Biomechanical simulation of chest-falsetto transitions and the influence of vocal tract resonators

Isao T. Tokuda* and Hanspeter Herzel†
* Japan Advanced Institute of Science and Technology, 923-1292 Ishikawa, Japan
E-mail: isao@jaist.ac.jp
† Humboldt University of Berlin, D-10115 Berlin, Germany

***Abstract*—Biomechanical modeling is carried out to study register transitions of the human voice. Registers and their transitions play an important role not only in the singing voice but also in the expressive speech. Our model has a body-cover structure, which is composed of fourth masses. A smooth geometry is realized by introducing a polygon shape to the vocal fold model. The simulation study shows that the model can reproduce many complex phenomena such as register jumps, hysteresis, subharmonics, and chaos, observed in excised larynx experiment as well as in vocalization of untrained singers. Influence of the subglottal and supraglottal resonances on the transition point of the registers is also investigated in detail.**

I. 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. Understanding of these phenomena provides a deeper insight not only into the singing voice but also into the expressive speech. To model the registers and their transitions, 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. Our model simulations reveal hysteresis at chest-falsetto transition, which is consistent with experimental data. We find that vocal tract resonances have a pronounced effect on the chest-falsetto transition.

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

Our model equations read

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

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 $r_i = 2\zeta_i\sqrt{m_i k_i}$ 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 $h_{1,0}$ ($x_0 \leq x \leq x_1$), $h_{2,1}$ ($x_1 \leq x \leq x_2$), $h_{3,2}$ ($x_2 \leq x \leq x_3$), and $h_{4,3}$ ($x_3 \leq x \leq x_4$), 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.

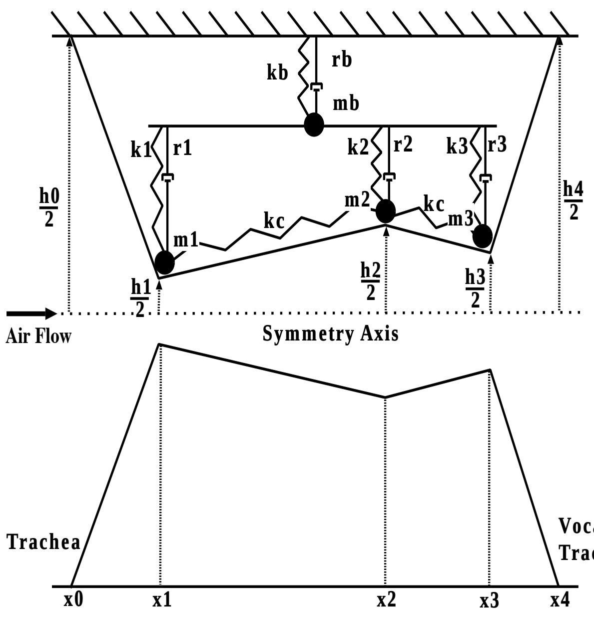

Fig. 1. Schematic illustration of the 4-mass polygon model.

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]. In this formula, the glottal flow is given by $U = \frac{a_g}{k_t}\{-(\frac{a_g}{A^*}) \pm [(\frac{a_g}{A^*})^2 + \frac{2k_t}{\rho c^2}(P_l + 2p_s^+ - 2p_e^-)]^{1/2}\}$, where $A^* = A_s A_e/(A_s + A_e)$ with A^s and A^e being the subglottal and supraglottal entry areas, respectively. k_t is a transglottal pressure coefficient set as 1.0. P_l stands for the lung pressure, whereas p_s^+ and p_e^- represent the incident partial wave pressures arriving from the subglottis and supraglottis, respectively. In the present study, subglottal and supraglottal entry areas were set to be equal to that of the last section of the subglottal system and that of the initial section of the supraglottal system, respectively. The lung pressure was set as $P_l = 0.012$ g/cm·ms^2.

III. 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.

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

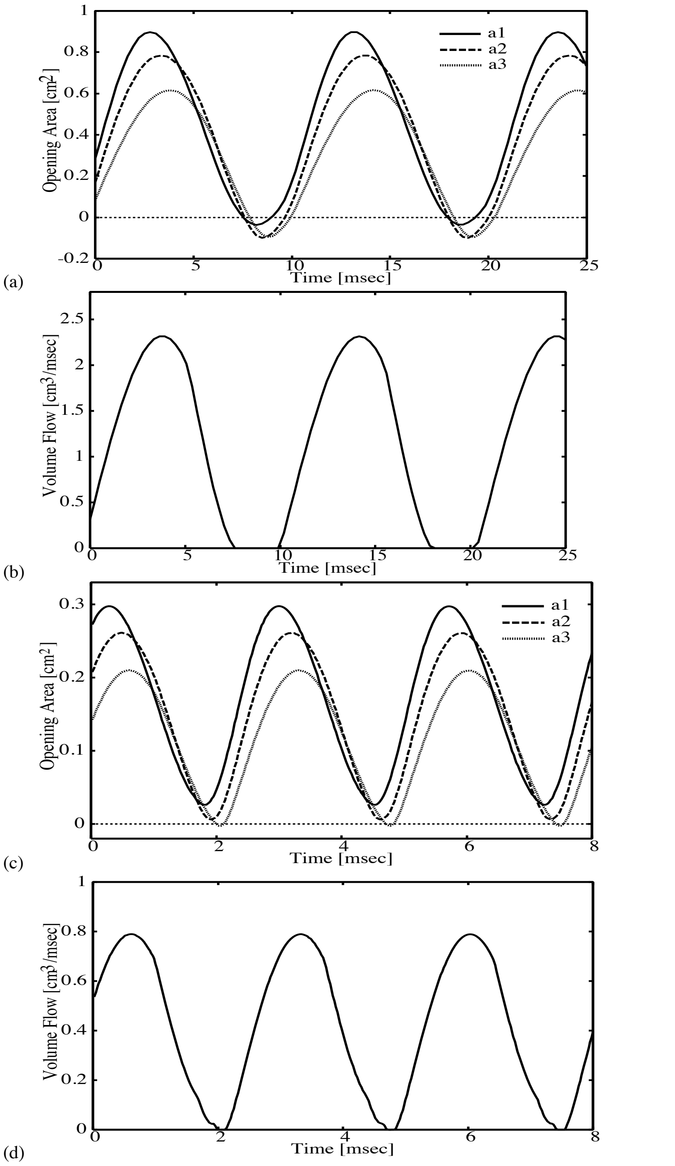

Fig. 2. (a), (c) Time series of glottal areas $h_i l$ (1st mass: dotted thin line, 2nd mass: dotted bold line, 3rd mass: solid line) for chest and falsetto registers, respectively. (b), (d): Glottal volume flow $U(t)$ corresponding to (a) and (c).

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.

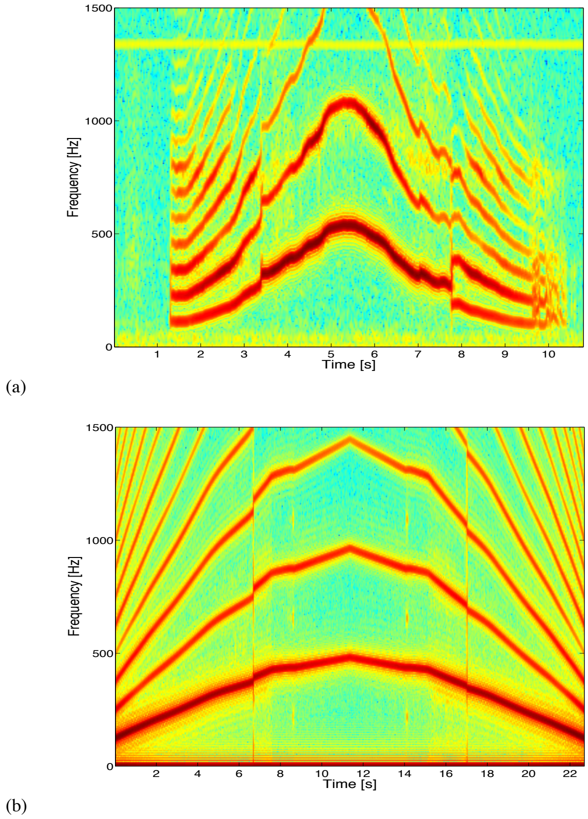

Fig. 3. (a) Spectrogram of human voice (subject: MF) with a gliding fundamental frequency (F_0). (b) Model simulation of the gliding F_0.

IV. Influence of Vocal Tract on Chest-Falsetto Transition

We induced register transitions in our model by changing the tension parameter Q gradually and measuring the fundamental frequency F_0, the amplitude and the number of the colliding masses. We observed a monotonous increase of F_0 and collision of all 3 cover masses at low F_0 and of only the top mass at high F_0. For simplicity, a binary classification is applied to draw the register transitions of Figures 4, 5, and 6 as follows: collision of 3 cover masses are termed chest, whereas collision of less masses are termed falsetto. If only the upper masses collide, the open quotient became large as known from measurement in singers [16]. In our bifurcation diagrams, with the tension Q as the bifurcation parameter, we plotted the fundamental frequency F_0 on the x-axis instead of Q, since this allows a direct comparison with glissando spectrograms.

Figure 4 compare register transitions of the isolated four-mass model with the ones of the complete model including sub- and supraglottal resonances. In both cases, we find a pronounced hysteresis ranging about 30-40 Hz. Most notable is the dramatic shift of the transition due to the coupling to vocal tract resonators. This observation reveals that the chest-falsetto transition depends sensitively on source-tract interactions.

In order to substantiate this finding, we varied the length of the sub- and supraglottal tubes. First, we decreased and increased the length of the subglottal tube by 25 %. It turned out that there are only minor effects on the register transition. The length changes led to shifts of the transition by 10-15 Hz (see Fig. 5). In contrast, the supraglottal resonance had a profound effect: changing the default length of 17.5 cm to 75 % or 125 % induced major shifts of the frequencies at which register transitions are observed (see Fig. 6). Note that these vocal tract lengths are within the physiological range. For all considered vocal tract length, subharmonics were observed slightly above the chest-falsetto transition. Hysteresis with a frequency difference of about 30-40 Hz and subharmonics were robust features of our simulation. The pitch of the register transition was, however, strongly affected by the formant frequencies. These numerical results lead to testable predictions. For instance, in glissando, the chosen vowel might influence the location of voice breaks. Furthermore, singing into a tube [17] should affect the chest-falsetto transition drastically.

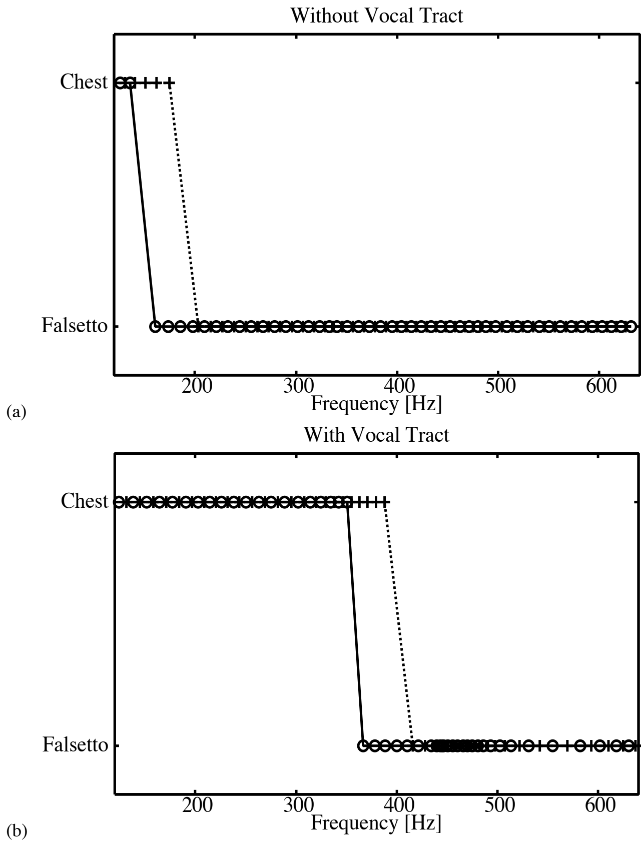

Fig. 4. Register transition of the vocal fold model without vocal tract (a) and with vocal tract (b). The default lengths for sub- and supraglottis are $L_{sub} = 24.7$ cm and $L_{sup} = 17.5$ cm, respectively. Frequency domains for chest and falsetto registers are drawn, where the curves were drawn by both increasing (dotted line with crosses) and decreasing (soild line with circles) the tension parameter Q.

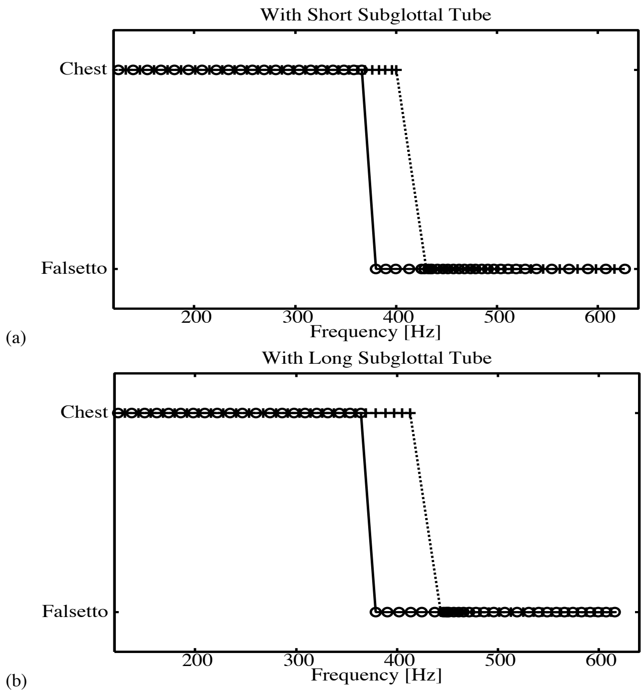

Fig. 5. Dependence of the register transition on the subglottal length. (a): $L_{sub} = 18.3$ cm; (b): $L_{sub} = 30.2$ cm.

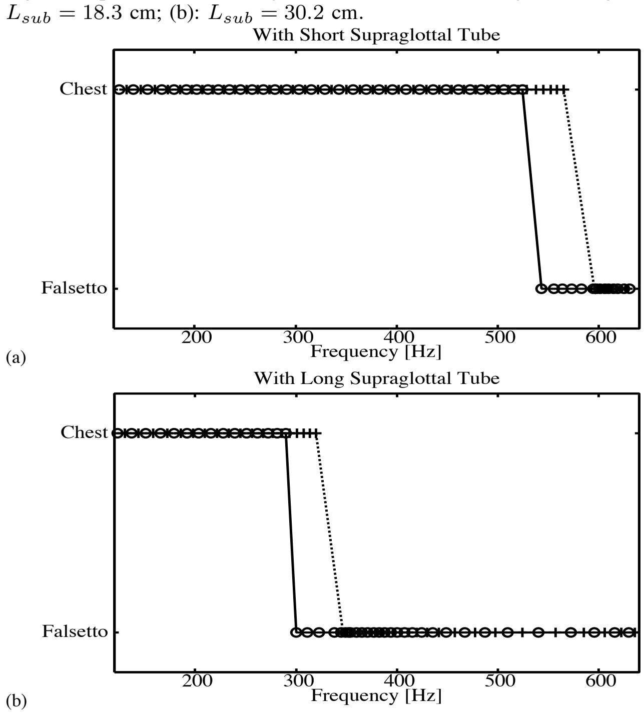

Fig. 6. Dependence of the register transition on the supraglottal length. (a): $L_{sup} = 13.125$ cm; (b): $L_{sup} = 21.875$ cm.

V. 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. It is also an impotant future work to model the register changes in the expressive speech.

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.*, "Videokymography in voice disorders: What to look for?" Ann. Otol. Rhinol. Laryngol. **116**, 172 (2007).

[3] D. A. Berry *et al.*, "Bifurcations in excised larynx experiments," J. Voice **10**, 129 (1996).

[4] J. Horáček *et al.*, "Bifurcations in excised larynges caused by vocal fold elongation," Proc. Int. Conf. Voice Phys. Biom., (Marseille, 2004), pp. 87.

[5] I. T. Tokuda *et al.*, "Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments," J. Acoust. Soc. Am. **122**, 519 (2007).

[6] B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. **97**, 1249 (1995).

[7] X. Pelorson *et al.*, "Theoretical and experimental study of quasi-steady flow separation within the glottis during phonation," J. Acoust. Soc. Am. **96**, 3416 (1994).

[8] N. J. Lous *et al.*, "A symmetrical two mass vocal fold model coupled to vocal tract and trachea, with application to prothesis design," Acta Acustica **84**, 1135 (1998).

[9] D. Sciamarella and C. d'Alessandro, "On the acoustic sensitivity of a symmetrical two-mass model of the vocal folds to the variation of control parameters," Acta Acustica **90**, 746 (2004).

[10] I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal fold model," J. Acoust. Soc. Am. **97**, 1571 (1995).

[11] J. L. Kelly and C. Lochbaum, "Speech synthesis," 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.*, "Influence of acoustic loading on an effective single mass model of the vocal folds," J. Acoust. Soc. Am. **121**, 1119 (2007).

[15] I. R. Titze, "Nonlinear source-filter coupling in phonation: Theory," J. Acoust. Soc. Am. **123**, 2733 (2008).

[16] N. Henrich *et al.*, "Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency," J. Acoust. Soc. Am. **117**, 1417 (2005).

[17] H. Hatzikirou, W. T. Fitch, and H. Herzel, "Voice instabilities due to source-tract interactions," Acta Acustica **92**, 468-475 (2006).