VOQUAL'03, Geneva, August 27-29, 2003

On the relationship between input parameters in the two-mass vocal-fold model with acoustical coupling and signal parameters of the glottal flow

Annemie Van Hirtum[1], *Ines Lopez*[2], *Mico Hirschberg*[2], *Xavier Pelorson*[1]

[1]ICP, UMR CNRS 5009, INPG-Université Stendhal, Grenoble, France
[2]Eindhoven University of Technology, Eindhoven, The Netherlands
`pelorson@icp.inpg.fr`

Abstract

In this paper the sensitivity of the two-mass model with acoustical coupling to the model input-parameters is assessed. The model-output or the glottal volume air flow is characterised by signal-parameters in the time-domain. The influence of changing input-parameters on the signal-parameters is quantified.

1. Introduction

The glottal volume air flow (Ug) and time derivative of Ug (dUg) are well known to be a fundamental issue in voiced sound production. Further research towards different phonation types as e.g. defined in [1], showed the major importance of Ug and dUg towards the study of voice quality and prosody. Consequently much effort is spent to predict, measure and model the glottal volume flow waveform Ug and dUg in time. Depending on the strived goal - analysis, synthesis or production of sound - different approaches and assumptions are applied to study Ug and dUg.

In literature both very complex and very simple models to estimate the waveforms of Ug and dUg are presented. In general complex models are characterised by numerous parameters attempting to provide an appropriate description of the acoustical and physiological reality. Striving to obtain a model of Ug and dUg allowing for physical insight the physiological reality is described by an oversimplified model of the main structures involved, i.e. vocal folds, sub- and supra-glottal cavities, in combination with a simplified flow model. In case sound analysis or synthesis is aimed, Ug and dUg waveforms are estimated from a generative function in time with an appropriate parameter set. Regardless the perspective multitudinous models are proposed. Well-known glottal flow generative models based on the reference models of Liljencrants-Fant (LF) or Rosenberg make use of 3 or 4 parameters to model the time evolution of Ug and dUg [2, 3]. Simplified models allowing for physical insight in the sound production account both for the mechanical behaviour of the vocal folds as for the airflow through the glottis. The mechanical behaviour is most commonly represented by a two-mass model of the vocal folds. Since a two-mass model of the vocal folds involves a severe oversimplification of reality, a simplified flow model capturing critical flow characteristics is favoured. The generative volume flow models implies only one type of parameters indicated with specification parameters. Knowledge of the specification parameters in combination with the generative function allows to compute Ug and dUg. Physical models like the distinct two-mass models however require different types of parameters. In this papers two types of parameters are distincted. Firstly both the flow and mechanical part include parameters which are directly related to physiological characteristics. Secondly there are parameters with the same functionality as described for the generative flow model, i.e. specification parameters, which are intrinsic to the model and are not directly depending on the physiology. The choice of the specification parameters has a major influence on the quality of the resulting Ug and dUg fit.

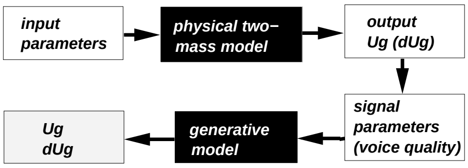


Figure 1: *General framework.*

Next to the parameter sets required to obtain the strived Ug and dUg waveforms, the mentioned relationship between on one hand the Ug and dUg waveforms and voice quality on other hand is quantitatively expressed by voice quality parameters. Those are calculated from the resulting waveforms.

Important work towards the outlined topics has been performed starting from as early as 1960. Consequently due to the fundamental importance of Ug and dUg in the diverse fields a whole bench of proposed parameter sets are searched with respect to the physical or generative models and to voice quality assessment. In this work we concentrate mainly on the parameters involved in the two-mass model described in [4] accounting for both sub- and supra-acoustical coupling. The influence of the distinct parameters on the estimated Ug and dUg and so on voice quality is expressed with simple time parameters. Finally the influence on the necessary specification parameteres, derived from the two-mass Ug and dUg estimate, in the generative 4-parameter LF model is assessed. The general framework is schematically outlined in Figure 1.

2. Parameter sets

2.1. Two-mass model with acoustical coupling: input parameters

The glottal flow Ug and differentiated flow dUg as outcome of a symmetrical two-mass model with sub- and supra-glottal coupling is assessed. A detailed description of the applied model, the different assumptions and the acoustical coupling is found in [4]. The following focusses on the different parameters involved in the model. Parameters are discussed in relation to physiology, mechanical behaviour, fluid flow and numerical value.

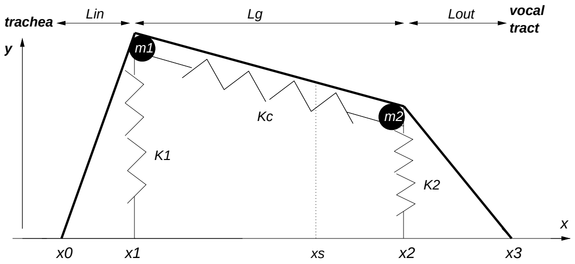


Figure 2: *2-dimensional 2-mass vocal fold model.*

2.1.1. physiology

The two-mass model consists of a severe oversimplification of the physiological reality with respect to important topics as neural control, lung geometry, tissue properties, intra-subject differences, etc. Instead the two-mass model needs a rough estimation of the necessary physiology critical for the further description of the mechanics and fluid flow. Therefore the physiological parameters in the two-mass model with acoustical coupling deal mostly with the relevant geometry and are in fact geometrical parameters allowing to approximate the shapes of the vocal folds and the surrounding sub- and supra glottal structures. The latest are concretely limited to the trachea and vocal tract. The geometry is assumed to be symmetrical above and below the flow direction i.e. from trachea to vocal tract. This means that both vocal folds are described by the same set of parameters. Furthermore a two dimensional description of the vocal folds is applied where the x coordinate acccounts for the flow direction and the y coordinate for the direction of vocal fold movement. The two-mass vocal fold model, explained further, is depicted in Figure 2. The third dimension is only accounted for considering the width of the vocal folds (L_w). The position of the two masses in the vocal folds is given by the positions x_1 and x_2. The y-coordinates y_1 and y_2 correspond to the initial equilibrium position along the y-axis. Remind that consequently during vocal fold simulation with the two-mass model the x-coordinate remains fixed while the y-coordinate, which is the coordinate of movement, varies in time. The distance $L_g = x_2 - x_1$ corresponds to the oscillating membranous portion of the vocal folds. The begin x_0 and end x_3 position of the vocal fold geometry determine the inlet $L_{in} = x_1 - x_0$ and outlet distances $L_{out} = x_3 - x_2$ to the membranous portion. The inlet heigth h_0 at the vocal fold entrance x_0 is estimated from the trachea area A as $h_0 = A/L_w$. The tracheal shape is approximated by the exponential-horn model [6]. The glottis heigth h_3 at the outlet of the vocal folds x_3 is determined by the vocal tract shape. Moreover the sub-and supra-glottal acoustical coupling depends on the estimated shape of the trachea and vocal tract. Besides the geometrical parameters the total vocal fold mass m_{tot} and the subglottal pressure P_{sub} are important physiological parameters.

2.1.2. mechanics

Although the total vocal fold mass m_{tot} is a physiological parameter the distribution of the total mass over the two point masses m_1 and m_2 with respectively position x_1 and x_2 in the two-mass model is of major importance for the mechanical behaviour and is as such a mechanical input parameter. To model vocal fold elasticity and damping with each point mass in the vocal fold (m_1, m_2) a spring constant (K_1, K_2) and damping coefficient ($\xi_1 < 1$, $\xi_2 < 1$) is associated. An additional coupling spring constant K_c is introduced to allow a coupling force between the two masses acting parallel to the y axis. Collision of both vocal folds is characterised by strongly non-linear behaviour. In the two-mass model this is expressed by introducing a constant critical heigth h_c. Whenever the distance between the two vocal folds, i.e. the distance between either the corresponding m_1 or m_2 masses in both vocal folds h_1 and h_2, is inferior to h_c the spring constants are stepwise increased to four times the normal value ($4 * K_1$ and $4 * K_2$) and the damping constants (ξ_1, ξ_2) are increased to a critical damping $\xi_c \approx 1$.

2.1.3. fluid flow

A pressure discontinuity of value P_{sub} is imposed at x_0 to approximate the driving lung pressure. A quasi-one-dimensional flow model is applied to calculate the aerodynamic forces excerted on the vocal folds. In the applied flow model flow separation is assumed to occur at a point x_s with $x_1 \leq x_s \leq x_2$. The position of x_s is estimated by applying Liljencrants ad-hoc experimental separation criterion [4]. So flow separation occurs where the distance h_s between the diverging part of the two vocal folds equals $h_s = s * h_{min}$ with $s = 1.2$ and h_{min} the minimum heigth of h_1 and h_2.

2.1.4. two-mass model input parameters

The previous subsections describe the set of input parameters required in the two-mass vocal fold model following [4]. Table 1 summarises the distinct input parameters and their reference values. The tabulated values of the input parameters with a physiological meaning are in accordance with 'in-vivo' measured values retrieved from literature. Since for non-pathological subjects there is no physiological reason for an asymmetry between the two masses m_1 and m_2, the two masses are treated alike. Consequently the input parameters with respect to m_1 and m_2 has the same value. Two different geometrical vocal tract shapes are considered corresponding to the area functions for the vowels /a/ and /i/. The measured vocal tract shapes are depicted in Figure 3. As can be seen from the figure accounting for /a/ and /i/ allows to study two complementary vocal tract configurations with respectively a constriction situated along the beginning and ending vocal tract portion.

Table 1: *Model input parameter reference values.*

L_w	10 mm	L_g	3 mm
L_{in}, L_{out}	0.0002 mm	h_0	12 mm
m_1, m_2	0.1 g	m_{tot}	0.2 g
K_1, K_2	40 N/m	K_c	20 N/m
ξ_1, ξ_2	0.1	ξ_c	1.1
y_1, y_2	0.01 mm	vocal tract, h_3	/a/, /i/
P_{sub}	1000 Pa	trachea	horn [6]

To study the strived influence of the input parameters of the two-mass vocal-fold model on the estimated glottal flow (Ug,dUg) and voice quality the input parameters are varied with respect to their reference values. In the current work we are mainly interested in the influence of the parameters specific for the model describing the mechanical behaviour. Therefore the greater part of the measurable physiological parameters is fixed at the reference values given in Table 1. On the contrary the input-parameters critical for the mechanical and fluid behaviour are varied slightly from their reference values. Concretely the two-mass model sensitivity towards the input-parameters $K = K_1 = K_2$, K_c, P_{sub} and $\frac{x_1}{x_2}$ is studied. Although P_{sub} is physiological measurable, the influence on the fluid and mechanical behaviour, and consequently on the model output Ug,

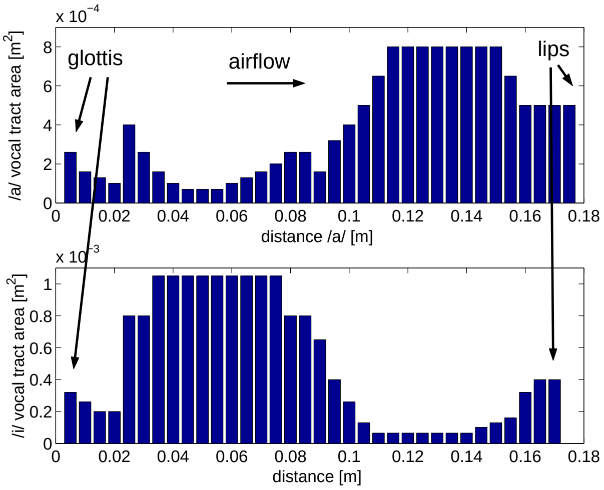


Figure 3: *Vocal tract shape for the vowels /a/ and /i/.*

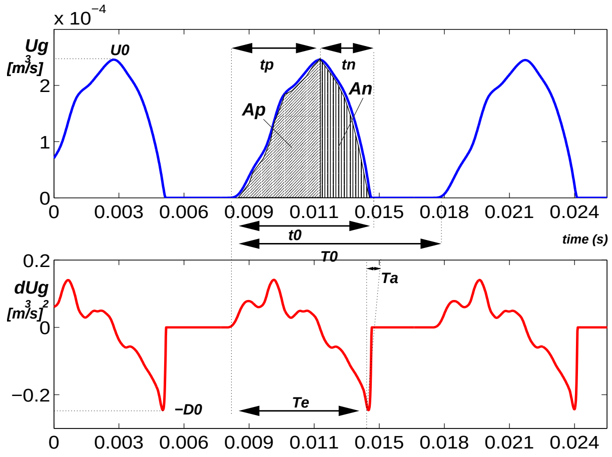


Figure 4: *Signal time-parameters illustrated on Ug and dUg. Ug is simulated with the two-mass model accounting for the vowel /a/. Input parameters are set to the values in Table 1.*

is too important to be neglected in this work. The value of the ratio $\frac{x_1}{x_2}$ is derived from the values of L_g, L_{in} and L_{out}. The reference values of Table 1 correspond with $\frac{x_1}{x_2} = 0.06$. Changing the ratio of x_1 and x_2 while maintaining x_0 and x_3 allows to evaluate the influence of the inlet and outlet depth to the oscillating membranous portion of the vocal folds.

2.2. Time-parameters from output signal Ug (dUg)

Assuming the idealized case of complete closure of the vocal folds, the following time-parameters are defined from important time events (instant of glottal flow peak, discontinuity at the instant of glottal closure,) to describe the glottal waveform Ug(t) and its derivative dUg(t) [3, 2]. For clearance the signal time-parameters on Ug and dUg are illustrated in Figure 4.

1) The open quotient $Q_o = \frac{t_0}{T_0}$ is defined as the ratio of the duration of the glottal open portion t_0 of the glottal waveform pulse Ug to the fundamental period T_0 of Ug or the full duration of a glottal cycle. So it describes the percentage of time in each period during which the vocal folds are open. This simple time-parameter is directly related to different phonation types as defined in [1] and therefore Q_o is an important parameter considering voice quality. The fundamental frequency or pitch $F_0 = \frac{1}{T_0}$ is derived from T_0. The pitch indicates the oscillatory frequency or the rhythm of glottal opening and closing.

2) The skewing ratio $Q_s = \frac{t_p}{t_n}$ is defined as the ratio of the duration of glottal opening t_p to the duration of glottal closing or return phase t_n. Glottal opening corresponds to the part of the glottal cycle during which the airflow rate is increasing and so the vocal folds are moving outwards. Consequently the slope of the Ug(t) fit is positive. Glottal closing corresponds to the part of the glottal cycle during which the airflow rate is decreasing and the vocal folds are moving inwards. Consequently the slope of the Ug(t) fit is negative. So the ratio of t_p to t_n informs on the symmetry in the open portion of the waveform.

3) The speed quotient $Q_r = \frac{A_p}{A_n}$ is defined as the ratio of the opening portion of the Ug waveform, with duration t_p, towards the closing portion, with duration t_n. So in addition to Q_s, Q_r contains additional information on the symmetry between the speed in glottal opening and closing.

4) The maximum glottal airflow U_0 is the flow peak value in a glottal cycle.

5) The absolute value of the minimum of the derivative of the glottal airflow in a glottal cycle is indicated with D_0. This value quantifies the maximum discontinuity. The instant of maximum discontinuity is indicated with T_e corresponding to the beginning of glottal opening.

6) Following the preceding definition of U_0 and D_0, their ratio $Q_m = \frac{U_0}{D_0}$, informs on the rate of glottal closing.

7) The parameter T_a defines the duration of a finite return phase following the instant T_e. T_a is determined by the intersection of the tangent to the differentiated glottal flow dUg immediately after T_e with the zero line. Remark that in this case the return phase is approximated with an exponentional. Increased duration of the return phase T_a is associated with a low pass spectral effect with cut-off frequency $\frac{1}{(2\pi T_a)}$ which influences the bandwidth of the first formant.

Among the above parameter items Q_o, Q_s, D_0 and T_a correspond to the four parameters in the reference model of LF [2]. In addition to the signal parameters calculated from Ug and dUg, h_{min} is considered which is related to the forming to the fluid flow and to the glottal area $A_g = h_{min} * L_w$.

2.3. Two-mass model sensitivity

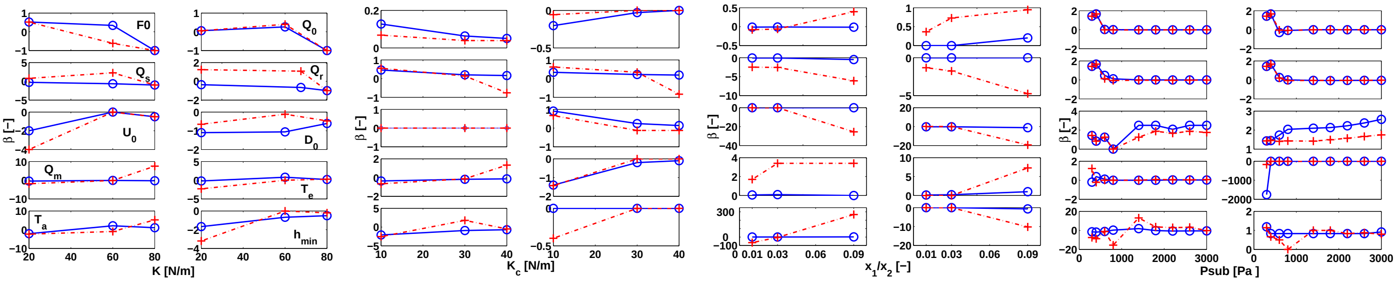
The influence of the input-parameters of the two-mass model on the above given signal-parameters is evaluated by performing a sensitivity analysis. The relative sensitivity is expressed by a dimensionless index of sensitivity $\beta = \frac{(S_{var}-S_{ref})/S_{ref}}{(I_{var}-I_{ref})/I_{ref}}$. The subscript $_{ref}$ denotes the values corresponding with the reference situation given in Table 1 and the subscript $_{var}$ corresponds to the varied values. For each of the changed input-parameters (I) and assessed signal-parameters (S) β is calculated. Since a change in the input parameter I results in a proportional change for the singal parameter S when β equals 1, an absolute value much less than 1 implies a low sensitivity of the two-mass model to that input parameter I.

3. Results and discussion

Table 2 presents the values of the signal-parameters given in 2.2 computed on the two-mass model output-signal Ug (dUg) with the input-parameters set to the values given in Table 1. The sensitivity analysis to the two-mass model input-parameters K, K_c, the ratio $\frac{x_1}{x_2}$ and P_{sub} is visualised in Figure 5 for the

Table 2: *Signal parameters computed on the two-mass model output Ug (dUg) for the input-parameters given in Table 1.*

	vocal tract	F_0 [Hz]	Q_0	Q_s	Q_r	U_0 $[m^3/s]$	D_0 $[m^3/s^2]$	$Q_m[s]$	T_e [s]	T_a [s]	h_{min} [m]
reference	/a/	108	0.69	1.78	1.37	0.0002	0.24	0.0022	0.0010	0.014	0.0006
reference	/i/	104	0.70	3.18	2.90	0.0002	0.29	0.0015	0.0007	0.0047	0.0005



(a) $\beta(K)$. (b) $\beta(K_c)$. (c) $\beta(\frac{x_1}{x_2})$. (d) $\beta(P_{sub})$.

Figure 5: Sensitivity analysis of the two-mass model input-parameters K, K_c, $\frac{x_1}{x_2}$ and P_{sub}. β-values for /a/ are indicated with (o) full, for /i/ with (+) dotted. The assessed signal-parameters are indicated in part 5(a) holds for all parts 5(b), 5(c) and 5(d).

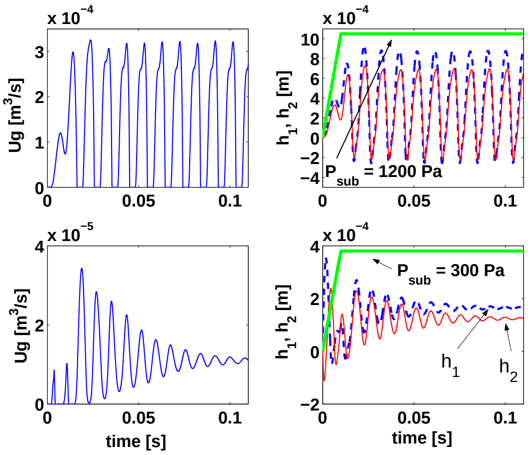


Figure 6: *Two-mass model output Ug at the left and h_1, h_2 at the right for the vowel /i/ different P_{sub} value of respectively 1200 Pa (top)and 300 Pa (bottom). The scaled pressure-steps are illustrated on the right side above h_1 (full) and h_2 (dotted).*

vowels /a/ and /i/ and all signal-parameters indicated in Table 2. From Figure 5(d) follows that except for the amplitudes U_0 and D_0 the two-mass model on both /a/ and /i/ isn't sensitive ($|\beta| < 1$) to changes in P_{sub} above 600 Pa. Below 600 Pa a meaningful influence is notable. Figure 6 shows that the increase in sensitivity corresponds to a severe change in physical behaviour of the model. The two-mass model respons Ug, h_1 and h_2 is shown to a pressure-step well above and well below the critical range, i.e. $P_{sub} = 1200$ Pa and $P_{sub} = 300$ Pa. For $P_{sub} = 300$ Pa the oscillation is damped and the vocal folds don't close. In agreement with the sensitivity analysis the oscillatory-threshold lies for both /i/ and /a/ at about 600 Pa. The used critical height in the collision model was set to $h_c = 0$. The sensitivity plots for $\frac{x_1}{x_2}$ in Figure 5(c) shows the impact of the vocal tract shape. The sensitivity to changes in the inlet and outlet depth or the ratio $\frac{x_1}{x_2}$ is major for the /i/ shape and minor for the /a/ shape. Figure 5(b) shows the influence of K_c on the parameters associated with the discontinuity T_e and T_a. Consequently changing K_c will alter the source spectrum. From Figures 5(a) and 5(b) follows that the two-mass model is more sensitive to changes in K compared to K_c. Once more model output obtained for the /i/ vocal tract shape is more critical to input-parameter values as the /a/ shape. Except for the influence on the discontinuity parameters T_a, T_e changing K has a major influence on the glottal Ug waveform U_0, on the glottal opening h_{min} and so on the fluid model through changing the position of x_s. The collision model needs further research. The large impact of the input-parameters on the signal-parameters in case of acoustical coupling is in agreement with the results presented in [5].

4. Conclusions

The influence of changing input-parameters of the two-mass model on signal-parameters of the glottal volume air flow and resonance behaviour is quantified for /a/ and /i/ vocal tract shapes. The /i/ vocal tract shape appears more sensitive to changes in the input-parameters as the /a/ vocal tract shape.

5. Acknowledgements

The research is supported by the Rhone-Alpes region, France.

6. References

[1] K. Johnson, *Acoustic and auditory phonetics.* Oxford: Blackwell, 1997.

[2] R. Veldhuis, "A computationally efficient alternative for the Liljencrants-Fant model and its perceptual evaluation," *JASA*, vol. 103, no. 1, pp. 566–571, 1998.

[3] A. Rosenberg, "Effect of glottal pulse shape on the quality of natural vowels," *JASA*, vol. 49, pp. 583–590, 1971.

[4] N. Lous, G. Hofmans, R. Veldhuis, and A. Hirschberg, "A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design," *Acta Acustica*, vol. 84, pp. 1135–1150, 1998.

[5] D. Sciamarella and C. d'Alessandro, "A study of the two-mass model in terms of acoustic parameters," in *Proc. ICSLP*, Denver, USA, 2002, pp. 2313–2316.

[6] K. Ishizaka, M. Matsudaira, and T. Kaneko, "Input acoustic-impedance measurement of the subglottal system," *JASA*, vol. 60, no. 1, pp. 190–197, 1976.