ON THE QUANTITATIVE RELATIONSHIP BETWEEN SUBGLOTTAL PRESSURE, VOCAL CORD TENSION, AND GLOTTAL ADDUCTION IN SINGING

B J Kröger

Phonetics Institute, University of Cologne, Germany

1. INTRODUCTION

The aim of this study is to establish rules for the control of a physiologically based glottis model for singing. The glottis model used is part of a physiologically based model of the human sound production mechanism also comprising a pulmonary model and a vocal tract model. It is able to simulate the aerodynamics and acoustics of human sound production.

The quantitative relationship between glottal flow parameters (fundamental frequency, average flow, peak flow etc.) and glottal control parameters (cord tension, glottal adduction, subglottal pressure) has been measured. Using the condition of constant glottal flow our model is able to produce the range of pitches of the modal tenor register. Increasing fundamental frequency is realized by increasing vocal cord tension, increasing subglottal pressure, and decreasing glottal adduction.

2. BACKGROUND

Quantitative relationships for acoustic and aerodynamic voice source parameters have been measured for speech and singing (Sundberg [1] and [2], Holmberg et al. [3], Tize [4], Titze and Sundberg [5]). In speech pulmonary pressure mainly changes with vocal loudness (5.9 cm H_2O for normal voice, 8.4 cm H_2O for loud voice 4.8 cm H_2O for soft voice, Holmberg et al. [3], p. 618) but is rather constant during an utterance. In singing pulmonary pressure also changes with fundamental frequency for a given loudness level (10 cm H_2O at 180 Hz to 35 cm H_2O at 330 Hz for a tenor voice, normal loudness, Sundberg [1], p. 36).

Assuming that all glottal control parameters are constant mean glottal airflow increases with subglottal pressure. But it has been found for singing that mean glottal flow remains rather constant during the whole pitch range of a register for a given loudness. Values of 0.1 cm^3/sec to 0.13 cm^3/sec have been measured during a descending glissando at normal loudness (Rubin et al. [6], Sundberg [1] , p. 42f). This condition of constant glottal flow can only be reached by a complex control strategy, i.e. by using a complex relationship between the phonatory control parameters (cord tension, glottal adduction and pulmonary pressure). It is the aim of this study to elucidate this relationship by means of voice source modelling since glottal control parameters are difficult to measure in vivo.

3. THE MODEL

The voice source model is part of a complete speech production model capable of generating articulator movements, vocal tract geometry, and the acoustic speech signal (Kröger [7] and [8]). The acoustics and aerodynamics of the vocal tract are calculated by a reflection type line analogue coupled with a pulmonary and trachea model, a glottis model, a friction noise generator and a radiation model (fig. 1).

As a simple biomechanical approximation of the vocal folds we used the two-mass model of Ishizaka and Flanagan [9] (fig. 2). In this model the upper and lower part of the vocal folds are represented by different harmonic oscillators (damped mass-spring-systems with mass m_i, spring constant s_i, and damping r_i; i=1,2). The masses are stiffness-coupled (k_c). The mechanical part is driven by pressure-induced forces acting on the inner surfaces of the vocal folds. During the open phase of the glottis the forces acting on the masses can be approximated by the pressure value p_g of the tube representing the glottal constriction. During the closed phase of the glottis they can be approximated by subglottal pressure p_{sub}. The control parameters of the mechanical part of the glottis model are *glottal aperture* GA and *vocal cord tension* CT. CT characterizes the longitudinal tension and GA the equilibrium position of the vocal folds, i.e. its degree of abduction or adduction. The output of the mechanical part of the glottis model is glottal area a_g which describes the instantaneous area of glottal constriction and hence the vocal fold oscillation. Glottal area a_g is the basis for the calculation of glottal flow u_g.

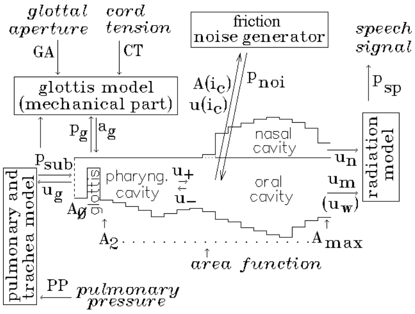


Figure 1 The aerodynamic-acoustic vocal tract model. Inputs and outputs in italics.

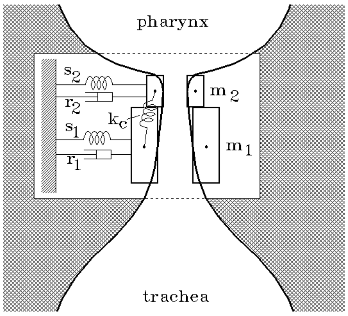


Figure 2 Cross-sectional frontal view of the glottis and its approximation by the two-mass model.

In our implementation the Ishizaka-Flanagan two-mass model has been extended by a aperture-dependent non-oscillating bypass in order to model the cartilaginous and membranous part of the vocal folds (Kröger [10]). The pulmonary and trachea model used is a simple functional aerodynamic model. The pulmonary system is modelled by a piston in a piston chamber. The piston is driven by an external force defining *pulmonary pressure* (PP). During sustained phonation (e.g. in singing) PP roughly equals subglottal pressure.

Basic relationships between the physiological control parameters CT and PP and the acoustic parameters fundamental frequency (F0) and sound pressure level (SPL) have been tested by Monsen et al. [11] indicating that the Ishizaka-Flanagan two-mass approximation models physiological data well: SPL is mainly determined by PP. Doubling PP increases SPL by 6 to 9 dB. F0 is mainly determined by CT. Increasing PP by 1 cm H_20 increases fundamental frequency only by less than 4 Hz. These relationships are reproduced by our model.

4. EXPERIMENTAL STUDIES

4.1 Range of oscillation

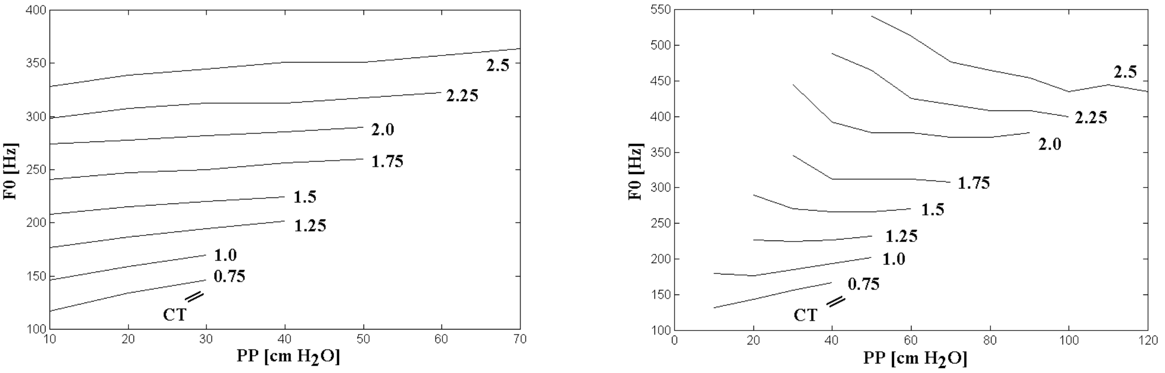
We started measuring glottal flow parameters (fundamental frequency F0, open quotient opqu, speed quotient spqu, average flow avfl, peak flow pkfl and maximum flow derivative mfdr) as function of the model control parameters (glottal aperture GA, cord tension CT, and pulmonary pressure PP) during the whole oscillation range of the model up to an average flow of 500 cm^3/sec. GA has been varied in steps of 1.5 mm^2 from 3 to -6 mm^2, CT in steps of 0.25 from 0.5 to 2.5, and PP in steps of 10 cm H_20 from 10 to 120 cm H_20. Table 1 indicates the oscillation range of the model by listing minimum and maximum values of PP as function of GA and CT (only the doubled step width is listed for GA and CT). It can be seen that the oscillation range is relatively limited for positive values of GA (abduction case: the vocal folds are not in contact in rest position) and oscillation only occurs for low values of CT and PP. If the vocal folds are adducted (GA≤0) the oscillation range is wider. In this case the minimum and maximum values for PP increase with increasing absolute values of GA (increasing adduction) and with increasing CT: The minimum values of PP increase in order to initiate oscillation with increasing CT and decreasing GA, the maximum values of PP increase in order to reach the limiting value for average flow.

CT [/]	**CT=0.5**	**CT=1.0**	**CT=1.5**	**CT=2.0**	**CT=2.5**
GA= 3 mm^2	10...20	10...30	20...30	/	/
GA= 0 mm^2	/	10...30	10...40	10...50	10...70
GA=-3 mm^2	/	10...40	10...50	20...70	30...90
GA=-6 mm^2	/	10...50	20...60	30...90	50...120

Table 1 Oscillation range of the model. Minimum and maximum values of pulmonary pressure (PP in cm H_2O) as function of glottal aperture (GA) and cord tension (CT).

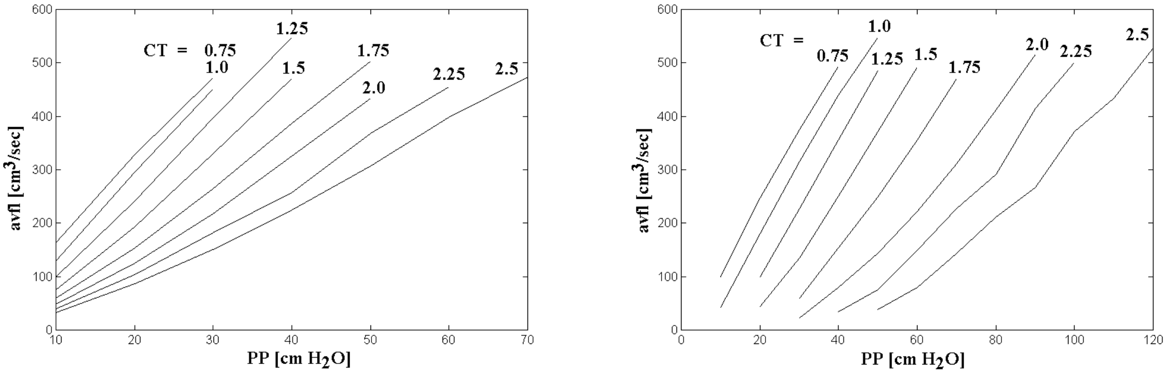
4.2 Fundamental frequency and average flow as function of control parameters

Figure 3 indicates that F0 is primarily controlled by CT. F0 increases only slightly with increasing PP for all values of CT (case GA = 0) while it decreases slightly with increasing PP for high CT values (case GA = -6 mm^2). Figure 4 indicates that flow is influenced by all three parameters: Average flow increases with PP but decreases with increasing CT and with increasing GA.


(a) GA = 0

(b) GA = -6 mm^2

Figure 3 Fundamental frequency (F0) as function of pulmonary pressure (PP) for eight levels of cord tension (CT) and for two levels of glottal aperture (GA). Fig. 3a: GA = 0, fig. 3b: GA = -6 mm^2.


(a) GA = 0

(b) GA = -6 mm^2

Figure 4 Average flow (avfl) as function of pulmonary pressure (PP) for eight levels of cord tension (CT) and for two levels of glottal aperture (GA). Fig. 4a: GA = 0, fig. 4b: GA = -6 mm^2.

4.3 The condition of constant flow

From physiological measurements of the singing voice it can be deduced that the flow is roughly constant during the whole pitch range of a register for a given loudness level (Sundberg [1]). Since flow varies considerably with all control parameters for our self-oscillating glottis model (fig. 2), we plotted the range of control parameter values for a defined range of flow values. Fig. 5 gives all combinations of control parameters leading to an average flow between 100 and 200 cm^3/sec. The thick solid lines indicate the AG-CT-planes for different levels of PP. The thick tube line gives a possible control strategy for avfl = 150 cm^3/sec. Its projection in the PP-CT-plane and the appertaining F0-values are also given for the tube line. The F0 values indicate that the pitch range of a tenor modal register (i.e. 130 Hz to 400 Hz) can be covered by this control strategy: increasing fundamental frequency is realized by increasing vocal cord tension CT, increasing pulmonary pressure PP, and decreasing glottal aperture GA (increasing glottal adduction).

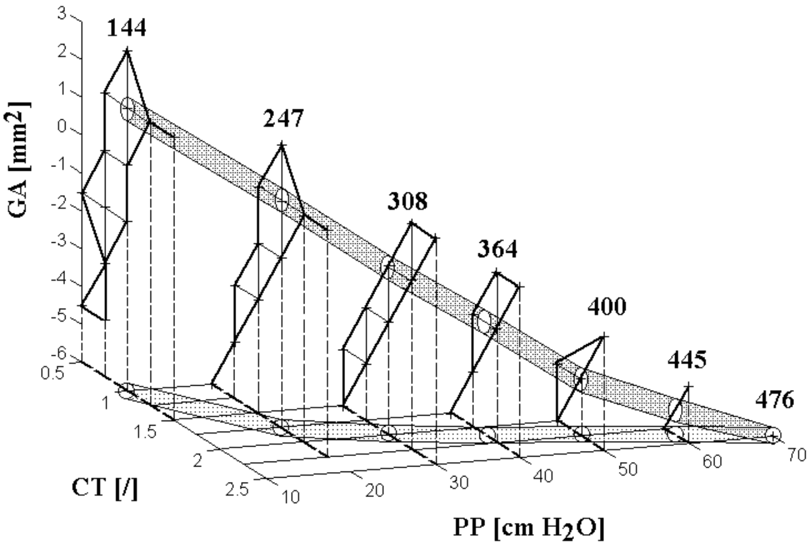
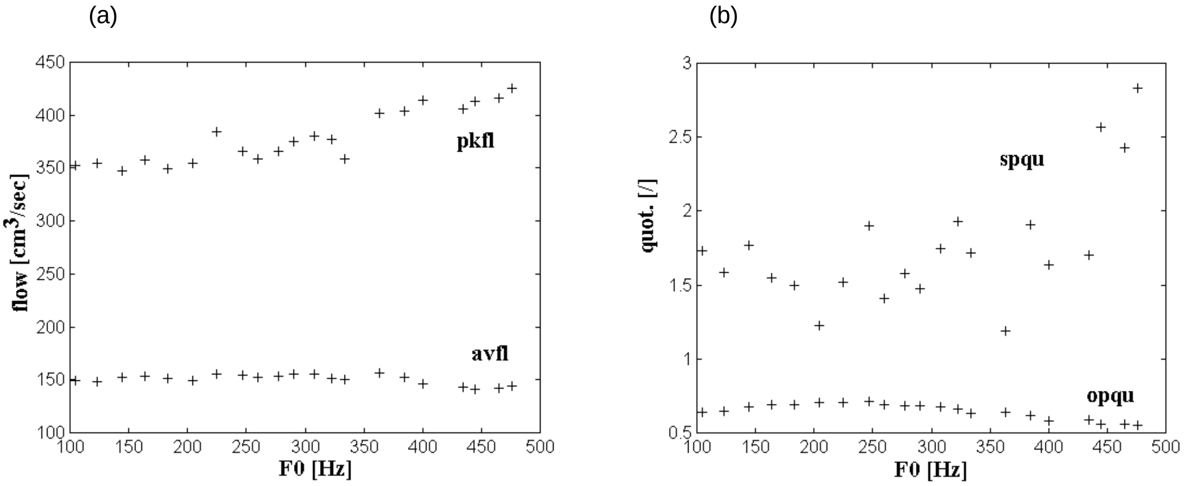


Figure 5 Oscillation range for the glottis model with average flow between 100 and 200 cm^3/sec (see text for further explanations).

Figure 6 gives glottal flow waveshape parameters as function of F0 for this control strategy indicating that they are in agreement with measured parameters. Average flow is roughly 150 cm^3/sec as was the definition for this control condition. Peak flow varies from 350 cm^3/sec to about 400 cm^3/sec with increasing fundamental frequency. Values around 200 cm^3/sec to 700 cm^3/sec are measured for peak flow during singing for trained tenors at 117 Hz, 234 Hz and 383 Hz by Titze ([4], p. 2930). The open quotient varies from 0.55 to 0.75 and the speed quotient varies from 1 to 2 within the pitch range of the tenor (up to 400 Hz) for our model as well as for trained tenors (Titze [4], p. 2931). Open quotient decreases and speed quotient increases with fundamental frequency, i.e. with pulmonary pressure. The maximum flow declination rate increases from 0.3 m^3/sec^2 to 1.8 m^3/sec^2 for our model as well as for trained tenors (Titze [4], p. 2932). This rate increases with increasing fundamental frequency, i.e. increasing pulmonary pressure. It indicates an increasing sound pressure level with increasing F0 as occurring naturally (Sundberg [1], p. 42f, Titze and Sundberg [5]). (The maximum flow declination rate can be taken as a measure for sound pressure level if the vocal tract geometry is constant since this rate sets the level of formant amplitudes (Fant [12], p. 12) and since sound pressure level depends on the amplitude of the strongest spectrum partial (Sundberg [1], p. 42).)



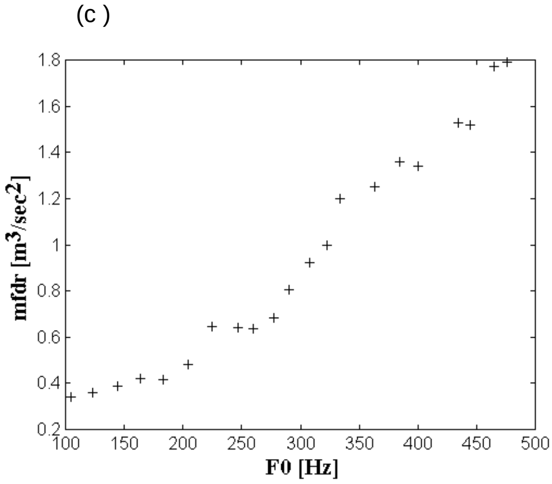


Figure 6 Peak flow (pkfl) and average flow (avfl) (fig. 6a), open quotient (opqu) and speech quotient (spqu) (fig. 6b), and maximum flow declination rate (mfdr) (fig. 6c) as function of fundamental frequency (F0) (for definition of these parameters see Holmberg et al. [3]).

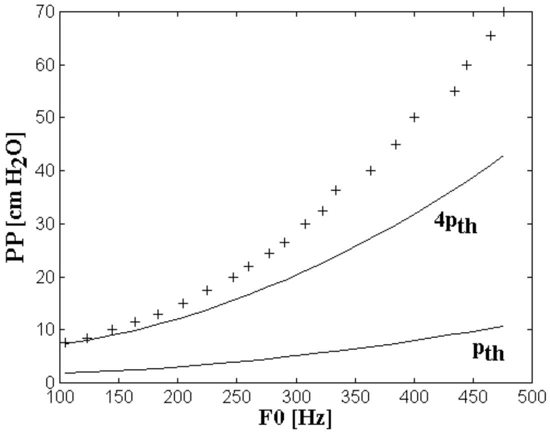


Figure 7 Pulmonary pressure (PP) for the model (condition of constant flow after fig. 5) (crosses) phonation threshold pressure (p_{th}) and $4p_{th}$ (lines) as function of fundamental frequency (F0). The area between the lines indicates the area of pulmonary pressure in natural singing.

It can be concluded that the shape of the glottal waveforms produced by our model is comparable to naturally produced waveforms during singing by trained singers. But in comparison to physiologically measured PP as function of F0 we found that our model needs very high PP. Fig. 7 indicates values of PP as needed by our model and as measured. Measured values of PP also increase with F0 and are located in the range of phonation threshold pressure to up to four times of phonation threshold pressure for a given F0 (Titze [4], p. 2933).

5. CONCLUSION

Using the condition of constant glottal airflow our model is able to produce the range of pitches of the modal tenor register. Increasing fundamental frequency is realized by increasing vocal cord tension, increasing pulmonary pressure, and increasing glottal adduction. Increasing pulmonary pressure is necessary in order to maintain the vocal cord oscillation. The glottal flow waveshape parameters describing the temporal features of the glottal flow pulses and thus describing the voice quality are in agreement with the values measured for trained tenor singers. But the comparison of pulmonary pressure with measured values indicates that the model needs too high pulmonary pressure for a given fundamental frequency. This indicates a lack in glottal efficiency for changing aerodynamic into acoustic energy. In further work this problem will be focused.

6. REFERENCES

[1] J SUNDBERG, *The science of the singing voice* (Dekalb: Northern Illinois University Press, 1987).
[2] J SUNDBERG, ‘Subglottal pressure behaviour in singing and speech“, *Proceedings of the 13th International Congress of Phonetic Sciences*, Vol. 3 (1995) 98-104.
[3] 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’, *Journal of the Acoustical Society of America* 84 (1988) 511-529.
[4] I R TITZE, ‘Phonation threshold pressure: A missing link in glottal aerodynamics’, *Journal of the Acoustical Society of America* 91 (1992) 2926-2935.
[5] I R TITZE & J SUNDBERG, ‘Vocal intensity in speakers and singers’, *Journal of the Acoustical Society of America* 91 (1992) 2936-2946.
[6] H J RUBIN, M LeCOVER & W VENNARD, ‘Vocal intensity, subglottic pressure, and airflow relationships in singers’, *Folia Phoniatrica* 19 (1967) 393-413.
[7] B J KRÖGER, ‘Minimal rules for articulatory speech synthesis’, in: J Vandewalle et al. (eds.): *Signal processing VI: Theories and applications* (Amsterdam: Elsevier, 1992) 331-334.
[8] B J KRÖGER, ‘A gestural production model and its application to reduction in German’, *Phonetica* 50 (1993) 213-233.
[9] K ISHIZAKA & J L FLANAGAN, ‘Synthesis of voiced sounds from a two-mass model of the vocal cords’, *The Bell System Technical Journal* 51 (1972) 1233-1268.
[10] B J KRÖGER, ‘Simulation of vocal fold oscillation behaviour by a self-oscillating glottis model’, *Journal de Physique* 4 (Colloque 5, 3ème Congrès français d'acoustique, Toulouse, 1994) 457-460.

[11] R B MONSEN, A M ENGEBRETSON & N R VEMULA, ‘Indirect assessment of the contributon of subglottal air pressure and vocal-fold tension to changes of fundamental frequency in English’, *Journal of the Acoustical Society of America* 64 (1978) 65-80.
[12] G FANT, ‘Some problems in voice source analysis’, *Speech Communication* 13 (1993) 7-22.