 1995); (b) The waveform is primarily intended for visual analysis, and it offers no automatic quantitative measurement for relevant signal fluctuations. A future version of the algorithm might offer an option to skip the normalization of the glottal period duration, thus revealing period perturbations such as jitter (Vieira et al., 1996; Vieira et al., 2002), shimmer or vibrato. Further research including experiments with recordings from dysphonic speakers is necessary in order to test the applicability of the waveform technique in pathologic voices.

Electroglottography is a relatively inexpensive and non-invasive technique. It holds a great potential for singing pedagogy (Herbst et al., 2010; Howard et al., 2004; Miller and Schutte, 1999; Rossiter et al., 1996), and possibly for detecting voice disorders (Baken and Orlikoff, 2000; Smith and Childers, 1983) and in voice therapy (Baken and Orlikoff, 2000; Fourcin et al., 1995). The novel waveform technique, presented in this paper, enables us to better tap the rich potential of the EGG waveform. Waveforms allow assessment and analysis of the EGG waveforms in more detail, and promise to enhance our understanding of the EGG signal and vocal fold vibration.

ACKNOWLEDGMENTS

This research was supported by the Grant Agency of the Czech Republic, project GAČR 101/08/1155, and by the

FIG. 9. Changes of the EGG and audio signals in case of female phonation with increasing fundamental frequency. (a) EGG (black) and DEGG (gray) waveforms representing glottal cycles extracted at \( t = 0.5 \) s, \( t = 1.5 \) s, \( t = 2 \) s, \( t = 2.25 \) s, \( t = 2.5 \) s, and \( t = 3.5 \) s, respectively; (b) EGG waveform graph; (c) DEGG waveform graph; (d) and (e) amplitude plot of audio and EGG signal; (f) fundamental frequency displayed in musical notation (ca. 208–415 Hz). The data reveals a smooth transition from chest to falsetto register, indicated by (i) a gradual change of the time-varying EGG waveform representing one glottal cycle; (ii) a lack of abrupt change of glottal closure duration in the waveform; (iii) a steady EGG signal amplitude; and (iv) no abrupt (involuntary) pitch changes.
ERC Advanced Grant “SOMACCA”. We thank the anonymous reviewers for their insights and their helpful feedback.

APPENDIX

When creating a wavegram, the start and end of a glottal cycle can be chosen arbitrarily. In doing so, it is preferable that the entire contact phase lies within one cycle, i.e., the contact phase is not spread over two consecutive glottal cycles. To achieve this, two options have been implemented in the wavegram algorithm: (a) by correlating an ideal EGG waveform template with the analyzed EGG signal (as described above); or (b) by computing local positive maxima within the first derivative of the EGG signal, i.e., the DEGG signal. (The region of interest is one glottal period, the duration of which is known from F0 extraction.) The start of each glottal cycle is located at an arbitrary offset of the (normalized) glottal cycle duration, preceding each computed positive maximum in the DEGG signal. Offsets of 15% to 20% proved to give visually the best results, ensuring that even EGG waveforms with a very long contact phase (80% of the cycle duration) could be displayed in one entire glottal cycle.

The DEGG strongest peak alignment method provides the most intuitive information on the duration of the contact phase, since the contacting events within each glottal cycle (as indicated by the strongest positive maximum of the DEGG signal) are aligned to form a straight dark horizontal line in the wavegram (see Fig. 10(c)). This approach fails, however, if there are multiple positive peaks in the glottal contact phase (i.e., the phase of a glottal cycle where the vocal fold area increases). This is particularly true if the amplitudes of the peaks vary (see for example the two EGG waveforms displayed in Fig. 8: at t = 10.8 s the second DEGG peak is the strongest peak, and at t = 11.2 s the first DEGG peak is stronger). In such a case, the alignment of the glottal cycles within the wavegram would abruptly change. Therefore, the “correlation method” has been chosen as the preferred default setting of the wavegram algorithm. In the current software, the “DEGG strongest peak method” has been implemented as a well, allowing us to explore both display options as needed.


Berg, R., and Gall, V. (1999a). “Physiological and pathophysiological findings in glottal segmentation of voice-onset and voice-offset,” in Commu-


