Biomechanical modeling of register transitions and the role of vocal tract resonators

Isao T. Tokuda\textsuperscript{b})
School of Information Science, Japan Advanced Institute of Science and Technology, Nomi-city, Ishikawa 923-1292, Japan

Marco Zemke
Institut für Informatik, Humboldt University Berlin, Rudower Chaussee 25, 12489 Berlin, Germany

Malte Kob
Erich Thienhaus Institute, University of Music Detmold, Neustadt 22, 23756 Detmold, Germany

Hanspeter Herzel
Institute for Theoretical Biology, Humboldt University Berlin, Invalidenstraße 43, 10115 Berlin, Germany

(Received 3 March 2009; revised 27 July 2009; accepted 31 July 2009)

Biomechanical modeling and bifurcation theory are applied to study phonation onset and register transition. A four-mass body-cover model with a smooth geometry is introduced to reproduce characteristic features of chest and falsetto registers. Sub- and supraglottal resonances are modeled using a wave-reflection model. Simulations for increasing and decreasing subglottal pressure reveal that the phonation onset exhibits amplitude jumps and hysteresis referring to a subcritical Hopf bifurcation. The onset pressure is reduced due to vocal tract resonances. Hysteresis is observed also for the voice breaks at the chest-falsetto transition. Varying the length of the subglottal resonator has only minor effects on this register transition. Contrarily, supraglottal resonances have a strong effect on the pitch, at which the chest-falsetto transition is found. Experiment of glissando singing shows that the supraglottis has indeed an influence on the register transition.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3299201]

PACS number(s): 43.70.Bk, 43.70.Gr [DAB]

I. INTRODUCTION

Voice registers have been introduced for perceptually distinct types of certain vocal qualities that can be maintained over some ranges of pitch and loudness (Titze, 2000). The perceptive classification can be accompanied by measurements of voice source parameters such as spectral slope or glottal open quotient (Henrich \textit{et al.}, 2005; Salomão and Sundberg, 2008). Characteristic features are vocal breaks at register transitions associated with pitch and amplitude jumps (Roubeau \textit{et al.}, 1987; Švec \textit{et al.}, 1999; Miller \textit{et al.}, 2000). It has been shown by Hirano \textit{et al.} (1970) that the thyroarytenoid (TA) muscle and the cricothyroid (CT) muscle regulate register transitions. Consequently, the perceptive aspects of registers are inherently related to laryngeal features of vocal fold vibrations. In chest phonation the vocal folds are thick and glottal closure is complete, whereas in falsetto only the vocal fold edges vibrate (Vilkman \textit{et al.}, 1995). Despite extensive experimental investigations using acoustic signals (Hollien, 1974; Sundberg and Gauffin, 1979), electromyography (Shipp and McGlone, 1971), electroglottography (Henrich \textit{et al.}, 2005), and videokymography (Švec \textit{et al.}, 2008), many questions regarding register transitions remain open: what determines the pitch of involuntary register transition? What is the role of sub- and supraglottal resonances? Is there hysteresis at register transition?

In order to address these problems, biomechanical modeling can complement experimental studies. Even though some register-like phenomena have been described in two-mass models (Sciamarella and d’Alessandro, 2004; Zacarelli \textit{et al.}, 2006), an appropriate representation of vibratory modes in chest and falsetto requires more advanced models such as body-cover model of Story and Titze (1995), two-dimensional model of Adachi and Yu (2005), and three-mass model of Tokuda \textit{et al.} (2007). Moreover, a smoothed glottal geometry improves the classical two-mass model of Ishizaka and Planagan (1972) considerably (Pelorson \textit{et al.}, 1994; Lous \textit{et al.}, 1998).

In this paper, we use a four-mass body-cover polygon model (Tokuda \textit{et al.}, 2008) to study register transitions and the influence of resonators. Sub- and supraglottal resonances are described using the wave-reflection model (Kelly and Lochbaum, 1962; Liljencrants, 1985; Story, 1995; Titze, 2006). Our model simulations reveal hysteresis at the phonation onset and at chest-falsetto transition, which is consistent with experimental data (Berry \textit{et al.}, 1996; Horáček \textit{et al.}, 2004). We find that vocal tract resonances have a pronounced effect on the chest-falsetto transition.

II. FOUR-MASS BODY-COVER POLYGON MODEL

There are complex high-dimensional models of vocal fold vibrations (Alipour \textit{et al.}, 2000; Titze, 2006; Gömmel \textit{et al.}, 1994; Story and Titze, 1995) to study register transitions. The perceptive classification can be accompanied by measurements of voice source parameters such as spectral slope or glottal open quotient (Henrich \textit{et al.}, 2005; Salomão and Sundberg, 2008). Characteristic features are vocal breaks at register transitions associated with pitch and amplitude jumps (Roubeau \textit{et al.}, 1987; Švec \textit{et al.}, 1999; Miller \textit{et al.}, 2000). It has been shown by Hirano \textit{et al.} (1970) that the thyroarytenoid (TA) muscle and the cricothyroid (CT) muscle regulate register transitions. Consequently, the perceptive aspects of registers are inherently related to laryngeal features of vocal fold vibrations. In chest phonation the vocal folds are thick and glottal closure is complete, whereas in falsetto only the vocal fold edges vibrate (Vilkman \textit{et al.}, 1995). Despite extensive experimental investigations using acoustic signals (Hollien, 1974; Sundberg and Gauffin, 1979), electromyography (Shipp and McGlone, 1971), electroglottography (Henrich \textit{et al.}, 2005), and videokymography (Švec \textit{et al.}, 2008), many questions regarding register transitions remain open: what determines the pitch of involuntary register transition? What is the role of sub- and supraglottal resonances? Is there hysteresis at register transition?

In order to address these problems, biomechanical modeling can complement experimental studies. Even though some register-like phenomena have been described in two-mass models (Sciamarella and d’Alessandro, 2004; Zacarelli \textit{et al.}, 2006), an appropriate representation of vibratory modes in chest and falsetto requires more advanced models such as body-cover model of Story and Titze (1995), two-dimensional model of Adachi and Yu (2005), and three-mass model of Tokuda \textit{et al.} (2007). Moreover, a smoothed glottal geometry improves the classical two-mass model of Ishizaka and Planagan (1972) considerably (Pelorson \textit{et al.}, 1994; Lous \textit{et al.}, 1998).

In this paper, we use a four-mass body-cover polygon model (Tokuda \textit{et al.}, 2008) to study register transitions and the influence of resonators. Sub- and supraglottal resonances are described using the wave-reflection model (Kelly and Lochbaum, 1962; Liljencrants, 1985; Story, 1995; Titze, 2006). Our model simulations reveal hysteresis at the phonation onset and at chest-falsetto transition, which is consistent with experimental data (Berry \textit{et al.}, 1996; Horáček \textit{et al.}, 2004). We find that vocal tract resonances have a pronounced effect on the chest-falsetto transition.
et al., 2008; Zhang, 2009) describing anatomical and physiological details. However, many parameters are not precisely known and a comprehensive bifurcation analysis is difficult (Berry et al., 1994). Consequently, we constructed a low-dimensional model that is consistent with the basic experimental observations.

Our model is based on the body-cover differentiation proposed by Story and Titze (1995), a three-mass representation of the cover (Tokuda et al., 2007), and a smooth vocal fold geometry as in Lous et al. (1998). Advantage of dividing the cover layer into the three masses is that they are suitable for representing the coexistence of different vibratory patterns, which may correspond to chest and falsetto registers (Tokuda et al., 2007). The chosen parameter values are similar to the values in these papers and they are consistent with muscle activation rules (Titze and Story, 2002). A detailed discussion of the modeling and a complete set of equations and parameters are given in Appendixes A and B and a recent thesis (Zemke, 2008).

Figure 1 visualizes our four-mass polygon model. The three cover masses allow wave-like vibrations of the whole vocal folds with a complete closure of the glottis in chest-register simulation. For other parameter sets, high-pitched oscillations with diminished closure of the glottis are simulated, which resemble falsetto register. Figure 2 shows chest-like vibrations (fundamental frequency of 96 Hz) for the default parameters listed in Appendix A. The phase shifts in the opening areas shown in the upper graph allow the energy transfer from the air flow to the masses and contribute to a skewing of the glottal pulses of the lower graph.

In order to simulate register transitions, we recall rules for controlling low-dimensional vocal fold models with muscle activation (Smith et al., 1992; Titze and Story, 2002). An active CT muscle decreases the vibrating mass and increases stiffness. We introduce a tension parameter $T$ which mimics the CT muscle (Steinecke and Herzel, 1995). Masses are divided by $T$ and stiffness parameters are multiplied by $T$ and thus the fundamental frequency of the model is roughly proportional to the parameter $T$. Increasing $T$ from 1 to 7 allows pitch variation from 100 to 600 Hz exhibiting register transitions (detailed simulations in Sec. IV). Figure 3 shows an example of simulating falsetto-like vocal fold oscillations (fundamental frequency of 369 Hz) with the tension parameter $T=3.6$. An almost sinusoidal glottal volume flow with weaker harmonics is observed.

III. MODELING VOCAL TRACT RESONANCES

For male speech, the linear source-filter theory (Fant, 1960) was quite successful (Stevens, 1999). However, as shown already in Ishizaka and Flanagan (1972), source-filter coupling is essential if the fundamental frequency $F_0$ is comparable to the formant frequencies (Titze, 2008; Titze et al., 2008). Register transition can be regarded as bifurcations of limit cycle oscillations (see Tokuda et al., 2007 for a detailed
discussed). It is known that bifurcations, i.e., sudden transitions due to slow parameter variations, might depend strongly on system parameters. Thus, it is possible that even medium effects of the resonances (see Titze et al., 2008 for data on source-filter interactions) can shift register transition drastically. It has been suggested that particular subglottal resonances govern involuntary register transitions (Titze, 2000; Zhang et al., 2006).

In order to study source-filter coupling, we implemented the wave-reflection model (Kelly and Lochbaum, 1962; Liljencrantz, 1985; Story, 1995; Titze, 2000). Details of the simulations are given in Appendix B. For simplicity we approximate the resonators by uniform tubes characterized by their length and area. This simplification gives direct insight on how resonance frequencies given by the tube lengths affect the location of register transitions.

IV. SIMULATION RESULTS

A. Simulation of gliding pitch

The analysis of register transitions and source-tract interaction is often studied using glissando singing (Henrich et al., 2005), experimental variation in vocal fold tension (Tokuda et al., 2007), or gliding of the fundamental frequency in biomechanical models (Titze, 2008). In Fig. 4, 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. 4(a) exhibits register transitions with frequency jumps around 3.3 and 7.8 s at slightly different pitches. There is an abrupt phonation onset at 1.2 s and a smoother offset with some irregularities. Glissando is simulated in Fig. 4(b) by varying our tension parameter $T$ from 1 to 5.5 and then back. We find a frequency jump at 6.7 s ($T=3.8,F_0=390$ Hz) and a backward transition at 17 s ($T=3.3,F_0=350$ Hz). These differences between chest-falsetto and falsetto-chest transitions are a landmark of hysteresis (see Tokuda et al., 2007 for a detailed discussion of bifurcations leading to 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 passage of trained singers.

In addition to register transitions, occasionally subharmonics are observed, e.g., at 8.7 and 14.1 s. It has been discussed earlier (Berry et al., 1996; Tokuda et al., 2007) 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 Fig. 4 are similar. The study of hysteresis at phonation onset/offset requires a more detailed Hopf bifurcation analysis.

B. Bifurcation diagrams of phonation onset

Phonation onset and offset can be studied in the context of Hopf bifurcations (see, e.g., Lucero, 1998; Mergell et al., 2000). Hopf bifurcation theory describes the onset of self-sustained oscillations due to parameter variations. Smooth oscillation onset is associated with a supercritical Hopf bifurcation whereas hysteresis and amplitude jumps indicate a subcritical Hopf bifurcation (see, e.g., Guckenheimer and Holmes, 1983 for details).

In case of phonation onset, an increasing subglottal pressure indicates vocal fold oscillations at overcritical values. In the simplified two-mass model (Steinecke and Herzel, 1995), no hysteresis was observed. In contrast, excised larynx experiments revealed a clear pressure difference between onset and offset values of about 0.2 kPa (Berry et al., 1996).

Such a hysteretic phonation onset/offset is shown in our model simulations in Fig. 5. Figure 5(a) refers to the model without vocal tract resonators. There are amplitude jumps at 0.52 kPa (onset) and at 0.34 kPa (offset). Differences between increasing and decreasing pressures indicate a subcritical Hopf bifurcation. On the other side, Fig. 5(b) shows that sub- and supercritical bifurcations can occur in the four-mass model, to which sub- and supraglottal resonators are attached. As discussed in Titze (1988), the resonators reduce
the threshold pressure. The phonation onset around 0.15 kPa seems rather smooth, implying a supercritical Hopf bifurcation. The generated phonation on the middle branch, however, has a relatively small amplitude and becomes unstable around 0.165 kPa, where it jumps to more stable one with larger amplitude on the upper branch. This jump creates hysteresis in the model simulations.

C. Influence of vocal tract on chest-falsetto transition

We induced register transitions in our model by changing the tension parameter $T$ gradually and by measuring the fundamental frequency $F_0$, the amplitude of the opening area, and the number of the colliding cover masses. We observed a steady increase in $F_0$ and collision of all three cover masses at low $F_0$ and collision of only the top mass at high $F_0$. For simplicity, a binary classification is applied to draw the register transitions of Figs. 6–8 as follows: collision of three cover masses are termed chest, whereas collision of less masses are termed falsetto. If only the upper masses collide, open quotient (OQ), defined as $OQ = \text{duration of the open phase of the glottis/pitch period}$, became large as known from measurement in singers (Henrich et al., 2005). In our bifurcation diagrams, with the tension $T$ as the bifurcation parameter, we plotted the fundamental frequency $F_0$ on the $x$-axis instead of $T$ since this allows a direct comparison with glissando spectrograms.

Figures 6 and 7 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 of about 30–40 Hz but relatively small jumps of the amplitudes. Most notable is the dramatic shift in 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 in the transition point by 10–15 Hz (no graphs shown). In contrast, the supraglottal resonance had a profound effect: changing the default length of 17.5 cm to 75% or 125% induced major shifts in the

![Fig. 5. Phonation onset. No vocal tract is attached to the vocal fold model in (a), but in (b) vocal tract is attached. Tension parameter is set as $T=4$. Local maxima of the opening area of the lower mass $a_{1}=h_{1}$ are plotted for both increasing (crosses) and decreasing (circles) subglottal pressure. The small graphs inside of each diagram represent the volume flow $U$ (cm$^3$/ms) corresponding to each branch of the onset curve.](image1.png)

![Fig. 6. Register transition of the vocal fold model without vocal tract. Frequency domains for chest and falsetto registers are drawn in the upper graph, whereas the corresponding bifurcation diagrams are drawn in the lower graph. The curves were drawn by both increasing (dotted line with crosses) and decreasing (solid line with circles) tension parameter $T$. In the bifurcation diagram, local maxima of the opening area of the lower mass $a_{1}=h_{1}$ were plotted.](image2.png)

![Fig. 7. Register transition of the vocal fold model with vocal tract. The default lengths for sub- and supraglottis are $L_{\text{sub}}=24.7$ cm and $L_{\text{sup}}=17.5$ cm, respectively.](image3.png)
frequencies at which register transitions are observed (see Fig. 8). 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.

V. EXPERIMENT

Our numerical study has shown that the supraglottal resonance has a strong influence on the pitch of the chest-falsetto transition. In order to examine this effect, we have carried out an experimental study of glissando on vowel /i/. This vowel has been chosen, since /i/ is one of the vowels that provide the lowest formant frequency $F_1$, thereby the $F_0$-$F_1$ interaction can be easily observed in the singing experiment. Four subjects were asked to perform $F_0$ gliding (from low $F_0$ to high $F_0$ and then back to low $F_0$). Two recordings were obtained from each subject. The subjects were all untrained males who have no evidence on laryngeal pathology. Both speech signal and electroglottographic (EGG) signal were simultaneously recorded.

Figure 9 shows an example of the recording data. The first formant $F_1$ was estimated from the speech signal by the conventional technique based on linear prediction analysis (McCandless, 1974). The spectrogram shows that, as the fundamental frequency $F_0$ increases and crosses the first formant $F_1=250$ Hz, the frequency jump is induced at $t=4.6$ s. The same frequency jump is observed, when the fundamental frequency decreases and crosses the first formant at $t=6.5$ s. As indicated by the OQ computed from the EGG signal by the method of Henrich et al. (2005), these frequency jumps are accompanied by the register change. The regime of falsetto register, characterized by high OQ, is clearly distinguished from the regime of chest register by dashed lines in Fig. 9(b). The timing of the register change coincides with the $F_0$-$F_1$ crossings quite well. This implies that the register transition is induced by the source-filter interaction, which is known to become strong when $F_0$ and $F_1$ are close to each other (Story et al., 2000; Titze, 2008). Among the four subjects, coincidence of the register transition and the $F_0$-$F_1$ crossing has been observed in three subjects, where the other subject showed register transitions with a pitch much higher than the first formant. Our experimental study therefore implies that the supraglottal resonance has indeed a strong influence on the register transitions. This influence of course depends on the individual characteristics of the subject, where well-trained singers should know how to sustain the chest register by avoiding the voice instability. Hence it is reasonable that not all subjects in our experiment showed a clear influence of the resonator on the registers.

To study how the different vowels affect the register transition, we further collected statistical data from the three subjects, who showed the influence of the resonator. Each subject was asked to perform $F_0$ gliding on both vowels /a/ and /i/, where ten recordings were obtained for each vowel. Average and standard deviations of the fundamental frequency $F_0$, at which the register transition takes place when

![Figure 8](https://example.com/figure8.png)

**FIG. 8.** Dependence of the register transition on the supraglottal length. The right and left graphs show the case of short supraglottis ($L_{sup}=13.125$ cm) and long supraglottis ($L_{sup}=21.875$ cm), respectively.

![Figure 9](https://example.com/figure9.png)

**FIG. 9.** (Color online) (a) Spectrogram of a male subject on vowel /i/ with a gliding fundamental frequency ($F_0$). The first formant ($F_1$) estimated from the speech signal is indicated by a solid line. (b) OQ computed from the EGG signal of (a). A region of high OQ, corresponding to falsetto register, is separated from low OQ regions by dashed lines.
increasing $F_0$, were computed from the ten data sets, as summarized in Table I. Because of the high variability of the register transitions, the standard deviation was estimated to be relatively large. According to Welch’s t-test, the mean frequency difference between /a/ and /i/ was statistically significant for subject I with a level of 1%. For the other two subjects, the difference was not significant.

We remark that Titze et al. (2008) carried out the same experimental framework in the context of voice instability induced by the source-tract coupling. Their main focus was, however, on the frequency jumps and not much attention has been paid to the register change. They found many frequency jumps induced by the $F_0$-/$F_1$ crossing accompanied by hysteresis, in particular, for male subjects. Our observation essentially agrees with their study.

**VI. SUMMARY AND DISCUSSION**

From the nonlinear dynamics point of view, voice registers are distinct types of limit cycle oscillations. In this context, phonation onset refers to a Hopf bifurcation and register transitions are associated with bifurcations of limit cycles. In Tokuda et al. (2007), we characterized register transition in excised larynx experiments and simulations by two-dimensional bifurcation diagrams. In that paper, we analyzed a simple three-mass cover model.

Here we introduced a more realistic four-mass polygon model coupled to sub- and supraglottal resonators. This model was used to study bifurcations at phonation onset and the chest-falsetto transition.

In Mergell et al. (2000), a smooth phonation onset was quantified using the normal form of a supercritical Hopf bifurcation. This model explained high speed glottographic data in a reasonable way. In excised larynx experiments, however, amplitude jumps and hysteresis were reported at phonation onset (Berry et al., 1996). The present simulation without vocal tract exhibits a subcritical Hopf bifurcation, where the associated hysteresis is in good agreement with the excised larynx experiments. Another simulation with vocal tract showed that the coupling to the resonators lowers the phonation onset threshold. It remains to be tested experimentally under which circumstances the phonation onset can be regarded as super- or subcritical Hopf bifurcation.

It is well known that register transitions in untrained singers are accompanied by vocal breaks (see, e.g., Švec and Pešák, 1994). In our simulations, we find indeed sudden jumps of pitch and amplitudes while varying the tension parameter $T$ smoothly. In computer simulations, the associated phenomenon of hysteresis can be studied more easily than in glissando of singers. It turns out that our model exhibits clear pitch differences between chest-falsetto and falsetto-chest transitions.

In order to analyze the role of sub- and supraglottal resonators on register transition, we varied the length of the tube by $\pm 25\%$. We find that the length of the subglottal tube has only minor effects on the register transition. Contrarily, the supraglottal resonator influences the pitch of the chest-falsetto transition strongly. To examine this effect, a simple experiment has been carried out based on the glissando singing on vowels /i/ and /a/. We have found a strong correlation between the register transition and the source-filter interaction for most of the subjects. Significant difference in the register transition point between the sung vowels was also detected from one subject despite high variability of the register breaks. These results provide a good indication that the supraglottal resonator has indeed an influence on the register transitions.

We remark that the present experiment is only preliminary and further investigations with more recording trials and with more subjects having various singing backgrounds are indispensable. To further study the influence of the resonators, the following experiments may also be of great interest. (i) The vocal tract anatomy can affect the register shift, i.e., vocal tract length might be correlated with the pitch of register transitions. (ii) Singing into tube (Hatzikirou et al., 2006) should shift register transitions. We finally note that comparable frequency jumps are also frequent in a variety of animal vocalizations (Wilden et al., 1998; Tembrock, 1996; Riede and Zuberbühler, 2003).

**ACKNOWLEDGMENTS**

We thank Markus Hess and Frank Müller for the opportunity of acoustic and EGG recordings. We are grateful to Tobias Riede for stimulating discussions. This work was supported by SCOPE (Grant No. 071705001) of Ministry of Internal Affairs and Communications (MIC), Japan.

**APPENDIX A: DETAILED FOUR-MASS MODEL**

Figure 1 shows a schematic illustration of the four-mass polygon model. Following the body-covered theory (Hirano, 1974; Hirano and Kakita, 1985) 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: $i=2$, and upper: $i=3$). Following the simplifications of (1) neglecting the cubic nonlinearities of the oscillators, (2) neglecting the additional pressure drop at inlet and considering the Bernoulli flow only below the narrowest part of the glottis (Steinecke and Herzl, 1995), and (3) assuming symmetry between the left and the right vocal folds, the model equations read as

$$m_1\ddot{y}_1 + r_1(\ddot{y}_1 - \ddot{y}_b) + k_1(y_1 - y_b) + \Theta(-h_1)c_1\left(\frac{h_1}{2}\right) + k_{1,2}(y_1 - y_2) = F_1,$$

where $r_i, y_i, h_i, k_i, c_i, \Theta$ are positive and $\Theta$ is a memory function.

For simplicity, the coupling between the vocal folds is modeled as a torsion spring. Following the simplifications of (1) neglecting the cubic nonlinearities of the oscillators, (2) neglecting the additional pressure drop at inlet and considering the Bernoulli flow only below the narrowest part of the glottis (Steinecke and Herzl, 1995), and (3) assuming symmetry between the left and the right vocal folds, the model equations read as
\[ m_3 \ddot{y}_2 + r_2 (\dot{y}_2 - \dot{y}_b) + k_2 (y_2 - y_b) + \Theta(-h_2)c_2 \left( \frac{h_2}{2} \right) \]
\[ + k_{1,2} (y_2 - y_1) + k_{2,3} (y_2 - y_3) = F_2, \]  
(A2)

\[ m_3 \ddot{y}_3 + r_3 (\dot{y}_3 - \dot{y}_b) + k_3 (y_3 - y_b) + \Theta(-h_3)c_3 \left( \frac{h_3}{2} \right) \]
\[ + k_{2,3} (y_3 - y_2) = F_3, \]  
(A3)

\[ m_b \ddot{y}_b + r_b \dot{y}_b + k_b y_b + r_1 (y_b - \dot{y}_1) + k_1 (y_b - y_1) \]
\[ + r_2 (\dot{y}_b - \dot{y}_2) + k_2 (y_b - y_2) + r_3 (\dot{y}_b - y_3) + k_3 (y_b - y_3) \]
\[ = 0. \]  
(A4)

The dynamical variables \( y_i \) represent displacements of the masses \( m_i \), where the corresponding glottal opening is given by \( h_i = h_0 + 2y_i \) \( (h_0 \) is prephonatory length; \( i = 1,2,3 \)). The constant parameters \( r_i, k_i \), and \( c_i \) represent damping, stiffness, and collision stiffness of the masses \( m_i \), respectively, whereas \( k_{ij} \) represents coupling strength between two masses \( m_i \) and \( m_j \). The stiffness is determined as \( r_i = 2\zeta_i \Omega m_i \) using the damping ratio \( \zeta_i \). 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 height \( h(x,t) \) is a piecewise linear function, composed of \( h_{i,0} \) \( (x_0 \leq x \leq x_1) \), \( h_{i,1} \) \( (x_1 \leq x \leq x_2) \), \( h_{i,2} \) \( (x_2 \leq x \leq x_3) \), and \( h_{i,4} \) \( (x_3 \leq x \leq x_4) \), which are determined as

\[ h_{i,i-1}(x,t) = \frac{h_i(t) - h_{i-1}(t)}{x_i - x_{i-1}} (x - x_{i-1}) + h_{i-1}(t), \]  
(A5)

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

\[ P_s = P(x,t) + \frac{\rho}{2} \left( \frac{U}{h(x,t)} \right)^2 = P_0 + \frac{\rho}{2} \left( \frac{U}{h_{\text{min}}} \right)^2, \]  
(A6)

where \( \rho \) represents the air density \( (\rho = 1.13 \text{ kg/m}^3) \), \( P_s \) is the subglottal pressure, \( P_0 \) is the supraglottal pressure, and \( l \) is the length of the glottis. In the case that no resonators are attached to the vocal folds, the subglottal pressure \( P_s \) is considered to be constant \( (P_s = 0.8 \text{ kPa}) \), whereas the supraglottal pressure is assumed zero \( (P_0 = 0 \text{ kPa}) \). This gives a simple formula for the glottal volume flow velocity as

\[ U = \sqrt{2P_s/\rho h_{\text{min}}} \Theta(h_{\text{min}}). \]  

The aerodynamic forces on the plates are induced by the pressure \( P(x,t) \) along the flow channel. As \( m_1, m_2, \) and \( m_3 \) support the plates, an aerodynamic force on point \( i \) \( (i = 1,2,3) \) is found to be

\[ F_i(t) = \int_{x_{i-1}}^{x_i} \frac{U}{l} \frac{x - x_{i-1}}{x_i - x_{i-1}} P(x,t) dx \]
\[ + \int_{x_{i-1}}^{x_{i+1}} \frac{x - x_i}{x - x_{i-1}} P(x,t) dx. \]  
(A7)

As shown by Lous et al. (1998), this integral can be solved analytically for the pressure distribution \( P(x,t) \) described by Eq. (A6).

Parameter values used as the default situation of the present study are summarized in Table II. These values have been carefully selected in accordance with the previous studies (Ishizaka and Flanagan, 1972; Story and Titze, 1995; Lous et al., 1998; Titze and Story, 2002). The observed phenomena were in general robust and did not show a strong dependence on the selected parameter values. To simulate the coexistence of the chest and falsetto registers, the damping ratio of the lower mass was set to be relatively small compared with the other masses. The small damping ratio activates the lower mass so that it leads to a large movement of all the cover masses to produce a chest-like register, which can easily coexist with a falsetto-like register.

The tension parameter \( T \) is also introduced to control the frequency of the four masses as

\[ m_i' = m_i/T, \]  
(A8)

\[ k_i' = k_i T \quad (i = 1,2,3,b). \]  
(A9)

The initial values for all simulations were set as \( x_1 = 0.02 \text{ cm}, x_2 = 0.015 \text{ cm}, x_3 = 0.01 \text{ cm}, x_4 = 0 \text{ cm}, \) and \( x_4 = x_2 = x_3 = 0 \text{ cm/s}. \) To integrate the four-mass model equations (A1)–(A4), Euler’s method was applied with an integration step of \( \Delta t = 11.4/8 \mu s \). The model equations were simulated also by using the MATLAB ODE solver.
(ODE45) and it was confirmed that essentially the same results can be obtained.

To draw the bifurcation diagrams of Figs. 5–7, 20 local maxima of the opening area of the lower mass \( a_1 = \ell h_1 \) were plotted after discarding the transients. For the next parameter values, the final state of the preceding simulation was used as the initial condition.

The spectrogram of Fig. 4(b) was computed using the minimum glottal area \( a_{\text{min}} = \ell h_{\text{min}} \) with the following parameters. Sampling rate of 44 kHz, window length of 8192 sample points, overlap of 496 sample points, and Hanning window.

**APPENDIX B: WAVE-REFLECTION MODEL FOR SUB- AND SUPRAGLOTTIS**

Sub- and supraglottal resonances were described by using the wave-reflection model (Kelly and Lochbaum, 1962; Liljencrans, 1985; Story, 1995; Titze, 2006), 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 of 3 cm² and length of 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. (2007). The area function is composed of 62 cylindrical sections. For both sub- and supraglottal systems, the section length \( \Delta z \) was set to 17.5/44 cm. This determines the sampling time interval as \( \Delta t = \Delta z/c = 11.4 \mu s \), where \( c = 350 \) m/s stands for the sound velocity. The corresponding sampling frequency is 88 kHz.

Attenuation factor for the resonators was approximated as \( a_k = 1 - 0.007(\pi/\lambda_k)^{1/2} \Delta z \) \( (A_k) \) is kth cylinder area). Radiation resistance and radiation inertia at the lip were

\[
R_s = \frac{128 \rho c}{9 \pi^2 A_L}, \quad I_s = \frac{8}{3 \pi^{3/2} \sqrt{A_L}}, \tag{B1}
\]

respectively, where the lip area \( A_L \) was set to be equal to the last section of the supraglottis.

To couple the sub- and supraglottal resonators to the vocal fold model, an interactive source-filter coupling was realized according to Titze (2006, 2008). In this formula, the glottal flow is given by

\[
U = \frac{a_k}{k} \left\{ - \frac{a_k}{A} \pm \left[ \left( \frac{a_k}{A} \right)^2 + \frac{2k_s}{\rho c^2} (P_s + P_{s}^{+} - 2P_{c}^{+}) \right]^{1/2} \right\}, \tag{B2}
\]

where \( \lambda' = A_A + A_s + A_t \) with \( A_A \) and \( A_t \) being the subglottal and supraglottal entry areas, respectively. \( k \) is a transglottal pressure coefficient set as 1.0. \( P_s \) stands for the lung pressure, whereas \( P_{s}^{+} \) and \( P_{c}^{+} \) 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 \( P_l = 1.2 \) kPa. Since subglottal pressure \( P_s \) is time dependent in this formula, the pressure value was averaged over a long-term simulation to plot \( P_s \) in Fig. 5(b).


