Biomechanical modeling of chest-falsetto registers and their transition

I. T. Tokuda[1], M. Zemke[2], M. Kob[3], and H. Herzel[2]
[1] Japan Advanced Institute of Science and Technology, Japan, Email: isao@jaist.ac.jp
[2] Humboldt University of Berlin, Germany; [3] University of Music Detmold, Germany

Introduction

The exact definition of registers in the human voice is still under debate [1, 2]. Especially the quantitative analysis of transitions between the registers have not been investigated in much detail yet. Excised larynx experiments show different kinds of voice instabilities that appear close to the transition from chest to falsetto register [3, 4]. These instabilities include abrupt jumps between the two registers exhibiting hysteresis, aphonic episodes, subharmonics and chaos. To model these phenomena, we started with a three-mass cover model, which was constructed by adding one more mass on top of the two-mass model [5]. Because of the additional mass, the upper part of the vocal folds can produce a small amplitude waveform, which resembles the falsetto register. This falsetto-like register can easily coexist with the chest-like register, giving rise to hysteresis phenomena. Near the register transitions, subharmonics and chaos are observed, which reproduce even details of the excised larynx experiment.

For a deeper understanding of the register transition in human voice, several extensions are indispensable. Introduction of the body-cover structure is important for physiologically more realistic modeling of the larynx [6]. Recent studies also showed that a smooth geometry in vocal folds is important for a precise computation of the aerodynamic force, that can produce distinct registers [7, 8, 9]. Influence of subglottal resonances as well as supraglotttal resonances may also play an important role. Therefore, we extend our model to a four-mass body-cover polygon model. Sub- and supraglottal resonances are also coupled to the vocal fold model. We make use of the bifurcation analysis to understand how small changes of parameter values can cause sudden qualitative changes in the dynamical behavior of the larynx.

4-Mass Model

Figure 1 shows a schematic illustration of the four-mass polygon model. This model is composed of a body part m_b and a cover part, which is divided into three masses m_i (lower: $i = 1$, middle: $i = 2$, upper: $i = 3$). Our basic modeling assumptions are the following:

1) Cubic nonlinearities of the oscillators are neglected.
2) Influence of vocal tract as well as subglottal resonances are not considered.
3) Additional pressure drop at inlet is neglected; the Bernoulli flow is considered only below the narrowest part of the glottis [10].

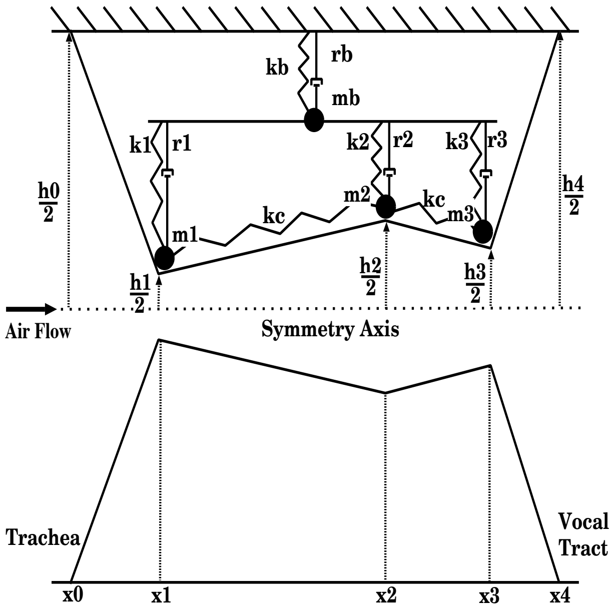


Figure 1: Schematic illustration of the 4-mass polygon model.

4) Symmetric motion between the left and the right vocal folds is assumed.

Our model equations read

$$\begin{aligned} m_1\ddot{y}_1 + r_1(\dot{y}_1 - \dot{y}_b) + k_1(y_1 - y_b) + \Theta(-h_1)c_1(\frac{h_1}{2}) \\ +k_{1,2}(y_1 - y_2) = F_1, \\ m_2\ddot{y}_2 + r_2(\dot{y}_2 - \dot{y}_b) + k_2(y_2 - y_2) + \Theta(-h_2)c_2(\frac{h_2}{2}) \\ +k_{1,2}(y_2 - y_1) + k_{2,3}(y_2 - y_3) = F_2, \\ m_3\ddot{y}_3 + r_3(\dot{y}_3 - \dot{y}_b) + k_3(y_3 - y_b) + \Theta(-h_3)c_3(\frac{h_3}{2}) \\ +k_{2,3}(y_3 - y_2) = F_3, \\ m_b\ddot{y}_b + r_b\dot{y}_b + k_by_b + r_1(\dot{y}_b - \dot{y}_1) + k_1(y_b - y_1) + r_2(\dot{y}_b \\ -\dot{y}_2) + k_2(y_b - y_2) + r_3(\dot{y}_b - \dot{y}_3) + k_3(y_b - y_3) = 0. \end{aligned}$$

The dynamical variables y_i represent displacements of the masses m_i, where the corresponding glottal opening length is given by $h_i = h_{0i} + 2y_i$ (h_{0i}: prephonatory length; $i = 1, 2, 3$). The constant parameters r_i, k_i, c_i represent damping, stiffness, and collision stiffness of the masses m_i, respectively, whereas $k_{i,j}$ represents coupling strength between two masses m_i and m_j. The stiffness is determined as $r_i = 2\zeta_i\sqrt{m_ik_i}$ using the damping ratio ζ_i. The collision function is approximated as $\Theta(\xi) = 0$ ($\xi \leq 0$); $\Theta(\xi) = 1$ $(0 < \xi)$.

The aerodynamic force, F_i, acting on each mass is derived as follows. First, the vocal fold geometry is described by a pair of four mass-less plates as shown in Fig. 1. The flow channel hight $h(x,t)$ is a piecewise linear function, composed of $h_{1,0}$ $(x_0 \le x \le x_1)$, $h_{2,1}$ $(x_1 \le x \le x_2)$, $h_{3,2}$ $(x_2 \le x \le x_3)$, and $h_{4,3}$ $(x_3 \le x \le x_4)$, which are determined as

$$h_{i,i-1}(x,t) = \frac{h_i(t) - h_{i-1}(t)}{x_i - x_{i-1}}(x - x_{i-1}) + h_{i-1}(t),$$

where $i = 1,2,3,4$ and h_0 and h_4 are constants. Assuming the Bernoulli flow, the pressure distribution $P(x,t)$ below the narrowest part of the glottis, $h_{min} = \min(h_1, h_2, h_3)$, is described as

$$P_s = P(x,t) + \frac{\varrho}{2}(\frac{U}{h(x,t)l})^2 = P_0 + \frac{\varrho}{2}(\frac{U}{h_{min}l})^2,$$

where ϱ represents the air density, P_s is the subglottal pressure, and l is the length of the glottis. The aerodynamic forces on the plates are induced by the pressure $P(x,t)$ along the flow channel. As m_1, m_2, m_3 support the plates, an aerodynamic force on point i $(i = 1,2,3)$ is found to be

$$\begin{aligned} F_i(t) &= \int_{x_{i-1}}^{x_i} l\frac{x - x_{i-1}}{x_i - x_{i-1}} P(x,t)dx \\ &+ \int_{x_i}^{x_{i+1}} l\frac{x_{i+1} - x}{x_{i+1} - x_{i-1}} P(x,t)dx. \end{aligned}$$

This integral can be solved analytically [8] for the pressure distribution $P(x,t)$ defined above.

As the default situation, the parameters are set as: $m_1 = 0.009$ g; $m_2 = 0.009$ g; $m_3 = 0.003$ g; $m_b = 0.05$ g; $k_1 = 0.006$ g/ms^2; $k_2 = 0.006$ g/ms^2; $k_3 = 0.002$ g/ms^2; $k_b = 0.03$ g/ms^2; $k_{1,2} = 0.001$ g/ms^2; $k_{2,3} = 0.0005$ g/ms^2; $c_1 = 3k_1$; $c_2 = 3k_2$; $c_3 = 3k_3$; $\zeta_1 = 0.1$; $\zeta_2 = 0.4$; $\zeta_3 = 0.4$; $\zeta_b = 0.4$; $h_{01} = 0.036$ cm^2; $h_{02} = 0.036$ cm^2; $h_{03} = 0.036$ cm^2; $x_0 = 0$ cm; $x_1 = 0.05$ cm; $x_2 = 0.2$ cm; $x_3 = 0.275$ cm; $x_4 = 0.3$ cm; $l = 1.4$ cm; $\varrho = 0.00113$ g/cm^3.

The tension parameter Q is introduced to control the frequency of the four masses as $m'_i = m_i/Q$, $k'_i = k_i \cdot Q$ $(i = 1,2,3,b)$.

Sub- and supraglottal resonances were described by using the wave-reflection model [11, 12, 13], which is a time-domain model of the propagation of one-dimensional planar acoustic waves through a collection of uniform cylindrical tubes. The supraglottal system was modeled as a simple uniform tube (area: 3 cm, length: 17.5 cm), which is divided into 44 cylindrical sections. The area function for the subglottal tract was based on the one proposed by Zañartu *et al.* [14]. The area function is composed of 62 cylindrical sections. For both sub- and supraglottal systems, the section length Δz was set to 17.5/44 cm.

To couple the sup- and supraglottal resonators to the vocal folds model, an interactive source-filter coupling was realized according to Titze [13, 15]. The lung pressure was set as $P_l = 0.012$ g/cm·ms^2.

Simulations

With the default parameter setting ($Q = 1$) without vocal tract but with constant subglottal pressure ($P_s = 0.008$ g/cm·ms^2), we observed a chest-like phonation as visualized in Fig. 2 (a). We observe the known phase advance of the lower mass and complete glottal closure leading to the slightly skewed volume flow shown in Fig. 2 (b). This chest-like waveform has a frequency of 96 Hz. If we change the tension parameter to $Q = 3.6$, qualitatively distinct vibration pattern appears. Fig. 2 (c) shows phase-shifted vibrations of the upper two masses, whereas the lowest mass is wide open. The observed frequency of 369 Hz is much higher than the chest-like vibration. Due to the lack of the vocal fold collision, the volume flow waveform in Fig. 2 (d) shows an almost sinusoidal waveform. In this way, falsetto-like vibrations can be simulated.

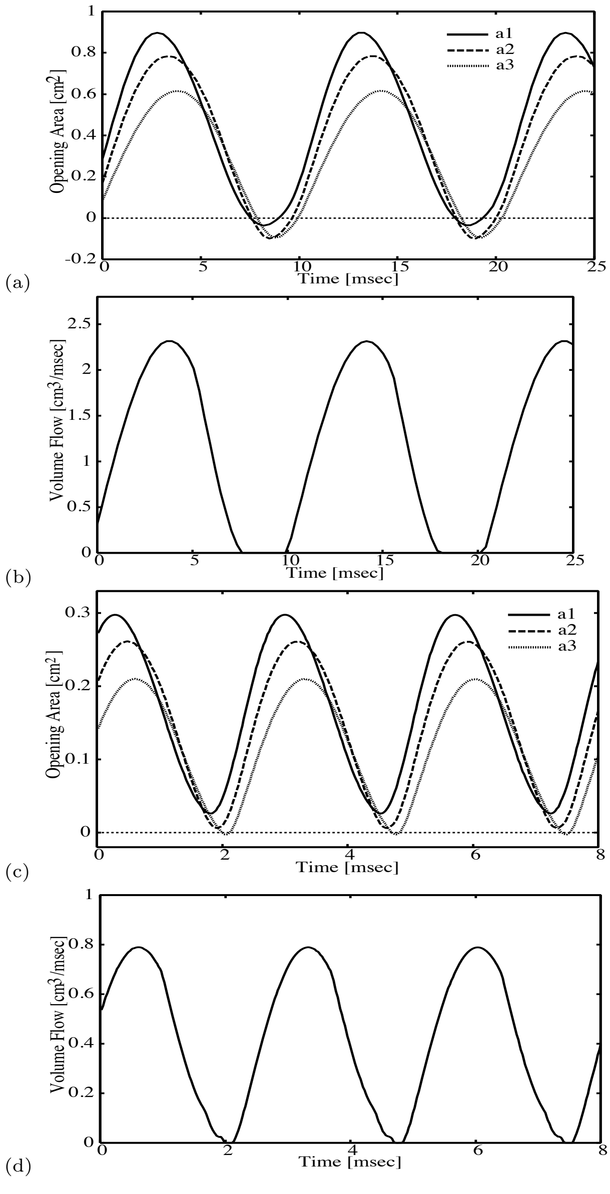
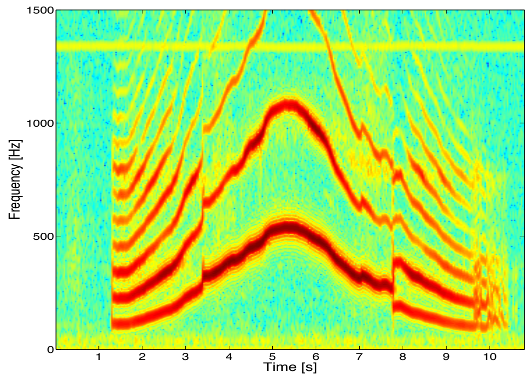


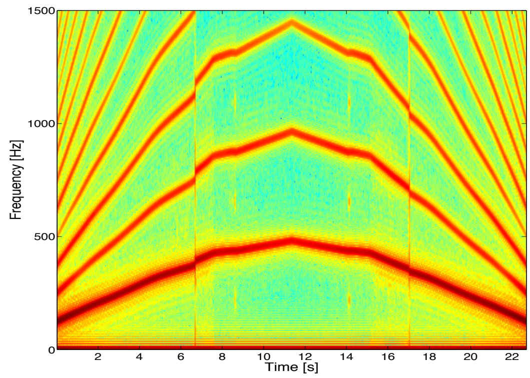
Figure 2: (a), (c) Time series of glottal areas $h_i l$ (1st mass: dotted thin line, 2nd mass: dotted bold line, 3rd mass: solid line) for chest and falsetto registers, respectively. (b), (d): Glottal volume flow $U(t)$ corresponding to (a) and (c).

Let us study the transition between the chest and the falsetto registers. The analysis of register transitions and source-tract interaction is often studied using glissando singing [16], experimental variation of vocal fold tension [5] or gliding of the fundamental frequency in biomechanical models [15]. In Figure 3, we compare a glissando of an untrained singer with simulations of a corresponding F_0 glide in our four-mass model coupled to sub- and supraglottal resonators. The singer's glissando in Fig. 3a exhibits register transitions with frequency jumps around 3.3 s and 7.8 s at slightly different pitches. There is an abrupt phonation onset at 1.2 s and a more smooth offset with some irregularities. Glissando is simulated in Fig. 3b by varying our tension parameter Q from 1 to 5.5 and then back. We find a frequency jump at 6.7 s ($Q = 3.8$, $F_0 = 390$ Hz) and a backward transition at 17 s ($Q = 3.3$, $F_0 = 350$ Hz). These differences between chest-falsetto and falsetto-chest transitions are a landmark of hysteresis. Hysteresis indicates that there are coexisting vibratory regimes ("limit cycles") for a range of parameters. Moreover, hysteresis implies that there are voice breaks instead of *passagi* of trained singers.

In addition to register transitions, occasionally subharmonics at 8.7 s and 14.1 s are observed. It has been discussed earlier [3, 5] that register transitions are often accompanied by nonlinear phenomena such as subharmonics and chaos. The gross features of the experimental and simulated F_0 glides in Figure 3 are similar.



(a)



(b)

Figure 3: (a) Spectrogram of human voice (subject: MF) with a gliding fundamental frequency (F_0). (b) Model simulation of the gliding F_0.

Summary

Our simulations reveal that the proposed 4-mass polygon model can reproduce coexistence of chest and falsetto registers as well as complex transitions between them as observed in vocalization of untrained singers. The register transitions exhibit pronounced hysteresis and near the frequency jumps subharmonics are observed. Combination of experimental studies with biomechanical modeling and bifurcation theory will lead to further insight into the debated field of voice registers.

Acknowledgements

This study was supported by SCOPE (071705001) of Ministry of Internal Affairs and Communications (MIC), Japan.

References

[1] I. R. Titze, *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, 1994).

[2] J. G. Švec *et al.*, Ann. Otol. Rhinol. Laryngol. **116**, 172 (2007).

[3] D. A. Berry *et al.*, J. Voice **10**, 129 (1996).

[4] J. Horáček *et al.*, Proc. Int. Conf. Voice Phys. Biom., (Marseille, 2004), pp. 87.

[5] I. T. Tokuda *et al.*, J. Acoust. Soc. Am. **122**, 519 (2007).

[6] B. H. Story and I. R. Titze, J. Acoust. Soc. Am. **97**, 1249 (1995).

[7] X. Pelorson *et al.*, J. Acoust. Soc. Am. **96**, 3416 (1994).

[8] N. J. Lous *et al.*, Acta Acustica **84**, 1135 (1998).

[9] D. Sciamarella and C. d'Alessandro, Acta Acustica **90**, 746 (2004).

[10] I. Steinecke and H. Herzel, J. Acoust. Soc. Am. **97**, 1571 (1995).

[11] J. L. Kelly and C. Lochbaum, Proc. 4th Int. Congress on Acoustics, Paper G42, pp. 1 (1962).

[12] J. Liljencrants, "Speech synthesis with a reflection-type line analog," Doctoral dissertation, Royal Institute of Technology, Stockholm (1985).

[13] I. R. Titze, *Myoelastic Arodynamic Theory of Phonation* (National Center for Voice and Speech, Iowa City, IA, 2006).

[14] M. Zañartu *et al.*, J. Acoust. Soc. Am. **121**, 1119 (2007).

[15] I. R. Titze, J. Acoust. Soc. Am. **123**, 2733 (2008).

[16] N. Henrich *et al.*, J. Acoust. Soc. Am. **117**, 1417 (2005).