Dynamics of a two-mass model of the vocal folds form men, women, and children

Jorge C. Lucero

Department of Mathematics
University of Brasília, Brasília 70910-900, Brazil
lucero@unb.br

Abstract

This paper analyzes how the oscillatory behavior of the vocal folds changes according to laryngeal size. A version of the popular two-mass model of the vocal folds is used, coupled to a two-tube approximation of the vocal tract in configuration for vowel /a/. The standard male configurations of these models are used as reference, and female and child configurations are derived by scaling the dimensions and biomechanical parameters of tissues. For simplicity, a single scaling parameter is used for all the dimensions of the vocal folds and vocal tract. Simulations of the vocal fold oscillation and oral output are produced by numerical solution of the equations, and for varying values of the size scaling parameter. The results show that the oscillation threshold conditions become more restricted for smaller larynges, such as the female and child larynges, in agreement with reported experimental results. They also show a clear hysteresis effect at voice onset-offset, with more severe threshold conditions to start the vocal fold oscillation than those to maintain it.

1. Introduction

The purpose of this paper is to analyze how the oscillatory behavior of the vocal folds at phonation changes according to laryngeal size, in cases of phonation by men, women and children. Since the biomechanical parameters of the vocal folds, the interaction between the vocal folds and airflow, and other terms of glottal aerodynamics depend on the anatomical dimensions of the glottis, variations of the oscillation dynamics as functions of those dimensions might be expected. Such variations might influence the strategies for controlling voicing onset and offset during speech by women vs. men, and might be important to understand the development of motor control of the larynx in children.

In recent works, a version of the two-mass model of the vocal folds [1] was used to simulate speech production of adults [2] and children [3] in the vicinity of an abduction gesture. The objective was to determine control strategies of voicing onset and offset used by speakers, and detect possible differences between female, male, and child speakers. There, an inverse dynamic approach was used, in which the model was fitted to collected speech records. The results showed that devoicing at the abduction-adduction gesture for /h/ is achieved by the combined action of vocal fold abduction, decrease of subglottal pressure, and increase of vocal fold tension. Each of these actions has the effect of inhibiting the vocal fold oscillation, suppressing it when reaching an offset threshold. Also, differences in oscillation regions between men and women were detected. Women had in general more restricted conditions for the vocal fold oscillation, probably as consequence of their smaller laryngeal size.

This paper intends to explore in more depth such relation of the vocal fold oscillation dynamics with laryngeal dimensions. The next sections will present the vocal fold model, and will explore its oscillatory behavior as a function of its size.

2. Models

The larynx is modeled using a modified version of the two-mass model of the vocal folds [1] coupled to a two-tube approximation of the vocal tract in configuration for vowel /a/ [4]. The two-mass model incorporates a nonlinear characteristics for the tissue damping [2,5], and a simplified description of flow separation within a divergent glottis [6].

The standard values of the models' parameters [1,4] are used as reference of a male configuration. To control their size, a single scaling factor β is used for all dimensions. This is a simplification of the actual size variations of the larynx and vocal tract. According to Titze [7], two main scaling factors for the size relation between male and female larynges may be identified, depending on the specific dimension. The relative lengths between the pharynx and oral cavity also differ for men, women, and children. Here, the single factor β is adopted as a convenient and simple way to control the overall size of the model. Thus, an adult female configuration would correspond to an approximate factor of β = 0.72 (according to data by Titze [7]), and a 5-year-old configuration to β = 0.64 (according to data by Goldstein [8]).

All linear dimensions are then scaled by multiplying by β. Masses are accordingly scaled by multiplying by β^3, to compensate the volume change. For the tissue stiffness, a constant elasticity modulus is assumed for all sizes. This assumption is again a simplification, since slightly stiffer tissue for females than for males has been reported [7], probably as a result of differences in tissue composition, Differences between child and adult tissues have also been found [9]. For a constant elasticity modulus, the stiffness coefficient is directly proportional to the cross-sectional area of the tissues, and inversely proportional to their length. Hence, the scaling of all dimensions by a factor β implies that stiffness is also scaled by this same factor. Finally, a constant damping ratio of tissues is assumed for all sizes.

3. Preliminary theory

The dynamics of the two-mass model in the vicinity of its rest position has been analyzed in previous studies (e.g., [10,11]). Let us briefly recall that the stability of that position may be determined by taking the linear part of the equations of motion in its vicinity. Simplifying those equations by neglecting losses by air viscosity, and assuming that the load presented by the vocal tract to the vocal folds is negligible, we find and equilibrium position at the rest position of the vocal folds. The linearized differential equations of the two-mass model around that position are:

$$\begin{cases} \left(\beta^3 m_{01}/Q\right)\ddot{x}_1 + \beta^2 r_1 \dot{x}_1 + \beta Q k_{01} x_1 + \beta Q k_{0c}(x_1 - x_2) \\ \qquad\qquad\qquad\qquad = 2\beta^2 d_1 l_g Ps / x_0, \\ \left(\beta^3 m_{02}/Q\right)\ddot{x}_2 + \beta^2 r_2 \dot{x}_2 + \beta Q k_{02} x_2 + \beta Q k_{0c}(x_2 - x_1) = 0, \end{cases} \quad (1)$$

where the scaling factor Q for the masses and stiffness coefficients [1], and the size scaling coefficient β have been introduced. The subindex 0 in the mass and stiffness coefficients denotes the male reference value. A detailed explanation of the above equations and meaning of their various parameters may be easily found in the literature (e.g., [1]), and hence is omitted here. Its characteristic equation is

$$\begin{aligned} & s^4 + \frac{Q}{\beta}\left(\frac{r_1}{m_{01}} + \frac{r_2}{m_{02}}\right)s^3 + \\ & \frac{Q^2}{\beta^2}\left(\frac{k_{01}+k_{0c}-\Gamma}{m_{01}} + \frac{k_{02}+k_{0c}}{m_{02}} + \frac{r_1 r_2}{m_{01}m_{02}}\right)s^2 + \\ & \frac{Q^3}{\beta^3 m_{01}m_{02}}\left[r_1\left(k_{02}+k_{0c}\right) + r_2\left(k_{01}+k_{0c}-\Gamma\right)\right] + \\ & \frac{Q^4}{\beta^4 m_{01}m_{02}}\left[\left(k_{01}+k_{0c}-\Gamma\right)\left(k_{02}+k_{0c}\right) - k_{0c}\left(k_{0c}-\Gamma\right)\right] = 0 \end{aligned} \quad (2)$$

where

$$\Gamma = \frac{2\beta d_1 l_g P_s}{x_0 Q} \quad (3)$$

Writing the above equation as $s^4 + a_1 s^3 + a_2 s^2 + a_3 s + a_4 = 0$, it may be shown that a pair of complex roots cross the imaginary axis from left to right when $a_1 a_2 a_3 - a_3^2 - a_1^2 a_4 = 0$ [10]. This fact signals the occurrence of a Hopf bifurcation, at which the rest position becomes unstable and a limit cycle is produced, and determines the onset threshold of the vocal fold oscillation.

Here, we want to analyze how the oscillation onset condition depends on the subglottal pressure P_s, the glottal half-width at the rest position of the vocal folds x_0, the Q factor, and the scaling factor β. The glottal half-width characterizes the degree of abduction-adduction of the vocal folds and is one of the major parameters for controlling their oscillation onset-offset. The Q factor represents the degree of tension of the vocal folds, and is the main control for the fundamental frequency of the vocal fold oscillation. The subglottal pressure is a main control for the amplitude of the oscillation, and thus for voice intensity.

Replacing the coefficient values from Eq. (2) into the onset threshold condition, it may be shown that it is independent of the individual values of the above four parameters, when keeping constant the relation

$$\frac{\beta P_s}{x_0 Q} = C \quad (4)$$

where C denotes a constant. For example, for the standard parameters it assumes the value $C = 6.83 \times 10^5$ N/m^3.

Eq. (4) tells that, for smaller larynges (smaller values of β), the threshold value of the subglottal pressure to start the vocal fold oscillation must be higher (larger P_s), the vocal folds must be driven closer together (smaller x_0, or larger adduction), and the tension of the tissues must be smaller (smaller Q). Let us note that this conclusion is not related to the existence of any losses due to air viscosity; in fact, such losses were neglected at the start of this analysis. Looking at the above equations, we may note that factor β appears in Eq. (4) due to the reduction in the medial surface of mass m_1, on which the air pressure acts (if this surface was constant, then the previous conclusions would be just the opposite). Hence, smaller larynges have more restricted phonation regions because their glottal surface is smaller, and so they absorb less energy from the airflow to fuel the vocal fold oscillation.

4. Simulation results

Fig. 1 shows plots of simulated oral airflow, as an example of the model's output. The simulations were obtain by varying the glottal half-width from 0.02 cm to 0.1 cm, and then back to the original value, following a sinusoidal pattern. This variation pattern imitates the glottal abduction-adduction gesture during the production of utterance /aha/ in running speech [2,3]. All other parameters were kept fixed at their standard values.

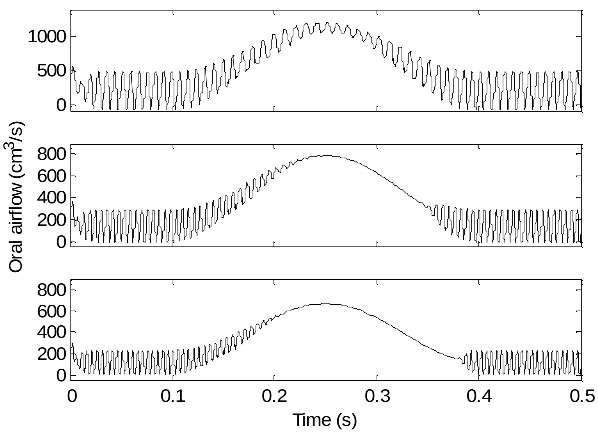

Figure 1: Oral airflow patterns during a vocal fold abduction-adduction gesture. Top: β = 1 (male adult), middle: β = 0.72 (female adult), bottom: β = 0.64 (5-year-old child).

Comparing the plots, we see that the male flow has larger amplitude and lower fundamental frequency (approximately 123 Hz, 165 Hz, 187 Hz, from top to bottom), as expected. In the female case, the glottal pulses stop at the peak abduction, and restart at the end of the following adduction. This is a clear oscillation hysteresis phenomenon, in which the vocal fold oscillation stops and starts at different values of the glottal width. Oscillation hysteresis is a classic nonlinear phenomenon which appears commonly in cases of flow-induced oscillations, and has been modeled by the combination of a cyclic fold and a subcritical Hopf bifurcation [12]. In the child case, the glottal pulses stop even earlier than the female case, at a lower value of the glottal width. The plots clearly show that the oscillation region becomes more restricted as the laryngeal size decreases. Similar results may be obtained when varying the other two main parameters, the subglottal pressure and the Q factor.

The next plots show the threshold conditions of parameters as function of the scaling factor β.

Let us consider first the subglottal pressure thresholds. Simulations of vocal fold oscillation were performed, while varying the subglottal pressure from 0 to 1000 Pa, and back to 0, following a sinusoidal curve, as shown in Fig. 2. The simulated glottal airflow was next low-pass filtered at a 50 Hz cut-off to eliminate glottal pulses, using a sixth order Butterworth filter. The AC flow component was next computed as the difference between the unfiltered airflow and the filtered result. From the AC component, we computed next the rms amplitude, using a zero-crossing algorithm with low pass filtering [13] to identify the individual cycles. The oscillation onset was determined as the instant of time at which the rms flow amplitude increased above a threshold value of 1 cm^3/s. Similarly, the offset was determined as the instant of time at which the rms flow decreased below 1 cm3/s. Identification of the onset threshold is in general an easy task, because the oscillation builds up quickly at that point. The offset threshold, on the other hand, is more difficult and imprecise, because the oscillation amplitude tends to vanish slowly, and it is not clear at which point the rest position has become a stable equilibrium point.

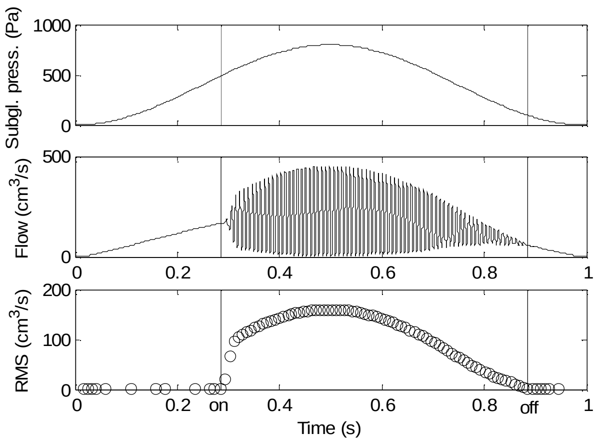

Figure 2: *Simulation results to compute the oscillation threshold value of subglottal pressure. Top: subglottal pressure, middle: glottal airflow, bottom: rms value of the AC glottal airflow. The left and right vertical lines mark the position of the oscillation onset and offset, respectively.*

Fig. 3 shows the computed values of the oscillation thresholds for the subglottal pressure. We can see that there are two different values of the thresholds, one for onset, and a lower value for offset. The existence of such two thresholds confirms again the occurrence of an oscillation hysteresis phenomenon. Both thresholds increase when the larynx size is reduced, as predicted by Eq. (4). For comparison, onset threshold values using this equation are also shown. The theoretical values are much lower that the simulation results, mainly because air viscosity was neglected in the previous analysis. Pressure losses for air viscosity have the consequence of leaving less energy to fuel the oscillation, and therefore of increasing the onset threshold.

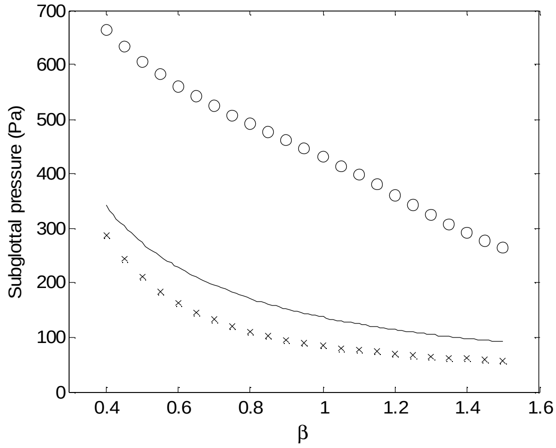

Figure 3: *Oscillation thresholds of subglottal pressure. Circles: onset, crosses: offset, full line: value predicted by Eq. (4).*

Fig. 4 shows the glottal half-width thresholds. Again, we note the existence of two thresholds, with the onset one smaller than the offset. This means that to start the oscillation, the vocal folds must be driven closer together than the position at which oscillation stops. Also, we note that both thresholds decrease with the laryngeal size.

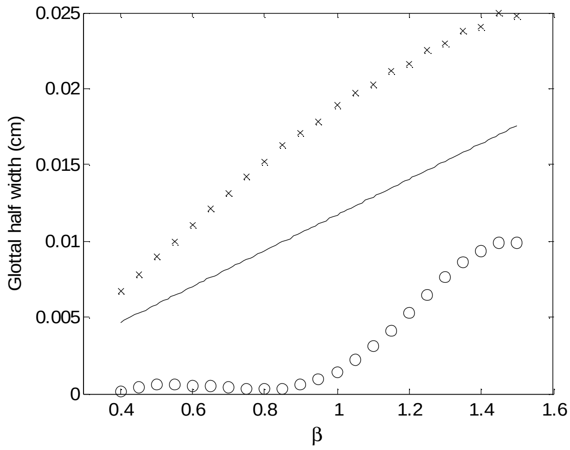

Figure 4: *Oscillation thresholds of glottal half width. Circles: onset, crosses: offset, full line: value predicted by Eq. (4).*

Finally, Fig. 5 shows the threshold values of *Q* factor, with similar results than the other parameters.

In the three plots above, the thresholds vary with laryngeal size always in the sense of restricting the vocal fold oscillation region for smaller sizes.

The final plot in Fig. 6 shows the oscillation thresholds of the glottal half width vs. the total period in which this parameter is varied, while keeping constant other parameters. Changing the period has the effect of changing the rate of variation of the glottal half width. We see that lower rates of variation (large *T*) decrease the offset threshold and increase the onset threshold, reducing then the oscillation hysteresis effect. Similar results are obtained when considering the subglottal pressure and *Q* factor.

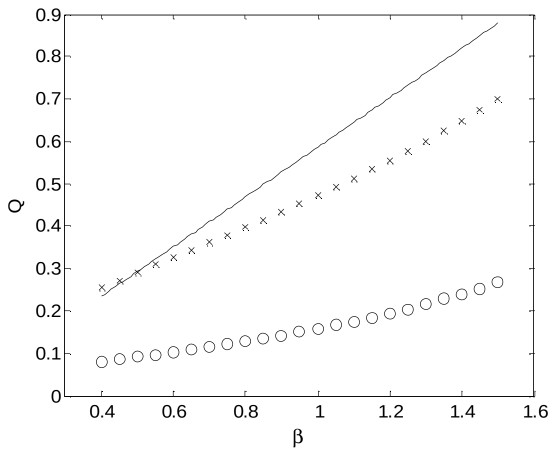

Figure 5: *Oscillation thresholds of Q factor. Circles: onset, crosses: offset, full line: value predicted by Eq. (4).*

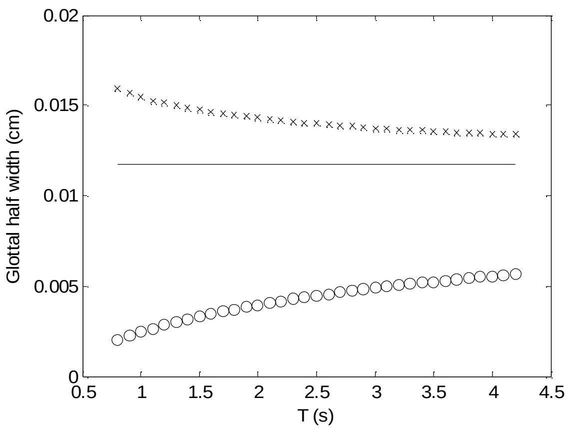

Figure 6: *Oscillation threshold values of glottal half-width, when varying its total period T of variation. . Circles: onset, crosses: offset, full line: value predicted by Eq. (4)*

5. Conclusions

In general, the results show that the dynamics of the vocal fold oscillation depends on laryngeal size, and thus vary for men, women, and children. For all parameters studied here, oscillation conditions become more restricted as the size is reduced. This restriction seems consequence of a reduction of the glottal area in contact with the airflow, on which energy is transferred from the flow to fuel the vocal fold oscillation. This result would explain the larger occurrence of devoicing in glottal abduction-adduction gestures, compared to men [14]. In the child cases, the restricted conditions would result in higher values of subglottal pressures, compared to adults, as observed experimentally. A hysteresis effect is always present at voice onset-offset, with threshold conditions to start the vocal fold oscillation more severe than those to maintain it. The hysteresis also depends on the rate of variation of the control parameters, with narrower onset-offset distances at lower rates.

6. Acknowledgments

This work was supported by CNPq and Finatec (Brazil).

7. References

[1] Ishizaka, K.; Flanagan, J.L., 1972. Synthesis of voiced sounds from a two-mass model of the vocal folds. *Bell System Technical Journal* 51, 1233-1268.

[2] Lucero, J.C.; Koenig, L.L., 2000. Simulations of voice onset-offset hysteresis in /aha/ utterances. *J. Acoust. Soc. Am.* 108, 2508.

[3] Lucero, J.C.; Koenig, L.L., 2003. Simulations of VhV sequences in children. *Proc. 15th ICPhS* (Barcelona, Spain), 2605-2608.

[4] Titze, I.R., 1994. *Principles of Voice Production.* Englewood Cliffs: Prentice-Hall.

[5] Lucero, J.C; Koenig, L.L., 2004. Simulations of oral airflow patterns in men and women using a two-mass model of the vocal folds. Submitted for publication to J. Acoust. Soc. Am.

[6] Pelorson, X.; Hirschberg, A.; van Hassel, R.R.; Wijnands, A.P.J.; Auregan, Y., 1994. Theoretical and experimental study of quasisteady flow separation within the glottis during phonation. Application to a modified two-mass model. *J. Acoust. Soc. Am.* 96, 3416-3431.

[7] Titze, I.R., 1989. Physiologic and acoustic differences between male and female voices. *J. Acoust. Soc. Am.* 85, 1699-1707.

[8] Goldstein, U. , 1980. An articulatory model for the vocal tracts of growing children. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

[9] Kurita, S.; Hirano, M.; Nakashima, T., 1983. Growth, development, and aging of human vocal folds. In *Vocal fold physiology: Contemporary research and clinical issues*, D. M. Bless; J. H. Abbs (eds.). San Diego: College Hill Press, 22-43.

[10] Lucero, J.C, 1993. Dynamics of the two-mass model of the vocal folds: Equilibria, bifurcations, and oscillation region. *J. Acoust. Soc. Am.* 94, 3104-3111.

[11] Steinecke, I.; Herzel, H., 1995. Bifurcations in an asymmetric vocal fold model. *J. Acoust. Soc. Am.* 97, 1878-1884.

[12] Lucero, J.C., 1999. A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset. *J. Acoust. Soc. Am.* 105, 423–431.

[13] Titze, I. R.; Liang, H., 1993. Comparison of F0 extraction methods for high-precision voice perturbation measurements. *J. Speech Hear. Res.* 36, 1120-1133.

[14] Koenig, L.L., 2000. Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *J. Speech Lang. Hear. Res.* 43, 1211–1228.