INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

Waveform patterns in pitch glides near a vocal tract resonance

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 time-domain model of vowel production is used to simulate fundamental frequency glides over the first vocal tract resonance. A vocal tract geometry extracted from MRI data of a female speaker pronouncing [i] is used. The model contains direct feedback from the acoustic loads to vocal fold tissues and the inertial effect of the full air column on the glottal flow. The simulations reveal that a perturbation pattern in the fundamental frequency, namely, a jump and locking to the vocal tract resonance, is accompanied by a specific pattern of glottal waveform changes.

Index Terms: speech production model, modal locking

1. Introduction

The source–filter theory of vowel production assumes that vocal fold vibrations occur independently of the vocal tract (VT) whose resonances modulate the vowel sound [1, 2]. Although this approach suffices for many purposes, some phenomena cannot be explained without filter-to-source feedback. Perturbations in vocal fold vibrations when the fundamental frequency f_o is near the lowest formant[1] F_1 are an example of such phenomena.

Although such perturbations have been reported in human speakers [3, 4, 5], excised larynges [6], physical laboratory models [7, 8, 9, 10], and computer simulations [4, 10, 11, 12], only a few speech modelling studies (namely, [4, 10, 12]) have given attention to the time-domain dynamics of f_o glides near F_1, and only limited details are reported on waveform changes, e.g., amplitudes or pulse shapes, in the glides. Yet time-domain simulations are befitting for studying the intrinsically time-domain speech production, and waveforms patterns can be used to compare speech models with *in vivo* measurements, e.g., [5], in greater detail than is possible with spectrograms.

The focus of this work is simulations of pitch glides over an isolated VT resonance. A VT geometry corresponding to [i] is used, and f_o is varied over the interval $[150\,\text{Hz},\ 310\,\text{Hz}]$ which contains the first VT resonance f_{R1}. A low-order time-domain model comprising a biomechanical submodel of the glottis and acoustic submodels of the VT and subglottal tract (SGT) is used to investigate changes in glottal waveforms during these glides. The model contains two mechanisms by which the VT and SGT loads affect the glottis model. First, the effect of the air column inertia on the glottal flow is included. This approach is related, but not identical, to the level 2 interaction explained in [12]. Second, feedback from the VT/SGT acoustics directly to the vocal fold tissues is included. To our knowledge, glide simulation studies have not previously included this mode of feedback even though it is an important factor during the closed phase of the glottal cycle.

[1]Here, conceptual difference between the lowest formant and the first VT resonance is neglected, i.e., $F_1 \approx f_{R1}$.

2. The model

The model used in this work is described in detail in [13, 14, 15] and outlined briefly below.

2.1. Vocal folds

The vocal folds are modelled as wedge-shaped vibrating elements of thickness L and length h (Figure 1). The oscillations of the elements are described using equivalent mass-spring-damper systems (Figure 2) with equations of motion

$$\begin{cases} M_1\ddot{W}_1(t) + B_1\dot{W}_1(t) + K_1W_1(t) = F_1(t), \\ M_2\ddot{W}_2(t) + B_2\dot{W}_2(t) + K_2W_2(t) = F_2(t), \end{cases} \tag{1}$$

where $W_j = [w_{j1} \quad w_{j2}]^T$ are the displacements of m_{j1} and m_{j2} in the y-direction for the j^{th} vocal fold ($j = 1, 2$) as shown in Figure 2. The load force pair $F_j(t) = [F_{j1}(t) \quad F_{j2}(t)]^T$ is described below. The mass, damping, and stiffness matrices M_j, B_j, and K_j, respectively, in (1) are

$$\begin{aligned} M_j &= \begin{bmatrix} m_{j1} + \frac{m_{j3}}{4} & \frac{m_{j3}}{4} \\ \frac{m_{j3}}{4} & m_{j2} + \frac{m_{j3}}{4} \end{bmatrix}, \\ B_j &= \begin{bmatrix} b_{j1} & 0 \\ 0 & b_{j2} \end{bmatrix}, \text{ and} \\ K_j &= \begin{bmatrix} l_1^2k_{j1} + l_2^2k_{j2} & l_1l_2(k_{j1} + k_{j2}) \\ l_1l_2(k_{j1} + k_{j2}) & l_2^2k_{j1} + l_1^2k_{j2} \end{bmatrix}, \end{aligned} \tag{2}$$

and the parameters can be found in Figures 1–2 and it has been assumed that $l_1 + l_2 = 1$. Symmetry is imposed on the vocal folds for this work by using $M = M_j$, $K = K_j$, and $B = B_j$, $j = 1, 2$, and by setting $F(t) = F_2(t) = -F_1(t)$. It is further assumed that $b = b_{j1} = b_{j2}$.

The glottal openings at the two ends of the vocal folds, denoted by ΔW_i, $i = 1, 2$, are related to (1) through

$$\begin{bmatrix} \Delta W_1 \\ \Delta W_2 \end{bmatrix} = W_2 - W_1 + \begin{bmatrix} g \\ H_0 \end{bmatrix}. \tag{3}$$

The rest gap parameters g and H_0 are shown in Figure 1.

When the vocal folds are closed, the load forces in (1) are produced by a nonlinear spring force with parameter k_H, accounting for the elastic collision of the vocal folds, and by the acoustic counter pressure $p_c = p_c(t)$ (see equation (11)) from the VT and SGT. Thus, the force pair for equations (1) during glottal closed phase is

$$F = F_H = \begin{bmatrix} k_H|\Delta W_1|^{3/2} - C_pp_c \\ C_pp_c \end{bmatrix} \quad \text{for} \quad \Delta W_1 < 0. \tag{4}$$

The geometric coefficient $C_p = C_p(t)$ accounts for the moment arms and areas on which p_c acts (see equation (12)).

http://dx.doi.org/10.21437/Interspeech.2017-875

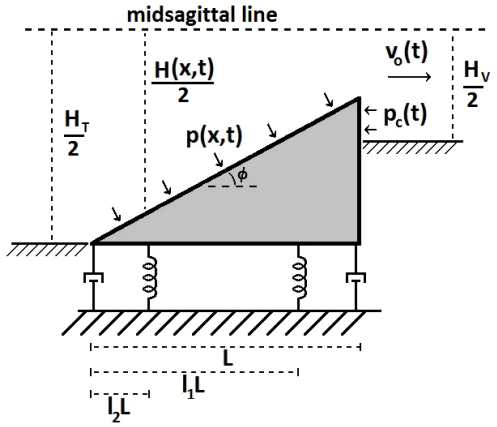

Figure 1: *Geometry of the one half of the glottis model with trachea on the left and vocal tract on the right.*

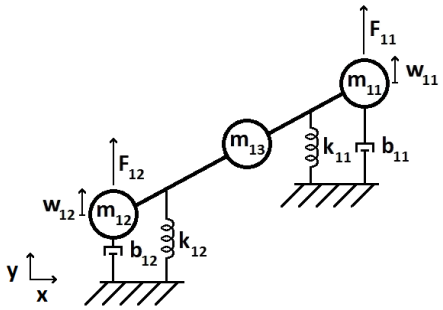

Figure 2: *The one vocal fold ($j = 1$) represented as a lumped-element model.*

2.2. Glottal flow and aerodynamic force

The main component of the air flow in the speech production system, to which the acoustic component acts as a perturbation, is assumed to be incompressible and one-dimensional, and to satisfy mass conservation and Newton's second law. The glottal velocity $v_o(t)$ of this flow component, through the control area hH_1 superior to the glottis, is then described by

$$\dot{v}_o(t) = \frac{1}{C_{iner} h H_1} \left(p_s - R_g(t) v_o(t) \right), \tag{5}$$

where p_s is an ideal driving pressure source, C_{iner} represents the inertance of the air column in the airways, and $R_g(t)$ regulates pressure losses in the glottis and hence the part of the driving pressure that is not available to accelerate the air column [13, pp. 8–10]. This simplified approach takes into consideration the translational inertial effect of the accelerating air column not described by the linear acoustic equations. The inertial effect represents, by Newton's third law, implicit feedback from the VT and SGT to the glottal velocity. Note that the glottal flow, i.e., volume velocity, is $U(t) = hH_1 v_o(t)$.

The pressure loss in the glottis consists of two components

$$R_g(t) = \frac{12 \mu H_1 L_g}{\Delta W_1(t)^3} + \frac{k_t \rho H_1^2 v_o(t)}{2 \Delta W_1(t)^2}. \tag{6}$$

The first term represents viscous pressure losses, and the second term accounts for entrance/exit effects (i.e., pressure loss at entrance and partial recovery at exit) which are not attributable to viscosity [15, p. 9]. The parameters L_g, k_t, μ, and ρ are described in Table 2.

The air flow (5) produces an aerodynamic force on the vocal folds [15, pp. 9–10]. When the vocal folds are open, the load force in (1) is hence $F = [F_{A,1} \quad F_{A,2}]^T$, and, for $\Delta W_1 > 0$,

$$\begin{aligned} F_{A,1} &= \frac{\rho h L H_1^2 v_o^2}{2 \cos^2 \phi} \left(-\frac{1}{\Delta W_1 (\Delta W_2 - \Delta W_1)} \right. \\ &\quad \left. + \frac{1}{(\Delta W_1 - \Delta W_2)^2} \ln \left(\frac{\Delta W_2}{\Delta W_1} \right) \right) - C_p p_c, \quad \text{and} \\ F_{A,2} &= \frac{\rho h L H_1^2 v_o^2}{2 \cos^2 \phi} \left(\frac{\sin^2 \phi \Delta W_2 + \cos^2 \phi \Delta W_1}{\Delta W_1 \Delta W_2 (\Delta W_2 - \Delta W_1)} \right. \\ &\quad \left. - \frac{1}{(\Delta W_1 - \Delta W_2)^2} \ln \left(\frac{\Delta W_2}{\Delta W_1} \right) \right) + C_p p_c, \end{aligned} \tag{7}$$

where ϕ is the angle of the inclined vocal fold surface from the horizontal (Figure 1), and p_c is given by (11).

2.3. Vocal and subglottal tract acoustics

The total air flow in the airways has a component due to compressibility that can be treated using acoustics. One way of deriving the acoustic equations is to start with the equations of fluid and thermodynamics, and then remove all terms not required for describing a radiative transfer mode of energy and momentum, propagating at the speed of sound [16, Section 2]. The generalised Webster's equation (see [17]) is used as acoustic loads to represent both the VT and the SGT

$$\begin{aligned} &\frac{1}{c^2 \Sigma^{(j)}(s)^2} \frac{\partial^2 \psi^{(j)}}{\partial t^2} + \frac{2\pi \alpha W^{(j)}(s)}{A^{(j)}(s)} \frac{\partial \psi^{(j)}}{\partial t} \\ &\quad - \frac{1}{A^{(j)}(s)} \frac{\partial}{\partial s} \left(A^{(j)}(s) \frac{\partial \psi^{(j)}}{\partial s} \right) = 0, \end{aligned} \tag{8}$$

where the parameters α, c are in Table 2, and $j =$"V" or "S" refers to the VT or the SGT, respectively. The solution $\psi^{(j)} = \psi^{(j)}(s,t)$ is the velocity potential of the acoustic field. The VT/SGT geometry is represented by the area function $A^{(j)}(s)$, the stretching factor $W^{(j)}(s)$, and the sound speech correction factor $\Sigma^{(j)}(s)$, all at distance $s \in [0, L^{(j)}]$ from the glottis, where $L^{(j)}$ is the length of the VT/SGT.

The following boundary conditions are used for (8)

$$\begin{cases} \frac{\partial \psi^{(j)}}{\partial s}(0,t) &= -C_v^{(j)} v_o(t), \\ \frac{\partial \psi^{(j)}}{\partial t}(L^{(j)},t) + \theta^{(j)} c \frac{\partial \psi^{(j)}}{\partial s}(L^{(j)},t) &= 0. \end{cases} \tag{9}$$

The first boundary condition couples the resonator to the glottal velocity (5) with scaling $C_v^{(V)} = hH_1/A^{(V)}(0)$ and $C_v^{(S)} = -hH_1/A^{(S)}(0)$. The normalised acoustic resistances $\theta^{(j)}$ at the other ends are given in Table 1.

2.3.1. VT and SGT geometries

The VT geometry of a healthy 26 years old female corresponding to a prolonged [i] at $f_o = 210$ Hz is used. MRI data was acquired using the method described in [18], and the computational geometries were extracted using the custom software described in [19, 20]. The VT geometry and area function are shown in Figure 3, and the related VT geometry dependent parameter values are given in Table 1. The inertia parameter in the glottal velocity (5) is taken to be $C_{iner} = C_{iner}^{(V)} + C_{iner}^{(S)}$.

Table 1: *Physical and physiological parameters dependent on the VT and SGT geometries.*

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

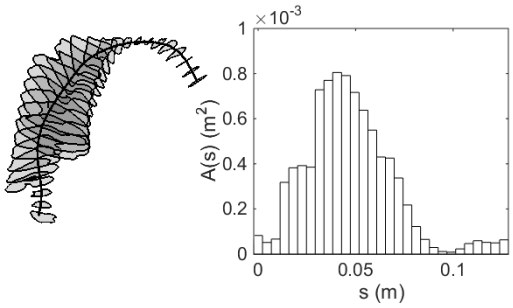

Figure 3: *VT intersections extracted from MRI data and the resulting area function at distance s from the glottis.*

A straight exponential horn is used as the subglottal area function

$$A^{(S)}(s) = A^{(S)}(0)e^{\epsilon s}, \quad \text{where} \quad \epsilon = \tfrac{1}{L^{(S)}} \ln\left(\tfrac{A^{(S)}(L^{(S)})}{A^{(S)}(0)}\right), \tag{10}$$

following [21]. The value for $A^{(S)}(0) = 200\,\mathrm{mm}^2$.

2.4. The acoustic counter pressure

The feedback from the VT/SGT acoustics to vocal fold surfaces is realised as a product of the acoustic counter pressure $p_c = p_c(t)$ and the geometric coefficient $C_p = C_p(t)$ in vocal fold load forces (4) and (7) above.

The counter pressure is the resultant of sub- and supraglottal pressure components

$$p_c(t) = Q_{pc}\rho\left(\frac{\partial\psi^{(V)}}{\partial t}(0,t) - \frac{\partial\psi^{(S)}}{\partial t}(0,t)\right), \tag{11}$$

where the parameter $Q_{pc} \in [0,1]$ enables scaling the magnitude of the feedback. This scaling is needed because determining the proportionality of the energy (i) in the underlying flow and (ii) in the superimposed acoustics is difficult to evaluate.

The geometric coefficient C_p is best understood in reference to the moment produced by p_c on the vocal fold wedge. Assuming p_c acts longitudinally on the vertical surface in Figure 1 and considering moments around the inferior edge of the wedge

$$C_p = \tfrac{h}{8L}(H_1 - \Delta W_1)(2H_0 - H_1 - \Delta W_1). \tag{12}$$

3. Numerical simulations

As a preparation of the pitch glides, the vocal fold mass matrix M is determined as detailed in [13, p. 14]. Long steady simulations are used to obtain an initial stiffness matrix K^0 that produces $f_o \approx 150\,\mathrm{Hz}$ and vocal fold damping that keeps the glottal flow amplitude unchanging, resulting in $b = 0.01$. Values for other constants are listed in Table 2.

Table 2: *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 [22])	730 N/m
glottal gap at rest g	0.4 mm
vocal fold length h (from [23])	10 mm
vocal fold thickness L (from [22])	6.8 mm
superior vocal fold spring location l_1	0.85
inferior vocal fold spring location l_2	0.15
control area height below glottis H_0	11.3 mm
control area height above glottis H_1	2 mm
viscous thickness L_g	1.5 mm
glottal entrance/exit coefficient k_t	0.4
subglottal pressure (over ambient) p_s^0	650 Pa

In all simulations, the equations of motion (1) are solved using a fourth order Runge-Kutta method, the glottal velocity (5) using implicit Euler method, and the Webster's equations (8) are discretised spatially using finite element method and temporally using Crank-Nicolson method (for further details, see [13, pp. 12–14 and 24–27]).

The pitch glides are achieved by changing K and p_s,

$$K(t) = 2.2^{2t/T}K^0 \quad \text{and} \quad p_s(t) = 2^{t/T}p_s^0. \tag{13}$$

Changing p_s has only a small impact on f_o in these glides but it ensures that, in the absence of VT/SGT feedback, increasing K does not lead to significant changes in glottal pulse shapes (i.e., type of phonation).

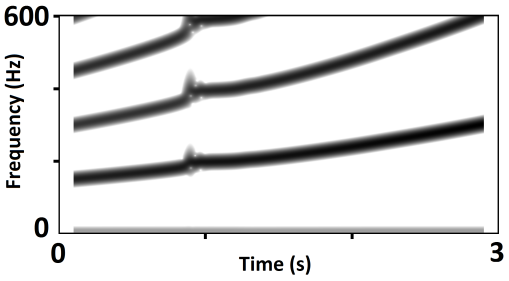

Figure 4: *A spectrogram of the glottal flow U.*

4. Simulation results

A spectrogram for the simulated glottal flow U in a glide with feedback level $Q_{pc} = 0.1$ is shown in Figure 4. A perturbation is clearly visible when f_o crosses f_{R1}. This pattern is absent if the same simulation is carried out without VT feedback (i.e., $Q_{pc} = 0$). Note that no subharmonics are observed in this glide.

Figure 5 (top panel) shows parameters which have been extracted from U pulse by pulse. Since subharmonics were not present, there was no ambiguity in the definition of a pulse.

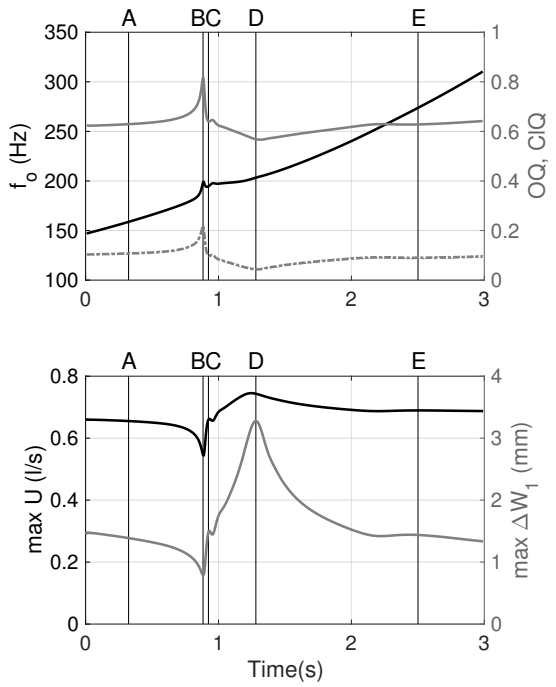

Figure 5: *Top: f_o (black) on left axis, and open quotient OQ (solid gray) and closing quotient ClQ (dashed gray) on right axis. Bottom: Maximum U in each pulse (black) on left axis and maximum glottal gap ΔW_1 in each pulse (gray) on right axis. Vertical lines indicate time instants shown in Figure 6.*

The f_o trajectory shows the perturbation pattern in greater detail than Figure 4: a rapid increase (latter referred to as a jump for clarity) from approximately 180 Hz to f_{R1}, a locking to a plateau at f_{R1} with small oscillations at the beginning, and a smooth release.

Figure 5 (top panel) also shows the open quotient OQ and the closing quotient ClQ, see [24, Fig. 4]. The lower panel of Figure 5 shows the amplitudes (i.e., maxima as glottis closes fully through the glide) of U and ΔW_1, and full pulse shapes at five instants are shown in Figure 6. The jump in f_o (instant B) is accompanied by a peak in the OQ and ClQ values, i.e., a momentary increase in breathiness, and a minimum in U and ΔW_1 amplitudes. The increased breathiness shows as a rounder and less skewed U pulse, and Figure 6 also displays more gradual opening and closing of the vocal folds and build-up and cut-off of the flow when compared to pre-perturbation instant A.

At the beginning of the locking plateau, OQ and ClQ drop while the amplitudes of U and ΔW_1 increase rapidly. By the instant C the pulse shapes have become very similar to the starting shapes (A). OQ and ClQ then continue to decrease (increased pressedness) and the amplitudes to increase through the locking event at fairly steady rates. The vocal fold oscillations amplitude reaches its maximum slightly after f_o starts to increase again (D). This is also when maximum pressedness of the phonation occurs. After instant D, OQ and ClQ start to increase gradually while amplitudes decrease, until well away from the perturbation (E), the pulse shapes are very similar to the start (A). The similarity of the pulse shapes at the beginning and end of the glide was achieved by using the p_s control described above.

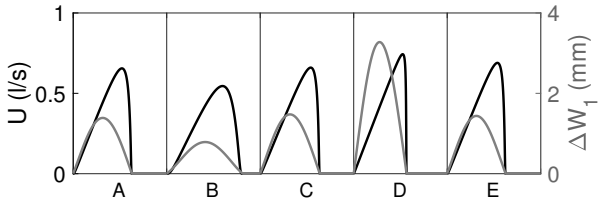

Figure 6: *Glottal flow U (black) on left axis and glottal gap ΔW_1 (gray) on right axis at five time instants in the glide.*

5. Discussion

The changes in glottal waveform around the f_o jump and locking can be interpreted by considering energy in the system. The total energy in the model is the sum of the acoustic, flow, and vocal fold oscillation energy. At the jump, i.e., when f_o first reaches f_{R1}, the energy contained in the vocal fold movements and the glottal flow is reduced to feed the resonating VT. Injected energy is needed to unlock the model to a frequency range different from the VT resonance. The glide controls (13) provide this energy injection into the glottal flow, and they also increase the storage potential of the vocal fold mass-spring system to accommodate a part of this energy injection at a frequency higher than the plateau value of f_o. Finally, the eiqenfrequency of the mass-spring system wins and f_o is released.

The pattern of waveform changes, accompanying the observed f_o locking pattern, is robust against variations in parameter values and some model modifications, such as changing the lip termination to a first-order impedance model approximation of the piston model [15, 25]. The existence of the f_o locking and the details of the f_o trajectory do, however, depend on the parameter values and model details. Besides the feedback parameter Q_{pc}, the loss parameters (b, L_g, k_t) also play a critical role in determining the f_o trajectory. A low b value, for example, can reduce or remove the jump and locking, while a high b has the opposite effect. This supports the energy interpretation as b regulates how much of the injected energy is stored in the vocal fold movements and hence available to unlock the system.

At the jump, the increased OQ reduces the duration of strong acoustics–vocal fold feedback in the glottal cycle, helping the system to settle at the locking plateau. In contrast, higher OQ increases the duration of the glottal cycle where feedback from the VT/SGT acoustics directly to the glottal flow (level 1 interaction in [12], not included here) is at its strongest. Waveform variations have not been reported for glides using this interaction mode, so questions remain open regarding whether such models exhibits similar self-liming dynamic behaviour and which effect would be dominant in systems containing both feedback modes.

6. Conclusions

The simulations indicate that the feedback from resonator acoustics directly to vocal fold tissues can cause significant perturbations to vocal fold dynamics when f_o crosses a VT resonance. These perturbations are visible in both f_o and the glottal waveforms, indicating a special energy transfer pattern between the vocal fold vibrations, glottal flow, and VT/SGT acoustics.

7. References

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

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

[3] I. R. Titze, T. Riede, and P. Popolo, "Nonlinear source-filter coupling in phonation: Vocal exercises," *The Journal of the Acoustical Society of America*, vol. 123, no. 4, pp. 1902–1915, 2008.

[4] 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.

[5] 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.

[6] S. F. Austin and I. R. Titze, "The effect of subglottal resonance upon vocal fold vibration," *Journal of Voice*, vol. 11, no. 4, pp. 391–402, 1997.

[7] Z. Zhang, J. Neubauer, and D. A. Berry, "The influence of subglottal acoustics on laboratory models of phonation," *The Journal of the Acoustical Society of America*, vol. 120, no. 3, pp. 1558–1569, 2006.

[8] N. Ruty, X. Pelorson, A. Van Hirtum, I. Lopez-Arteaga, and A. Hirschberg, "An in-vitro setup to test the relevance and the accuracy of low-order vocal folds models," *The Journal of the Acoustical Society of America*, vol. 121, no. 5, pp. 479–490, 2007.

[9] N. Ruty, X. Pelorson, and A. Van Hirtum, "Influence of acoustic waveguides lengths on self-sustained oscillations: Theoretical prediction and experimental validation," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 3121–3121, 2008.

[10] 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.

[11] H. Hatzikirou, W. T. Fitch, and H. Herzel, "Voice instabilities due to source-tract interactions," *Acta Acustica united with Acustica*, vol. 92, pp. 468–475, 2006.

[12] 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.

[13] 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.

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

[15] A. Aalto, T. Murtola, J. Malinen, D. Aalto, and M. Vainio, "Modal locking between vocal fold and vocal tract oscillations: Simulations in time domain," arXiv:1506.01395, Tech. Rep., 2017.

[16] A. Hannukainen, T. Lukkari, J. Malinen, and P. Palo, "Vowel formants from the wave equation," *The Journal of the Acoustical Society of America Express Letters*, vol. 122, no. 1, pp. EL1–EL7, 2007.

[17] T. Lukkari and J. Malinen, "Webster's equation with curvature and dissipation," arXiv:1204.4075, Tech. Rep., 2016.

[18] 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.

[19] A. Kivelä, "Acoustics of the Vocal Tract: MR image segmentation for modelling," Master's thesis, Aalto University School of Science, Department of Mathematics and Systems Analysis, 2015.

[20] A. Ojalammi and J. Malinen, "Automated segmentation of upper airways from MRI: Vocal tract geometry extraction," in *BIOIMAGING 2017 – 4th International Conference on Bioimaging, February 21–23, Porto, Portugal, Proceedings*, 2017, pp. 77–84.

[21] P. Birkholz, D. Jackel, and B. Kröger, "Simulation of losses due to turbulence in the time-varying vocal system," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 4, pp. 1218–1226, 2007.

[22] 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.

[23] 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.

[24] P. Alku, "Glottal inverse filtering analysis of human voice production - A review of estimation and parameterization methods of the glottal excitation and their applications," *Sadhana*, vol. 36, no. 5, pp. 623–650, 2011.

[25] T. Hélie and X. Rodet, "Radiation of a pulsating portion of a sphere: application to horn radiation," *Acta Acustica united with Acustica*, vol. 89, pp. 565–577, 2003.