Interspeech 2018
2-6 September 2018, Hyderabad

Interaction mechanisms between glottal source and vocal tract in pitch glides

Tiina Murtola[1], *Jarmo Malinen*[2]

[1]Department of Signal Processing and Acoustics, Aalto University, Finland
[2]Department of Mathematics and Systems Analysis, Aalto University, Finland
tiina.murtola@aalto.fi, jarmo.malinen@aalto.fi

Abstract

A computational model for vowel production has been used to simulate rising pitch glides in the time domain. Such glides reveal multi-faceted nonlinear system behaviour when the fundamental frequency f_o is near the first vocal tract resonance f_{R1}. There are multiple physical mechanisms for how the acoustic field in the vocal tract can interact with vocal fold dynamics causing this behaviour. The model used in this work includes the direct impact of the acoustic pressure on the transversal plane of the vocal folds and an acoustic perturbation component to the glottal flow. Simulations indicate that both of these mechanisms, when applied separately, cause similar perturbations in phonation parameters when f_o crosses f_{R1}. Enabling both mechanisms simultaneously tends to make the separately emerging features more prominent. In simulated glottal flow waveforms, the tendency towards a formant ripple increases when acoustic feedback to glottal flow is enabled, whereas the phenomenon occurs more rarely as a result of the direct acoustic pressure to vocal folds. In all cases, the formant ripple is more pronounced for frequencies below f_{R1}.

Index Terms: vowel production model, modal locking, source–filter interaction

1. Introduction

Physics-based models are used to investigate speech production phenomena where direct observation and measurements are infeasible or highly invasive. Although advances in high-performance computing have made it possible to model the speech production system as a whole (e.g., [1, 2]), these models are still computationally too heavy to be practical when a high number of iterations is required, for example, to find control parameters that produce desired output. Hence, low-order models, such as [3, 4, 5] for vowel production, remain in use and are subject to further development.

One of the key choices for any vowel production model is the implementation of interaction mechanisms between the glottal source (i.e., vocal fold oscillations and/or glottal flow) and the vocal tract (VT) and subglottal tract (SGT) loads. The classic source–filter theory [6, 7] assumes the source affects the loads but there is no feedback from the loads to the source. There are, however, multiple physical mechanisms for how the acoustic loads can affect the glottis. Modelling phenomena such as pitch glides, where the fundamental frequency f_o crosses the first VT resonance f_{R1}, requires that this feedback is accounted for. Such glides are valuable for benchmarking vowel production models as comparable measurement data from humans is obtainable (see, e.g., [8, 9]).

Interaction mechanisms for pitch glides include feedback from the load acoustics to the glottal flow [10, 11], including the acoustic pressures in the force driving vocal fold oscillations [12, 13, 14], and allowing the VT reactance to affect vocal fold vibrations [10]. The two main mechanisms, acoustics-to-flow and acoustics-to-tissue feedback, have not been implemented in the same model, however, and hence comparisons remain scarce. The aim of this work is to investigate the effect of these two mechanisms in rising pitch glides. A computational time-domain model is used to simulate glides over f_o range [145 Hz, 310 Hz] which includes f_{R1} of the used VT geometry.

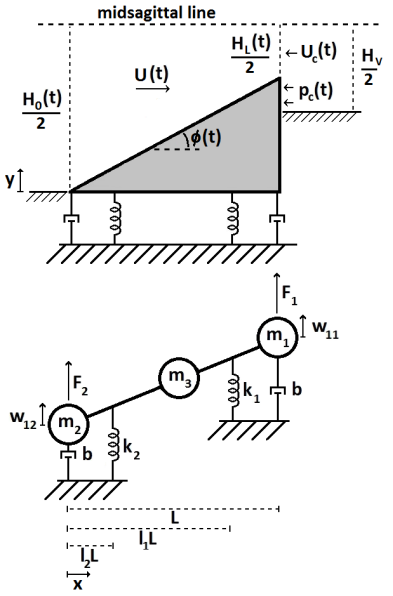

Figure 1: *Model vocal fold geometry and equivalent mass-spring-damper system.*

2. Model

The model proposed has been used to investigate pitch glides previously in [13, 14], and it is outlined briefly below. Two modifications have been made to the model: (i) driving pressure difference has been added to the force acting on the vocal folds when the glottis is closed, and (ii) a perturbation flow caused by the load acoustics has been added to the total glottal flow. The purpose of this article is to discuss the consequences of the latter modification in pitch glides.

The vocal folds are represented by wedge-shaped elements of thickness L and length h whose dynamics is modelled using mass-spring-damper systems (Figure 1). For this work, symmetry of left and right vocal folds is assumed. The equations of motion for the vocal folds are

$$\begin{cases} M\ddot{W}_1(t) + B\dot{W}_1(t) + KW_1(t) = -F(t), \\ M\ddot{W}_2(t) + B\dot{W}_2(t) + KW_2(t) = F(t), \end{cases} \quad (1)$$

10.21437/Interspeech.2018-1827

where $W_j = [w_{j1}\ w_{j2}]^T$ are the displacements of m_1 and m_2 for the j^{th} vocal fold ($j = 1, 2$), and

$$M = \begin{bmatrix} m_1 + \frac{m_3}{4} & \frac{m_3}{4} \\ \frac{m_3}{4} & m_2 + \frac{m_3}{4} \end{bmatrix}, \quad B = \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix},$$
$$\text{and} \quad K = \begin{bmatrix} l_1^2 k_1 + l_2^2 k_2 & l_1 l_2 (k_1 + k_2) \\ l_1 l_2 (k_1 + k_2) & l_2^2 k_1 + l_1^2 k_2 \end{bmatrix}. \tag{2}$$

The parameters in (2) can be found in Figure 1, and the assumption $l_1 + l_2 = 1$ has been used to obtain K. The load force pair $F(t) = [F_1(t)\ F_2(t)]^T$ depends on the open state of the vocal folds, i.e., whether the glottal gap at $x = L$, $H_L(t)$, (see Figure 1) is positive

$$F(t) = \begin{cases} F_A(t) + F_{pc}(t), & \text{when } H_L(t) > 0, \\ F_H(t) + F_{pc}(t), & \text{when } H_L(t) \leq 0. \end{cases} \tag{3}$$

The aerodynamic force pair $F_A(t) = [F_{A,1}(t)\ F_{A,2}(t)]^T$ arises from the airflow $U(t)$ between the vocal folds (see (6)–(8)) when the glottis is open (i.e., $H_L(t) > 0$):

$$F_{A,1} = \frac{hL}{2\cos^2\phi} \left(-\frac{\rho U^2}{h^2 H_L (H_0 - H_L)} + \frac{\rho U^2}{h^2 (H_L - H_0)^2} \ln\left(\frac{H_0}{H_L}\right) + p_s - p_r \right) \quad \text{and}$$
$$F_{A,2} = \frac{hL}{2\cos^2\phi} \left(\frac{\rho U^2 \left(H_0 \sin^2\phi + H_L \cos^2\phi\right)}{h^2 H_L H_0 (H_0 - H_L)} - \frac{\rho U^2}{h^2 (H_L - H_0)^2} \ln\left(\frac{H_0}{H_L}\right) + \cos(2\phi)\,(p_s - p_r) \right). \tag{4}$$

Here, ρ is the density of air, and the pressure difference $p_s - p_r$ is between an ideal driving pressure $p_s = p_s(t)$ and the local ambient pressure at which the vocal folds are at equilibrium p_r. Time dependencies of both glottal gaps $H_0 = H_0(t)$ and $H_L = H_L(t)$, as well as $U = U(t)$ and angle of the vocal fold surface $\phi = \phi(t)$ have been suppressed in (4) for readability.

When the vocal folds are closed (i.e., $H_L(t) \leq 0$), there is no airflow between them, and thus $F_A(t)$ is not enabled. Instead, the driving pressure difference and the collision of the vocal folds produce a force pair

$$F_H(t) = \begin{bmatrix} k_H |H_L|^{3/2} + \frac{hL}{2\cos^2\phi}\,(p_s - p_r) \\ \frac{hL\cos(2\phi)}{2\cos^2\phi}\,(p_s - p_r) \end{bmatrix}, \tag{5}$$

where vocal fold collision is modelled with a nonlinear spring force with parameter k_H.

The force pair $F_{pc}(t)$ in (3) accounts for the feedback force from VT and SGT acoustics to the vocal fold tissues. This force is enabled regardless of glottal open state, and it is defined in (12) and discussed further in Section 3.1.

The total glottal flow $U(t)$ is considered to consist of two components, an incompressible main component $U_m(t)$ and a perturbation caused by the acoustic loads $U_c(t)$; that is

$$U(t) = U_m(t) + Q_{uc} U_c(t), \tag{6}$$

where the control parameter Q_{uc} enables scaling of the feedback from the acoustics to the glottal flow. The perturbation term is a new addition to the model used in [13, 14], and it is discussed further in Section 3.2.

The main flow component is described by

$$\dot{U}_m(t) = \frac{1}{I_l} \left(p_s(t) - R_g(t) U_m(t) \right), \tag{7}$$

where I_l is the inertance of the loads and $R_g(t)$ controls the pressure loss within the glottis

$$R_g(t) = \frac{12\mu L_v}{H_L(t)^3} + \frac{k_t \rho U_m(t)}{2h^2 H_L(t)^2}. \tag{8}$$

The first term represents viscous losses and it is regulated via the thickness parameter L_v while μ is the kinetic viscosity of air. The second term accounts for entrance/exit effects, and it is controlled via the coefficient k_t.

The acoustic loads of the VT and SGT are represented using generalised Webster's equation

$$\frac{1}{c^2 \Sigma(s)^2} \frac{\partial^2 \psi}{\partial t^2} + \frac{2\pi\alpha W(s)}{A(s)} \frac{\partial \psi}{\partial t} - \frac{1}{A(s)} \frac{\partial}{\partial s} \left(A(s) \frac{\partial \psi}{\partial s} \right) = 0, \tag{9}$$

where c denotes the speed of sound, α regulates the energy dissipation through air/tissue interface, and the solution $\psi = \psi(s, t)$ is the velocity potential of the acoustic field. The VT/SGT geometry is represented by the area function $A(s)$, the stretching factor $W(s)$, and the sound speed correction factor $\Sigma(s)$, all at distance $s \in [0, L_l]$ from the glottis, where L_l is the length of the VT/SGT.

At the glottal end, the boundary condition

$$\frac{\partial \psi}{\partial s}(0, t) = -\frac{\tilde{U}_m(t)}{A(0)} \tag{10}$$

is used for (9). Here, the DC component has been removed from the glottal flow before coupling using averaging over two pulses, i.e., $\tilde{U}_m(t) = U_m(t) - \frac{f_o}{2} \int_{t-2/f_o}^{t} U_m(\tau)\, d\tau$. A resistive boundary condition is used at the lips and the lungs

$$\frac{\partial \psi}{\partial t}(L_l, t) + \theta c \frac{\partial \psi}{\partial s}(L_l, t) = 0, \tag{11}$$

where θ is the normalised acoustic resistance.

3. Interaction mechanisms

3.1. Acoustics-to-tissue feedback

The feedback from load acoustics to vocal fold tissues is represented by the force $F_{pc}(t)$ in (3), and it is given by

$$F_{pc}(t) = Q_{pc} \rho \frac{\partial \psi^V}{\partial t}(0, t) \begin{bmatrix} -A_{pc} \\ A_{pc} \end{bmatrix} + Q_{sg} \rho \frac{\partial \psi^S}{\partial t}(0, t) \begin{bmatrix} A_{sg} \\ -A_{sg} \end{bmatrix}, \tag{12}$$

where the solutions ψ^V and ψ^S for (9) correspond to the VT and SGT, respectively, and $p_c(t) = \rho \frac{\partial \psi^V}{\partial t}(0, t)$ in Figure 1. Unlike in [13, 14], the acoustic VT and SGT pressure forces are controlled separately with parameters Q_{pc} and Q_{sg} in order to compare interaction mechanisms for VT while maintaining SGT feedback constant. However, following [13, 14] $A_{pc} = A_{sg} = \frac{h}{8L} \left(H_V - H_L(t)\right) \left(2H_0 - H_V - H_L(t)\right)$ which is the moment adjusted area on which the pressures act.

3.2. Acoustics-to-flow feedback

The feedback from the VT acoustics to the glottal flow is represented by the perturbation flow $U_c(t)$ and its control parameter Q_{uc} in (6). When the glottis is open,

$$U_c(t) = \frac{H_L(t)}{H_L(t) + H_V} \frac{\rho}{I_g(t) T} \int_{t-T}^{t} \psi^V(0, \tau) d\tau, \tag{13}$$

where the inertance of the glottis is given by

$$I_g(t) = \frac{\rho L}{h\left(H_0(t) - H_L(t)\right)} \ln\left(\frac{H_0(t)}{H_L(t)}\right). \qquad (14)$$

Equations (13)–(14) have been obtained from a low-frequency approximation of (9) within the glottis (with $\alpha = 0$ and $\Sigma(s) = 1$). The approximation is accurate for the frequency range of interest since speech wavelengths are several orders of magnitude larger than the vocal fold thickness L.

The scaling coefficient $H_L(t)/\left(H_L(t) + H_V\right)$ in (13) can be derived using an argument where two acoustic waveguides of different constant characteristic impedances are coupled, and an extra acoustic resistance term is included in addition to the reflection and transmission of acoustic power at the interface. The acoustic pressure loss in the extra resistance is inversely proportional to the square of glottal opening area as in [15, Eq. (1)].

Averaging of ψ^V in (13) is carried out over 350 samples in simulations ($T \approx 3.5$ ms) to remove high-frequency components where (13) becomes physically unrealistic. This averaging also serves to reduce parasitic oscillations that occur at certain parameter values and which would result in counter-factual simulation outcomes.

The contribution from the SGT acoustics to the glottal flow could be added to the model in a similar manner. However, the lowest SGT resonance is more than an octave above the frequency interval used in the glides, and hence the effect of this addition is negligible in this study.

4. Numerical simulations

Details of the numerical implementation of the model can be found in [16, pp. 12–14 and 24–27]). The values of constants used in simulations are listed in Table 1. The VT geometry required by (9) is extracted from MRI data of a female speaker producing a prolonged [i] at $f_o = 210$ Hz (see [14, Sec. 3.1–3.2] and [17, 18]), and an exponential horn is used to represent the SGT. Key parameters describing the VT and SGT geometries can be found in Table 2.

Before glides, long steady simulations are carried out with all parameters held constant. The word "steady" is understood here in the sense of quasi-stability, i.e., slowly changing amplitudes and pulse shapes. In order to achieve full glottal closure during these steady simulations under all the presented feedback configurations, the total inertance in (6) is increased to $I_l = 4790$ kg/m^4 (24% increase from the VT and SGT air column inertances).

The glides are achieved by dynamic control of two parameters as in [13, 14]:

$$K(t) = 2.2^{2t/T} K^0, \quad p_s(t) = 2.2^{t/T} p_s^0, \qquad (15)$$

where K^0 is the initial vocal fold stiffness matrix, which is tuned to achieve $f_o \approx 145$ Hz in steady simulations, and p_s^0 is the initial subglottal pressure (Table 1). The control strategy (15) produces glides approximately over the range [145 Hz, 310 Hz]. The total duration of the glides is $T = 3$ s.

Table 1: *Physical and physiological constants.*

Parameter	**Value**
Speed of sound in air c	343 m/s
Density of air ρ	1.2 kg/m^3
Kinematic viscosity of air μ	18.27 μN s/m^2
VT/SGT boundary loss coefficient α	$7.6 \cdot 10^{-7}$ s/m
Spring constant in contact k_H (from [19])	730 N/m
Equilibrium glottal gap at $x = L$	0.5 mm
Equilibrium glottal gap at $x = 0$	11 mm
Control area height above glottis H_V	2 mm
Vocal fold length h (from [20])	10 mm
Vocal fold thickness L (from [19])	6.8 mm
Superior vocal fold spring location l_1	0.85
Inferior vocal fold spring location l_2	0.15
Viscous thickness L_v	1.5 mm
Glottal damping b	0.009
Glottal entrance/exit coefficient k_t	0.4
Subglottal pressure (over ambient) p_s^0	700 Pa
Reference pressure (over ambient) p_r	350 Pa
Subglottal counter pressure scale Q_{sg}	0.1

Table 2: *Physical and physiological parameters dependent on the VT and SGT geometries.*

Parameter	**VT** [i]	**SGT**
Acoustic resistance θ	0.014	1
Termination area $A(L_l)$	66 mm^2	1000 mm^2
Inertance	2820 kg/m^4	1040 kg/m^4
Tract length L_l	136 mm	350 mm
1$^{\text{st}}$ resonance f_{R1}	199 Hz	500 Hz
2$^{\text{nd}}$ resonance f_{R2}	2798 Hz	1000 Hz

Table 3: *Steady phonation in different feedback cases*

Case	Q_{pc}	Q_{uc}	f_o	**ClQ**	**OQ**
I	0	0	144.9042	0.1382	0.7450
II	0.1	0	144.9100	0.1459	0.7589
III	0	0.04	144.5146	0.1600	0.7789
IV	0.1	0.04	144.6296	0.1728	0.8023

5. Results

Table 3 shows phonation parameters in steady simulations in four cases I–IV with varying feedback configurations. Case I corresponds to no feedback, case II to acoustics-to-tissue feedback only, case III to acoustics-to-flow feedback only, and case IV to enabling both the interaction mechanisms simultaneously. Stable phonation is more difficult to maintain when Q_{uc} is increased compared to an increase in Q_{pc}, and hence lower Q_{uc} values are generally used.

The difference in the output parameter values in Table 3 are small since f_o is far from load resonances. It can be seen, however, that when compared with case I, f_o increases slightly in case II but decreases in cases III and IV. Enabling any feedback (cases II–IV) tends to increase open quotient (OQ), i.e., the proportion of the glottal cycle when glottis is open, and closing quotient (ClQ), i.e., the proportion of the glottal cycle when the glottis is closing. Increases in both of these values indicate more breathy mode of phonation.

The feedback configuration has a notable impact on simulated glottal output in pitch glides. Figure 2 shows the trajectories of f_o, OQ, and ClQ during the glides for all cases I–IV described above. These parameters have been computed pulse-by-pulse from $U_m(t)$ which has single-peaked pulses that are easy to identify and parametrise automatically. Figure 2 also shows the envelopes of $U(t)$ and $H_L(t)$.

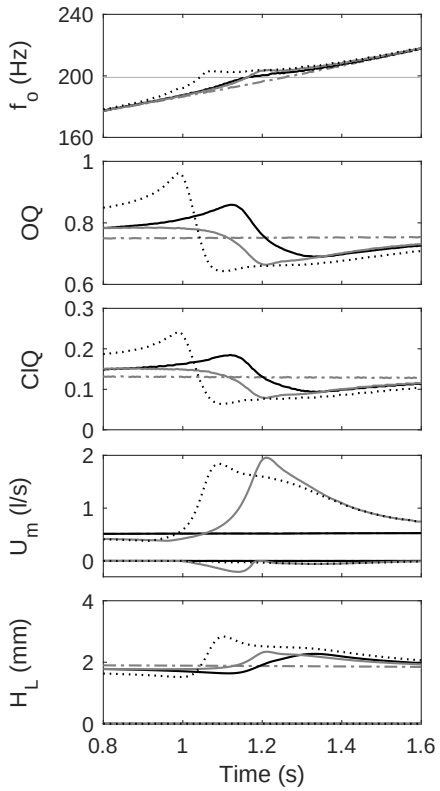

Figure 2: *Fundamental frequency f_o, open quotient OQ, closing quotient ClQ, and envelopes of total glottal flow U and glottal gap H_L in pitch glides under different feedback configurations: case I (dashed gray), case II (solid black), case III (solid gray), and case IV (dotted black). Thin horizontal line in top panel is f_{R1}.*

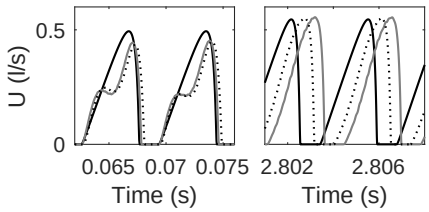

Figure 3: *Total glottal flow pulses at the beginning (left) and end (right) of pitch glides. Different feedback configurations are shown: case II (solid black), case III (solid gray), and case IV (dotted black).*

Perturbations in the output parameters occur only in the vicinity of f_{R1} and only when at least one kind of feedback from the load to the source is enabled. Both interaction mechanisms produce a f_o locking pattern (c.f., [14]) but f_o locks to a frequency slightly above f_{R1} in cases III and IV, whereas the locking frequency is closed to f_{R1} in case II. Both feedback mechanisms produce changes in pulse shapes during the locking. The clearest difference is that cases II and IV produce a peak in OQ and ClQ values at the beginning of the locking. This peak is not present in the other cases although case III shows an overall increase in OQ and ClQ values before the perturbations begin.

When both interaction mechanisms are enabled, they appear to reinforce each other: the perturbations in f_o, pulse shapes, and vocal fold oscillation amplitude start earlier in the glide, locking lasts longer, and changes in OQ and ClQ values during the locking become larger. Only the envelope of $U(t)$ shows signs of the two mechanisms counteracting each other.

Cases III and IV produce glottal flow pulses with a formant ripple (also known as an interaction ripple in [21]) when f_o is below f_{R1} in the glides (Figure 3). This ripple disappears gradually during and after the locking. The formant ripple was not observed in cases I and II in these glides, but it has been reported to occur with some parameter value combinations under acoustic-to-tissue feedback when f_o is below f_{R1} [22, Fig. 3.4].

6. Discussion

Simulations of rising pitch glides indicate that both acoustics-to-flow and acoustics-to-tissue feedback causes perturbations at the glottal source when f_o crosses f_{R1}. When both interaction mechanisms are enabled simultaneously, they appear to reinforce one another leading to higher prominence of the locking pattern and the associated waveform changes. The similarity of the perturbation patterns under all investigated feedback configurations suggests that the redistribution of stored energy between the main subsystems (i.e., underlying flow, VT/SGT acoustics, and vocal fold dynamics) is largely independent of the energy transfer mechanisms near f_o–f_{R1} crossings.

The introduction of the acoustics-to-flow feedback loop makes the model more prone to parasitic oscillations, i.e., quasi-periodic changes in waveform amplitudes and shapes which are not supported by evidence from human speakers. The simplifications that make the model computationally light enough to be feasibly parametrisable for, e.g., pitch glides also leave out some phenomena which are expected to damp these oscillations. For example, viscous losses in the acoustic subsystems would naturally attenuate high frequencies.

The tendency towards parasitic oscillations is controlled in the present study by (i) time averaging the acoustic velocity potential in (13) before it is used to compute the perturbation flow $U_c(t)$, (ii) using the main component of the glottal flow $U_m(t)$ as the flow driving the acoustic resonators in (10) instead of the total flow $U(t)$, (iii) using a resistive mouth load in (11) instead of a reactive one, since the latter has been observed to also cause instability in the model output [14], and (iv) carefully choosing the parameter values. The last step is essential in maintaining phonation throughout the glide. Without it the system may not recover from passing bifurcation points in the glides, and its behaviour becomes chaotic instead.

7. Conclusions

Simulations indicate that feedback from VT acoustics to both glottal flow and vocal fold tissues have only a small effect on phonation parameters when f_o is far from a VT resonance, although feedback may cause a formant ripple in the glottal flow waveform. However, the effect on vocal fold dynamics is significant when f_o crosses a VT resonance, and both interaction mechanisms produce similar perturbation patterns in these cases.

8. Acknowledgements

This study was funded by the Academy of Finland (projects no. 284671 and 312490).

9. References

[1] J. Horáček, V. Uruba, V. Radolf, J. Veselỳ, and V. Bula, "Airflow visualization in a model of human glottis near the self-oscillating vocal folds model," *Applied and Computational Mechanics*, vol. 5, pp. 21–28, 2011.

[2] N. C. Degirmenci, J. Jansson, J. Hoffman, M. Arnela, P. Snchez-Martn, O. Guasch, and S. Ternstrm, "A unified numerical simulation of vowel production that comprises phonation and the emitted sound," in *Interspeech 2017*, Stockholm, Sweden, August, 20–24 2017, pp. 3492–3496.

[3] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two mass model of the vocal cords," *Bell System Technical Journal*, vol. 51, pp. 1233–1268, 1972.

[4] B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," *The Journal of the Acoustical Society of America*, vol. 97, no. 2, pp. 1249–1260, 1995.

[5] I. R. Titze, "The physics of small-amplitude oscillation of the vocal folds," *The Journal of the Acoustical Society of America*, vol. 83, no. 4, pp. 1536–1552, 1988.

[6] T. Chiba and M. Kajiyama, *The Vowel, Its Nature and Structure*. Tokyo: Phonetic Society of Japan, 1941.

[7] G. Fant, *Acoustic Theory of Speech Production*. The Hague: Mouton, 1960.

[8] M. Zañartu, D. D. Mehta, J. C. Ho, G. R. Wodicka, and R. E. Hillman, "Observation and analysis of in vivo vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study," *The Journal of the Acoustical Society of America*, vol. 129, no. 1, pp. 326–339, 2011.

[9] L. Wade, N. Hanna, J. Smith, and J. Wolfe, "The role of vocal tract and subglottal resonances in producing vocal instabilities," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 1546–1559, 2017.

[10] I. R. Titze, "Nonlinear source-filter coupling in phonation: Theory," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 2733–2749, 2008.

[11] I. T. Tokuda, M. Zemke, M. Kob, and H. Herzel, "Biomechanical modeling of register transitions and the role of vocal tract resonators," *The Journal of the Acoustical Society of America*, vol. 127, no. 3, pp. 1528–1536, 2010.

[12] J. C. Lucero, K. G. Lourenço, N. Hermant, A. Van Hirtum, and X. Pelorson, "Effect of source-tract acoustical coupling on the oscillation onset of the vocal folds," *The Journal of the Acoustical Society of America*, vol. 132, no. 1, pp. 403–411, 2012.

[13] T. Murtola and J. Malinen, "Waveform patterns in pitch glides near a vocal tract resonance," in *Interspeech 2017*, Stockholm, Sweden, August, 20–24 2017, pp. 3487–3491.

[14] T. Murtola, A. Aalto, J. Malinen, D. Aalto, and M. Vainio, "Modal locking between vocal fold oscillations and vocal tract acoustics," *Acta Acoustica united with Acoustica*, vol. 104, no. 2, pp. 323–337, 2018.

[15] J. C. Ho, M. Zañartu, and G. R. Wodicka, "An anatomically based, time-domain acoustic model of the subglottal system for speech production," *The Journal of the Acoustical Society of America*, vol. 129, no. 3, pp. 1531–1547, 2011.

[16] A. Aalto, "A low-order glottis model with nonturbulent flow and mechanically coupled acoustic load," Master's thesis, Helsinki University of Technology, Department of Mathematics and Systems Analysis, 2009. [Online]. Available: http://speech.math.aalto.fi/pdf/diplomityo_atte_aalto.pdf

[17] D. Aalto, O. Aaltonen, R.-P. Happonen, P. Jääsaari, A. Kivelä, J. Kuortti, J.-M. Luukinen, J. Malinen, T. Murtola, R. Parkkola, J. Saunavaara, T. Soukka, and M. Vainio, "Large scale data acquisition of simultaneous MRI and speech," *Applied Acoustics*, vol. 83, pp. 64–75, 2014.

[18] A. Ojalammi and J. Malinen, "Automated segmentation of upper airways from MRI: Vocal tract geometry extraction," in *BIOIMAGING 2017*, Porto, Portugal, 2017, pp. 77–84.

[19] J. Horáček, P. Šidlof, and J. G. Švec, "Numerical simulation of self-oscillations of human vocal folds with Hertz model of impact forces," *Journal of Fluids and Structures*, vol. 20, no. 6, pp. 853–869, 2005.

[20] I. R. Titze, "Physiologic and acoustic differences between male and female voices," *The Journal of the Acoustical Society of America*, vol. 85, no. 4, pp. 1699–1707, 1989.

[21] M. Bvegrd and G. Fant, "Notes on glottal source interaction ripple," *STL-QPSR*, vol. 35, no. 4, pp. 63–78, 1994.

[22] T. Murtola, "Modelling vowel production," Licentiate thesis, Aalto University School of Science, Department of Mathematics and Systems Analysis 2014.