Interactive Processes Underlying the Production of Voice

Tokihiko Kaburagi*
* Kyushu University, Fukuoka, Japan
E-mail: kabu@design.kyushu-u.ac.jp Tel/Fax: +81-92-553-4572/4520

***Abstract*—This paper presents recent progress in research examining the mechanisms of voice generation and a method for physiologically-based speech synthesis. The overarching goal of this research is the precise modeling of interactions among physical systems involved in the processes underlying voice generation. The interaction between the voice generation system (glottal flow and the vocal folds) and the vocal-tract filter has been found to involve nonlinear factors in speech, such as the skewing of glottal flow, unsteadiness of vocal fold oscillations, and transition of the voice register. The results of a quantitative analysis also revealed frequency-dependent effects of source-filter interaction.**

I. Introduction

In the production of voiced speech, glottal sound sources are generated by oscillations of the vocal folds, which travel through the acoustic filter of the vocal tract. The vocal tract acts as a resonator, which forms the peaks of the speech spectrum called formants. In many models or analysis methods of speech, these two sub-systems, i.e., the source generating system and the vocal tract filter, are thought to be independent because of the high acoustic impedance of the glottis. Based on this assumption, linear source-filter theory [1] may explain the nature of the speech production mechanism in humans.

However, when glottal sound sources travel through the vocal tract, they form a sound field in the vicinity of the glottis. This acoustic pressure can influence the flow though the glottis and the movements of the vocal folds, a phenomenon known as source-filter interaction (SFI). This interaction has various nonlinear effects, such as flow skewing [2], enhancement or suppression of vocal-fold oscillations [3], [4], lowering oscillation threshold pressure [5], and discontinuity of changes in the fundamental frequency caused by continuous changes in vocal fold parameters [6], [7].

Time-domain computer simulation has been an important tool for investigating these effects, and results from these studies have been analyzed in the frequency domain as an interrelation between the fundamental frequency of the voice and the resonant and antiresonant frequencies of the vocal tract [6]. It has been shown that the effect becomes more intense when the fundamental frequency approaches the resonant frequency. In addition, vocal fold oscillation becomes unstable when the fundamental frequency is between the resonant and antiresonant frequencies. Titze [4] interpreted this effect in terms of the reactance of impedances arising from the acoustic load of the sub- and supra-glottal tracts. The results revealed that the harmonic component of the glottal volume flow waveform becomes weak when the reactance is negative at that frequency (level one interaction). In addition, the author proposed that the reactance of vocal-tract impedance should be inertive, while that of the sub-glottal tract should be compliant for the oscillation of the vocal folds (level two interaction).

In the current paper, the processes underlying the generation of voice are investigated, and a mathematical model is presented by focusing on the interactive behavior of three physical sub-systems, i.e., the glottal flow, the vocal folds, and the vocal tract. With respect to source-filter interaction, the acoustic tubes of the sub- and supra-glottal tracts are connected to the intermediate part of the glottis, and the frequency response of the acoustic pressures near the glottis is expressed relative to the acoustic component of the glottal flow. The results of a quantitative analysis are presented, including an examination of frequency responses related to SFI and a time-domain simulation of voice production.

II. Flow-structure interaction

Myoelastic-aerodynamic theory is widely considered the basic principle underlying the production of voice. This theory was proposed by Helmholtz and Müller in the 19th century, and confirmed experimentally by van den Berg in the 1950s. Based on the interaction between air flow passing through the glottis and visco-elastic movements of the vocal folds, the theory proposes that the vocal folds are driven by aerodynamic pressure of glottal flow. Their movements in turn open and close the glottis and control the volume flow that gushes out of the glottis. Acoustic sound sources are then generated as a result of the self-sustained oscillation of the vocal folds.

By assuming that the flow through the glottis is quasi-steady, incompressible and inviscid, its behavior can be described by the conservation law of energy, known as the Bernoulli relation, as

$$p(x) + \frac{1}{2}\rho v(x)^2 = \text{const.}, \quad (1)$$

where $p(x)$ is the pressure, $v(x)$ is the velocity, and ρ is the air density. The x axis is taken in many voice production models as the axis of symmetry between both vocal folds, as shown in Fig. 1. The flow velocity can be written as

$$v(x) = \frac{u_g}{S(x)} = \frac{u_g}{h(x)l_g}, \quad (2)$$

where u_g is the volume flow rate, $S(x) = h(x)l_g$ is the sectional area of the glottal channel, and l_g is the length of the

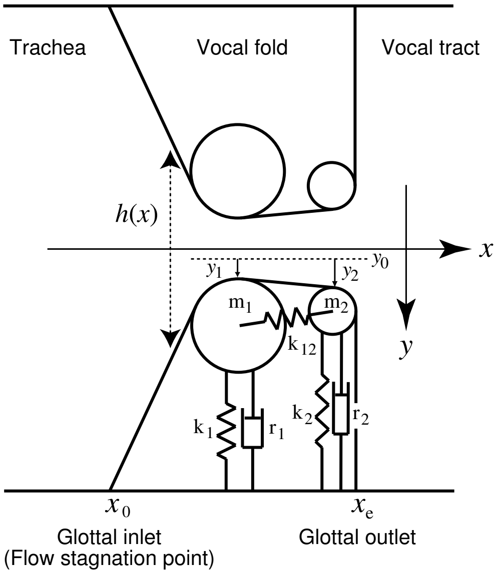

Fig. 1. Representation of the one-dimensional glottal channel and the mass-spring model of the vocal folds. The x axis is taken as the symmetrical axis of both folds. $h(x)$ is the channel height.

vocal fold which is perpendicular to the $x-y$ plane shown in the figure. In the tracheal region below the glottis, the channel is widely opened and the velocity and kinetic energy are very small compared with those in the glottis. At the outlet of the glottis, on the other hand, the flow pressure is almost equal to the atmospheric pressure, and hence can be set to zero. From the Bernoulli relation, the following relation is then obtained for the sub-, intra-, and supra-glottal regions:

$$\underbrace{p_{F0}}_{\text{trachea}} = \underbrace{p(x) + \frac{1}{2}\rho\frac{u_g^2}{S(x)^2}}_{\text{glottis}} = \underbrace{\frac{1}{2}\rho\frac{u_g^2}{S_s^2}}_{\text{vocal tract}}, \tag{3}$$

where p_{F0} is the sub-glottal pressure and S_s is the sectional area of the glottal jet formed by the separation of flow from the wall of the channel. If the point of flow separation, x_s, is given, S_s can be written as $S_s = h(x_s)l_g$ by assuming that the sectional area of the jet is constant.

The flow separation point was fixed to the glottal exit in the two-mass model of the vocal folds[6]. However in reality, this strongly depends on the shape of the glottis. It is also known that the glottis takes both the divergent and convergent shapes during a period of the vocal fold oscillation. Recently, the behavior of viscous flow near the surface of the vocal fold was analyzed using the boundary-layer theory [8], [9]. In addition to the flow separation point, this analysis method can estimate characteristic quantities of the boundary layer, such as the displacement thickness, momentum thickness, and wall shear stress. The *effective* flow channel is then determined by excluding the displacement thickness, $\delta(x)$, from the geometric shape of the channel as

$$S(x) = \{h(x) - 2\delta(x)\}l_g. \tag{4}$$

The *effective* flow velocity is also determined as

$$v(x) = \frac{u_g}{\{h(x) - 2\delta(x)\}l_g}. \tag{5}$$

With these effective values, the glottal volume flow is estimated from (3) as

$$u_g = S_s\sqrt{\frac{2p_{F0}}{\rho}} \tag{6}$$

and the pressure within the glottis as

$$p(x) = p_{F0} - \frac{1}{2}\rho\frac{u_g^2}{S(x)^2}, \tag{7}$$

allowing the driving force of the vocal folds to be determined. The effective channel area at the separation point is now given as

$$S_s = \{h(x_s) - 2\delta(x_s)\}l_g \tag{8}$$

by considering the influence of the boundary layer.

Movement of the vocal fold is modeled using the modified two-mass model [10] in this paper. Each vocal fold is constructed by two cylinders, located at the lower and upper parts of the fold, and three plates, which connect the inlet and outlet of the glottis and the cylinders. In Fig. 1, masses m_1 and m_2 are assigned to each cylinder, and are connected to the fixed wall by dampers of resistance r_1 and r_2, and linear springs with Hooke's constants k_1 and k_2. The two masses are joined by another linear spring of constant k_{12}. The equation of motion is then given as

$$m_1\frac{d^2y_1}{dt^2} + r_1\frac{dy_1}{dt} + k_1y_1 + k_{12}(y_1 - y_2) = f_1 \tag{9}$$

and

$$m_2\frac{d^2y_2}{dt^2} + r_2\frac{dy_2}{dt} + k_2y_2 + k_{12}(y_2 - y_1) = f_2, \tag{10}$$

where y_1 and y_2 are the displacements of the masses perpendicular to the x axis. The absolute mass position is $y_1 + y_0$ and $y_2 + y_0$, giving the channel height

$$h(x) = 2(y_1 + y_0) \tag{11}$$

and

$$h(x) = 2(y_2 + y_0) \tag{12}$$

at the mass positions, where y_0 is the common rest position. As such, vocal fold movement determines the height of the glottal channel and gives the boundary condition to the flow analysis problem. The actual expression of the driving forces, f_1 and f_2, is explained below, but is basically given by integrating the pressure $p(x)$ in (7) along the x axis and by multiplying the vocal fold length l_g.

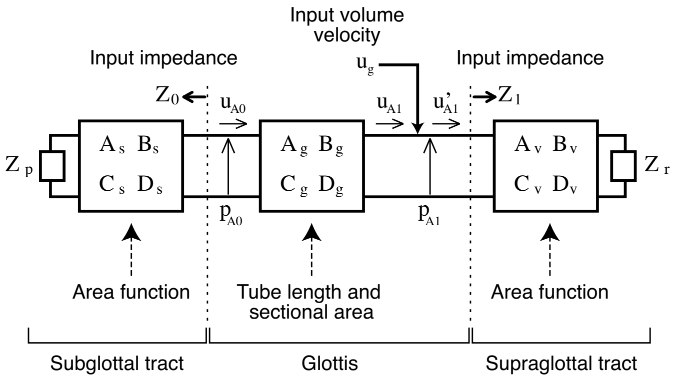

Fig. 2. Acoustic tube model connecting the sub- and supra-glottal tracts via the glottis.

III. SOURCE-FILTER INTERACTION (SFI)

When the glottal sound source is generated as a result of the flow-structure interaction, it travels through the vocal tract and forms a sound field in the vicinity of the glottis. This sound field causes a pressure gradient between the sub- and supra-glottal regions and induces additional volume flow. In addition, the instantaneous sound pressure in the glottis can push or pull the vocal folds. As such, the acoustic field can influence the glottal flow and vocal fold movements, a phenomenon known as source-filter interaction (SFI). This section describes how this interaction can be formulated in a model of voice generation [11].

A. *Specific acoustic pressures related to SFI*

SFI represents an acoustic influence of pressure near the glottis on the voice generation system of the vocal folds and glottal flow. Two specific pressure values are used here to represent these effects. One is the difference between the pressures at the inlet of the glottis (p_{A0}) and at the outlet (p_{A1})

$$\Delta p_A = p_{A0} - p_{A1}. \tag{13}$$

The other is the mean acoustic pressure inside the glottis

$$\bar{p}_A = \frac{p_{A0} + p_{A1}}{2}. \tag{14}$$

To discriminate between these temporal waveforms and their Fourier transforms, time-domain variables are denoted by lower-case letters while frequency-domain variables are denoted by capital letters in the following equations.

The values of these specific acoustic pressures are determined on the basis of the model shown in Fig. 2. From this model, relationships

$$P_{A0} = -Z_0 U_{A0}, \tag{15}$$

$$P_{A1} = Z_1 U'_{A1}, \tag{16}$$

and

$$U'_{A1} = U_{A1} + U_g \tag{17}$$

can be obtained, where Z_0 and Z_1 represent the input impedance of the sub- and supra-glottal tracts, respectively. The input impedance of the glottis seen from the vocal tract, Z_g, can be expressed as

$$Z_g = -\frac{P_{A1}}{U_{A1}} = -\frac{(B_g - A_g Z_0)}{(D_g - C_g Z_0)}, \tag{18}$$

where A_g, B_g, C_g, and D_g are elements of the transmission matrix of the glottis approximated as a tube with a uniform cross section.

From these relationships, Δp_A and $\bar{p}_A$ are each obtained as the frequency responses of a filter having the acoustic component of the glottal volume flow as its input:

$$Z_\Delta = \frac{\Delta P_A}{U_g} = \frac{P_{A0} - P_{A1}}{U_g} = \frac{\{B_g - (A_g - 1)Z_0\}Z_1}{Z_D} \tag{19}$$

and

$$Z_M = \frac{\bar{P}_A}{U_g} = \frac{P_{A0} + P_{A1}}{2U_g} = -\frac{\{B_g - (A_g + 1)Z_0\}Z_1}{2Z_D}, \tag{20}$$

where

$$Z_D = (D_g - C_g Z_0)Z_1 - (B_g - A_g Z_0). \tag{21}$$

The input impedances of the tracts can be obtained as

$$Z_0 = -\frac{(A_s Z_p + B_s)}{(C_s Z_p + D_s)} \tag{22}$$

and

$$Z_1 = \frac{(D_v Z_r - B_v)}{(A_v - C_v Z_r)}, \tag{23}$$

where A_s, B_s, C_s, and D_s are the elements of the transmission matrix for the sub-glottal tract and A_v, B_v, C_v, and D_v those for the supra-glottal tract. The values of these elements are calculated by applying an acoustic tube model of the lossy vocal tract [12], taking into account data regarding vocal-tract area function [13]. Z_p and Z_r are the terminal impedance for the lungs and lips, respectively.

B. *Total voice production model including SFI*

Fig. 3 shows the framework of the voice production model [11] to examine the influence of SFI on the generation of voice. The frequency responses related to SFI, Z_Δ and Z_M, can be determined from the area function data as explained in the preceding subsection. Next, the inverse Fourier transforms of the frequency responses, z_Δ and z_M, are convolved with the waveform of glottal volume flow, u_g, and the values of the specific acoustic pressures, Δp_A and $\bar{p}_A$, are calculated.

The volume flow is estimated from the specific acoustic pressure, Δp_A, and the static lung pressure, p_{F0}. The total pressure difference between the sub- and supra-glottal regions is determined by summing both pressures as

$$\Delta p = p_{F0} + \Delta p_A = p_{F0} + z_\Delta * u_g. \tag{24}$$

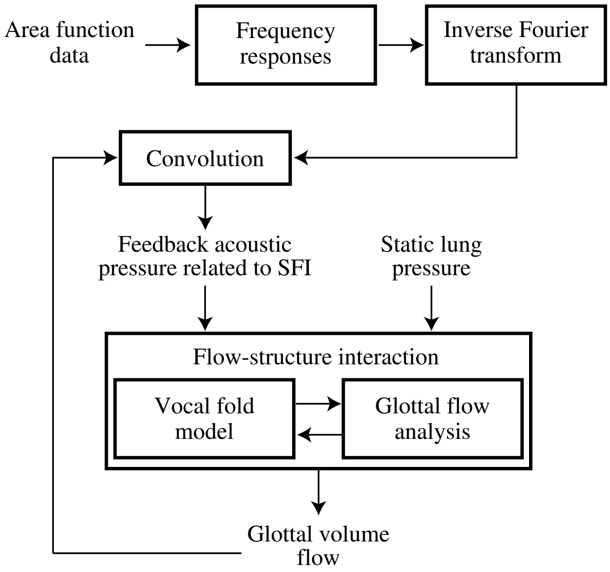

Fig. 3. Framework of the voice production model incorporating the effect of source-filter coupling.

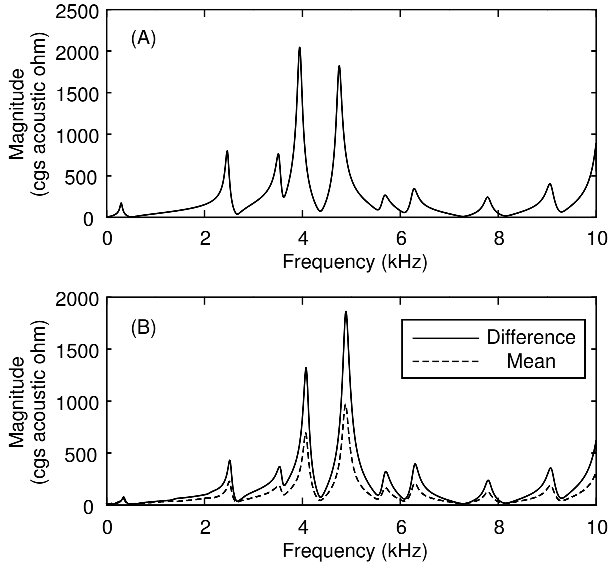

Fig. 4. Computed results of (A) vocal-tract input impedance and (B) frequency responses of the acoustic pressure difference Z_Δ (solid) and mean acoustic pressure Z_M (broken) for the vowel /i/.

Here, the convolution is computed only for the latest time instant. Considering the Bernoulli relation given in (3), a new pressure-flow relation can be obtained:

$$\Delta p = \frac{1}{2}\rho\frac{u_g^2}{S_s^2}. \tag{25}$$

By eliminating Δp, these relationships can be rewritten in a discrete-time form as

$$p_{F0} + \sum_{k=0}^{K-1} z_\Delta(k)u_g(n-k) = \frac{1}{2}\rho\frac{u_g(n)^2}{S_s^2}, \tag{26}$$

where K is the length of $z_\Delta(k)$, and n is the index corresponding to the latest time instant. The volume flow is then estimated as

$$u_g(n) = \frac{z_\Delta(0)S_s^2 + S_s\sqrt{(z_\Delta(0)S_s)^2 + 2\rho P}}{\rho} \tag{27}$$

where

$$P = p_{F0} + \sum_{k=1}^{K-1} z_\Delta(k)u_g(n-k). \tag{28}$$

These equations indicate that the time history of the glottal flow waveform affects the flow estimation recursively under the influence of SFI.

Next, we consider the influence of SFI on vocal fold movements. In the current study, the driving force was exerted only on the lower, first mass, and f_2 was set to zero [10]. The entire driving pressure is given by summing the aerodynamic and acoustic pressures as

$$p_{total}(x) = p(x) + \bar{p}_A = p_{F0} - \frac{1}{2}\rho\frac{u_g^2}{S(x)^2} + z_M * u_g. \tag{29}$$

The convolution is computed here for the latest time instant. The driving force for the lower mass is then given as

$$f_1 = \lambda l_g \left\{ \int_{x_0}^{x_s} p(x)dx + (x_e - x_0)\bar{p}_A \right\}. \tag{30}$$

x_0 is the inlet of the glottis, x_e is the outlet of the glottis, x_s is the flow separation point, and λ is a parameter specifying the area upon which the pressure is exerted. If the glottis is closed, f_1 is set to $f_1 = \lambda l_g(x_e - x_0)p_{F0}$. By substituting the value of f_1 into the equation of motion of the vocal fold model, the displacements of the lower and upper masses, y_1 and y_2, are calculated. These values determine the glottal shape $h(x)$, and the entire procedure is repeated for the dynamic simulation of voice production.

IV. Numerical Results

A. Frequency responses representing the effect of SFI

The frequency responses of the specific acoustic pressures, Z_Δ and Z_M, were computed from data related to vocal-tract area function taken for the vowel /i/ [13] and data related to the sub-glottal tract [14]. The length of the glottal tube (depth of the glottis) was set to 0.3 cm, and the cross-sectional area was set to 0.014 cm^2. The results are plotted in Fig. 4.

From the top plot, it can be seen that the input impedance of the vocal tract, Z_1, exhibited a significant peak at approximately 4 kHz. This peak was much stronger than those corresponding to the first to third formants. An investigation confirmed that this high-frequency peak was due to the resonance of the epiglottal tube, i.e., a local, small acoustic cavity above the glottis.

The figure also shows that each peak in the input impedance forms a peak of the same frequency in the magnitude of Z_Δ and Z_M. Concerning the peaks seen in the input impedance,

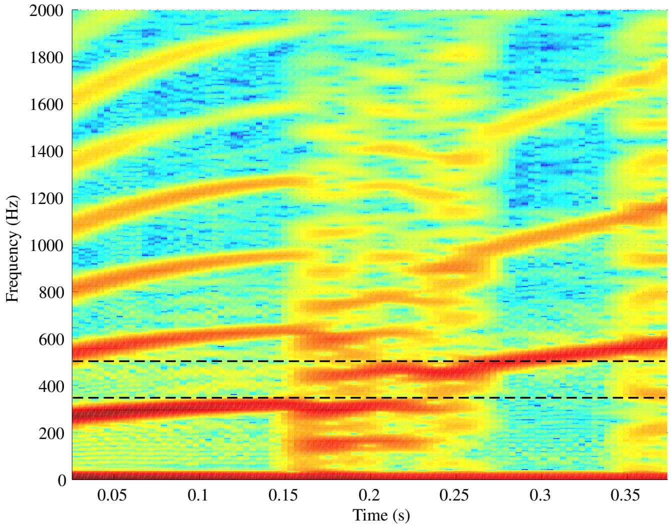

Fig. 5. Spectrogram for the temporal pattern of the glottal volume flow for the vowel /i/. The tension parameter of the vocal-fold model was linearly changed in time from two to five in this simulation.

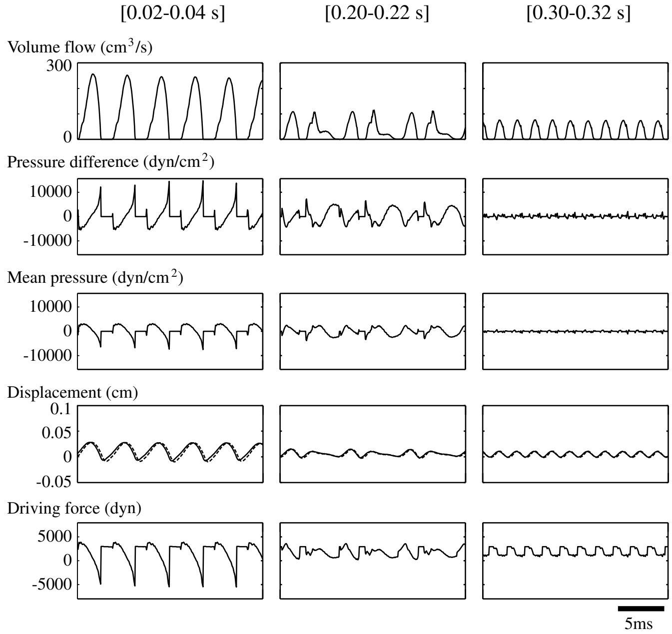

Fig. 6. Results of a time-domain simulation of voice production for the vowel /i/. The plots show the changes over time of the glottal volume flow, acoustic pressure difference across the glottis, mean acoustic pressure in the glottis, mass displacements, and driving force on the vocal folds. The displacements of the lower and upper masses are shown by solid and broken lines, respectively.

Titze [4] clarified that source-filter interaction can be enhanced by reducing the cross-sectional area of the epiglottal tube. It should also be noted that the magnitude of frequency responses is interpreted as the frequency-dependent strength of source-filter interaction, because they determine the amplitude of the pressure values, Δp_A and $\bar{p}_A$, resulting from the glottal volume flow.

B. Dynamic simulation

Dynamic simulation was conducted for the vowel /i/, and the results are plotted in Figs. 5 and 6. In the simulation, the parameters of the vocal-fold model were set to $m_1 = 0.125$ g, $m_2 = 0.025$ g, $k_1 = 80000$ dyn/cm, $k_2 = 8000$ dyn/cm, $k_{12} = 55000$ dyn/cm, $r_1 = 0.2\sqrt{m_1 k_1} = 23.3$ dyn·s/cm, and $r_2 = 1.2\sqrt{m_2 k_2} = 18.6$ dyn·s/cm. For the glottis in a closed state, the values of the stiffness parameters were increased such that $k_1 = 320000$ and $k_2 = 32000$ dyn/cm. The dampers were also increased such that $r_1 = 257$ and $r_2 = 49.6$ dyn·s/cm. The initial displacement was $y_0 = 0.00014$ cm and the vocal-fold length was $l_g = 1.2$ cm. λ was determined so that the effective depth of the lower mass, $\lambda(x_e - x_0)$, was approximately 0.27 cm. The sub-glottal pressure was 8 cmH_2O and the sampling frequency was 20 kHz.

To control the fundamental frequency, the masses were divided by a parameter T and the spring constants were multiplied by T. The natural frequency of each mass-spring system was then changed such that $\sqrt{k_1 T/(m_1/T)} = \sqrt{k_1/m_1}T$. The value of T was changed from two to five linearly with time in this simulation. The fundamental frequency was initially below the first formant of /i/, and above the first formant and the first dip at the end of the simulation.

Figure 5 shows the spectrogram of the glottal flow, u_g, within the frequency range 0-2 kHz. The horizontal broken lines around 350 and 500 Hz correspond to the first peak and dip, respectively, of $Z_\triangle$ and Z_M (see Fig. 4). The frequency region enclosed by the broken lines is therefore unstable for phonation [4], [6]. The fundamental frequency reaches this peak frequency at a time of approximately 0.15 s. Between 0.15 and 0.28 s, the spectrogram indicates the appearance of subharmonic and unstable behavior. The harmonic pattern then stabilizes when, at 0.28 s, the fundamental frequency agrees with the dip frequency of approximately 500 Hz. As such, the effect of SFI is apparent in the unstable region between 350 and 500 Hz. In addition, it is noteworthy that the voice register changes when the fundamental frequency crosses this specific region [15]: the vocal folds oscillate in the chest register when the fundamental frequency is below the unstable region. In contrast, when the fundamental frequency is above the region, the vocal folds oscillate in the falsetto register.

Figure 6 also shows waveforms of the main variables for the same simulation. In the first time section ($0.02-0.04$ s), the fundamental frequency is outside the unstable frequency region. In addition, the glottal flow waveform appeared skewed, and a stable positive-to-negative change of driving force was observed. The acoustic pressure difference is positive when the glottis is closing, which induces a forward flow and a sharpening of the waveform of the glottal flow. On the other hand, the mean pressure and driving force to the vocal folds is negative when the glottis is closing. The glottis is then effectively closed and the vocal fold oscillation is maintained.

In the second section ($0.20-0.22$ s), the vocal fold oscillation is unstable. In the waveform of the glottal flow and other physical quantities, we can observe an additional periodicity occurring for each of two pitch periods, resulting in the existence of subharmonic behavior in the spectrogram. Finally, over 0.30−0.32 s, regular vocal-fold oscillation is recovered, and the voice register is now changed to the falsetto. Note that the fundamental frequency is close to the dip frequency of Z_Δ and Z_M. The feedback acoustic pressures and the effect of SFI are therefore weak. In the falsetto register, the phase difference of the oscillatory movements of the upper and lower masses is not apparent, indicating that a convergent-divergent change in the glottal shape is not attained. The driving force to the vocal folds is mainly given by the aerodynamic pressure of the glottal flow, and from the figure, it is clear that the force is always positive.

V. Conclusion

The production process of the voice was examined by focusing on the interactions of three physical systems, i.e., the glottal flow, the vocal folds, and the vocal tract. The flow-structure interaction between the glottal flow and vocal folds plays an essential role in the production of voice. Regarding the influence of the geometry of glottal channel on the behavior of flow, a detailed analysis result was reported in [8] and [9]. In the present study, the source-filter interaction between the voice production system (glottal flow and the vocal folds) and the acoustic cavity of the vocal tract was modeled using the frequency response of two types of specific acoustic pressures in the vicinity of the glottis. Quantitative investigations revealed that these frequency responses exhibited dominant peaks in the high frequency region at approximately 4-5 kHz. In addition, SFI was typically observed when the fundamental frequency of the vocal fold oscillation approached the resonant frequency of the vocal tract. The strength of the interaction was also found to increase near the resonant frequency, and simulation results revealed instability in phonation and changes of the voice register.

This research was partly supported by the Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (Grant No. 19103003).

References

[1] G. Fant, *Acoustic theory of speech production, with calculations based on X-ray studies of Russian articulations (2nd printing)*, The Hague: Mouton, 1970.

[2] M. Rothenberg, “Acoustic interaction between the glottal source and the vocal tract,” in *Vocal Fold Physiology*, Stevens and Hirano Eds. Tokyo: University of Tokyo Press, 1981, pp. 305–328.

[3] M. Zañartu, L. Mongeau, and G. R. Wodicka, “Influence of acoustic loading on an effective single mass model of the vocal folds,” J. Acoust. Soc. Am., vol. 121, pp. 1119–1129, 2007.

[4] I. R. Titze, “Nonlinear source-filter coupling in phonation: Theory,” J. Acoust. Soc. Am., vol. 123, pp. 2733–2749, 2008.

[5] R. W. Chan and I. R. Titze, “Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics,” J. Acoust. Soc. Am., vol. 119, pp. 2351–2362, 2006.

[6] K. Ishizaka and J. L. Flanagan, “Synthesis of voiced sounds from a two-mass model of the vocal cords,” Bell Syst. Tech. J., vol. 51, pp. 1233–1268, 1972.

[7] J. G. Švec, H. K. Schutte, and D. G. Miller, “On pitch jumps between chest and falsetto registers in voice: Data from living and excised human larynges”, J. Acoust. Soc. Am., vol. 106, pp. 1523–1531, 1999.

[8] T. Kaburagi, “On the viscous-inviscid interaction of the flow passing through the glottis,” Acoust. Sci. Tech., vol. 29, pp. 167-175, 2008.

[9] T. Kaburagi and Y. Tanabe, “Low-dimensional models of the glottal flow incorporating viscous-inviscid interaction,” J. Acoust. Soc. Am., vol. 125, pp. 391–404, 2009.

[10] X. Pelorson, A. Hirschberg, R. R. van Hassel, A. P. J. Wijnands, and Y. Auregan, “Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model,” J. Acoust. Soc. Am., vol. 96, pp. 3416–3431, 1994.

[11] T. Kaburagi, “Voice production model integrating boundary-layer analysis of glottal flow and source-filter coupling,” J. Acoust. Soc. Am., vol. 129, pp. 1554–1567, 2011.

[12] M. M. Sondhi and J. Schroeter, “A hybrid time-frequency domain articulatory speech synthesizer,” IEEE Trans. Acoust., Speech & Signal Process., vol. ASSP-35, pp. 955-967, 1987.

[13] B. H. Story, I. R. Titze, and E. A. Hoffman, “Vocal tract area functions from magnetic resonance imaging,” J. Acoust. Soc. Am., vol. 100, pp. 537–554, 1996.

[14] E. R. Weibel, *Morphometry of the human lung*, Berlin: Springer, 1963, pp. 136–140.

[15] I. Tokuda, M. Zemke, M. Kob, and H. Herzel, “Biomechanical modeling of register transitions and ’the role of vocal tract resonators,” J. Acoust. Soc. Am., vol. 127, pp. 1528–1536, 2010.