Title:

Contact quotient versus closed quotient: a comparative study on professional male singers

Abstract

Objectives: The term "closed quotient" is frequently used for data derived both from inverse filtering and from electroglottography. In the former case it is defined as the ratio between the closed phase and the period, as measured in flow glottograms (FLOGG), while in the latter case it is defined as the time interval between the falling and rising parts of the electroglottogram (EGG), measured at some percentage of the peak-to-peak amplitude. The study aims at analyzing differences between EGG and FLOGG based closed quotients and their relationships with voice source parameters.
Study Design: Comparative study
Methods: FLOGG based and EGG based measures collected from five professional male singers were compared, under different pitch and loudness conditions.
Results: As compared with the FLOGG based quotient, the EGG based quotient (i) varied more between subjects, (ii) presented greater values, (iii) varied less with subglottal pressure, (iv) varied less with the normalized amplitude quotient (i.e. the ratio between the flow pulse amplitude and the product of period and maximum flow declination rate), and (v) varied less with the relative amplitude of the voice source fundamental.
Conclusions: Although positively related, FLOGG based and EGG based closed quotients differ and must not be confused.

Key words: Closed quotient; Contact Quotient; Flow glottogram; Electrolaryngography; Voice source analysis; Inverse filtering; Trained male voices

Introduction

Closed or open quotient (Q_{Closed}, $Q_{\text{Open,}}$ respectively) have been used extensively in voice research as a standard measure of assessing vibratory, clinically relevant voicing patterns [1; 2]. Mostly it is determined by either of two methods: applying inverse filtering or using electroglottography. The first method implies filtering of the audio or the flow signal by the inverse of the vocal tract transfer function [3]. This procedure offers a flow glottogram (FLOGG) showing transglottal airflow versus time. The electroglottography technique, first developed by Fabre (1957), offers measures of vocal fold vibration characteristics [4]. The signal is obtained from two electrodes placed on each side of the thyroid cartilage and modulated by the transglottal conductivity [5].

The term Q_{Closed} is commonly used for values obtained both from FLOGG and EGG data. However, values derived from these two methods, henceforth $Q_{\text{Closed FLOGG}}$ and $Q_{\text{Closed EGG}}$, tend to differ. The former refers to the ratio between the closed phase of the glottal vibration cycle, while the latter is defined as the time interval between the rising and falling parts of the EGG waveform at a given percentage of the peak-to-peak signal amplitude. However, different percentages are used [6-8], which limits the possibilities to compare $Q_{\text{Closed EGG}}$ data reported in different studies.

The relationship between $Q_{\text{Closed FLOGG}}$ and $Q_{\text{Closed EGG}}$ has been studied in previous research. For example, in one study $Q_{\text{Open FLOGG}}$ and $Q_{\text{Open EGG}}$ values (1-$Q_{\text{Closed FLOGG}}$ and 1-$Q_{\text{Closed EGG}}$, respectively) were compared; the FLOGG data were obtained from automatized linear predictive coding of the flow signal [9]. The subjects were ten adult females, none of whom had any voice training. They were asked to produce the syllable /pa/ at four sound pressure levels (SPL) covering a range of 15 dB. The $Q_{\text{Open EGG}}$ was measured using a Kay Elemetrics Glottograph. A 50% and a 20% criterion were applied to both signals. As illustrated in Figure 1, the results showed that the $Q_{\text{Closed FLOGG}}$ were greater than $Q_{\text{Closed EGG}}$. In addition, the former increased more with SPL than the latter.

<Please insert Figure 1 about here>

Henrich and associates compared the relative duration of the open phase as measured from flow glottograms with those obtained by applying different criteria to the EGG waveform [10]. Data from five male speakers' productions of the vowels /a, i, u/ were taken from Childers (2000). $Q_{\text{Open EGG}}$ was extracted, applying three different amplitude criteria - 50%, 35%, 3/7 - and the so-called DECOM method, based on the derivative of the EGG signal. $Q_{\text{Open FLOGG}}$ was derived from inverse filtering the acoustic signal by means of a pitch synchronous covariant linear prediction method. Depending on the amplitude criterion applied, the correlation between $Q_{\text{Open FLOGG}}$ and $Q_{\text{Open EGG}}$ varied ($0.429 < R^2 < 0.933$). Of the amplitude criteria, the 35% criterion showed the best correlation with $Q_{\text{Open FLOGG}}$ but the DECOM method yielded an even better agreement.

Herbst and Ternström (2006) compared $Q_{\text{Closed EGG}}$ with the closed quotient derived from simultaneous videokymographic imaging in two subjects who varied register and glottal adduction [11]. The best agreement between the two methods occurred by applying and $Q_{\text{Closed EGG}}$ criterion of 0.2 or 0.25. They also found reasons to assume that the contacting and de-contacting moments in EGG signals might not refer to the same physical events as the beginning and cessation of airflow. They therefore stressed the importance of distinguishing between the closed phase and the vocal fold contact phase. Following their advice, we will refer to $Q_{\text{Closed FLOGG}}$ as Q_{Closed} and to $Q_{\text{Closed EGG}}$ as Q_{Contact}.

As can be seen in Figure 1, the data of Sapienza and associates showed that the relationship between Q_{Closed} and Q_{Contact} varied with SPL. This parameter is closely related to subglottal pressure, which, in both trained and untrained voices, has strong effects also on other voice source parameters [12]. This raises the question whether the relationship between Q_{Closed} and Q_{Contact} is affected also by such parameters. The question seems relevant as, at least in untrained voices, an increase of vocal loudness is often combined with an increase of glottal adduction, which affects both Q_{Closed} and Q_{Contact} [13]. In addition, fundamental frequency, which has not been varied in previous comparisons of Q_{Closed} and Q_{Contact}, may be another relevant parameter. Hence, there seems to be reasons to analyze also how the relationship between Q_{Closed} and Q_{Contact} varies with different voice source parameters.

The aim of the present investigation is to analyze, in more detail, the differences between $Q_{Contact}$, obtained with the 3/7 criterion, and Q_{Closed}, considering also effects of voice source parameters.

Method

As Q_{Closed} varies with Psub [12, 14], it was important to measure Q_{Closed} and $Q_{Contact}$ over a wide range of vocal loudness. In untrained voices, soft phonation is often associated with breathy voice quality and weak high partials, which makes inverse filtering difficult or even impossible. Also, in such voices, vocal loudness variation is often associated with F0 variation; the softer the phonation, the lower the pitch. For these reasons, we used classically trained male singers as subjects, 2 tenors and 3 baritones.

The subjects were asked to sing a set of repetitions of the syllable /pae/, sung as diminuendos on different pitches. For the recordings, made in a sound treated room at Aveiro University, Portugal, and at KTH, Sweden, a hybrid system was used, combining a Laryngograph microprocessor and a Glottal Enterprise MS-110 computer interface. This equipment allows simultaneous recording of audio, electrolaryngograph (ELG), pressure and flow signals. Audio was picked up at a measured distance from the mouth by a head-mounted electret microphone (Knowles EK3132). A 1 kHz sine wave was used to calibrate sound level; its sound pressure level was measured in dB(C) next to the recording microphone by means of a sound level meter and the value was announced on the recording. Oral pressure and flow were measured by means of pressure and flow transducers, respectively. Both signals were calibrated using Glottal Enterprises calibration devices. Intraoral pressure during /p/ occlusion was determined as an estimation of subglottal pressure (Psub).

The four signals were recorded using Laryngograph software, Speech Studio (Laryngograph©) and stored as wav files. The recorded wav files were converted into smp files and analyzed by SoundSwell workstation software (Hitech, Solna Sweden). FLOGG was analysed by means of inverse filtering, using the custom made Decap software (by Svante Granqvist, KTH). This program can display flow waveform and spectrum in separate windows (see Figure 2).

<Please insert Figure 2 about here>

The frequencies and bandwidths of the inverse filters are set manually and the program applies the classical equations for calculating the transfer function corresponding to the given combination of formant frequencies and bandwidths [15]. The resulting FLOGG and corresponding spectrum are displayed in quasi-real-time. Provided that the filters were correctly set, the output then displays the waveform and spectrum of the transglottal airflow. The program can also display the derivative of the ELG signal (dELG), with an adjustable time delay corresponding to the delay of the acoustic signal relative to the ELG. The formant frequencies and bandwidths were adjusted according to three criteria: (1) ripple free closed phase; (2) voice source spectrum envelope as void as possible of peaks and valleys near formants; and (3) synchrony between the negative dELG peak and the maximum declination rate of transglottal flow during closure [16].

The FLOGG was analyzed by means of the custom made SNAQ software (Svante Granqvist), displaying the FLOGG and its time derivative (see Figure 3).

<Please Insert Figure 3 about here>

After marking both period and closed phase in the waveform, it offers values of fundamental frequency (F0), maximum flow declination rate (MFDR), normalised amplitude quotient (NAQ), dominance of the fundamental (H1-H2) and closed quotient (Q_{Closed}). An example of is displayed in Table 1.

<Please Insert Table 1 about here>

Contact quotient data (CQ) were derived from the ELG signal by means of the QAnalyses option, included on the Speech Studio Software. It allows measurements of several vocal parameters (i.e. vocal fold contact area, amplitude of vibration, time of vibration and speed of execution). The CQ ratio, determined as a 3/7 (43%) of the peak-to-peak amplitude of the ELG waveform was used as $Q_{Contact}$ value (see Figure 4). Measures were taken at the same time coordinates as the ones used for FLOGG Q_{Closed} measurements.

<Please insert Figure 4 about here>

Results

Figure 5 shows, for each of the five participants, the median and the interquartile ranges for Q_{Closed} and $Q_{Contact}$. For all singers, Q_{Closed} was smaller than $Q_{Contact}$; the latter varied considerably between subjects and presented higher number of outliers when compared with Q_{Closed}.

<Please insert Figure 5 about here>

Before testing if differences between Q_{Closed} and $Q_{Contact}$ were significant for this group of singers, a Shapiro-Wilk test was run to determine if data were normally distributed. The results revealed a skewed distribution ($p<0.05$): a pronounced positive and a smaller negative skewedness were found for Q_{Closed} (0.501) and for $Q_{Contact}$ (-0.151).

Given this result, a non-parametric Wilcoxon signed rank test was carried out. It revealed that the median and interquartile range values were significantly lower for Q_{Closed} than for $Q_{Contact}$ [$z= -9.8$; $p<0.01$], see Figure 6.

<Please insert Figure 6 about here>

The causes of this difference were further analyzed: Q_{Closed} and $Q_{Contact}$ were compared for identical periods, measured for all tones sung by all singers. The result is illustrated in Figure 7. A systematic relationship was revealed, specified in the following equation (Eq. 1):

<Please insert Figure 7 about here>

$Q_{Contact} = 0.705 * Q_{Closed} + 0.20$ [$R^2 = 0{,}531$, standard error of estimate 0.07] (Eq. 1)

As according to previous research, Q_{Closed} is correlated with various phonatory parameters, such as Psub, NAQ, H1-H2 [12], it seemed worthwhile to run a Spearman's rank correlation coefficient (1) for Q_{Closed} and each of Psub, NAQ, H1-H2, and (2) for $Q_{Contact}$ and each of these same parameters. Table 2 lists the results.

<Please insert Table 2 about here>

The Q_{Closed} to $Q_{Contact}$ correlation was quite strong ($r_s = 0.714$). Also some of the other parameters showed a strong correlation. For both measures, the correlation was strongest and negative for H1-H2, followed by a positive one for Psub. Also NAQ showed a relatively strong negative correlation, particularly for Q_{Closed}. For F0, no significant correlation was found.

Another relevant question is whether one can predict values of Q_{Closed} and $Q_{Contact}$ from these same parameters (Psub, F0, NAQ, and H1-H2). A multiple regression analysis was carried out, using the z-scores of all measures (zPsub, zF0, zNAQ, and zH1-H2). The results (see Figure 8) suggest that Q_{Closed} and $Q_{Contact}$ can be predicted by means of the following two equations (Eq. 2 and Eq. 3).

$Q_{Closed} \approx 0.344 + 0.014 * zPsub - 0.017 * zNAQ - 0.0067 * zF0 - 0.076 * zH1\text{-}H2$ (Eq. 2)
[$R^2 = 0.775$]

$Q_{Contact} \approx 0.445 + 0.007 * zPsub - 0.019 * zNAQ + 0.018 * zF0 - 0.062 * zH1\text{-}H2$ (Eq. 3)
[$R^2 = 0.589$]

<Please insert Figure 8 about here>

As expected, the prediction was more accurate for Q_{Closed}; H1-H2 was the strongest predictor ($p<0.001$), followed by Psub and NAQ ($p<0.05$). With regard to $Q_{Contact}$, H1-H2 was the only predictor that showed significance ($p <0.001$). In other words, Q_{Closed} is more related to the above-mentioned voice source parameters. Using Eq. 2 and Eq. 3, the difference between $Q_{Contact}$ and Q_{Closed} can be expressed in terms of Eq. 4:

$Q_{Contact} - Q_{Closed} \approx 0.025 * zF0 + 0.006 * z(H1 - H2) - 0.007 * zPsub - 0.002 * zNAQ$ (Eq. 4)

Thus, the difference between Q_{Closed} and $Q_{Contact}$ tends to increase with rising F0, and the dependence of H1-H2, Psub and NAQ is much smaller.

Discussion

The present results corroborate earlier findings that Q_{Closed} is smaller than $Q_{Contact}$ [9] and they also shed some light on the reasons. The explanation is illustrated in Figure 9. At 20% of the peak-to-peak amplitude, the distance between the rising and falling parts of the inverted ELG waveform is much larger for Q_{Closed} than for $Q_{Contact}$; however, at 50% and at 43% (3/7), Q_{Closed} and $Q_{Contact}$ values are more similar, at least in the example shown in this Figure. Thus, the difference between Q_{Closed} and $Q_{Contact}$ obviously is heavily influenced not only by the percentage criteria used for defining the former, but also by the ELG waveform. Note that in Figure 9, in contrast to a previous study [9], the EGG waveform has not been inverted.

<Please insert Figure 9 about here>

Q_{Closed} and $Q_{Contact}$ may differ also for another reason, illustrated in Figure 10. It shows synchronized FLOGG, ELG and ELG derivative (dELG) for the same periods, obtained after delaying the ELG and dELG signals by 1 ms (corresponding to the travel time of the sound from the glottis to the microphone placed 17 cm from the mouth). Vocal fold contact is clearly depicted by the sharp positive dELG spike. It appears at the moment when the transglottal airflow ceases as indicated by the knee of the FLOGG. The opening of the glottis is clearly marked by the sudden increase of transglottal airflow in the FLOGG waveform; however, the knee in the ELG waveform appears somewhat later in time. This suggests that part of the glottis opens somewhat before complete contact between the vocal folds is lost. Such effect will obviously cause $Q_{Contact}$ to be greater than Q_{Closed}.

<Please insert Figure 10 about here>

The analysis revealed that both Q_{Closed} and $Q_{Contact}$ were significantly correlated with Psub, NAQ and H1-H2. However, the correlations were stronger for Q_{Closed}, which suggests that Q_{Closed} is more informative regarding voice source characteristics than $Q_{Contact}$. On the other hand, determining Q_{Closed} requests inverse filtering, a rather time consuming method. $Q_{Contact}$, by contrast, can be obtained from generally available voice analysis software, thus making it a more convenient measure for clinical purposes. Yet, it is limited by a greater inter-individual variation than Q_{Closed}. On the other hand, Q_{Closed} has a low inter-rater reliability [17, 18].

The novelty of this study lies in the fact that, for the first time, the relationships between both Q_{Closed} and $Q_{Contact}$ and voice source parameters have been analyzed. Also the voice source parameters that best predict Q_{Closed} and $Q_{Contact}$ were investigated. An important condition for these analyses was that professional singers were used as subjects; such subjects are trained to independently control voice source parameters, e.g. F0, loudness, and phonation type. This allowed analysis of a great number of Psub values, widely spaced over large F0 and pressure ranges.

With respect to the relationships between both Q_{Closed} and $Q_{Contact}$ and the voice source parameters, neither of them showed any significant correlation with F0. This finding may reflect the subjects' phonatory habits; some singers may produce high F0 with a more pressed phonation than they produce low F0. It would be worthwhile to examine to what extent similar findings would be obtained from females and from untrained subjects [6, 14].

With regard to other FLOGG parameters, Psub, NAQ and H1-H2 were significantly correlated with both Q_{Closed} and $Q_{Contact}$, as mentioned. An increase of Psub and a decrease of NAQ and H1-H2 were found to be associated with an increase of Q_{Closed} and $Q_{Contact}$. This corroborates findings reported elsewhere [11]. The closed phase is lengthened as a response to a rise of Psub; it is lengthened also by an increase of glottal adduction, which, in turn, causes a decrease of both NAQ and H1-H2. It may be mentioned that the marking of the closed phase in the SNAQ analysis affects neither NAQ nor H1-H2.

H1-H2 seems to be the best predictor for both Q_{Closed} and $Q_{Contact}$ measures. This is in accordance with previous studies, which have shown a strong correlation between Q_{Closed} and the strength of the voice source fundamental [19]. The lower predictor strength of this parameter for $Q_{Contact}$ would reflect the fact that this measure is less strongly associated with the voice source than Q_{Closed}. This is another reason for carefully distinguishing between Q_{Closed} and $Q_{Contact}$.

Conclusion

This study analyzed the relationship between FLOGG based and ELG based closed quotient, Q_{Closed} and $Q_{Contact}$, respectively. The Q_{Closed} values tend to be lower than those for $Q_{Contact}$. In addition, the latter showed a greater variation between subjects and was less strongly related with Psub, NAQ, and H1-H2. Both Q_{Closed} and $Q_{Contact}$ could be best predicted by H1-H2 values, although the former with a higher accuracy. Thus, although $Q_{Contact}$ and Q_{Closed} are related, they differ and should not be confused.

Acknowledgements

The authors wish to thank the participants, the Foundation Gomes Teixeira, and Bayer Ltd for the support on equipment acquisition.

References

1. Baken, RJ & Orlikoff, RF. Clinical Measurement of Speech Voice (2nd ed.). San Diego: Singular Publishing Group; 2000.

2. Peterson, KL, Barkmeier, JM, Verdolinimarston, K & Hoffman, HT. Comparison of Aerodynamic and Electroglottographic Parameters in Evaluating Clinically Relevant Voicing Patterns. Ann Oto. Rhinol Laryngol, 1994: 103(5), 335-346.

3. Rothenberg, M. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. J Acoust Soc Am, 1973: 53(6), 1632-1645.

4. Baken, RJ. Electroglottography. J Voice*,* 1992: *6*(2), 98–110.

5. Abberton, ERM, Howard, DM & Fourcin, AJ. Laryngographic assessment of normal voice: a tutorial. Clin Linguist Phon, 1989: 3(3), 281-296.

6. Howard, DM. Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. J Voice, 1995: 9(2), 163-172.

7. Howard, DM, Lindsey, GA & Allen, B. Toward the quantification of vocal efficiency. J Voice, 1990: 4(3), 205-212.

8. Rothenberg, M & Mashie, JJ. Monitoring Vocal Fold Abduction Through Vocal Fold Contact Area. J Speech Lang Hear Res, 1988: 31, 338-351.

9. Sapienza, CM, Stathopoulos, ET & Dromey, C. Approximations of open quotient and speed quotient from glottal airflow and egg waveforms : Effects of measurement criteria and sound pressure level. J Voice, 1998: 12(1), 31–43.

10. Henrich, N, d'Alessandro, C, Doval, B & Castellengo, M. On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. J Acoust Soc Am, 115(3), 2004: 1321-1332.

11. Herbst, CT, Ternstrom, S & Svec, JG. Investigation of four distinct glottal configurations in classical singing-A pilot study. J Acoust Soc Am, 2009: 125(3), EL104-EL109.

12. Sundberg, J, Andersson, A & Hultqvist, C. Effects of subglottal pressure variation on professional baritone singers' voice sources. J Acoust Soc Am, 1999: 105, 1965-1971.

13. Scherer, RC. Laryngeal function during phonation. In JS. Rubin, RT. Sataloff, GS. Korovin & WJ. Gould (Eds.) Diagnosis and Treatment of Voice Disorders. New York: Igaku-Shoin; 1995: 86-104.

14. Sundberg, J, Fahlstedt, E & Morell, A. Effects on the glottal voice source of vocal loudness variation in untrained female and male subjects. J Acoust Soc Am, 2005: 117, 879-885.

15. Fant, G. Vocal-tract determinants of resonance frequencies and bandwidths J Acoust Soc Am, 1976: 59, S70-S70.

16. Sundberg, J, Lã, FMB & Gill, BP. Professional male singers' formant tuning strategies for the vowel /a. Logopedics Phoniatrics Vocology, 2011: 36(4), 156-167.

17. Lehto, L, Airas, M, Bjorkner, E, Sundberg, J & Alku, P. Comparison of two inverse filtering methods in parameterization of the glottal closing phase characteristics in different phonation types. J Voice, 2007: 21(2), 138-150.

18. Mecke, AC, Sundberg, J, Granqvist, S & Echternach, M. Comparing closed quotient in children singers' voices as measured by high-speed-imaging, electroglottography, and inverse filtering. J Acoust Soc Am, 2012: 131(1), 435-441.

19. Salomão, GL, & Sundberg, J. Relation between perceived voice register and flow glottogram parameters in males. J Acoust Soc Am, 2008: 124(1), 546-551.

FIGURE CAPTIONS

Figure 1. Data on $Q_{\text{Closed FLOOG}}$ (circles) and $Q_{\text{Closed EGG}}$ (diamonds) derived from Sapienza et al. (1998), after conversion of Q_{Open} to Q_{Closed}. Left and right panels refer to results obtained by applying a 20% and 50% criterion, respectively. The dashed lines and the equations represent a linear approximation of the data points.

Figure 2. Decap display showing waveform and derivative of the ELG (upper panel) and the spectrum of the input audio and of the filtered flow (lower panel). Formant bandwidths are given on an arbitrary scale along the ordinate. The arrows show the formant frequencies and bandwidths used for the inverse filtering. The two parallel curves in the lower pannel represent realistic bandwidths according to Fant (1970). (after Sundberg, Lã & Gill, 2011).

Figure 3. SNAQ display showing FLOGG and its derivative (upper and lower window, respectively). Vertical markers show the manually set closed phase and period.

Figure 4. Example of an ELG waveform with contact and open quotients (CQ and OQ, respectively) illustrating application of the 3/7 criterion of the peak-to-peak amplitude for measuring Q_{Contact}.

Figure 5. Median and interquartile ranges for both Q_{Closed} (left) and Q_{Contact} (right) for each participant.

Figure 6. Box-plots showing the median and interquartile range values for Q_{Closed} and Q_{Contact}.

Figure 7. Comparison of all measured Q_{Contact} and Q_{Closed} values for identical samples, produced by the five participants.

Figure 8. Scatter plot comparing observed and predicted values of Q_{Closed} (open circles) and Q_{Contact} (asterisks)

Figure 9. Illustration of one reason why Q_{Closed} and Q_{Contact} (upper and lower curves) tend to differ, if both are measured applying the criterion used by Sapienza et al. (1998), i.e. as the time interval between the rising and falling parts of the waveforms at 20% and 50% of the peak-to-peak amplitude.

Figure 10. Synchronised FLOGG, ELG and dELG waveforms (upper, middle and lower curves, respectively) of three periods taken from one of the recorded singers.