# Phonatory Strategies of Male Vocalists in Singing Diatonic Scales With Various Dynamic Shapings

**Allan Vurma,** *Tallinn, Estonia*

**Summary:** This work aimed to investigate glottal adduction in the case of the singing of diatonic D major ascending scales with different dynamic shapings on vowel [aː] by 10 classically trained male singers. The parameters obtained by inverse filtering—closed quotient (CQ), normalized amplitude quotient (NAQ), and the level difference between the first two partials of the glottal flow waveform spectrum (H1–H2)—all indicated statistically significant systematic changes when singers sang (1) with habitual dynamics without intentional dynamic changes; (2) with *sempre crescendo* from *piano* to *forte*; and (3) with *sempre diminuendo* from *forte* to *piano*. Adduction on average was lower if the tone with the same pitch was sung *piano* compared with when it was sung *forte*. The *piano* tones with the limit of one and the same dynamic task were sung with lower adduction on average compared with one octave higher or lower tones sung *forte*, although here the agreement between different parameters was poorer and the individual variability was greater. There was only a moderate negative correlation between CQ and NAQ ($r = -0.41$) but a strong negative correlation between CQ and H1–H2 ($r = -0.8$). CQ showed a moderate positive correlation with sound level but no significant correlation with pitch. NAQ showed a strong and significant positive correlation with pitch but not with sound level (except when it was calculated only at constant pitch). Some of the most experienced singers used quite high glottal adduction levels, questioning the concept of whether *flow phonation* is always the favorite classical style.

**Key Words:** glottal adduction–dynamics–closed quotient–inverse filtering–singing voice.

## INTRODUCTION

Professional vocalists must perform different musical tasks at wide pitch and dynamic ranges, while pronouncing the text and maintaining control over the overall timbre of their voices. According to the source-filter theory of voice production, the acoustical properties of the voice depend on the working regime of the vocal folds, as well as the shape of the vocal tract, which acts as the resonator.[1] It is possible to produce a voice with the same pitch and loudness with different phonatory characteristics, although this ability is usually better developed in the case of trained vocalists compared with vocally untrained persons.[2] There are three main aspects of phonation (voice source) in addition to pitch that can be controlled by the singer: (1) the degree of glottal adduction, (2) the voice register, and (3) subglottal pressure.[2] Glottal adduction is the force with which the vocal folds are pressed together.[2] The degree of adduction is important because (1) it is related to the level of sound energy that the singer is able to produce and to the efficiency of airflow energy conversion into sound energy; (2) adduction affects the level of balance between spectral partials, and thus the timbre of the voice; (3) the level of expiratory flow during phonation is dependent on it, ie, how long the phrases are that can be sung in one breath; (4) the stability of phonation is related to the strength of glottal adduction; and (5) overadduction can be detrimental to voice health.[3]

The degree of glottal adduction depends on the activity of the lateral cricoarytenoids, which approximate the vocalis processes of the arytenoid cartilage, as well as the activity of the posterior interarytenoids, which close the gap at the posterior cartilaginous part of the glottis.[4] In addition, greater activity of the thyroarytenoids (TAs) increases the adduction at the medial part of the glottis.[3,5] The strength of glottal adduction is also affected by other factors, such as the caudally directed force on the larynx that widens the glottis,[6] the lung volume (as the overall glottal abduction force, ie, the force that is directed to separate the vocal folds, which tends to be greater at high lung volumes[7]), and the nonlinear impact of the subglottal and supraglottal cavities on the voice source, which can either enhance or impede glottal vibration.[8] The degree of such nonlinear interaction can be controlled by the width of the epilarynx tube and the shape of the vocal tract, but it also depends on the mutual position of the fundamental (F0) and formant frequencies along the frequency axis, particularly whether the F0 is lower or higher than the frequency of the first formant.[8] Vibration is only possible to sustain if the vocal folds are not too widely separated or too tightly approximated.[3]

The details of glottal activity can be described with the help of various measurable parameters. For example, the closed quotient (CQ) shows the proportion of time in a vibrational cycle when the glottis is closed[9] (see also Figure 1). Stronger glottal adduction is correlated with higher CQ: voice production with $CQ > 0.6$ tends to be perceived as "pressed" and $CQ < 0.3$ as "breathy".[3] According to the formula proposed by Titze and Sundberg,[10] the shape of the glottal flow waveform that has a CQ between 0.4 and 0.5 promotes the generation of maximum sound energy at the same maximum flow declination rate (MFDR; a more detailed description of this parameter is provided later). This type of phonation is sometimes referred to as "flow phonation" and is often considered to be the most typical in the classical style.[1–3] However, the conversion of airflow energy into sound is better if the value of the CQ is high.[3]

Björkner et al[11] found that vocally untrained persons are inclined to increase glottal adduction with pitch and loudness, but vocally trained persons tend to be less affected by such automatic reactions. The aim of the present work was to investigate whether this assumption holds in the case of classically trained

---

Accepted for publication June 28, 2016.

From the Estonian Academy of Music and Theatre, Tatari 13, 10116 Tallinn, Estonia.

Address correspondence and reprint requests to Allan Vurma, Estonian Academy of Music and Theatre, Tatari 13, 10116 Tallinn, Estonia. E-mail: vurma@ema.edu.ee

Journal of Voice, Vol. 31, No. 2, pp. 254.e17–254.e29
0892-1997

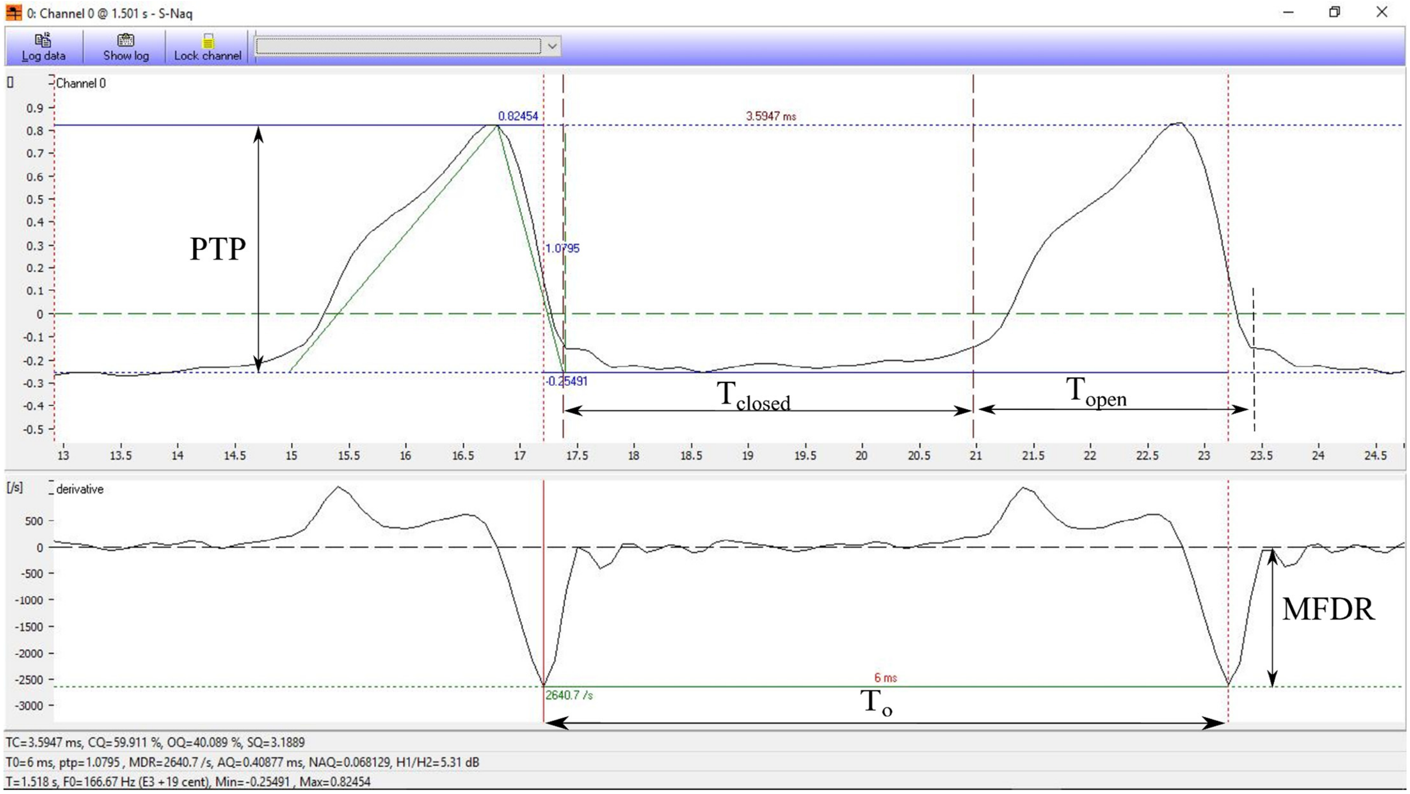
http://dx.doi.org/10.1016/j.jvoice.2016.06.018

male vocalists when they sing one octave diatonic scales with various dynamic shapings, as well as the degree of similarity in the phonatory strategies of different singers. We hypothesized that classically trained singers would tend to keep stable (without any systematic trends) the parameters that could indicate the strength of glottal adduction (CQ, normalized amplitude quotient [NAQ], and H1–H2), regardless of changes in F0 and in intensity. We also hypothesized that classically trained singers would tend to maintain the value of the CQ between approximately 0.4 and 0.5, which, as has been reported, is possibly the optimum for producing maximum sound energy.[3,10]

In this work we use a recognized and noninvasive method—the inverse filtering—to determine the CQ.[12] In this method, the inversed vocal tract transfer function is applied to the microphone-recorded signal. The influence of the vowel formants is thus canceled, and the resulting flow glottogram (FLOGG) shows the airflow volume through the glottis. Ideally, the recording should be performed in a free field in order to avoid contamination from room reflections.[13] However, inverse filtering as a method is not entirely free from shortcomings as the determination of the precise moments when the airflow through the glottis starts and stops is sometimes difficult because of some conceptual as well as methodological reasons.[13]

The FLOGG curve obtained by inverse filtering also allows for the determination of some amplitude domain parameters in which the determination of the instants of glottal closing and opening is not necessary (Figure 1). Its peak-to-peak amplitude (PTP) is strongly correlated with the amplitude of the fundamental component of the FLOGG spectrum.[14,15] The MFDR indicates the maximal speed of the airflow change in the glottal closing phase, and its numerical value corresponds to the level of the negative peak of the FLOGG's derivative (DFLOGG). MFDR is related to vocal intensity,[16] sound pressure level (SPL),[14] and subglottal pressure.[17] The amplitude quotient (AQ) is defined as the ratio of PTP to MFDR (AQ = PTP/MDR). Because the AQ depends on frequency—it tends to decrease with increasing pitch—its frequency normalized value (NAQ = AQ/F0) is often used instead of AQ. A smaller NAQ (and a smaller AQ at constant pitch) corresponds to a more adducted phonation type[18] and to the increased perceived pressedness of the vocal sound.[19] The value of CQ that is obtained from an inverse-filtered signal depends greatly on the decisions of the operator regarding how he or she determines the instants of glottal opening and closing. Because the values of AQ and NAQ do not depend on the time domain, the inter-rater reliability in the case of later parameters is much higher than with the CQ.[20]

The FLOGG spectrum allows for the determination of the levels of its harmonics. The level difference between the fundamental component and the second harmonic (H1–H2) partially determines how strong the fundamental component is in the voice spectrum (but the shape of the vocal tract's transfer function also plays an important role here). A reduced level of the fundamental could refer to the increased pressedness and/or elevated larynx position.[2] More indicative than the absolute values is the



```
TC=3.5947 ms, CQ=59.911 %, OQ=40.089 %, SQ=3.1889
T0=6 ms, ptp=1.0795 , MDR=2640.7 /s, AQ=0.40877 ms, NAQ=0.068129, H1/H2=5.31 dB
T=1.518 s, F0=166.67 Hz (E3 +19 cent), Min=-0.25491 , Max=0.82454
```

**FIGURE 1.** The flow glottogram (FLOGG, upper panel) and its derivative (DFLOGG, lower panel) as obtained by inverse filtering with *Sopran 1.0.10.* On the graphs are indicated the following: the period of the glottal flow waveform ($T_o$), closed phase ($T_{closed}$), open phase ($T_{open}$), peak-to-peak amplitude of the glottal flow pulse (PTP), and maximum flow declination rate (MFDR).