Comparison of parametrization methods of electroglottographic and inverse filtered acoustic speech pressure signals in distinguishing between phonation types

Dong Liu [1, 2], Elina Kankare [3], Anne-Maria Laukkanen [4,*], Paavo Alku [5]

[1] CAS Key Laboratory of Microscale Magnetic Resonance and Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China
[2] Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China
[3] Ear and Oral Diseases, Department of Phoniatrics, Tampere University Hospital, Tampere, Finland
[4] Speech and Voice Research Laboratory, University of Tampere, Tampere, Finland
[5] Department of Signal Processing and Acoustics, Aalto University, Helsinki, Finland
*Corresponding author: Anne-Maria.Laukkanen@uta.fi

ABSTRACT

This study compared for the electroglottographic (EGG) signal how well six earlier presented and two new parameters distinguish between normal, breathy and pressed phonation and how well they correlate with perceptual evaluation. The results were compared with those obtained for nine parameters describing the glottal flow waveform obtained through inverse filtering of the acoustic speech pressure signal. Acoustic and dual-channel EGG signals were recorded for twenty female and twenty male subjects with healthy voices phonating sustained samples of the vowel [a:] in their habitual normal voice and in simulated breathy (hypofunctional) and pressed (hyperfunctional) phonation. The samples were perceptually evaluated by five voice specialists and rated for firmness of phonation. The best examples from 12 females and 12 males were used for the analyses. Few earlier studies have ranked the behavior of this many EGG and glottal flow parameters from this large speech data.

Although the parameters differed in their ranking order, contact quotient calculated with a criterion level at 50% both from the EGG and the inverse filtered signal was strong in correlating with perception and in distinguishing phonation types in cases where fundamental frequency and sound pressure level also varied. When this variation was taken into account, the normalized amplitude quotient NAQ still had an effect in predicting voice quality. The results will have applicability in voice training and therapy and in development of machine learning -based classification methods.

Key Words: Contact quotient (CQ), normalized amplitude quotient (NAQ), perceived firmness, voice training and therapy.

1. Introduction

How the voice is produced - i.e., phonation quality - is important, as it essentially affects how the voice sounds and functions in communication and how it resists vocal loading. Thus, phonation quality also has an important role in the prevention of potential traumas related to vocal overloading. Of special interest are non-invasive methods of studying phonation quality. Electroglottography (EGG) is a non-invasive method where a high frequency, low voltage current is fed through the larynx to study changes in the vocal fold contact area during phonation, based on the electrical impedance changes the varying contact causes [1]. EGG has been used to study phonation types [2-6], vocal differences between genders [7-10], ages [11-13], vowels [5, 14-17], emotional expressions [18-21], vocal registers [e.g., 22-27] and vibrato or tremor [28, 29], and the effects of voice training [e.g., 30, 31], vocal exercises [32-38], and voice disorders [16, 39-43]. The correspondence between the EGG signal waveform and physiologic and acoustic measures has been investigated and found to be relatively good [44-50]. Due to difficulties in pointing to the exact beginning and ending of the opening and closing times of the glottis when using the EGG, the method has been regarded as more suitable for analyzing the duty cycle [51]. The relative contact time (contact quotient, CQ, i.e., contact time divided by period time) has been extensively focused on. It has been found to distinguish between registers [26, 27] and vocal expressions of emotions [19], to differentiate healthy and disordered voices at least in some cases [43], and to correlate with perception of voice quality [6, 52] and even to some extent with the impact stress (force per unit area) in vocal fold vibration [53]. The impact stress (IS) is regarded as the main loading factor during phonation [54]. Since IS is difficult to measure in humans [55, 56], methods for non-invasive estimation of IS are important.

CQ has been measured using different peak-to-peak amplitude-based criterion levels from 10 to 80% [4, 6, 8, 11, 51, 57-58] because it is problematic to place the exact beginning and ending of the glottal closing events in the EGG signal. The choice of criterion level has been made on the basis of the sample and signal types or based on another method used for comparison with the EGG signal (such as stroboscopy, high-speed filming, inverse filtering, videokymography, or modeling of vocal fold vibration). Hacki suggests the use of area-based contact quotient (CQA) to study disordered voices [41]. The results of Higgins and Schulte showed that gender effects become visible for criterion levels from 55% upward [8]. In male singers, a CQ with a criterion level at 25% (CQ25%) seems to fit best with videokymographic images [47]. According to Kania et al. [59], the criterion level of a CQ higher than 25% is more affected by F0 and intensity than phonation type in male voices. Furthermore, Kankare et al. found that in the female speaking voice, a CQ with criterion levels at 25% and 35% (CQ35%) correlates best with perceived phonation type, and CQ25% is least affected by F0 and sound pressure level (SPL) but seems to reflect phonation type best [52].

To avoid the difficulty of choosing the glottal opening instant (GOI) and glottal closing instant (GCI) in the EGG signal [49, 60], the first derivative of the EGG signal (DEGG) has been used to detect the GOI and GCI, as shown in Figure 1. One problem related to the use of DEGG is that the signal is vulnerable to noise, and it may be very difficult to pinpoint its opening instant in particular. Therefore, a hybrid parameter for the calculation of the CQ has been proposed [30, 61]. The opening instant is obtained by using a criterion level of about 42% (three-sevenths) of the peak-to-peak amplitude of the EGG, and the closing instant is defined by the maximum peak of the DEGG.

The amplitude of the maximum peak of the DEGG (MDEGG) reflects the glottal closing speed. Therefore, it should also correlate with SPL, F0, and phonation type. Results by Kankare et al. showed that MDEGG correlates with the perceived firmness of phonation in female speakers [62]. As far as the authors know, the parameter has not been systematically studied. For example, the parameter's behavior has not been tested for male voices.

Inverse filtering (of either flow or the speech pressure signal) is another non-invasive method for studying voice quality, and many methods have been applied to parametrize the resulting volume velocity waveform. Relative glottal closing speed (closing quotient; ClQ; i.e., closing time divided by period time) calculated from the volume velocity waveform is known to increase with SPL and a stronger adduction [63]. It thus seems to be well suited for estimating vocal loading, since IS also increases together with these factors [64]. The pulse asymmetry parameter speed quotient (SQ; i.e., glottal opening time divided by closing time) has also been found to increase with loudness, especially in males [63, 65]. Furthermore, it has been reported to correlate with perceived effort of voice production [60]. On the other hand, both SQ and ClQ are most sensitive to abnormalities of the glottal flow, for example due to lesions of the vocal folds [66]. The normalized amplitude quotient (NAQ_{inv}) from the glottal volume waveform has been found to distinguish between phonation types [67]. In the present paper, NAQ is for the first time calculated for the EGG. Glottal spectrum-based parameters like the harmonic level difference between the first two harmonics (DH12), the harmonic richness factor (HRF), and the parabolic spectrum parameter (PSP) have also been shown to correlate with voice quality [68-70].

The glottal inverse filtering has several advantages, such as the method's ability to estimate the voice source non-invasively from the microphone signal, and the possibility to implement an analysis in an automatic manner for modern applications such as parametric speech synthesis [71]. Glottal inverse filtering, however, suffers from a few drawbacks, the most severe of which is poor estimation accuracy in the analysis of high-pitched speech due to the biasing of the formant estimates by the sparse harmonics in the spectra of high-pitched speech. In addition, most of the inverse filtering algorithms are not capable of modeling non-linearities in speech production because the methods are built on the assumption of linearity between the source and the tract. For more details on the pros and cons of glottal inverse filtering, see two recent review articles [72, 73].

In summary, for the EGG signal, DEGG should be more accurate in reflecting the glottal opening and closing events than the threshold-based methods [49], and MDEGG should reflect perceived firmness of phonation well, at least for females voices [62]. So far MDEGG has not been tested for male voices. The hybrid parameter CQ3/7 combines threshold-based and derivative-based approaches to provide a more accurate and robust method [30]. However, derivative of the EGG signal is vulnerable to noise. CQ % < 55 should be less affected by gender [8], and CQA should be robust enough to suit even for pathological voices [41], but CQ35% should correlate best with perceived voice quality and CQ25% should distinguish phonation type best in samples where F0 and SPL also have variation [52, 59]. NAQ_{inv} and several spectrum-based parameters calculated for the inverse filtered acoustic speech pressure waveform have been found to correlate with voice quality [67-70]. However, their performance has not been extensively tested against the more traditional time-based parameters and against each other. Furthermore, NAQ has not been calculated for the EGG signal before. Additionally, in most of the previous studies, simulated phonation type has been investigated keeping F0 and SPL constant, which is an unnatural situation.

Due to the above mentioned reasons there seems to be a need to test the performance of various parametrization methods of EGG and inverse filtered signal for the same speech sample. The present study compared a set of eight EGG parametrization methods and nine glottal flow parameters (thus 17 parameters in total). The parameters chosen were for the EGG: CQ25%, CQ35%, CQ50%, CQA, CQDEGG, CQ3/7, NAQ and MDEGG. The inverse filtered signal parameters were CQ, $CQ50\%_{inv}$, CQA_{inv}, ClQ, SQ, NAQ_{inv}, DH12, HRF, and PSP. F0 and SPL were allowed to vary naturally in the samples presenting three phonation types. The questions of interest are: 1) Do the parameters differentiate between phonation types, and 2) Which of the parameters correlate best with perception of phonation quality. To the best of our knowledge, no study so far has extensively ranked the behavior of this many parametrization methods of the EGG and glottal flow signal in reflecting the phonation quality of the same sound samples.

2. Materials and methods

2.1. Subjects and recordings

Twenty females and twenty males with healthy voices volunteered as subjects. They phonated at their habitual conversational pitch and loudness on the vowel [a:] in three ways: habitual voice, breathy voice, and pressed voice. The duration of each vowel sample was approximately five seconds. The samples were recorded in a sound-treated studio using a dual-channel EGG (Glottal Enterprises; low frequency limit set to 20 Hz) and a headset microphone (AKG C477) at a distance of 6 cm from the corner of the subject's mouth. The samples were recorded on a PC through an external sound card (M-Audio, MobilePre USB) using SoundForge software. The sampling rate was 44.1 kHz and the amplitude quantization was 16 bits. The samples were calibrated for SPL measurements by using a buzzer (Boss TU-120) and a sound level meter (Bruel & Kjaer 2206).

2.2. Perceptual evaluation

The sound samples were evaluated by five experienced voice trainers for perceived firmness of phonation. Visual Analogous Scale of Judge software (Svante Granqvist) was used. The scale, ranging from 0 to 1000 units, was labeled to show either very low firmness (0 = breathy phonation) or high firmness (1000 = pressed phonation). The samples were listened to through headphones (Sony Stereo MDR-CD480). The reliability coefficient of the perceptual evaluation was high (Cronbach's alpha 0.96).

On the basis of this evaluation, the most successful samples from twelve females (mean age 36 years, range 19-52) and twelve males (mean ages 35 years, range 21–65) were selected for the analyses. The success of the samples meant that the samples that were intended to be, for example, "breathy" were also clearly rated as breathy.

2.3. Parametrization of the EGG

To make the samples comparable between subjects, the EGG signal was normalized in amplitude by using the formula:

$$\text{Normalized EGG Signal} = \frac{\text{EGG Signal} - \text{Min(EGG Signal)}}{\text{Max(EGG Signal)} - \text{Min(EGG Signal)}}. \quad \text{(Eq. 1)}$$

Within this formulation, the original EGG signal was converted into a real amplitude value between 0 and 1.

Using a Matlab script written by Dong Liu, the EGG signal was analyzed in eight different ways:

(a) *Criterion level-based methods*: CQ25%, CQ35%, and CQ50%

The beginning and the end of the contacting event of the vocal folds are defined by a criterion level of 25%, 35%, and 50% from the peak-to-peak amplitude of the normalized EGG signal, as shown in Figure 1.

(b) Area-based method: CQA

In this method, an imaginary line is placed on the EGG waveform so that the area left above and below the line is equal. The crossings of this line then are used to identify the beginning and ending of the vocal fold contact phase, as shown in Figure 1.

(c) Derivative-based method: CQDEGG

This method interprets the positive peak of the DEGG signal as the beginning of the contacting event and the subsequent negative peak in the same cycle as the end of the contacting event.

(d) Hybrid method: CQ3/7

The beginning of the contact is defined as in (c), but its end is determined as the instant when the EGG signal reaches the value of 3/7 of its maximum.

(e) NAQ

This method parameterizes the EGG signal waveform by using two amplitude domain measurements: the peak-to-peak amplitude is divided by the maximum of the first derivative.

(f) MDEGG

This method uses the maximum peak in the DEGG signal (see [62] for more details).

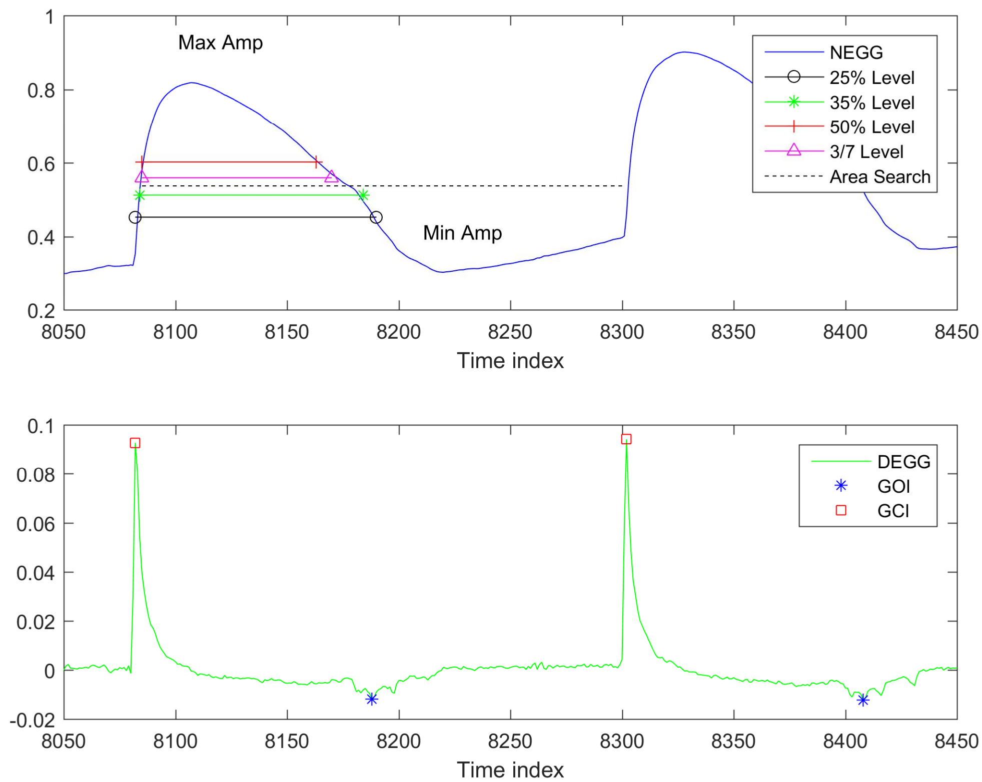

Fig. 1: Top: Illustration of setting the criterion level in the six time domain-based methods for measuring the CQ in the normalized EGG (increasing contact upwards). Bottom: The DEGG signal. GOI and GCI denote the instants of fastest glottal opening and closing, respectively. Dotted line (Area search) shows period time. Circle marks the beginning and end of vocal fold contact time when the threshold level is 25 % from the signal amplitude (difference between maximum and minimum amplitude). CQ is calculated as contact time / period time. Derivative-based CQDEGG is measured: Time from GCI to GOI / period time (time from GCI to the next GCI).

2.4. Sound pressure level analysis

The calibrated acoustic signals were analyzed for sound pressure level (SPL) using Praat signal analysis software.

2.4.1. Inverse filtering

Estimation of the glottal flow was computed by inverse filtering the recorded speech pressure waveforms with iterative adaptive inverse filtering (IAIF) [74]. IAIF estimates the glottal flow automatically by using a straightforward two-stage procedure in which the spectral contribution of the glottal source is first modelled pitch-asynchronously from a speech frame using low-order all-pole modeling. By first canceling the estimated glottal contribution, IAIF computes an all-pole model for the vocal tract, the inverse of which is finally used in removing the effect of the vocal tract resonances from speech to obtain the estimated glottal flow. As an all-pole modeling method, IAIF can be computed either with conventional linear prediction [75] or with discrete all-pole (DAP) modeling [76]. In the current study, DAP was used because it has been shown [77] to reduce formant ripple, a time domain artefact in the glottal flow estimates caused by poorly estimated vocal tract resonances.

IAIF analysis was conducted with Aparat, an interactive glottal inverse filtering and parameterization tool [78]. Aparat enables the semiautomatic and user-friendly estimation of glottal flow with IAIF using the following procedures. First, the user imports the recorded speech signal into the system, and selects a section of the time-domain waveform to be inverse filtered. In the current study, this analysis frame was always 50 ms in duration, and it was positioned in the middle of the recorded utterance. In addition, the sampling frequency was reduced to 8 kHz, and the signal was high-pass filtered to remove frequencies below 70 Hz. After this, the Aparat system automatically computes glottal flow estimates for the selected frame by varying the IAIF parameters (model order of the vocal tract, lip radiation coefficient) and displays the resulting waveforms in the time domain on the computer screen. The user is then given an opportunity to subjectively compare the different waveforms, and to select, by mouse, the best one for further analyses (for details, see [78]). In this study, the following selection criteria were adopted: the best estimate was the one that showed the maximally flat closed phase for the glottal flow waveform and a minimal remaining formant ripple. These criteria have been widely used in previous glottal inverse filtering studies [e.g., 79-81]. After the best estimate has been selected, Aparat automatically parameterizes the signal using a multitude of parameters. In this study, the following glottal flow parameters were selected for further analysis: Closed quotient (CQ1, CQ2) and speed quotient (SQ1, SQ2) measured by using the primary and secondary glottal opening as the landmarks (see Figure 2), $CQ50\%_{inv}$, CQA_{inv}, ClQ, NAQ_{inv}, DH12, PSP, and HRF.

The calculation of DH12, HRF, and PSP from the glottal flow spectrum is illustrated in Figure 3 [82]. Previous studies indicate that when the phonation type changes from breathy to normal and then further to pressed, the values of NAQ_{inv} [67], ClQ [67, 69], OQ (open quotient; i.e., the inverse of CQ) [69], and PSP [70] generally follow a monotonically declining trend, whereas the opposite trend is typically observed for the values of SQ [69] and HRF [69].

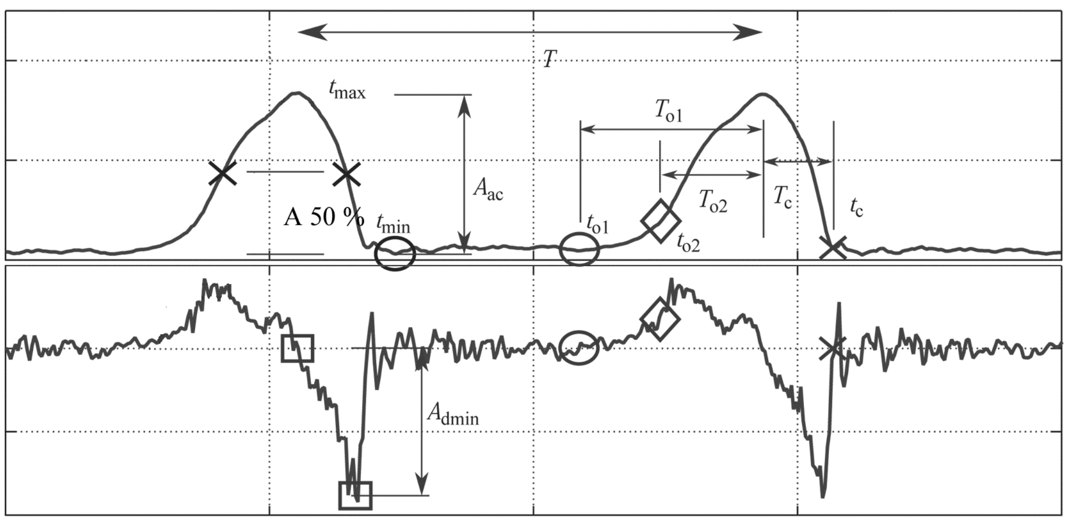

Fig. 2. Instants used to calculate time and amplitude domain parameters. Glottal flow waveform (top) and its first derivative (bottom). t_{01} and t_{02} denote the primary and secondary openings of the glottis. T is the period length. T_{o1} and T_{o2} are alternative opening times of the glottis. T_c is the closing time of the glottis. CQ1= T-(T_{o1}+T_c), CQ2= T-(T_{o2}+T_c). SQ1= T_{o1}/T_c. SQ2= T_{o2}/T_c. t_{max}-t_{min} is the flow signal amplitude A_{ac}. A_{dmin} = minimum of the first derivative. NAQ_{inv}= (A_{ac}/A_{dmin})/T. The crosses (X) denote 50% of the flow signal amplitude, which was used to calculate $CQ50\%_{inv}$ (Adapted from [78]).

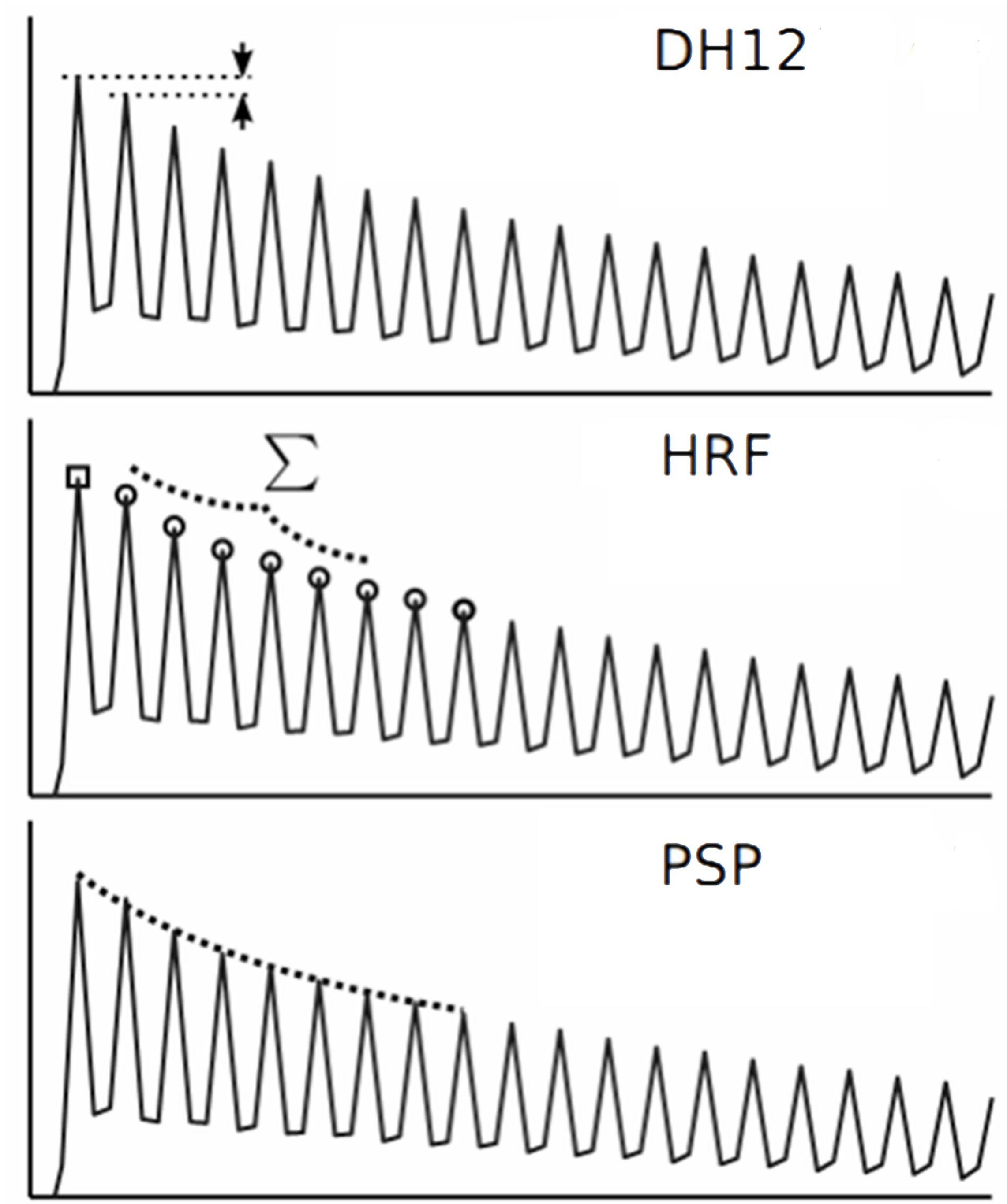

Fig. 3. Illustration of the calculation of DH12, HRF, and PSP from the glottal flow spectrum (Adapted from [82]).

2.4.2 Statistical analyses

Mean and standard deviation (SD) values were computed to describe parameter values in phonation types, and *t*-tests (repeated measures analysis of variance; RM ANOVA) were calculated to study differences between the types. Relations between the

EGG and inverse filtering parameters and firmness of phonation, F0, and SPL were studied with Spearman's rho and linear regression analysis. Analyses were made using SPSS21.

3. Results

Table 1 shows the mean parameter values and the *t*-test results, Tables 2 and 3 show the results of the correlation analyses, and Tables 4 and 5 show the regression analysis results.

Table 1(a). Means (and standard deviations) of parameters measured from the EGG and inverse filtered signals for different phonation types in 12 males and 12 females. X marks the parameters for which *none* of the phonation types differed significantly from each other, i.e., $p = \geq 0.05$. (RM-ANOVA, pairwise comparisons with Sidak adjustment for multiple comparisons.)

		FEMALES				MALES		
EGG	Breathy	Normal	Pressed		Breathy	Normal	Pressed	
F0	173.5 (20.9)	190.7 (26.8)	219 (40.2)		111.1 (21.9)	115.2 (23.2)	168.7 (61.4)	
SPL	80.4 (6.1)	85.9 (3.9)	94.9 (4.2)		79.1 (4.1)	86.0 (4.6)	95.1 (8.2)	
Firmness	142.7 (118.7)	504 (48.1)	770.6 (84.7)		145.3 (83.2)	561.9 (44.6)	803.0 (70.9)	
CQDEGG	0.34 (0.10)	0.41 (0.08)	0.58 (0.13)		0.33 (0.08)	0.49 (0.07)	0.61 (0.04)	
CQ25%	0.52 (0.09)	0.53 (0.08)	0.64 (0.10)		0.43 (0.06)	0.53 (0.07)	0.65 (0.06)	
CQ35%	0.44 (0.09)	0.47 (0.07)	0.60 (0.11)		0.37 (0.05)	0.48 (0.06)	0.60 (0.06)	
CQ50%	0.35 (0.09)	0.40 (0.07)	0.53 (0.12)		0.30 (0.04)	0.41 (0.06)	0.53 (0.06)	
CQA	0.41 (0.08)	0.44 (0.07)	0.57 (0.11)		031 (0.11)	0.45 (0.05)	0.57 (0.06)	
CQ3/7	0.37 (0.09)	0.43 (0.05)	0.53 (0.07)		0.35 (0.05)	0.45 (0.04)	0.53 (0.04)	
NAQ	0.44 (0.12)	0.29 (0.11)	0.28 (0.14)		0.25 (0.16)	0.17 (0.13)	0.28 (0.22)	X
MDEGG	0.06 (0.01)	0.08 (0.03)	0.10 (0.05)		0.06 (0.03)	0.07 (0.02)	0.09 (0.03)	X
		FEMALES				MALES		
INVERSE	Breathy	Normal	Pressed		Breathy	Normal	Pressed	
CQ1	0.15 (0.17)	0.08 (0.09)	0.32 (0.16)		0.03 (0.05)	0.17 (0.14)	0.29 (0.20)	
CQ2	0.27 (0.20)	0.23 (0.15)	0.47 (0.18)		0.14 (0.08)	0.38 (0.17)	0.54 (0.08)	
CQ50%inv	0.53 (0.12)	0.57 (0.06)	0.71 (0.05)		0.45 (0.05)	0.60 (0.05)	0.71 (0.07)	
CQAinv	0.42 (0.14)	0.33 (0.08)	0.52 (0.11)		0.32 (0.07)	0.40 (0.08)	0.56 (0.10)	
CIQ	0.35 (0.10)	0.34 (0.07)	0.24 (0.09)	X	0.41 (0.08)	0.26 (0.05)	0.24 (0.07)	X
SQ1	1.48 (0.31)	1.80 (0.45)	2.04 (0.68)	X	1.40 (0.40)	2.30 (0.38)	2.20 (0.94)	X
SQ2	1.11 (0.32)	1.30 (0.23)	1.27 (0.49)	X	1.15 (0.41)	1.46 (0.65)	1.03 (0.43)	X
NAQinv	0.19 (0.06)	0.16(0.02)	0.11 (0.02)		0.20 (0.05)	0.12 (0.01)	0.11 (0.04)	X
DH12	12.7 (10.8)	13.0 (2.7)	7.00 (3.25)	X	16.0 (7.9)	9.76 (1.89)	4.08 (4.91)	
PSP	0.41 (0.26)	0.24 (0.05)	0.17 (0.04)		0.36 (0.07)	0.18 (0.04)	0.25 (0.20)	X
HRF	-3.50 (8.3)	-5.02 (2.53)	2.31(3.26)	X	-4.96 (4.80)	-2.09 (1.60)	4.45 (4.95)	

Table 1(b). Significance of differences (*p*-values) between phonation types (RM ANOVA, pairwise comparisons with Sidak adjustment for multiple comparisons. NS = non-significant, i.e., $p = \geq 0.05$)

		FEMALES				MALES	
EGG	Breathy/ Normal	Breathy /Pressed	Normal/Pressed		Breathy/ Normal	Breathy /Pressed	Normal/Pressed
F0	0.009	0.001	0.021		ns	0.011	0.009
SPL	0.001	< 0.001	< 0.001		< 0.001	< 0.001	0.001
Firmness	< 0.001	< 0.001	< 0.001		< 0.001	< 0.001	< 0.001
CQDEGG	ns	0.002	0.002		< 0.001	< 0.001	0.002
CQ25%	ns	ns	0.023		0.004	< 0.001	0.001
CQ35%	ns	0.029	0.011		0.001	< 0.001	< 0.001
CQ50%	ns	0.009	0.004		0.001	< 0.001	< 0.001
CQA	ns	0.012	0.006		0.015	< 0.001	< 0.001
CQ3/7	ns	0.004	0.009		0.001	< 0.001	< 0.001
NAQ	0.008	0.006	ns		ns	ns	ns
MDEGG	0.028	0.013	ns		ns	ns	ns
		FEMALES				MALES	
INVERSE	Breathy/ Normal	Breathy /Pressed	Normal/Pressed		Breathy/ Normal	Breathy /Pressed	Normal/Pressed
CQ1	ns	0.022	0.008		0.023	0.004	ns
CQ2	ns	ns	0.019		0.001	< 0.001	0.03
CQ50%inv	ns	< 0.001	0.001		< 0.001	< 0.001	0.001
CQAinv	ns	0.01	0.003		0.043	0.001	0.001
ClQ	ns	ns	ns		ns	ns	ns
SQ1	ns	ns	ns		ns	ns	ns
SQ2	ns	ns	ns		ns	ns	ns
NAQinv	ns	< 0.001	0.007		ns	ns	ns
DH12	ns	ns	ns		ns	0.004	ns
PSP	ns	ns	< 0.001		ns	ns	ns
HRF	ns	ns	ns		0.006	0.001	ns

Table 1 (a) shows that F0 and SPL increased together with the firmness of phonation. All CQ parameters and MDEGG from the EGG signal increased together with the firmness of phonation, while NAQ decreased, as could be expected, but only for females. For males, instead, NAQ was smaller in normal phonation than in breathy or pressed phonation, and larger in pressed than in breathy phonation. Parameters from the glottal volume waveform signal showed even more variations in the pattern. CQ1, CQ2, and CQA_{inv} did not behave linearly for females. In males, the average values of SQ1 and SQ2 were highest in normal phonation and lowest either in breathy or pressed phonation. In females, SQ2, DH12, and HRF showed a similar nonlinear pattern. ClQ and NAQ_{inv} decreased with increasing firmness of phonation in both genders, and DH12 decreased and HRF increased in males.

According to RM ANOVA results, most EGG parameters and three out of nine glottal waveform parameters differentiated all phonation types from each other for males, while for females only F0, SPL, and perceived firmness distinguished all types statistically significantly (see Table 1 (b)). For males, NAQ and MDEGG from the EGG parameters and ClQ, SQ1, SQ2, NAQ_{inv}, and PSP from the glottal waveform parameters did not distinguish any phonation types from each other. Additionally, CQ1 and HRF did not distinguish normal from pressed, and DH12 distinguished only breathy from pressed. F0 did not differ significantly between breathy and normal in males. For females, NAQ from EGG and MDEGG did not distinguish normal from pressed, CQ25% (from EGG) did not distinguish breathy from the other phonation types, and CQDEGG, CQ35%, CQ50%, CQA, and CQ3/7 from the EGG did not distinguish breathy from normal. Furthermore, of the glottal waveform parameters for females, ClQ, SQ, DH12, and HRF did not distinguish significantly any of the three phonation types. PSP and CQ2 did not distinguish breathy from the other types, while CQ1, $CQ50\%_{inv}$, CQA_{inv}, and NAQ_{inv} did not distinguish breathy from normal.

To sum up, SPL and perceived firmness differentiated all phonation types and CQDEGG, CQ35%, CQ50%, CQA, and CQ3/7 differentiated either all or 2/3 of the phonation types in both genders. The same was found for $CQ50\%_{inv}$, CQA_{inv}, and either CQ1 or CQ2 from the glottal waveform parameters. For females, CQDEGG, CQ50% and CQ3/7 distinguished the phonation types best, and CQ35%, CQ3/7 and CQ50% for males. From inverse filtered signal, $CQ50\%_{inv}$ distinguished best in both genders. See Table 1 (b).

For both genders, the hybrid parameter CQ3/7 seemed to correlate best with firmness of phonation (Table 2, Figure 4). The second and third best were CQDEGG and CQ50% in females and CQA and CQ50% for males. NAQ from EGG showed the weakest correlations out of all the parameters studied in both genders. Of the glottal waveform parameters (see Table 3, Figure 5), PSP, NAQ_{inv}, and $CQ50\%_{inv}$ correlated best with firmness of phonation in females, whereas for males the best parameters were $CQ50\%_{inv}$ and DH12. SQ1 and SQ2 showed the weakest correlations of the parameters in both genders. The correlation results should be interpreted with caution due to the relatively small number of subjects.

Table 2. Correlations (Spearman's rho) between perceived voice quality ('firmness'), SPL, F0 and *EGG parameters* in females and males.

Females			Firmness Mean	SPL	F0	CQDEGG	CQ25%	CQ35%	CQ50%	CQA	CQ3/7	NAQ	MDEGG
Spearman 's rho	Firmness Mean		1,000	,696	,603	,692	,508	,598	,626	,616	,713	-,508	,618
				,000	,000	,000	,002	,000	,000	,000	,000	,002	,000
			36	36	36	36	36	36	36	36	36	36	36
	SPL		,696	1,000	,237	,609	,357	,467	,532	,491	,539	-,368	,292
			,000		,164	,000	,033	,004	,001	,002	,001	,027	,084
			36	36	36	36	36	36	36	36	36	36	36
	F0		,603	,237	1,000	,258	,206	,223	,213	,235	,372	-,105	,492
			,000	,164		,129	,228	,191	,212	,169	,026	,541	,002
			36	36	36	36	36	36	36	36	36	36	36

Males			Firmness Mean	SPL	F0	CQDEGG	CQ25%	CQ35%	CQ50%	CQA	CQ3/7	NAQ	MDEGG
Spearman 's rho	Firmness Mean		1,000	,745	,474	,856	,806	,856	,877	,880	,881	,087	,490
				,000	,004	,000	,000	,000	,000	,000	,000	,615	,002
			36	36	36	36	36	36	36	36	36	36	36
	SPL		,745	1,000	,628	,561	,606	,641	,662	,653	,679	-,208	,482
			,000		,000	,000	,000	,000	,000	,000	,000	,223	,003
			36	36	36	36	36	36	36	36	36	36	36
	F0		,474	,628	1,000	,336	,426	,460	,517	,515	,487	,199	,424
			,004	,000		,045	,010	,005	,001	,001	,003	,245	,010
			36	36	36	36	36	36	36	36	36	36	36

Table 3. Correlations (Spearman's rho) between perceived voice quality ('firmness'), SPL, F0 and *glottal volume waveform* parameters in females and males.

Females			Firmness Mean	SPL	F0	CQ1	CQ2	CQ50%inv	CQainv	ClQ	SQ1	SQ2	NAQinv	DH12	PSP	HRF
Spearman 's rho	Firmness Mean		1,000	,696	,603	,451	,429	,677	,347	-,451	,340	,143	-,707	-,525	-,720	,538
				,000	,000	,008	,013	,000	,048	,008	,053	,428	,000	,002	,000	,001
			36	36	36	33	33	33	33	33	33	33	33	33	33	33
	SPL		,696	1,000	,237	,439	,406	,601	,296	-,481	,361	,252	-,523	-,506	-,691	,501
			,000		,164	,011	,019	,000	,095	,005	,039	,158	,002	,003	,000	,003
			36	36	36	33	33	33	33	33	33	33	33	33	33	33
	F0		,603	,237	1,000	,053	,044	,295	,091	-,074	,215	,085	-,389	-,195	-,351	,187
			,000	,164		,769	,807	,096	,615	,684	,230	,637	,025	,278	,046	,297
			36	36	36	33	33	33	33	33	33	33	33	33	33	33

Males			Firmness Mean	SPL	F0	CQ1	CQ2	CQ50%inv	CQainv	ClQ	SQ1	SQ2	NAQinv	DH12	PSP	HRF
Spearman 's rho	Firmness Mean		1,000	,745	,474	,623	,782	,866	,768	-,691	,355	-,088	-,706	-,789	-,544	,768
				,000	,004	,000	,000	,000	,000	,000	,034	,612	,000	,000	,001	,000
			36	36	36	36	36	36	36	36	36	36	36	36	36	36
	SPL		,745	1,000	,628	,611	,692	,692	,649	-,674	,383	,048	-,641	-,626	-,394	,655
			,000		,000	,000	,000	,000	,000	,000	,021	,779	,000	,000	,017	,000
			36	36	36	36	36	36	36	36	36	36	36	36	36	36
	F0		,474	,628	1,000	,559	,515	,425	,425	-,346	,022	-,040	-,230	-,306	-,043	,330
			,004	,000		,000	,001	,010	,010	,039	,897	,816	,177	,069	,804	,049
			36	36	36	36	36	36	36	36	36	36	36	36	36	36

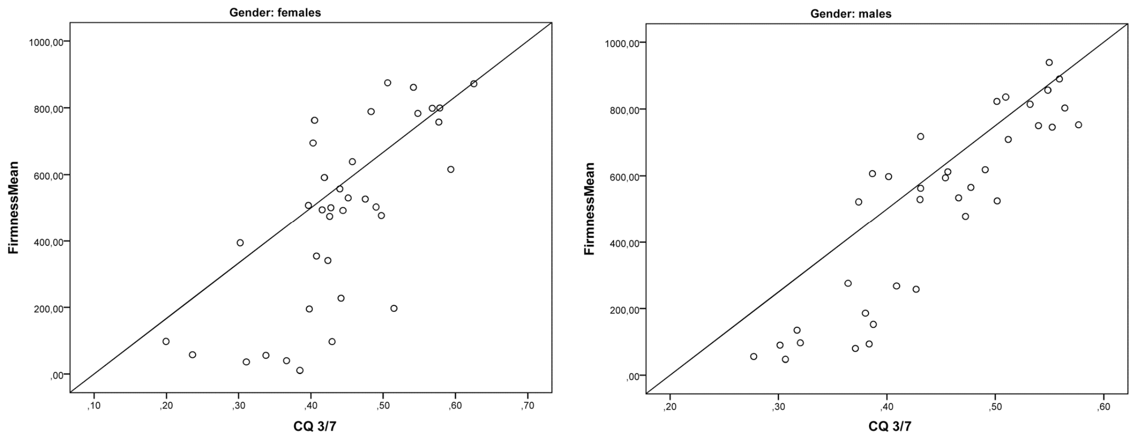

Fig. 4. Scatterplots for CQ3/7 from the EGG versus perceived firmness.

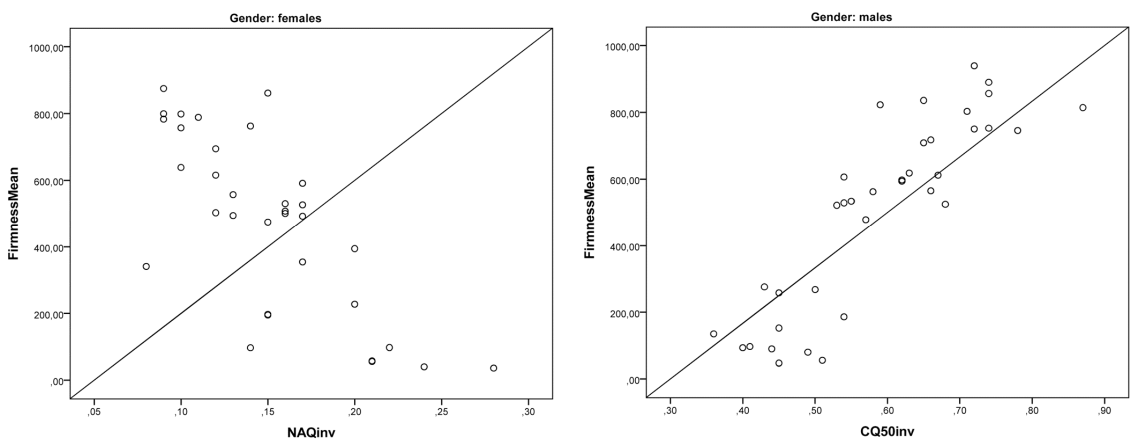

Fig. 5. Scatterplots for NAQ_{inv} and $CQ50\%_{inv}$ versus perceived firmness.

In order to study further the interrelations between firmness, F0, SPL and the EGG and glottal waveform parameters, regression analyses were carried out. Table 4 shows that the regression model for F0, SPL, and the EGG parameters explained 71% of the variation for perceived firmness in females and about 90% in males. The strongest predictors were F0, SPL, and NAQ in females and F0, SPL, CQDEGG, NAQ, and MDEGG in males.

Table 4. Linear regression analysis results for EGG parameters in females (gender 1) and males (gender 2).

Model Summary[a]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,860[b]	,739	,715	146,72382

a. gender = 1,00 b. Predictors: (Constant), NAQ, F0, SPL

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1286,144	331,600		-3,879	,000
	F0	3,256	,734	,416	4,436	,000
	SPL	15,857	3,577	,442	4,433	,000
	NAQ	-755,956	185,811	-,392	-4,068	,000

a. gender = 1,00 b. Dependent Variable: Firmness

Model Summary[a]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,954[b]	,911	,896	91,53740

a. gender = 2,00 b. Predictors: (Constant), MDEGG, NAQ, CQDEGG, F0, SPL

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1662,092	221,669		-7,498	,000
	F0	-1,073	,512	-,178	-2,096	,045
	SPL	17,609	3,230	,546	5,452	,000
	CQDEGG	1191,539	153,120	,561	7,782	,000
	NAQ	404,156	116,517	,250	3,469	,002
	MDEGG	1636,668	650,011	,169	2,518	,017

a. gender = 2,00 b. Dependent Variable: Firmness

Table 5. Linear regression analysis results for glottal waveform parameters in females (gender 1) and males (gender 2).

Model Summary[a]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,859[b]	,738	,711	143,18295

a. gender = 1,00 b. Predictors: (Constant), CQ50inv, F0, NAQinv

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-579,988	407,411		-1,424	,165
	SPL	12,149	3,795	,354	3,202	,003
	F0	2,090	,752	,278	2,778	,010
	NAQinv	-2715,207	631,161	-,485	-4,302	,000

a. gender = 1,00 b. Dependent Variable: Firmness

Model Summary[a]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,930[b]	,864	,851	109,25673

a. gender = 2,00 b. Predictors: (Constant), CQ50inv, F0, NAQinv

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-27,821	166,662		-,167	,868
	F0	1,468	,483	,244	3,039	,005
	NAQinv	-2176,681	457,531	-,436	-4,757	,000
	CQ50inv	1104,036	239,052	,480	4,618	,000

a. gender = 2,00 b. Dependent Variable: Firmness

The regression model for glottal waveform parameters explained 71% of variation in perceived firmness in females and 85% in males (Table 5). The strongest predictors for females were F0, SPL, and NAQ_{inv}, and the strongest predictors for males were F0, NAQ_{inv}, and $CQ50\%_{inv}$.

In total, the regression results suggest that phonation type-related differences in F0 and SPL strongly affect most of the EGG and glottal waveform parameters. In both genders, NAQ and NAQ_{inv} seem to have a significant effect on perceived firmness, even when the effect of F0 and SPL is taken into account.

4. Discussion

The present study aimed to find the best EGG and glottal waveform parameters to describe phonation type along the axis from breathy (hypofunctional) to pressed (hyperfunctional). F0 and SPL were allowed to vary as they normally do when the phonation type is changed. This may have affected the results somewhat and may explain the discrepancies with the results of previous investigations where pitch and loudness have been kept the same as much as possible. However, this procedure was chosen in order to allow as natural voice production as possible and to reveal the most usable and robust parameters out of the 17 that were studied.

CQ50% seemed to distinguish best the phonation types for both genders and derived from both the EGG and glottal waveform. Additionally, CQA_{inv} was a good differentiator for the glottal waveform. These parameters have been found to be robust. For example, according to Higgins and Schulte [8], CQ is affected by gender only when the criterion level is above 50%. CQA from the EGG in turn has also been found to be suited for differentiating between normal and pathological voices [41].

NAQ did not differentiate phonation types in males, and pressed and normal did not differ from each other in females (Table 1). NAQ – i.e., the amplitude of the EGG divided by MDEGG – may be affected by changes in pitch and loudness, as the amplitude of the EGG reflects vocal fold contact area, which in turn is supposed to diminish if F0 rises sufficiently or increase when SPL is raised. On the other hand, contrary to earlier findings [67], NAQ_{inv} did not distinguish breathy and normal phonation in females and distinguished none of the phonation types in males. Thus, similarly, NAQ_{inv} from the glottal waveform may have been affected by variation in F0 and SPL in the present study. However, as F0 and SPL have been taken into account in the regression model, NAQ retained its predictive power both when it was derived from the EGG and when it was calculated from the glottal waveform.

The other new EGG parameter, MDEGG, distinguished other phonation types except for pressed and normal in females, but none of the phonation types in males. Earlier results [62] had shown a relatively good correlation between MDEGG and perceived firmness of phonation in females. In the present study, the correlation was good for females but weak for males. The phonation types were simulated in the present study, whereas the ordinary phonation of different females was studied in the previous study.

CQDEGG and CQ3/7 correlated well with the perceived firmness of phonation. However, the derivative of the EGG is known to be vulnerable to noise, and, for example, the study by Kankare [62] had to exclude circa 22% of potential subjects, mainly due to the EGG derivative being too noisy. The results obtained by Herbst et al. [83] using super-high-speed filming also showed that peaks in the first derivative of the EGG do not always coincide with the exact moments of glottal opening and closing.

The waveform-reflecting parameter, SQ, and spectral parameters calculated for the glottal waveform seemed to weakly distinguish phonation types in the present material. This may be related to the fact that simulated sound samples with additional variation in F0 and SPL were studied. Even though increasing firmness of phonation has been found to result in increased SQ and decreased PSP [67, 69], a simultaneous increase in F0 may cause difficulties for inverse filtering as such. Additionally, increased F0 may lead to a more symmetric waveform (lower SQ) with a steeper spectral slope (higher PSP), and in contrast, a more whispery voice quality with noise components in the breathy phonation type of signal may result in an erroneously gentle spectral slope (lower PSP). On the other hand, there was a good correlation between spectral parameters (PSP in females and DH12 in males) and perception of voice quality. These parameters thus seem to follow grades on firmness rather well.

In line with earlier findings [52], CQ25% from the EGG seemed to co-vary less with SPL and F0 than CQ measured with a higher criterion level. The normative values that we are able to give from our results for the most robust parameters are as follows: CQ50% from EGG: 30% (breathy), 41% (normal), and 53% (pressed) in males; and 35%, 40%, and 53% in females, respectively. Similarly, the CQA values for males were 31%, 45%, and 57%, and for females 41%, 44%, and 57%. $CQ50\%_{inv}$ resulted in the mean values of 45%, 60%, and 71% in males and 53%, 57%, and 71% in females.

Future study should focus on comparison between different machine learning methods, such as Gaussian mixture models (GMMs) or support vector machines (SVMs), in distinguishing between phonation types. The results of the present study will potentially serve also the field of phonation type classification in proposing new parameterization methods to be used with advanced data driven back ends.

5. Conclusions

In order to study the most suitable parameters to distinguish between phonation types (breathy, normal, pressed) and to correlate with their perception, this study tested eight parameters describing the EGG signal and nine parameters describing the glottal flow waveform.

From the EGG signal, CQDEGG, CQ50%, and CQ3/7 distinguished the phonation types best in females, and CQ35%, CQ3/7, and CQ50% distinguished the phonation types best in males. From the inverse filtered signal, $CQ50\%_{inv}$ distinguished the best for both genders.

The hybrid parameter CQ3/7 from the EGG showed the best correlation with perceived voice quality, and CQ50% ranked among the best three. Of the glottal flow waveform parameters, PSP, NAQ_{inv}, and $CQ50\%_{inv}$ correlated best with perceived phonation quality in females and $CQ50\%_{inv}$ and DH12 correlated best with perceived phonation quality in males.

Most parameters, especially CQ, showed correlations between F0 and SPL. When their effect was taken into account in a regression model, the NAQ from both the EGG and the glottal volume waveform retained its effect as a predictor of voice quality. Additionally, in males, CQDEGG and MDEGG from the EGG and $CQ50\%_{inv}$ from the glottal flow waveform also remained as predictors of voice quality.

The normative values for the most robust parameters are as follows: For the EGG, CQ50% obtained mean values of 30% (breathy), 41% (normal), and 53% (pressed) in males, and 35%, 40%, and 53% in females, respectively. $CQ50\%_{inv}$ obtained mean values of 45%, 60%, and 71% in males, and 53%, 57%, and 71% in females, respectively.

Disclosure statement

The authors have no conflicts of interest to report.

Acknowledgements

This study was supported by the Academy of Finland (grants No. 1128095, 134868 and 284671). We thank Ms Päivi Svärd for her assistance with the figures and Mr Matthew James for the language correction.

References

[1] R.J. Baken, R.F. Orlikoff. Clinical measurement of speech and voice. Singular, San Diego, 2000.

[2] A. Fourcin, E. Abberton. First application of a new laryngograph. Medical and biological illustration 1971, 21:172-182.

[3] J.H. Esling. Laryngographic study of phonation type and laryngeal configuration. J Int Phon Ass 14 (1984) 56-73.

[4] R. Scherer, V. Vail, B. Rockwell. Examination of the laryngeal adduction measure EGGW. National Center for Voice and Speech Status and Progress Report 5 (1993) 73-82.

[5] K.L. Peterson, K. Verdolini-Marston, J.M. Barkmeier, H.T. Hoffman: Comparison of aerodynamic and electroglottographic parameters in evaluating clinically relevant voicing pattern. Ann Otol Rhinol Laryngol 103 (1994) 335-346.

[6] K. Verdolini, D.G. Druker, P.M. Palmer, H. Samawi. Laryngeal adduction in resonant voice. J Voice 12 (1998) 315-327.

[7] M.P. Robb, J.O. Simmons. Gender comparisons of children’s vocal fold contact behavior. J Acoust Soc Am 88 (1990) 1318-1322.

[8] M.B. Higgins, L. Schulte L. Gender differences in vocal fold contact computed from electroglottographic signals: The influence of measurement criteria. J Acoust Soc Am 111 (2002) 1865-1871.

[9] M. Pűtzer, M. Multiparametrische stimmqualitätserfassung männicher und weiblicher normalstimmen. Folia Phoniatr Logop 53 (2001) 73-84.

[10] Y. Chen, M.P. Robb, H.R. Gilbert. Electroglottographic evaluation of gender and vowel effects during modal and vocal fry phonation. J Speech Lang Hear Res 45 (2002) 821-829.

[11] M.B. Higgins, J.H. Saxman. Inverse-filtered air flow and EGG measures for sustained vowels and syllables. J Voice 7 (1993) 47-53.

[12] E.P. Ma, A.L. Love. Electroglottographic evaluation of age and gender effects during sustained phonation and connected speech. J Voice 24 (2010) 146-152.

[13] H.D. Mautner. A Cross-System Instrumental Voice Profile of the Aging Voice: With Considerations of Jaw Posture Effects. Ph.D. thesis- Communication Disorders Department University of Canterbury, Christchurch, New Zealand, 2011.

[14] M.B. Higgins, R. Netsell, L. Schulte. Vowel-related differences in laryngeal articulatory and phonatory function. J Speech Lang Hear Res 41 (1998) 712-724.

[15] K. Marasek, M. Pűtzer. Electroglottographical differentiation of pathological voice qualities. Larynx 1997, Marseille, France June 16-18 1997.

[16] M. Lim, E. Lin, P. Bones. Vowel Effect on Glottal Parameters and the Magnitude of Jaw Opening. J Voice 20 (2006) 46-54.

[17] N. Paul, S. Kumar, I. Chatterjee, B. Mukherjee. Electroglottographic parameterization of the effects of gender, vowel and phonatory registers on vocal fold vibratory patterns: an Indian perspective. Indian J Otolaryngol. 63 (2011) 27 – 31.

[18] P.J. Murphy, A.-M. Laukkanen. Electroglottogram Analysis of Emotionally Styled Phonation. In: Anna Esposito, Amir Hussain, Maria Marinaro, Raffaele Martone (eds.) Multimodal signals: Cognitive and Algorithmic Issues. Berlin Heidelberg: Springer, 2009 (a);pp 264-270.

[19] P.J. Murphy, A.-M. Laukkanen. Investigation of Normalised Time of Increasing Vocal Fold Contact as a Discriminator of Emotional Voice Type In: Anna Esposito and Robert Vich (eds.) Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions SpringerVerlag 2009 (b);pp 90-97.

[20] P. Murphy, A.-M. Laukkanen. Analysis of emotional voice using electroglottogram based temporal measures of vocal fold opening. In: Anna Esposito, Nick Campbell, Carl Vogel, Amir Hussain, Anton Nijholt (eds.) Development of multimodal interfaces: active listening and synchrony. Heidelberg, Germany: Springer, 2010, pp. 286-293.

[21] T. Waaramaa, E. Kankare. Acoustic and EGG analyses of emotional utterances. Log Phon Vocol38 (2013) 11-18.

[22] P. Kitzing. “Photo- and electroglottographical recording of the laryngeal vibratory pattern during different registers,” Folia Phoniatr. 34 (1982) 234–41.

[23] H.K. Shutte, W.W. Seidner. Registerabhängige Differenzierung von Elektroglottogrammen. Sprache-Stimme-Gehör 12 (1988) 59-62.

[24] N. Henrich, B. Roubeau, M. Castellengo. On the use of electroglottography for characterization of the laryngeal mechanisms. In proceedings of the Stockholm Music Acoustics Conference 2003, (SMAC 03), Stockholm, Sweden.

[25] B. Roubeau, N. Henrich, M. Castellengo. Laryngeal vibratory mechanisms: The notion of vocal register revisited. J Voice 23 (2009) 425-438.

[26] G.L. Salomão, J. Sundberg. What do male singers mean by modal and falsetto register? An investigation of the glottal voice source. Log Phon Vocol 34 (2009) 73-83.

[27] C. Herbst. Evaluation of various methods to calculate the EGG contact quotient. Diploma Thesis in Music Acoustics, Stockholm 2004. Kungliga Tekniska Högskolan, Department of Speech, Music and Hearing.

[28] A.-M. Laukkanen, E. Vilkman, U.K. Laine. Aspect of the physiological sources of vocal vibrato. A study of fundamental period-synchronous changes in electroglottographic signals obtained from one singer and two excised human larynges. Scand J Log Phon 17 (1992) 87-93.

[29] A.-M. Laukkanen, E. Vilkman. Tremor in the light of sound production with excised human larynges. In: Dejonckere PH, Hirano M, Sundberg J, editors. Vibrato. San Diego (California): Singular Publishing Group, Inc., 1995:93-110.

[30] D.M. Howard, G.A. Lindsey, B. Allen. Toward the quantification of vocal efficiency. J Voice 4 (1990) 205-212.

[31] D.M. Howard. Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. J Voice 9 (1995) 163-172.

[32] A.-M. Laukkanen. About the so called "resonance tubes" used in Finnish voice training practice. An electroglottographic and acoustic investigation on the effects of this method on the voice quality of subjects with normal voice. Scand J Log Phon 17 (1992)151-161.

[33] A.-M. Laukkanen. Voiced bilabial fricative /B:/ as a vocal exercise. Scand J Log Phon17 (1992) 181-189.

[34] I.R. Titze, E.M. Finnegan, A.-M. Laukkanen, S. Jaiswal. Raising Lung Pressure and Pitch in Vocal Warm-ups: The Use of Flow-Resistant Straws. J Singing 58 (2002) 329-338.

[35] A.-M. Laukkanen, H. Pulakka, P. Alku, E. Vilkman, S. Hertegård, P.A. Lindestad, H. Larsson, S. Granqvist. High-speed registration of phonation-related glottal area variation during artificial lengthening of the vocal tract. Log Phon Vocol 32 (2007) 157-164.

[36] A.-M. Laukkanen, I.R. Titze, H. Hoffman, E.M. Finnegan. Effects of a semi-occluded vocal tract on laryngeal muscle activity and glottal adduction in a single female subject. Folia Phon Log 60 (2008) 298-311.

[37] C.S. Gaskill, M.L. Erickson. The Effect of a Voiced Lip Trill on Estimated Glottal Closed Quotient. J Voice 22 (2008) 634-643.

[38] C.S. Gaskill, D.M. Quinney. The effect of resonance tubes on glottal contact quotient with and without task instruction: A comparison of trained and untrained voices. J Voice 26 (2012) e79-e93

[39] D.G. Childers, A.M. Smith, G.P. Moore. Relationships between electroglottograph, speech and vocal cord contact. Folia phon 36 (1984) 105-118.

[40] P.H. Dejonckere, J. Lebacq. Electroglottography and vocal nodules: An attempt to quantify the shape of the signal. Folia phon 337 (1985) 195-200.

[41] T. Hacki. Klassifizierung von Glottisdysfunktionen mit Hilfe der Elektroglottographie. Folia phon 41 (1989) 43-48.

[42] G. Motta, U. Cesari, M. Iengo, G. Jr Motta. Clinical application of electroglottography. Folia phon 42 (1990) 111-117.

[43] O. Zagolski, E. Carlson, G. Murty, P. Carding, P. Kelly, R.L. Lancaster, R.L. Plant. Electroglottographic measurements of glottal function in vocal fold paralysis in women; The effect of frequency on combined glottography; Aerodynamics of the Human Larynx During Vocal Fold Vibration Clinical Otolaryngology & Allied Sciences; The Laryngoscope 27 (2002) 246-253.

[44] T. Baer, A. Löfqvist, N.S: McGarr. Laryngeal vibrations: A comparison between high-speed filming and glottographic techniques. J Acoust Soc Am 73 (1983) 1304-1308.

[45] E.B. Holmberg, R.E. Hillman. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. J Speech Hear Res 38 (1995) 1212-1224.

[46] S. Hertegård, J. Gauffin. Glottal area and vibratory patterns studied with simultaneous stroboscopy, flow glottography, and electroglottography. J Speech Hear Res 38 (1995) 85-101.

[47] C. Herbst, S. Ternström. A comparison of different methods to measure the EGG contact quotient. Log Phon Vocol 31 (2006) 126-138.

[48] P. Alku. An automatic inverse filtering method for the analysis of glottal waveforms. Doctoral thesis, Helsinki University of Technology, 1992.

[49] N. Henrich, C. d'Alessandro, B. Doval, M. Castellengo. On the use of the derivative of electroglottographic signal for characterization of nonpathological phonation. J Acoust Soc Am 115 (2004) 1321-1332.

[50] R.F. Orlikoff, M.E. Golla, D.D. Deliyski. Analysis of longitudinal phase differences in vocal-fold vibration using synchronous high-speed videoendoscopy and electroglottography. J Voice 26 (2012) e13-816.e20.

[51] M. Rothenberg, J. Mashie. Monitoring vocal fold abduction through vocal fold contact area. J Speech Lang Hear Res 31 (1988) 338-351.

[52] E. Kankare, A.-M. Laukkanen, I. Ilomäki, A. Miettinen, T. Pylkkänen. Electroglottographic contact quotient in different phonation types using different amplitude threshold levels. Log Phon Vocol 37 (2012) 127-132.

[53] K. Verdolini, R. Chan, I.R. Titze, M. Hess, W. Bierhals. Correspondence of electroglottographic closed quotient to vocal fold impact stress in excised canine larynges. J Voice, 12 (1998) 415-423.

[54] I.R. Titze. Mechanical stress in phonation. J Voice 8 (1994) 99-105.

[55] C. Reed, E. Doherty, T. Shipp. Direct measurement of vocal fold medial forces. Am Speech Hear Ass Rep 34 (1992) 131 (A).

[56] K. Verdolini, M.M. Hess, I.R. Titze, W. Bierhals, M. Gross. Investigation of vocal fold impact stress in human subjects. J Voice 13 (1999) 184-202.

[57] K. Marasek. Electroglottographic description of voice quality. Arbeitspapiere des Instituts für maschinelle Sprachverarbeitung, Stuttgart, 3(2),1997

[58] R.F. Orlikoff. Assessment of the dynamics of vocal fold contact from the electroglottogram: data from normal male subjects. J Speech Hear Res 34 (1991) 1066-1072.

[59] R.E. Kania, S. Hans, D.M. Hartl, P. Clement, L. Crevier-Buchman, D.F. Brasnu. Variability of electroglottograpic glottal closed quotients: Necessity of standardization to obtain normative values. Arch Otolaryngol Head Neck Surg 130 (2004) 349-352.

[60] D. Childers, C. Lee. Vocal quality factors: Analysis, synthesis, and perception. J Acoust Soc Am 90 (1991) 2394-2410.

[61] P. Davies, G.A: Lindsey, H. Fuller, A. Fourcin. Variation in glottal open and closed phases for speakers of English. In proceedings of the Institute of Acoustics 8 (1986) 539-546.

[62] E. Kankare, D. Liu, A.-M. Laukkanen, A. Geneid. EGG and acoustic analyses of different voice samples: Comparison between perceptual evaluation and voice activity and participation profile. Folia Phon Log 65 (2013) 98-104.

[63] E.B. Holmberg, R.E. Hillman, J.S. Perkell. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal and loud voice. J Acoust Soc Am 82 (1988) 511-1787.

[64] J.J. Jiang, I.R. Titze. Measurement of vocal fold intraglottal pressure and impact stress. J Voice 8 (1994) 132-144.

[65] A. Sulter, H. Wit. Glottal volume velocity waveform characteristics in subjects with and without vocal training, related to gender, sound intensity, fundamental frequency, and age. J Acoust Soc Am 100 (1996) 3360-3373.

[66] R.E. Hillman, E.B. Holmberg, J.S. Perkell, M. Walsh, C. Vaughan. Objective assessment of vocal hyperfunction: An experimental framework and initial results J. Speech Hear Res. 32 (1989) 373-392.

[67] P. Alku, T. Bäckström, E. Vilkman. Normalized amplitude quotient for parametrization of the glottal flow. J Acoust Soc Am 112 (2002) 701-710.

[68] A. NiChasaide, C. Gobl. Voice source variation. In: Hardcastle,W., Laver, J. (Eds.), The Handbook of Phonetic Sciences. Blackwell Publishers, Oxford, 1997, pp. 1–11.

[69] P. Alku, E. Vilkman. A comparison of glottal voice source quantification parameters in breathy, normal and pressed phonation of female and male speakers. Folia Phon Log 48 (1996) 240-254.

[70] P. Alku, H. Strik, E. Vilkman. Parabolic spectral parameter – A new method for quantifiction of the glottal flow. Speech Commun 22 (1997) 67–79.

[71] T. Raitio, A. Suni, J. Yamagishi, H. Pulakka, J. Nurminen, M. Vainio, P. Alku. HMM-based speech synthesis utilizing glottal inverse filtering. IEEE Transactions on Audio, Speech, and Language Processing 19 (2011) 153-165.

[72] P. Alku P. Glottal inverse filtering analysis of human voice production – A review of estimation and parameterization methods of the glottal excitation and their applications. (Invited article). Sadhana – Academy Proceedings in Engineering Sciences. Vol. 36, Part 5, pp. 623-650, 2011.

[73] T. Drugman, P. Alku, A. Alwan, B. Yegnanarayana. Glottal source processing: from analysis to applications. Computer Speech and Language 28 (2014) 1117-1138.

[74] P. Alku. Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. Speech Comm. 11 (1992) 109–118.

[75] J. Makhoul. Linear prediction: A tutorial review. Proc. IEEE 63 (1975) 561-580.

[76] A. El-Jaroudi, J. Makhoul. Discrete all-pole modeling. IEEE Trans Signal Proc. 39 (1991) 411–423.

[77] P. Alku, E. Vilkman. Estimation of the glottal pulseform based on Discrete All-Pole modeling. Proc. Int. Conf. on Spoken Language Processing, pp. 1619-1622, Yokohama, Japan, 1994.

[78] M. Airas. TKK Aparat: An environment for voice inverse filtering and parameterization. Log Phon Vocol 33 (2008) 49-64.

[79] J. Gauffin-Lindqvist. Studies of the voice source by means of inverse filtering. Speech Transmission Laboratory, Quarterly Progress and Status Report 2 (1965) 8–13.

[80] L. Lehto, M. Airas, E. Björkner, J. Sundberg, P. Alku. Comparison of two inverse filtering methods in parameterization of the glottal closing phase characteristics in different phonation types. J Voice 21(2007) 138–150

[81] M. Rothenberg. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. J Acoust. Soc. Am. 53 (1973) 1632–1645.

[82] M. Airas. Methods and studies of laryngeal voice quality analysis in speech production. Teknillinen korkeakoulu, 2008.

[83] C.T. Herbst, J. Lohscheller, J.G. Svec, N. Henrich Bernardoni, G.W. Weissengruber T. Fitch. Glottal opening and closing events investigated by electroglottography and super-high-speed video recordings. J Exp Biol 217 (2014) 955-963.