

On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation

Nathalie Henrich^{a)}

LAM (UPMC, CNRS, Ministère de la culture), 11 rue de Lourmel, 75015 Paris, France
and LIMSI-CNRS, BP 133, F91403 Orsay, France

Christophe d'Alessandro and Boris Doval

LIMSI-CNRS, BP 133, F91403 Orsay, France

Michèle Castellengo

LAM (UPMC, CNRS, Ministère de la culture), 11 rue de Lourmel, 75015 Paris, France

(Received 26 March 2003; revised 28 October 2003; accepted 8 December 2003)

Electroglottography is a common method for providing noninvasive measurements of glottal activity. The derivative of the electroglottographic signal, however, has not attracted much attention, although it yields reliable indicators of glottal closing instants. The purpose of this paper is to provide a guide to the usefulness of this signal. The main features that are to be found in this signal are presented on the basis of an extensive analysis of a database of items sung by 18 trained singers. Glottal opening and closing instants are related to peaks in the signal; the latter can be used to measure glottal parameters such as fundamental frequency and open quotient. In some cases, peaks are doubled or imprecise, which points to special (but by no means uncommon) glottal configurations. A correlation-based algorithm for the automatic measurement of fundamental frequency and open quotient using the derivative of electroglottographic signals is proposed. It is compared to three other electroglottographic-based methods with regard to the measurement of open quotient in inverse-filtered derived glottal flow. It is shown that agreement with the glottal-flow measurements is much better than most threshold-based measurements in the case of sustained sounds. © 2004 Acoustical Society of America. [DOI: 10.1121/1.1646401]

PACS numbers: 43.75.Yy, 43.70.Jt, 43.75.Rs [SM]

Pages: 1321–1332

I. INTRODUCTION

Electroglottography (EGG), invented by Fabre in 1956 (Fabre, 1957, 1958, 1959, 1961), is a common, widespread technique that enables the investigation of vocal-fold contact area in phonation in an easy and noninvasive way. A high-frequency modulated current ($F \approx 1$ MHz) is sent through the neck of the subject, as shown in Fig. 1. Between the electrodes, electrical admittance varies with the vibratory movements of the vocal folds, increasing as the vocal folds increase in contact. For a more detailed review, we refer the reader to Childers and Krishnamurthy (1985), Colton and Conture (1990), and Orlikoff (1998).

In order to analyze the EGG signal, several comparative studies have been conducted using stroboscopic photography (Fourcin, 1974; Lecluse *et al.*, 1975; Pedersen, 1977; Teaney and Fourcin, 1980), videostroboscopy (Anastaplo and Karnell, 1988; Karnell, 1989), high-speed cinematography (Baer *et al.*, 1983a; Childers *et al.*, 1990; Childers and Krishnamurthy, 1985; Childers and Larar, 1984; Childers *et al.*, 1983b, 1984), photoglottography (Baer *et al.*, 1983a,b; Berke *et al.*, 1987; Dejonckere, 1981; Gerratt *et al.*, 1988; Kitzing, 1977, 1983; Kitzing *et al.*, 1982; Titze *et al.*, 1984), subglottal pressure measurements (Kitzing *et al.*, 1982) and inverse filtering (Childers *et al.*, 1983b; Fourcin, 1981; Rothenberg, 1981; Rothenberg and Mahshie, 1988). All of these studies

confirm that the EGG signal is related to the vocal-fold contact area: the larger the contact surface, the larger the measured admittance. Lecluse and Brocaar (1977) and Childers and Krishnamurthy (1985) have proposed a schematic description of the EGG signal during one glottal period produced by laryngeal mechanism 1 on the basis of comparisons with glottal visualization (we refer the reader to the Appendix for details concerning laryngeal mechanisms and the terminology used in this paper). We can draw on their observations to provide a schematic description of the EGG signal in the case of laryngeal mechanisms 1 and 2, as shown in Fig. 2.

During a vocal-fold vibratory cycle, the corresponding EGG signal can be described by four main phases.

①–③: Closing phase. Contact is initiated along the lower margins of the vocal folds (① to ②), then propagating to the upper margins (② to ③). As closing is generally faster than opening, this phase is characterized by a steep slope in the EGG signal. The instant of maximum slope can be found at ②, which corresponds to a strong positive peak in the derivative of the EGG signal (DEGG signal).

③–④: Closed phase. The vocal folds are in full contact, preventing air from passing through the glottis. The slight increase and decrease observed in the signal could be due to the elastic collision of tissue (Childers and Krishnamurthy, 1985).

^{a)}Electronic mail: henrich@lam.jussieu.fr

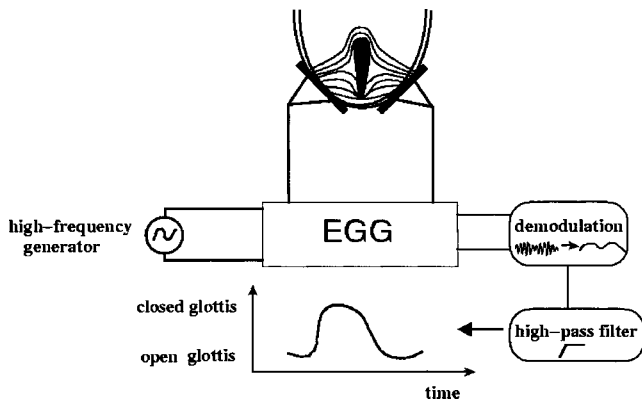


FIG. 1. Principle of an electroglottograph (EGG), using the “vocal fold contact area” (VFCA) convention in which the EGG signal is represented as a function of vocal-fold contact.

- ④–⑥: Opening phase. The lower margins of the vocal folds begin to separate gradually (④ to ⑤), followed by separation along the upper margins (⑤ to ⑥). The instant of maximum slope can be found at ⑤, which corresponds to the negative peak in the DEGG signal.
- ⑥–①: Open phase. The vocal folds are apart. A relatively flat signal is observed, as there is little variation in the admittance.

The closing and opening phases are illustrated in Fig. 3 by simultaneous visualization of high-speed images,¹ the corresponding electroglottographic signal, and its derivative during nonpathological male phonation (mechanism 1, $f_0 = 110$ Hz). The physiological correlates of the peaks observed on the DEGG signal in ② (“closing peak”) and ⑤ (“opening peak”) were mainly studied by Childers *et al.*, using simultaneous and synchronized measurements of EGG and DEGG signals, inverse-filtered derived glottal flow and glottal area measured from ultrahigh-speed cinematography (Childers *et al.*, 1990; Childers and Krishnamurthy, 1985; Childers and Larar, 1984; Childers *et al.*, 1983a,b), along with a theoretical study of the relation between the EGG/DEGG signals and the glottal contact area calculated from a physical model (Childers *et al.*, 1986). In the case of non-pathological productions, they showed that these peaks are related to the instants of glottal opening and closing, defined as the instants of initialization and termination of glottal area

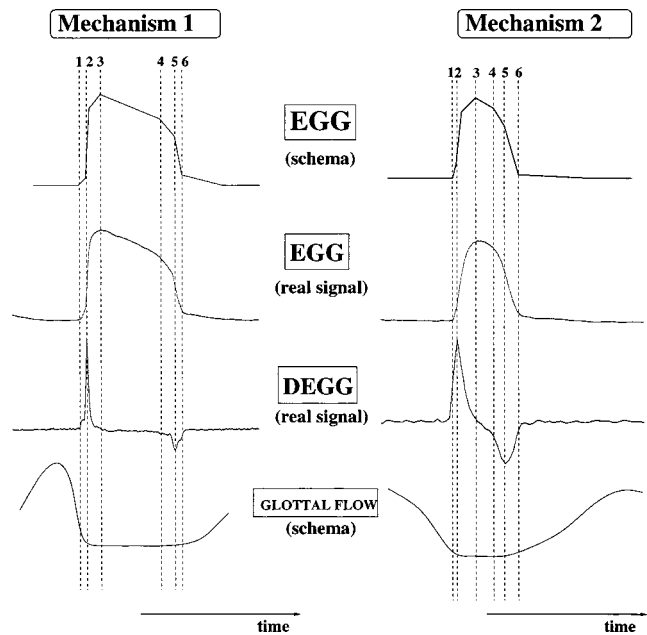


FIG. 2. Illustration of a glottal duty cycle as seen on an EGG and a DEGG signal. (1)–(3) closing phase; (3)–(4) closed phase; (4)–(6) opening phase; (6)–(1): open phase. As a comparison, one glottal flow period is presented.

variation. But, this is only valid for a normal male voice in mechanism 1. A special case was brought to light by these authors: in the case of a mucus strand bridging the glottis during the opening phase, the opening peak occurs at the instant when the mucus strand breaks, rather than at the instant of glottal opening. Colton and Conture (1990) showed that the presence of a mucus bridge affected the EGG signal but also the inverse-filtered glottal flow. The glottal flow does not increase much until the mucus strand breaks, and thus the opening peak occurs when glottal flow starts to increase greatly.

Other studies confirmed these results. Berke *et al.* (1987) made simultaneous measurements of EGG, photoglottography, subglottal pressure, and stroboscopic photography “*in vivo*” in dogs, and they observed that slope variations in the EGG signal were correlated with the opening of the vocal folds along their upper margin and with the complete closure of the lower margin. Anastaplo and Karnell (1988) obtained similar results concerning the glottal open-

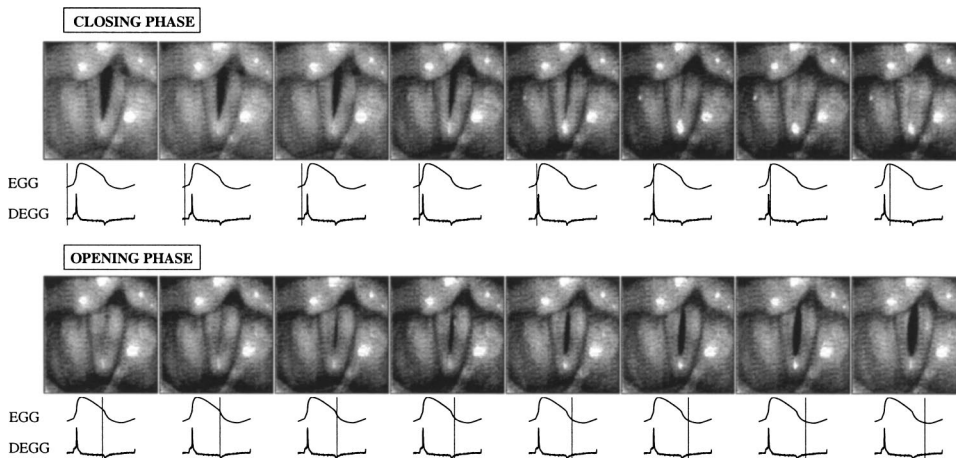


FIG. 3. Visualization of the closing and opening phases by simultaneous high-speed and electroglottographic recordings. Vertical bars indicate the moment in time at which the visual image occurs with respect to the EGG and DEGG signals. This example corresponds to a nonpathological male phonation in mechanism 1, with $f_0 = 110$ Hz. The EGG sampling frequency is 44 444 Hz and the high-speed camera sampling frequency is 3704 frames/s.

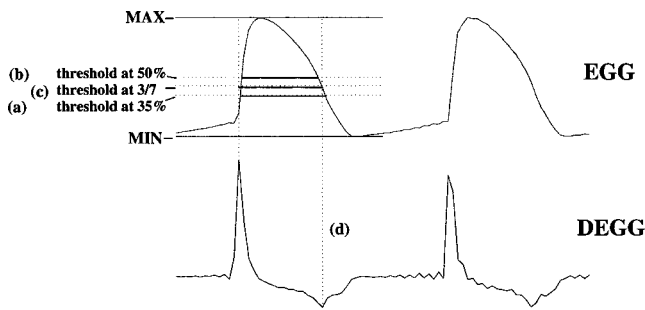


FIG. 4. Illustration of four methods for measuring the open quotient (or its equivalent, the closed quotient) on an EGG signal and its derivative: (a) on the EGG signal using a threshold detection (35% of the difference between the minimum and maximum value of the signal over a glottal period); (b) same method as (a) with a threshold of 50%; (c) in combining an EGG-based threshold method (3/7) with a detection of glottal closing instants on the DEGG signal; (d) on the DEGG signal by the detection of opening and closing peaks.

ing instant, using EGG and synchronized videostroboscopy in 4 subjects (2 males, 2 females). By comparing EGG, photoglottographic, and inverse-filtered derived glottal flow in V-C-V samples, Cranen (1991) showed that DEGG peaks were related to the peaks observed in the derivative of photoglottographic signals and to the opening and closing instants measured on the glottal flow and its derivative.

On the substantial basis of these studies, DEGG peaks may be considered as reliable indicators of glottal opening and closing instants, the latter being defined by reference to the glottal flow, as the instants when the flow starts to increase greatly from the baseline (opening) and decrease greatly to the baseline (closing). As glottal closing is usually abrupt, the closing peak is often very strong and precise. The opening can be less precise, which is reflected in a weaker opening peak. This may explain the reservations expressed by Baken (1992) on the use of these peaks as a measure of glottal opening instant.

The first part of the present study offers a characterization of DEGG signals from an extensive database of 18 trained singers. The recording procedure is briefly presented in Sec. II A; its main features are brought out in Sec. II B and discussed in Sec. II C.

The second part of this study concerns the use of DEGG signals for the measurement of fundamental frequency and open quotient. This is one major application of the detection of glottal closing and opening instants. The duration between two consecutive glottal closing instants corresponds to a fundamental period; its inverse gives the fundamental frequency of the voice. In the same manner, the duration between the glottal opening instant and the consecutive glottal closing instant corresponds to the open time. The open quotient can then be derived from these two measures as the ratio between open time and fundamental period.

In the literature, the main electroglottographic-based methods for detecting the glottal closing and opening instants use the EGG signal and not its derivative. As illustrated by cases (a) and (b) in Fig. 4, a threshold method is generally applied (Rothenberg and Mahshie, 1988): a level line is chosen either as a percentage of the amplitude between the minimum and maximum of the signal over a glot-

tal period (50% for a normal to pressed voice and 35% for a relaxed voice). The crossing points between the level line and the EGG signal are approximated as the instants of glottal closing and opening, and these are used for the open quotient measurement. These methods are very convenient for medical purposes, as they are robust and can be applied even on noisy or weak signals. Nevertheless, the results of such methods are by their nature imprecise and can be inaccurate, as compared to what would be measured on a glottal area signal or on a glottal flow signal (Rothenberg and Mahshie, 1988).

An improvement over such a method has recently been proposed by Howard [Howard (1995); Howard *et al.* (1990)]: using the DEGG closing peak for detecting the glottal closing instant, and an EGG-based threshold method for detecting the glottal opening instant. This is illustrated by case (c) in Fig. 4. This method should give more accurate results, because of reliable detection of the glottal closing instants. In the present study, we propose a method based on the DEGG signals only, in the case of precise opening and closing peaks. This correlation-based method is described in Sec. III. The measurements of open quotient made on EGG, DEGG, and glottal-flow signals are compared and discussed in Sec. III D.

II. DESCRIPTION OF THE DEGG SIGNALS EXTRACTED FROM A SINGING DATABASE

A. Material

A corpus of sung items designed for the purpose of open quotient measurements was recorded by 18 trained singers. The singers were classically trained and most of them were professional singers, insofar as they earn their living from singing. Seven were baritones (subjects B1 to B7), 2 were tenors (T1 and T2), 3 were countertenors (CT1 to CT3), 3 were mezzosopranos (MS1 to MS3), and 3 were sopranos (S1 to S3). The database comprises several parts; the present study only draws on the part where the singers were asked to produce glissandos, crescendos, and sustained vowels for different pitches and at different loudness levels. Three vowels were used: (French) [a], [e], and [u]. When necessary, the singers were asked to indicate in which mechanism (M1 or M2) they were singing.

The singers were placed in a soundproof booth. The acoustic and electroglottographic signals were recorded simultaneously on the two channels of a DAT recorder (PORTADAT PDR1000), with a sampling frequency of 44.1 or 48 kHz. A two-channel electroglottograph [EG2, Rothenberg (1992)] was used for the electroglottographic recordings. The low-frequency cutoff was set to 20 Hz and no high-frequency limit was applied. The technical computing environment MATLAB was used for digital signal processing.

More details about the recording procedure, the subjects, and the tasks which are not relevant for the understanding of this study are presented in a subsequent paper concerning the results of open quotient measurements in singing (Henrich *et al.*, submitted).

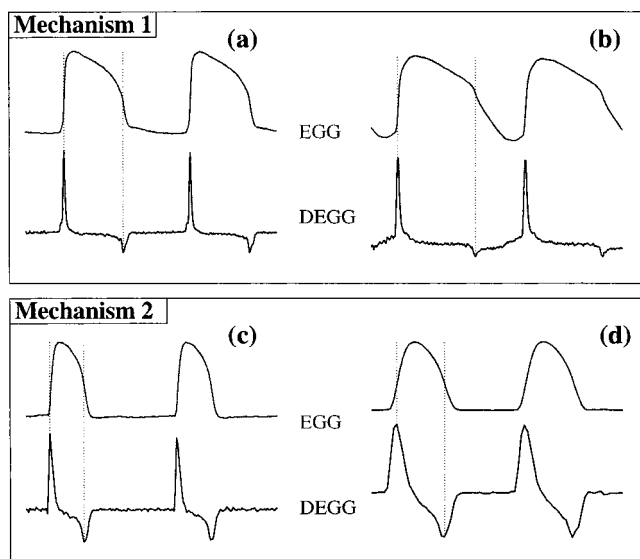


FIG. 5. Typical examples of two-period EGG and DEGG signals, extracted from the singing voice database. In these cases, the opening and closing peaks are precisely defined. Case (a): subject B2, vowel [u], C4 (260 Hz); Case (b): subject B4, vowel [e], E4 (330 Hz); Case (c): subject S2, vowel [a], A4 (440 Hz); Case (d): subject S3, vowel [a], C6 (1040 Hz).

B. Main features

Visual exploration of the database revealed that a DEGG signal presents recurrent features that can easily be categorized. These categories will now be defined and illustrated through typical examples. A major distinction is made between the cases where both opening and closing peaks are well defined and the cases where one (or both) is not well defined.

1. Well-defined peaks

Typical cases when the opening and closing peaks are precisely defined are shown in Fig. 5.

Cases (a) and (b) correspond to voiced productions in mechanism 1. They are taken from crescendos sung by two baritones (subjects B2 and B4). A long closed phase is noticeable (relative to the fundamental period), represented by the distance between two successive vertical broken lines. Cases (c) and (d) correspond to voiced production in mechanism 2. They are extracted from crescendos sung by two sopranos (subjects S2 and S3). The closed phase is shortened as compared to the previous examples.

The laryngeal mechanism differences can be seen in the shape of both the EGG and DEGG signals. The EGG signal is usually more rounded and symmetrical in mechanism 2. The corresponding DEGG signal presents less abrupt peaks. The amplitude of the closing peak is comparable to that of the opening peak, whereas in mechanism 1, the closing peak is much stronger than the opening peak.

2. Double or imprecise peaks

Typical cases where the opening or closing peaks are double or imprecise are presented in Fig. 6. The observed features can be divided into three main categories: imprecise opening peaks, double closing peaks, and double opening peaks.

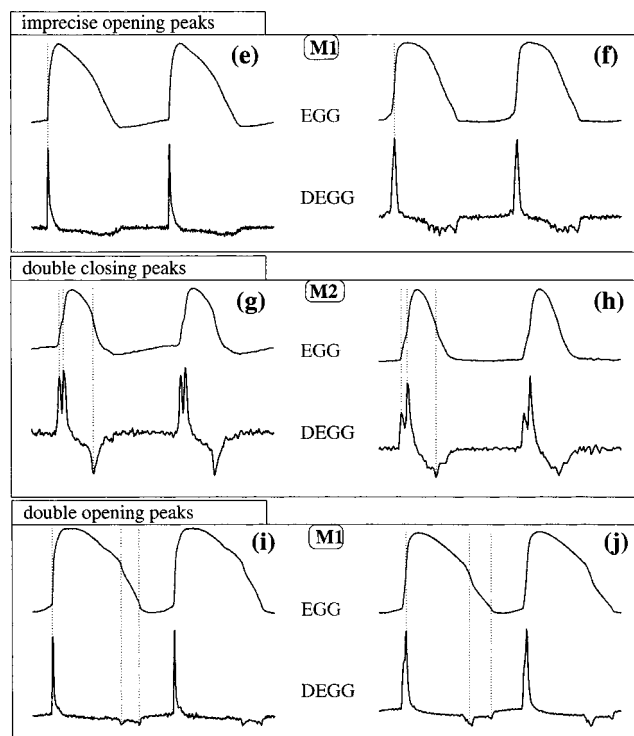


FIG. 6. Typical examples of two-period EGG and DEGG signals, where the opening or closing peaks are double or imprecise. Case (e): subject MS2, vowel [a], F3# (185 Hz); Case (f): subject B2, vowel [u], C4 (260 Hz); Case (g): subject CT3, vowel [a], E4 (330 Hz); Case (h): subject S2, vowel [a], A4 (440 Hz); Case (i): subject B4, vowel [a], G3 (196 Hz); Case (j): subject B6, vowel [a], B3 (245 Hz).

The first typical case is the absence or imprecision of the opening peak, as illustrated at the top of Fig. 6. In case (e), no opening peak is detected. In case (f), numerous peaks can be observed, none of them standing out clearly. Such lack of precision, which was often observed with regard to the opening peaks, was not found in the case of closing peaks in the present material.

Either the closing peak is extremely precise and unique, as we have seen previously, or there are two visible peaks, as illustrated in Fig. 6 for cases (g) and (h). The latter phenomenon, hereafter called “double closing peak,” can also be found at opening, as illustrated in cases (i) and (j), although “double opening peaks” are less strong and less close to each other than “double closing peaks.” In cases (g) and (i), both peaks have the same amplitude, whereas the peak amplitude differs greatly in cases (h) and (j). It should be noted that the opening or closing peaks are seldom tripled or quadrupled.

3. Occurrences

What is the proportion of precise single peaks as compared to imprecise or double peaks? Occurrences of single opening peaks and single closing peaks were measured by applying an automatic method for single peak detection (see further description in Sec. III). Table I presents the results as a percentage of the voiced measures, 100% meaning that only precise peaks were detected during the voicing part and 0% meaning that not a single precise peak was found.

TABLE I. Detection of single closing peaks ($N_c=1$) and single opening peaks ($N_o=1$), expressed as a percentage of the voiced measures. The third row gives the percentage of the voiced measures for which the closing and the opening peaks were detected as single.

Singer	B1	B2	B3	B4	B5	B6	B7	T1	T2	CT1	CT2	CT3	MS1	MS2	MS3	S1	S2	S3
$N_c=1$	51	85	66	75	76	68	16	81	65	81	82	35	86	59	61	68	89	90
$N_o=1$	34	49	21	56	41	50	19	64	31	49	53	64	63	44	46	52	82	80
$\left\{ \begin{array}{l} N_c=1 \\ \text{and} \\ N_o=1 \end{array} \right.$	31	48	19	56	40	43	8	61	25	46	46	30	62	44	45	35	81	78

These results indicate that the double-peak feature is not uncommon, for opening as well as for closing. The rate of single closing peaks ranges from 16% (baritone B7) to 90% (soprano S3) and the rate of single opening peaks from 19% (baritone B7) to 82% (soprano S2). Except for two singers (B7 and CT3), there are more single closing peaks than single opening peaks. As an extreme example, only 8% of the peaks detected for baritone B7 were single at both opening and closing.

It appears that for a given singer, peak doubling may be consistently associated with either soft or loud production, and with either low or high pitches. As an example, case (f) in Fig. 6 and case (a) in Fig. 5 are extracted from the same crescendo sung by baritone B2. The onset of the crescendo is characterized by imprecise opening peaks; this imprecision disappears when the singer sings loudly. Another example is given by case (c) in Fig. 5 and case (h) in Fig. 6, which are extracted from a crescendo sung by soprano S2. In this case, the increase in vocal intensity is accompanied by a doubling of closing peaks. No evidence was found, however, relating the double-peak feature in any general way to a change in fundamental frequency or vocal intensity. Variability across singers was found to be rather large, e.g., opening peak doubling is related to an increase of vocal intensity for singers B3, B4, and T2 and to a decrease of vocal intensity for singers B2, CT2, and S1.

C. Discussion

These results highlight the difficulty of using the DEGG signal to detect the closing and opening instants. This may be

part of the reason why this signal is not more widely used in the literature. Nevertheless, these difficulties should not prevent its use, as the DEGG signal provides an accurate detection of glottal opening and closing instants, at least in the case of single peaks. Moreover, these peaks may offer visual clues to some characteristic features of the vocal folds' vibratory movement. This hypothesis will be illustrated below by the analysis of the transition between laryngeal mechanisms.

1. Visualization of the laryngeal mechanism transition on the DEGG signal

The studies conducted by Askenfelt *et al.* (1980); Kitzing (1982); Lecluse (1977); Lecluse and Brocaar (1977); Roubeau (1993); Roubeau and Castellengo (1993); Roubeau *et al.* (1987) have shown that a transition between laryngeal mechanisms affects the amplitude and shape of an EGG signal. A rapid change in shape occurs within two voiced periods (Roubeau *et al.*, 1987; Svec *et al.*, 1999), whereas amplitude changes more slowly within about 130–150 ms (Roubeau *et al.*, 1987). A loss of pitch control is often observed (Miller, 2000; Roubeau *et al.*, 1987; Svec *et al.*, 1999). These amplitude variations can also be seen on the DEGG signals, as illustrated in Fig. 7. Thus, observation of a DEGG signal can complement that of an EGG signal for the detection of laryngeal mechanism transitions. Moreover, the slight changes in shape can be seen more easily on the DEGG than on the EGG signal. Figure 7(b) shows the transition from mechanism 1 to mechanism 2 and the bottom panels (c) zoom in on two periods before the transition

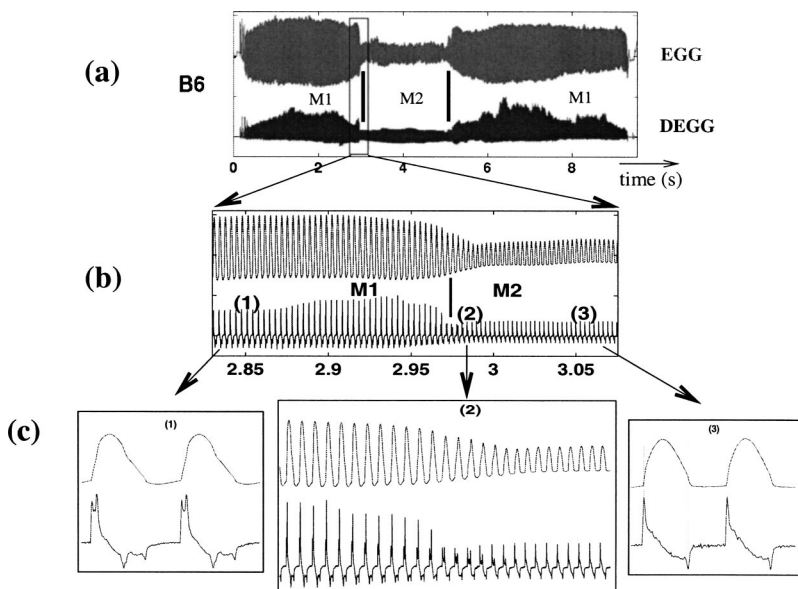


FIG. 7. EGG and DEGG signals, in the case of a glissando sung by the baritone B6. The top panel (a) presents the shape of both signals over the whole glissando. The transitions between mechanism 1 (M1) and mechanism 2 (M2) correspond to the instants of major amplitude change (around 3 s for M1 to M2 transition and 5 s for M2 to M1 transition). The middle panel (b) highlights the first transition (M1–M2) and the bottom panels (c) present several EGG and DEGG periods just before (1), during (2), and just after (3) the transition.

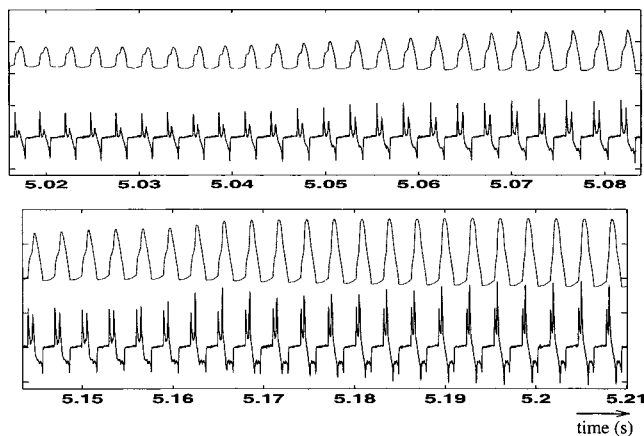


FIG. 8. The second transition (M2–M1) located around 5 s corresponding to the case of the glissando sung by the baritone B6 presented in Fig. 7.

[panel c(1)], on a few periods during the transition [panel c(2)], and on two periods after the transition [panel c(3)]. Right before the transition [panel c(1)], both the closing and opening peaks are doubled. During the transition [panel c(2)], the amplitude of one peak decreases, whereas the amplitude of the other increases. This can be observed for the closing peaks as well as for the opening peaks. After the transition [panel c(3)], both closing and opening peaks are single and the characteristic shape of a DEGG signal in mechanism 2 is observed.

The transition from mechanism 2 to mechanism 1 presents the same features in reverse succession, as shown in Fig. 8: single peaks in mechanism 2, doubling and amplitude change during the transition, and double peaks in mechanism 1. The analysis of the glissandos in the whole database reveals that these features are neither “singer-specific” (specific to this individual) nor a common characteristic found at any transition between laryngeal mechanisms.

It is clear from this example that a DEGG signal can highlight tiny changes in shape that would have gone unnoticed in the EGG signal. For instance, if the transition between laryngeal mechanisms appears to be very sudden (a couple of periods) when looking at the amplitude of the EGG and DEGG signals, the peak changes over several periods indicate that there may be some slower adjustments in the vocal-fold contact process. These observations call for research combining electroglottography with some kind of visualization. A characterization of DEGG features using high-speed cinematography is currently under way. At present, the following preliminary hypothesis to explain the double-peak feature is proposed.

2. Hypothesis for explaining double peaks

It is speculated that the double-peak feature observed on a DEGG signal is related to either a difference in how opening (or closing) takes place over the thickness of the vocal folds or to a time-lag opening (or closing) at two different parts of the glottis. A time-lag opening over the length of the glottis was observed by Karnell (1989), who compared synchronized videostroboscopy and electroglottography. For one subject, glottis-anterior opening occurred before posterior opening; for another, posterior opening occurred first. These

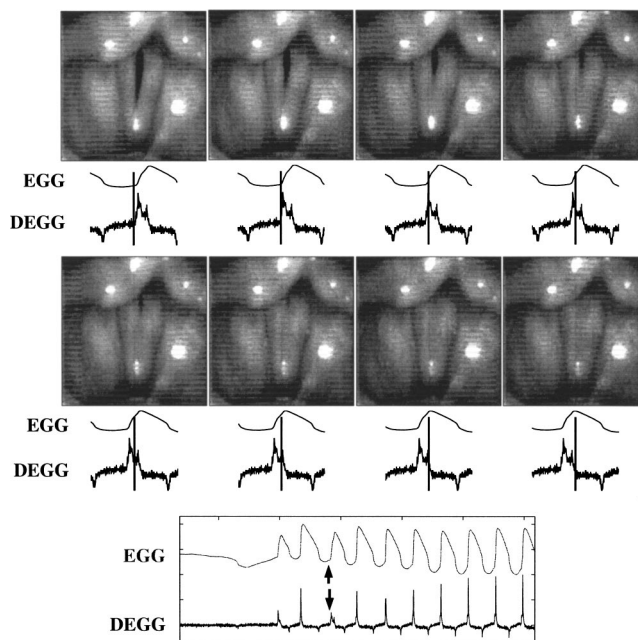


FIG. 9. Visualization of a double peak by simultaneous high-speed and electroglottographic recordings. Vertical bars indicate the moment in time at which the visual image occurs with respect to the EGG and DEGG signals. This example corresponds to the voice onset (bottom panel) of a nonpathological male phonation in mechanism 1, with $f_0=110$ Hz. The EGG sampling frequency is 44 444 Hz and the high-speed camera sampling frequency is 3704 frames/s.

glottal opening behaviors are not uncommon, having also been observed by Anastaplo and Karnell (1988) and recently by Hess and Ludwigs (2000). They could be related to voice quality: according to Hess and Ludwigs (2000), a relaxed voice would be produced with a posterior-to-anterior opening, whereas a pressed voice would rather be produced with an anterior-to-posterior opening. This corresponds to expectations, as arytenoid tension is greater in the case of pressed phonation, preventing the air from passing through the posterior part, whereas the posterior part of the glottis is not firmly closed during relaxed phonation (it can even be slightly open, as in the case of a glottal chink).

The double-peak feature is illustrated in Fig. 9, by simultaneous visualization of high-speed images and electroglottographic signal at voice onset. The data come from the same male phonation previously used to describe the closing and opening phases. The first closing peak corresponds to vocal-fold contact over two-thirds of the glottis length, as shown on the four top images. A glottal chink remains in the posterior part of the glottis. The second closing peak seems to be related to the closing of this glottal chink, as shown on the four bottom images.

D. Conclusion

These data warrant the assertion that DEGG signals provide information on glottal opening and closing for M1 and M2. These features may be related to some properties of vocal-fold contact that call for an extensive study that would combine this signal with other means of exploration such as high-speed visualization and glottal-flow estimation. When the peaks observed on the DEGG signal are single, they can

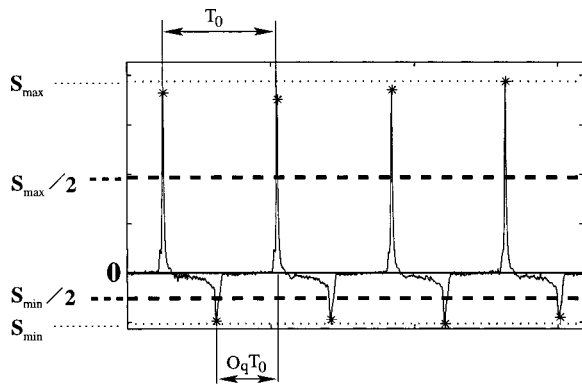


FIG. 10. Direct detection of closing and opening peaks by using a threshold. The values of the fundamental period T_0 and the open quotient O_q can be derived directly from the peak detection.

be related to the instants of glottal opening and closing. In that case, they can be used to measure fundamental frequency and open quotient. A method for such measurement is described in the following section.

III. MEASUREMENTS OF FUNDAMENTAL FREQUENCY AND OPEN QUOTIENT ON A DEGG SIGNAL

A. Introduction

In the previous section, it has been shown that the DEGG signal can present precise single peaks at glottal opening instants. In such a case, it can be assumed that a method which would take these peaks into account should give more accurate open quotient measurements than any threshold-based method. In Sec. III B, a new DEGG-based method (DECOM method) is proposed that can be applied to the case of quasisteady voiced sounds. The performance is compared to the above-mentioned methods in Sec. III D.

B. A DEGG-based method: DECOM

As illustrated in Fig. 10, the easiest way to measure the fundamental period and the open quotient using a DEGG signal consists of detecting the maxima located above a given positive threshold and the minima located below a given negative threshold. This threshold can be chosen as a percentage of the maximum (respectively, minimum) amplitude of the windowed signal.

This method yields good results when the closing and opening peaks are very precise. When the opening peaks are less precise, it can no longer be applied. Thus, a more general and robust method which is based on the determination of peak distance by correlation was developed: the DECOM method. The designation DECOM comes from “DEgg Correlation-based method for Open quotient Measurement.” This method is rooted in the standard methods for fundamental frequency estimation of a voiced signal.

The principle of the DECOM method is to measure fundamental frequency and open quotient using a correlation-based method to estimate the distance between two consecutive closing peaks and the distance between an opening peak

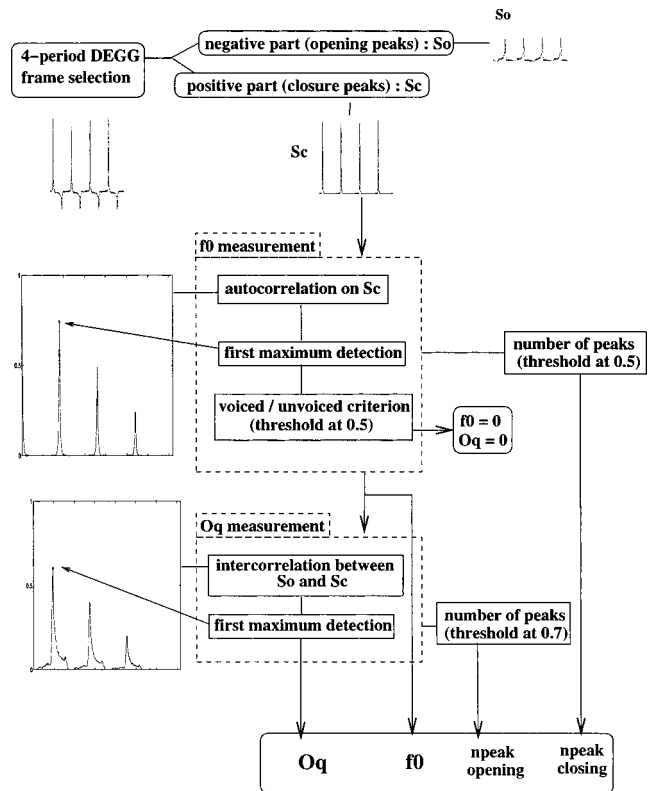


FIG. 11. Schematic description of the algorithm for the measurement of fundamental frequency (f_0) and open quotient (O_q). In addition, an automatic detection of the number of peaks at closing (n_{peak} closing) and at opening (n_{peak} opening) is performed.

and the consecutive closing peak. The corresponding algorithm is presented in Fig. 11. A step-by-step description follows.

1. Frame selection

The selected frames are four periods long, assuming that the fundamental period is already known. During the analysis process, an approximate value of the fundamental period can be deduced from the preceding frame. If not, as in the case of voicing onset or absence of voicing, the rectangular window has a fixed length of 40 ms for male voice and 20 ms for female voice. The analysis step is two fundamental periods or 5 ms if the fundamental period is unknown.

2. Division of the signal into positive and negative parts

The signal is separated into two parts: a positive part S_c , which contains information concerning glottal closing peaks, and a negative part S_o , which contains information concerning glottal opening peaks. The signal S_c is derived from the original signal by replacing any negative value by zero. In the same way, the signal S_o is derived from the original signal by replacing positive values by zero. Moreover, it is inverted, resulting in a positive signal. The original signal can then be reconstructed as the difference between the two signals: $S_{original} = S_c - S_o$.

3. Fundamental frequency measurement

A normalized biased autocorrelation function is calculated on signal S_c . The fundamental period is given by the position of the first maximum, which corresponds to the time between two consecutive closing peaks. The maxima of the autocorrelation function are detected as the inflection points for which the first derivative's sign inverts from positive to negative. Only the inflection points whose amplitude is above a given threshold of 50% of the maximal inflection points amplitude are selected, and the first of these points is kept as the first maximum of the autocorrelation function. It should be noted that no subharmonic error is likely to occur, i.e., a false selection of higher-order maxima, because of the quasi-impulsive form of signal S_c , which implies that the first maximum always has the maximal amplitude.

A criterion for discrimination between voiced and unvoiced parts is applied. The amplitude of the detected maximum of the normalized autocorrelation function is compared to a threshold value, empirically set at 0.5. If the amplitude value is less than the threshold value, the frame is considered to correspond to an unvoiced part, f_0 and O_q are set to zero, and the algorithm stops. If the amplitude is greater than the threshold value, the frame is considered to correspond to a voiced part. The measurement of the fundamental period is refined by applying a cubic-spline interpolation between five samples chosen around the first maximum of the autocorrelation function. The interpolation step is related to the uncertainty ΔT_0 of T_0 measurement. The uncertainty Δf_0 of f_0 depends on the uncertainty of T_0 by: $\Delta f_0/f_0 = \Delta T_0/T_0$. Thus, a given uncertainty Δf_0 of f_0 measurement requires the uncertainty of T_0 to be $\Delta T_0 = \Delta f_0/f_0^2$. The interpolation step is chosen so that the uncertainty of f_0 is 0.5 Hz and is calculated using an approximated value of f_0 .

4. Open quotient measurement

A normalized biased intercorrelation function is calculated between signals S_o and S_c . The first maximum of this function corresponds to the time between an opening peak and the consecutive closing peak. It is detected in the same way as in the case of the autocorrelation function, but only the inflection points whose amplitudes are above a given threshold of 80% of the maximal inflection points amplitude are now selected. If we note T_1 , the corresponding time, the open quotient O_q can be deduced from this time and the previous measure of fundamental period T_0 by: $O_q = T_1/T_0$. The uncertainty of O_q depends on the uncertainty of T_1 and T_0 by: $\Delta O_q/O_q = (\Delta T_1/T_1) + (\Delta T_0/T_0)$. As previously done for the f_0 measurement, the measurement is refined by applying a cubic-spline interpolation between five samples chosen around the first maximum of the intercorrelation function. The interpolation step is chosen to be equal to the one calculated for f_0 measurement, which implies that $\Delta T_1 = \Delta T_0$. As $T_1 = O_q T_0$, $\Delta O_q = (\Delta T_0/T_0)(1 + O_q) < 2(\Delta T_0/T_0) = 2(\Delta f_0/f_0)$. For the considered frequency range ($f_0 \geq 100$ Hz), the uncertainty of O_q is less than 0.01.

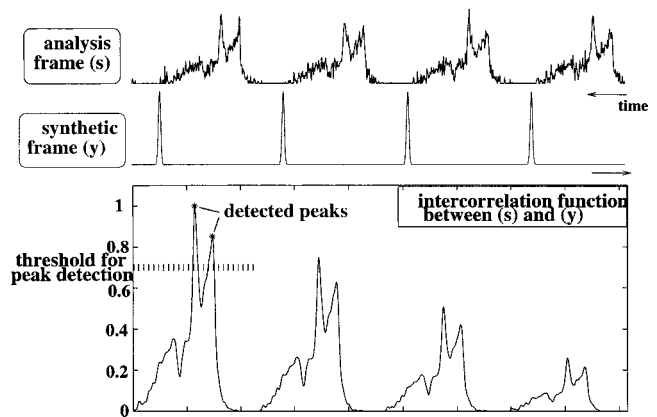


FIG. 12. Method for the detection of double- or imprecise peaks in the DEGG signal by intercorrelation between the signal to be analyzed and a Gaussian comb.

5. Automatic detection of peak doubling

The final part of the algorithm is dedicated to the automatic detection of double or undefined peaks. As illustrated in Fig. 12, a biased intercorrelation function is calculated between the signal to be analyzed (either S_c or S_o) and a synthetic signal of identical length and period, composed of Gaussian-modulated sinusoidal pulses. The pulses are given by the following equation:

$$y(t) = e^{-(\pi^2 t^2) / [2 \ln(2) * (1/BWf_0)]} \cos(2\pi f_0 t).$$

The fractional bandwidth BW is chosen to be equal to 40 Hz, so that the Gaussian pulses are very narrow as compared to the fundamental period. Thus, this synthetic signal is similar to an S_c signal for which the peaks would be very precise. The detection is made over a window of length T_0 , centered on the first maximum. Only the maxima which are located above a given threshold are taken into account, to allow for peak detection even when there is a great difference in amplitude between peaks. The threshold is set to 50% of the maximum in the case of closing and 70% in the case of opening.

6. Result of the algorithm

The algorithm returns the values of the fundamental frequency and open quotient measured for the given frame, as well as the estimated number of peaks (n_{peak}) at closing and opening ($n_{peak} = 1$ for a single peak, $n_{peak} = 2$ for a double peak; if $n_{peak} > 2$, the peak is undefined). This last information will be taken into account in the experimental results, as the measure of open quotient is only valid if the opening and closing peaks are precise and single.

C. Illustration

This correlation-based method is applied to the study of open quotient in singing in a companion paper (Henrich *et al.*, submitted). Figure 13 illustrates the analysis of a sustained vowel sung by a mezzosoprano in mechanism 1. A doubling of the opening peak is detected between $t = 0.2$ and $t = 0.65$ s and between $t = 1.65$ and $t = 2$ s ("o" marks). The second opening peaks, which have the greatest amplitude, have been chosen for open quotient estimation in this case.

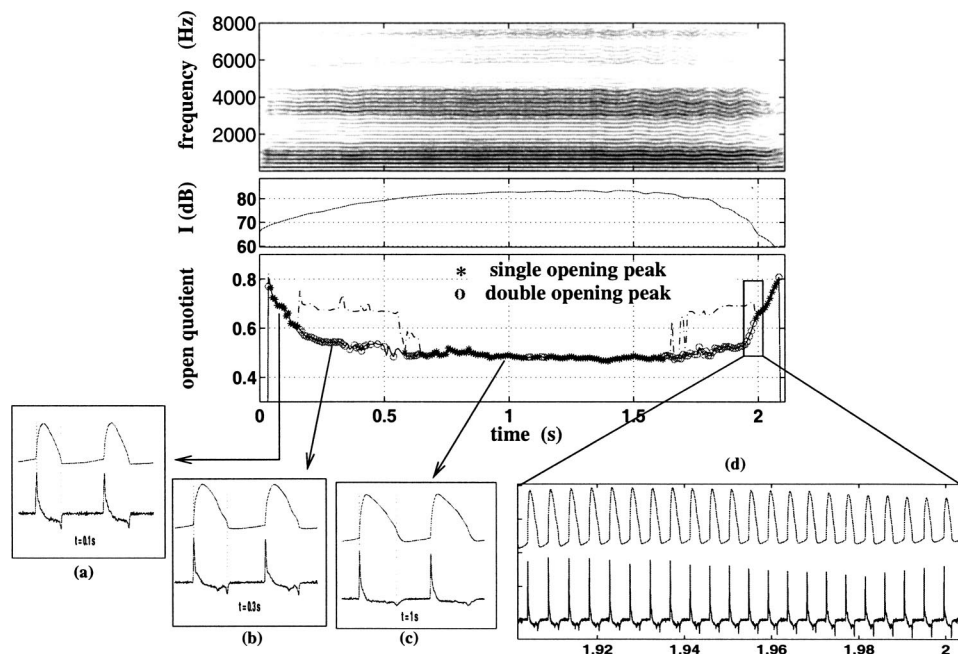


FIG. 13. Analysis of a sustained vowel sung by the mezzo-soprano MS2 in mechanism 1, with $f_0=220$ Hz (A3). The top panel presents the sound spectrogram, the vocal intensity (I) curve, and the open quotient measures plotted against time. The measures related to single-peak detection are marked with “*” and the ones related to double-peak detection are marked with “o”. In the latter case, the second opening peak which has the maximum amplitude has been chosen for the measure of open quotient. The dash-dot lines show the measures that would have resulted if the first opening peaks had been chosen (between 0.2 and 0.65 s and between 1.65 and 2 s). The bottom panels (a), (b), (c), and (d) present the EGG and DEGG signals on several periods at a given time.

The open quotient values resulting from the choice of the first opening peaks are also plotted (broken lines). In the case of a double peak, the choice of one peak or another leads to markedly different values of the open quotient, the difference ranging between 0.1 and 0.2 in this case. It is clear from this example that the open quotient cannot be accurately measured on a DEGG signal unless the closing and opening peaks are single and precise.

D. EGG, DEGG, and glottal-flow measurements comparison

A comparison of the different methods for open quotient measurement mentioned in Sec. III A has been conducted on

a database of five male speakers provided by Childers (2000) (subjects M01 to M05). The open quotient was measured on the EGG or DEGG signal using the following four methods: EGG thresholds at 50% and 35%, Howard’s method with a threshold at 3/7, and the DECOM method. The open quotient was also estimated on the acoustic signal by a pitch-synchronous covariant linear prediction method proposed by Ljungqvist (1986). The error criterion was minimized with regard to the KLGLOTT88 glottal flow model (Klatt and Klatt, 1990).

Table II presents the results of the acoustic and electroglottographic measurements for vowels [a], [i], and [u]. The table on the left gives the mean open quotient values mea-

TABLE II. Comparison of five different methods for open quotient measurement, for 5 speakers (M01 to M05, Childers, 2000) and for the 3 vowels [a], [i], and [u]: threshold methods based on the EGG signal (50% and 35%), on a combination of EGG and DEGG signals (3/7), on the DEGG signal (DECOM), and on the inverse-filtered glottal flow (gf). The left table gives the mean value of open quotient in each case and the right table the mean difference in percent between the open quotient measured on the EGG or DEGG signal and the open quotient measured on the inverse-filtered glottal flow. For each case, the minimum difference is in bold.

	Mean open quotient					Mean difference (%)				
	50%	35%	3/7	DECOM	gf	50%	35%	3/7	DECOM	
vowel [a]										
M01	0.62	0.58	0.59	0.60	0.59	2.99	1.25	0.79	0.99	
M02	0.53	0.47	0.50	0.47	0.46	7.37	2.02	3.75	1.56	
M03	0.55	0.50	0.51	0.52	0.50	4.71	1.08	1.39	2.26	
M04	0.54	0.46	0.53	0.55	0.57	3.30	10.89	3.81	2.39	
M05	0.44	0.36	0.40	0.37	0.37	6.03	1.45	2.44	1.25	
vowel [i]										
M01	0.61	0.58	0.59	0.59	0.58	3.26	1.10	1.12	0.93	
M02	0.53	0.47	0.50	0.51	0.52	1.45	4.86	1.84	0.96	
M03	0.53	0.46	0.48	0.49	0.47	6.47	2.08	3.06	2.98	
M04	0.50	0.45	0.49	0.48	0.49	2.00	4.03	1.59	1.45	
M05	0.43	0.38	0.39	0.40	0.41	4.84	4.70	4.13	2.47	
vowel [u]										
M01	0.66	0.62	0.63	0.64	0.63	3.52	1.01	0.83	1.29	
M02	0.51	0.47	0.49	0.48	0.49	2.37	1.97	0.97	1.58	
M03	0.51	0.46	0.47	0.48	0.46	5.22	1.70	2.39	2.82	
M04	0.50	0.43	0.49	0.50	0.52	2.60	8.55	3.00	2.60	
M05	0.39	0.34	0.35	0.33	0.33	5.39	1.21	1.89	1.10	

sured with the five different methods; that on the right gives the differences in percent between the measures using electroglottography and the measures conducted on the estimated glottal flow. Threshold methods (either 50% or 35%) show variable differences compared with the ones made on glottal flow, ranging from 1% to 10%. The 35% method seems to yield better results. When the DEGG signal is taken into account, the results are even more similar to the results obtained for the glottal flow. The differences range from 0.79% to 4.13% in the case of Howard's method (3/7) and from 0.93% to 2.98% in the case of the DECOM method presented in this paper. As expected, the open quotient measurements using the DECOM method are, on average, in much better agreement with the glottal-flow measurements than most threshold-based methods for sustained phonation when double pulsing is rejected.

These results confirm that there is a good correspondence between the open quotient measured on a glottal-flow signal and that measured on an electroglottographic signal. The correspondence improves if the derivative of the EGG signal is taken into account. The method developed by Howard gives good results. Yet, in the case of precise opening and closing peaks, the DECOM method seems more appropriate. It is envisioned, as a further improvement of EGG-based open quotient measurement methods, to combine the DECOM method with Howard's method when the peaks are double or imprecise.

IV. CONCLUSION

The use of the differentiated EGG signal to measure the open quotient has been controversial. Yet, little research has been devoted to characterizing this signal in terms of glottal opening and closing instants. One purpose of this paper was to describe the various features which can be observed in this signal by analyzing an extensive corpus of sung items recorded by 18 trained singers.

A DEGG signal presents strong and weak peaks which can accurately be related to the glottal closing and opening instants, respectively, in the case where these peaks are single and precise. A correlation-based algorithm, DECOM, is proposed for measuring fundamental frequency and open quotient using the DEGG signal rather than the EGG signal. The DECOM method is compared to threshold-based methods. It appears that open quotient measurements based on electroglottography are in better agreement with the measurements based on inverse-filtered glottal flow if the DEGG signal is taken into account. The application of the DECOM method to the study of open quotient variations in singing is addressed in a companion paper (Henrich *et al.*, submitted).

Exploration of the database shows that the DEGG can present other typical features, such as a doubling of the opening or of the closing peaks. In these cases, the DEGG signal is not suitable for the purpose of open quotient measurement, but could still complement the EGG signal, as small changes in vocal-fold contact are not easily noticeable on an EGG signal, whereas they appear much more clearly on a DEGG signal. The DEGG signal may provide interesting clues for the detection of peculiar glottal configurations. Yet, the features that can be found on such a signal have to be charac-

terized with regard to vocal-fold physiology and glottal vibratory movement. Comparison with high-speed visualization would significantly advance the understanding of the characteristics of a DEGG signal and is the aim of continuing research. Further research could also be conducted on the possible auditory effect associated with double pulsing.

ACKNOWLEDGMENTS

This work is part of a Ph.D. thesis defended at the Université Pierre & Marie Curie (Paris, France) in November 2001 (Henrich, 2001). The authors would like to thank Jürgen Neubauer and the team from the Department of Phoniatrics and Pediatric Audiology of the University of Erlangen-Nuremberg (Germany) for kindly providing a data file containing the EGG signal and high-speed images of male phonation. They wish to thank Adrian Fourcin and Svante Granqvist for their valuable comments, which resulted in significant improvements of the manuscript. They are grateful to Alexis Michaud, Brian Katz, and Evelyn Abberton for English corrections.

APPENDIX: DEFINITION OF THE LARYNGEAL MECHANISMS

Human vocal production is characterized by the use of four distinctive vocal-fold configurations which differ in respect to the length and thickness of the vocal folds and the muscular laryngeal tensions involved in the vocal process (Roubeau, 1993). Evidence in support of this classification is provided by the noticeable transitions, which can for instance be heard on a glissando and detected on an EGG signal (Henrich *et al.*, 2003b). Many different terms have been used in the literature to describe these laryngeal configurations, which leads to confusing terminology. So as to avoid this confusion, which is related to the notion of register in singing, and to bring out the differences between change in the glottis configuration (laryngeal mechanism) and adjustment of the vocal tract (resonance register), we have chosen to use the term *laryngeal mechanism* instead of *register*, and to number laryngeal mechanisms, following Roubeau (1993).

- (1) mechanism 0, or M0, corresponds to the so-called "vocal fry," "pulse," or "strobass" register;
- (2) mechanism 1, or M1, corresponds to the so-called "chest" or "modal" register for male and female and male "head" register;
- (3) mechanism 2, or M2, corresponds to the so-called "falsetto" or "loft" male register and to the "head" female register;
- (4) mechanism 3, or M3, corresponds to the so-called "whistle," "flageolet," or "flute" register.

A rationale for explaining this terminology and its choice in further detail is formulated in Henrich *et al.* (2003b).

In this paper, we focus on the two main laryngeal mechanisms, M1 and M2, which are commonly used in speech and singing. In mechanism 1, the vocal folds are thick, leading to vertical phase differences in vibration, a

longer closing and opening phase as compared to mechanism 2, where the vocal folds are thin and vibrate without any vertical phase difference (Vennard, 1967).

¹For the purpose of this illustration, a datafile was kindly provided by a team of the Department of Phoniatrics and Pediatric Audiology of the University of Erlangen-Nuremberg (Germany). It contains a 2-s-long recording of high-speed images and the corresponding EGG signal in the case of a nonpathological male phonation. This recording was made by J. Neubauer, M. Tigges, T. Wittenberg, and U. Eysholdt.

- Anastaplo, S., and Karnell, M. P. (1988). "Synchronized videostroboscopic and electroglottographic examination of glottal opening," *J. Acoust. Soc. Am.* **83**, 1883–1890.
- Askenfelt, A., Gauffin, J., Sundberg, J., and Kitzing, P. (1980). "A comparison of microphone and electroglottograph for the measurement of vocal fundamental frequency," *J. Speech Hear. Res.* **23**, 250–273.
- Baer, T., Löfqvist, A., and McGarr, N. S. (1983a). "Laryngeal vibrations: A comparison between high-speed filming and glottographic techniques," *J. Acoust. Soc. Am.* **73**, 1304–1308.
- Baer, T., Titze, I. R., and Yoshioka, H. (1983b). "Multiple simultaneous measures of vocal fold activity," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by A. Bless (College Hill, San Diego), pp. 229–237.
- Baken, R. J. (1992). "Electroglottography," *J. Voice* **6**, 98–110.
- Berke, G. S., Moore, D. M., Hantke, D. R., Hanson, D. G., Gerratt, B. R., and Burstein, F. (1987). "Laryngeal modeling: Theoretical, *in vitro*, *in vivo*," *Laryngoscope* **97**, 871–881.
- Childers, D. G. (2000). *Speech Processing and Synthesis Toolboxes* (Wiley, New York).
- Childers, D. G., Hicks, D. M., Moore, G. P., and Alsaka, Y. A. (1986). "A model for vocal fold vibratory motion, contact area, and the electroglottogram," *J. Acoust. Soc. Am.* **80**, 1309–1320.
- Childers, D. G., Hicks, D. M., Moore, G. P., Eskenazi, L., and Lalwani, A. L. (1990). "Electroglottography and vocal fold physiology," *J. Speech Hear. Res.* **33**, 245–254.
- Childers, D. G., and Krishnamurthy, A. K. (1985). "A critical review of electroglottography," *CRC Crit. Rev. Biomed. Eng.* **12**, 131–161.
- Childers, D. G., and Larar, J. N. (1984). "Electroglottography for laryngeal function assessment and speech analysis," *IEEE Trans. Biomed. Eng.* **BME-31**, 807–817.
- Childers, D. G., Moore, G. P., Naik, J. M., Larar, J. N., and Krishnamurthy, A. K. (1983a). "Assessment of laryngeal function by simultaneous, synchronized measurement of speech, electroglottography and ultra-high speed film," in *Transcripts of the Eleventh Symposium: Care of the Professional Voice*, edited by V. Lawrence (Voice Foundation, New York), pp. 234–244.
- Childers, D. G., Naik, J. M., Larar, J. N., Krishnamurthy, A. K., and Moore, G. P. (1983b). "Electroglottography, speech and ultra-high speed cinematography," in *Vocal Fold Physiology and Biophysics of Voice*, edited by I. Titze and R. Scherer (Denver Center for the Performing Arts, Denver), pp. 202–220.
- Childers, D. G., Smith, A. M., and Moore, G. P. (1984). "Relationships between electroglottograph, speech, and vocal cord contact," *Folia Phoniatr.* **36**, 105–118.
- Colton, R. H., and Conture, E. G. (1990). "Problems and pitfalls of electroglottography," *J. Voice* **4**, 10–24.
- Cranen, B. (1991). "Simultaneous modelling of EGG, PGG, and glottal flow," in *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*, edited by H. B. Gauffin J. (Singular, San Diego), pp. 57–64.
- Dejonckere, P. (1981). "Comparison of two methods of photoglottography in relation to electroglottography," *Folia Phoniatr.* **33**, 338.
- Fabre, P. (1957). "Un procédé électrique percutané d'inscription de l'accolement glottique au cours de la phonation: Glottographie de haute fréquence," *Bulletin de l'Académie Nationale de Médecine*, pp. 66–69.
- Fabre, P. (1958). "Etude comparée des glottogrammes et des phonogrammes de la voix humaine," *Ann. Oto. Lar. (Paris)* **75**, 767–775.
- Fabre, P. (1959). "La glottographie électrique en haute fréquence: Particularités de l'appareillage," *C. R. Soc. Biol. (Paris)* **153**, 1361–1364.
- Fabre, P. (1961). "Glottographie respiratoire," *Ann. Oto. Lar. (Paris)* **78**, 814–824.
- Fourcin, A. J. (1974). "Laryngographic examination of vocal fold vibration," in *Ventilatory and Phonatory Function*, edited by B. Wyke (Oxford University Press, London), pp. 315–326.
- Fourcin, A. J. (1981). "Laryngographic assessment of phonatory function," *ASHA Rep.* **11**, 116–124.
- Gerratt, B. R., Hanson, D. G., and Berke, G. S. (1988). "Laryngeal configuration associated with glottography," *Am. J. Otolaryngol.* **9**, 173–179.
- Henrich, N. (2001). "Etude de la source glottique en voix parlée et chantée (Study of the glottal source in speech and singing)," Ph.D. thesis, Université Paris 6, France.
- Henrich, N., d'Alessandro, C., Castellengo, M., and Doval, B. "Open quotient measurements in speech and singing," *J. Acoust. Soc. Am.* (submitted).
- Henrich, N., Roubeau, B., and Castellengo, M. (2003b). "On the use of electroglottography for characterisation of the laryngeal mechanisms," in *Stockholm Music Acoustics Conference* (Stockholm, Sweden).
- Hess, M. H., and Ludwigs, M. (2000). "Strobophotoglottographic transillumination as a method for the analysis of vocal fold vibration patterns," *J. Voice* **14**, 255–271.
- Howard, D. M. (1995). "Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers," *J. Voice* **9**, 163–172.
- Howard, D. M., Lindsey, G. A., and Allen, B. (1990). "Toward the quantification of vocal efficiency," *J. Voice* **4**, 205–212.
- Karnell, M. P. (1989). "Synchronized videostroboscopy and electroglottography," *J. Voice* **3**, 68–75.
- Kitzing, P. (1977). "Methode zur kombinierten photo- und elektroglottographischen registrierung von stimmklippenschwingungen," *Folia Phoniatr.* **29**, 249–260.
- Kitzing, P. (1982). "Photo- and electroglottographical recording of the laryngeal vibratory pattern during different registers," *Folia Phoniatr.* **34**, 234–241.
- Kitzing, P. (1983). "Simultaneous photo- and electroglottographic measurements of voice strain," in *Vocal Fold Physiology*, edited by I. R. Totze and R. C. Scherer (The Denver Center for the Performing Arts, Denver), pp. 221–229.
- Kitzing, P., Carlborg, B., and Löfqvist, A. (1982). "Aerodynamic and glottographic studies of the laryngeal vibratory cycle," *Folia Phoniatr.* **34**, 216–224.
- Klatt, D., and Klatt, L. (1990). "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *J. Acoust. Soc. Am.* **87**, 820–857.
- Lecluse, F. (1977). "Elektroglottografie," Thesis Rotterdam, Drukkerij Elinkwijk, Utrecht.
- Lecluse, F., and Brocaar, M. (1977). "Quantitative measurements in the electroglottogram," 17th International Congress of Logopedics and Phoniatrics.
- Lecluse, F. L. E., Brocaar, M. P., and Verschuure, J. (1975). "The electroglottography and its relation to glottal activity," *Folia Phoniatr.* **27**, 215–224.
- Ljungqvist, M. G. (1986). "Speech analysis-synthesis based on modeling of voice source and vocal-tract characteristics," Ph.D. thesis, University of Tokyo.
- Miller, D. G. (2000). "Registers in singing," Ph.D. thesis, Rijksuniversiteit Groningen, the Netherlands.
- Orlikoff, R. F. (1998). "The uses and abuses of electroglottography," *Phonoscope* **1**, 37–53.
- Pedersen, M. F. (1977). "Electroglottography compared with synchronized stroboscopy in normal persons," *Folia Phoniatr.* **29**, 191–199.
- Rothenberg, M. (1981). "Some relations between glottal air flow and vocal fold contact area," *ASHA Rep.* **11**, 88–96.
- Rothenberg, M. (1992). "A multichannel electroglottograph," *J. Voice* **6**, 36–43.
- Rothenberg, M., and Mahshie, J. J. (1988). "Monitoring vocal fold abduction through vocal fold contact area," *J. Speech Hear. Res.* **31**, 338–351.
- Roubeau, B. (1993). "Mécanismes vibratoires laryngés et contrôle neuro-musculaire de la fréquence fondamentale," thèse, Université Paris XI, Orsay.
- Roubeau, B., and Castellengo, M. (1993). "Revision of the notion of voice register," XIXth International CoMeT Congress, Utrecht.
- Roubeau, B., Chevré-Muller, C., and Arabia-Guidet, C. (1987). "Electroglottographic study of the changes of voice registers," *Folia Phoniatr.* **39**, 280–289.
- Svec, J. G., Schutte, H. K., and Miller, D. G. (1999). "On pitch jumps

between chest and falsetto registers in voice: Data from living and excised human larynges," *J. Acoust. Soc. Am.* **106**, 1523–1531.

Teaney, D., and Fourcin, A. J. (1980). "The electrolaryngography as a clinical tool for the observation and analysis of vocal fold vibration," The Voice Foundation.

Titze, I. R., Baer, T., Cooper, D., and Scherer, R. (1984). "Automatic ex-

traction of glottographic waveform parameters and regression to acoustic and physiologic variables," in *Vocal Fold Physiology: Contemporary Research Clinical Issues*, edited by A. J. Bless DM (College Hill, San Diego), pp. 146–154.

Vennard, W. (1967). *Singing: The Mechanism and the Technic* (Carl Fischer, New York).