ARCHIVES OF ACOUSTICS
36, 1, 15–27 (2011) *DOI: 10.2478/v10168-011-0002-3*

A One-Mass Physical Model of the Vocal Folds with Seesaw-Like Oscillations

Tahir Mushtaq QURESHI, Khalid Saifullah SYED

Bahauddin Zakariya University
Center for Advance Studies in Pure and Applied Mathematics
Bosan Road, 60800, Multan, Pakistan
e-mail: tahmush@hotmail.com

(*received September 8, 2010; accepted January 10, 2011*)

A low-dimensional physical model of small-amplitude oscillations of the vocal folds is proposed here. The model is a simplified version of the body-cover one in which mucosal surface wave propagation has been approximated by the seesaw-like oscillation of the vocal fold about its fulcrum point whose position is adjustable in both the horizontal and vertical directions. This approach works for 180 degree phase difference between the glottal entry and exit displacements. The fulcrum point position has a significant role in determining the shape of the glottal flow. The vertical position of the fulcrum point determines the amplitude of the glottal exit displacement, while its horizontal position governs the shape and amplitude of the glottal flow. An increment in its horizontal position leads to an increase in the amplitude of the glottal flow and the time period of the opening and closing phases, as well as a decrease in the time period of the closed phase. The proposed model is validated by comparing its results with the low-dimensional mucosal surface wave propagation model.

Keywords: body-cover model, seesaw-like oscillation, glottal flow, vocal folds, spring-mass system, mucosal wave propagation.

1. Introduction

Vocal folds are known to be the source of sound and their motion has fundamental position in speech production. Research on glottal source modeling has a great significance in using voice source within speech synthesis systems and studying specific aspects of the mechanics of vibration. Parametric and physical models are the most widely known approaches for modeling of vocal folds.

In a parametric model the glottal flow is parameterized with the piecewise analytical functions. But in physical modeling approach the oscillatory characteristics are represented through numerical simulation based on the physical description of the vocal folds. Physical models provide realistic excitation signals that are easily incorporated into articulatory models of the vocal apparatus (Sondhi, 2002). The simplest lumped-element model is a one-mass model which simulates flow-induced oscillation of the vocal folds (Flanagan, Landgraf, 1968). In this model the vocal tract is not considered, which makes it impossible to achieve a self-sustained oscillation. The limitations of a one-mass model lead to two-mass model approach (Ishizaka, Flanagan, 1972), where two masses are coupled by stiffness which provides a basis for a two degrees of freedom model. In this model the interaction of the vocal tract is also considered, which causes several natural effects such as oscillatory ripples and skewing of the glottal flow. Several variations have been proposed in the literature after the introduction of one- and two-mass models (Ishizaka, Flanagan, 1977; Koizumi *et al.*, 1987; Titze, 1988; Liljencrants, 1991; Pelorson *et al.*, 1994; Story, Titze, 1995; Lous *et al.*, 1998; Titze, 1973; Avanzini *et al.*, 2001). Among the lumped models, multi-mass model has been proposed in the literature to increase the degrees of freedom for naturalness and accuracy of the glottal model (Titze, 1973). In the spring-mass system a difference in the hydrodynamic forces on the vocal folds induces them to oscillation. In order to simplify the fluid dynamics, Bernoulli equation is used to relate the pressure and flow within the glottis. Some literature is also available on non-steady behavior and collision of vocal folds (Pelorson *et al.*, 1994; Deverge *et al.*, 2003; de Vries *et al.*, 2002; Vilain *et al.*, 2004).

In 1974 Hirano proposed that a vocal fold comprises two tissue layers: a body layer and a cover one (Hirano, 1974). In view of this observation, the motion of a vocal fold may be regarded as the propagation of a surface wave along the body-cover from the bottom to the top of the glottis. Body-cover models based on this hypothesis were presented in (Titze, 1988; Story, Titze, 1995; Drioli, 2002).

The present work focuses on a simplified version of the body-cover model in which a transmission line is approximated by seesaw-like rotary motion about its fulcrum point adjustable in both the horizontal and vertical directions. The seesaw-like oscillation of glottis makes a 180 degrees phase delay between the entrance and the exit of the glottis. A mass-spring system is also attached to the glottis at its entrance to make it oscillate with the displacement of the mass. This work is similar to those proposed in (Liljencrants, 1991; Avanzini *et al.*, 2001; Drioli, 2002). In the present work, the introduction of the fulcrum point position is responsible for the control over the amplitude and the shape of the glottal flow.

The present section is followed by four more sections. In Sec. 2 we present the model with its geometrical view and derive the equation of motion with aerodynamic equations for the proposed model. Section 3 describes the method opted for the numerical solution of this system. Section 4 is devoted for the results and discussion. We draw a specific shape of the glottal flow with specific values of its parameters. The fulcrum point position has been shown to play a key role in the shape of the glottal flow. Section 5 contains conclusions.

2. The model

The glottal model presented here is a simplified version of the body-cover model in line with the model proposed in (Drioli, 2002). In the present model the propagation of mucosal wave is highly approximated by the seesaw-like rotary motion about its fulcrum point $P_0(X_0, Y_0)$ adjustable in both the horizontal and vertical directions. A single mass-spring system with mass M, spring constant K and damper B is attached to the glottis at its entrance as shown in Fig. 1. The model is assumed to be symmetric about its central line so that the left side of the vocal fold is the same as its right one. The glottal entry displacement is represented by X_1, while X_2 corresponds to the glottal exit displacement. The vertical component of the fulcrum point divides the glottis in the ratio s:t.

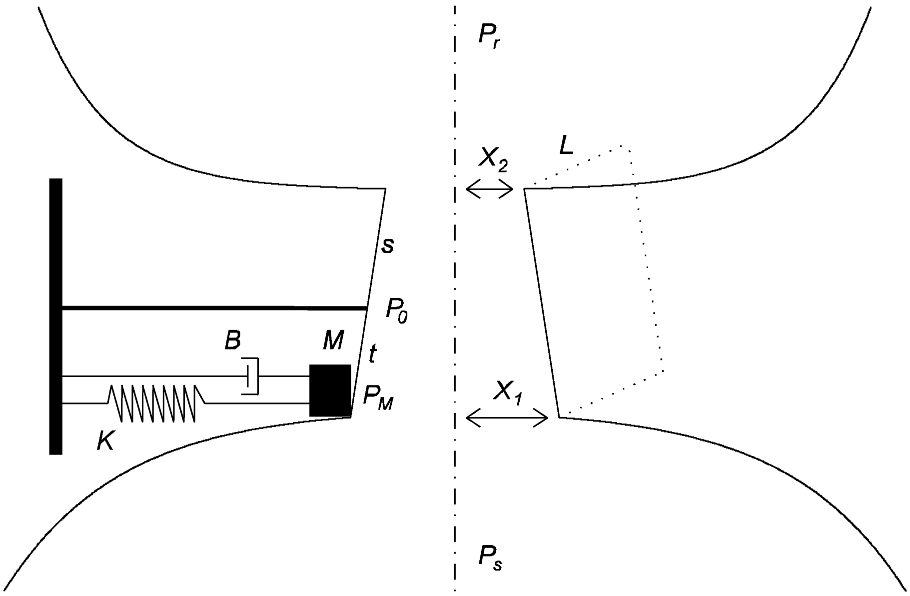

Fig. 1. A simplified body-cover model of the vocal folds.

The mass M is allowed to move only in the horizontal direction at the entrance of the glottis. Due to the oscillation of mass M the variation of the glottal entry displacement results into the variation of the glottal exit displacement triggered by seesaw-like oscillation of the glottis about its fulcrum point. The fulcrum point in the seesaw-like oscillation of the glottis has three implications. The first one is that the glottal exit displacement X_2 depends on the glottal entry displacement X_1, the horizontal component of the fulcrum point P_0 and the ratio in which the vertical component of the fulcrum point divides the glottis. A detailed description can be found in the Subsec. 2.1. The second one is that the amplitude of the glottal exit displacement can be controlled by the vertical position of the fulcrum point on the glottis as illustrated in Fig. 2. For example, when the fulcrum point is in the middle of the glottis both the glottal entry and exit displacements will have the same amplitude. If the fulcrum point position is nearer to the entrance of the glottis in such a way that it divides the glottis in the ratio $t{:}s = 1{:}2$, then the amplitude of the glottal exit displacement will be two times that of the glottal entry displacement. The third one corresponds to controlling the extreme values of the minimum cross-section area which can never be greater than the cross-sectional area at the fulcrum point position. This implication results into having control over the shape of the glottal pulse. It may be noted that the minimum cross-section area will always occur at the entrance when the glottis is in a divergent configuration and at the exit when the glottis is in a convergent one.

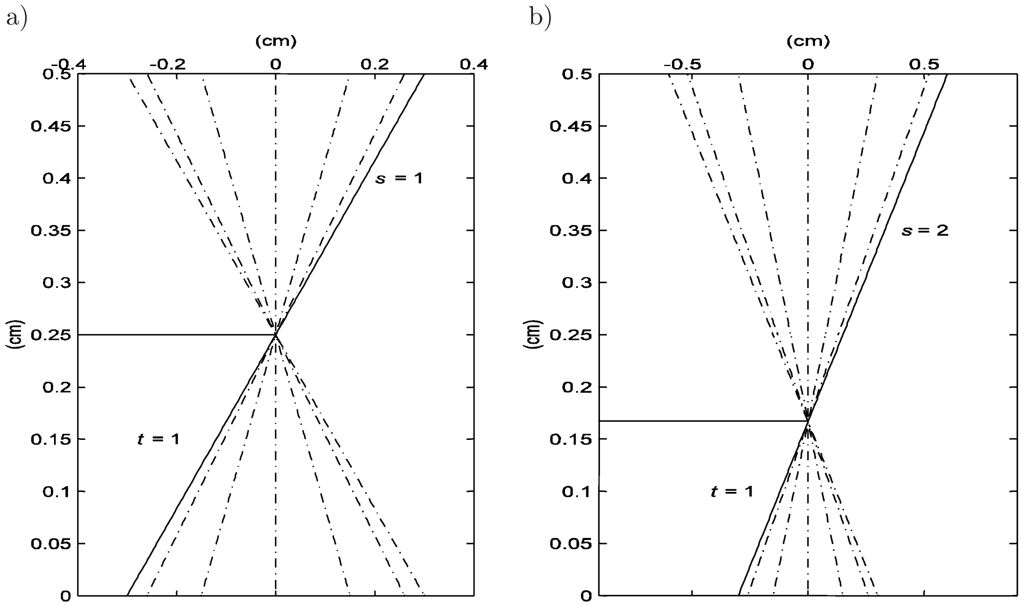

Fig. 2. Oscillation of the seesaw-like single vocal fold a) when the fulcrum point position is in the middle of the glottis surface, b) when the fulcrum point position divides the glottis surface in the ratio $t{:}s = 1{:}2$.

2.1. Equation of motion

As the oscillations of the vocal folds are assumed symmetric only half of the glottal region is modeled. The equation of motion for a single vocal fold is described by the spring-mass oscillator equation:

$$M\ddot{X}_1 + B\dot{X}_1 + KX_1 = F_M, \tag{1}$$

where M, B, K and X_1 are as described earlier, and F_M is the driving force for the oscillator.

We assume that the upper and lower edges of the seesaw-like glottis move only in the horizontal direction and remain constant in the vertical direction as shown in Fig. 2. In seesaw-like oscillation the fulcrum point is the mean position of its oscillation in the direction of motion. We suppose that the vertical component of the fulcrum point divides the glottis in the ratio $t{:}s = 1{:}1$ and zero is the mean position of the oscillation as depicted in Fig. 2a. Then the glottal exit displacement in terms of delayed version of the entrance of the glottis can be expressed as

$$X_2(t) = -X_1(t). \tag{2}$$

If the vertical component of the fulcrum point divides the glottis in the ratio $t{:}s$ or $r = s/t$, then the amplitude of the glottal exit displacement is r times the amplitude of the glottal entry displacement as demonstrated in Fig. 2b. In the present case, we have

$$X_2(t) = -rX_1(t). \tag{3}$$

Further we suppose that the mean position of the oscillation is X_0, which is specified by the horizontal position of the fulcrum point. The glottal exit displacement can then be represented as

$$X_2(t) = -r(X_1(t) - X_0). \tag{4}$$

Finally, the area of the entrance and the exit of the glottis can be expressed respectively as follows:

$$A_1(t) = \begin{cases} 2L(X_{01} + X_1(t)) & \text{if} \quad X_1(t) > -X_{01}, \\ 2LX_{01} & \text{otherwise;} \end{cases}$$

$$A_2(t) = \begin{cases} 2L(X_{02} - r(X_1(t) - X_0)) & \text{if} \quad X_2(t) > -X_{02}, \\ 2LX_{02} & \text{otherwise;} \end{cases} \tag{5$_1$}$$

or

$$A_2(t) = \begin{cases} 2L(X_{02} + X_2(t)) & \text{if} \quad X_2(t) > -X_{02}, \\ 2LX_{02} & \text{otherwise,} \end{cases} \tag{5$_2$}$$

where L is the length of the glottis, X_{01} and X_{02} are the rest positions of the fold at the entrance and the exit of the glottis.

2.2. Aerodynamic equations for small-amplitude oscillation of the vocal folds

In multi-mass models driving forces are computed for each mass separately but sometimes a mean glottal pressure is provided as the net driving pressure for the entire vocal tissue (Titze, 1988). We assume that fluid mechanical forces are exerted only on the first mass, while other masses are driven by the displacement of the first one (Pelorson *et al.*, 1994). Let P_M be the mean glottal pressure exerted over the fixed surface area S_M at the entrance of the glottis. The driving force F_M can then be written as

$$F_M = S_M \cdot P_M. \tag{6}$$

Let A_s and A_r be the constant sub- and supra-glottal areas respectively and the glottal flow is assumed to be incompressible, quasi-steady and one-dimensional so that the pressure distribution can be approximated with the help of Bernoulli equation. A mean glottal pressure P_M for linear glottis can be computed by the relation (Titze, 1988):

$$P_M = P_r + P_{k2}\left(1 - \frac{A_2}{A_1} - k_e\right), \tag{7}$$

where k_e is the pressure recovery coefficient (Ishizaka, Matsudaira, 1972), A_1 and A_2 are the glottal entry and exit areas respectively, P_r is the supra-glottal pressure and P_{k2} is the kinetic pressure at the glottal exit which is defined as (Titze, 1988)

$$P_{k2} = \frac{\rho}{2}\left(\frac{U}{A_2}\right)^2. \tag{8}$$

Here ρ is the air density and U is the volume velocity of the airflow.

The trans-glottal pressure can be approximated by the following relation (Titze, 1988):

$$P_s - P_r = k_t P_{k2}, \tag{9}$$

where P_s is the sub-glottal pressure and k_t is the trans-glottal pressure coefficient (Scherer, Titze, 1983).

By using Eqs. (7), (8) and (9) in Eq. (6), the driving force F_M can be expressed as

$$F_M = S_M\left[P_r + \left(\frac{P_s - P_r}{k_t}\right)\left(1 - \frac{A_2}{A_1} - k_e\right)\right]. \tag{10}$$

By substituting Eq. (8) into Eq. (9), the flow can be represented as

$$U = \sqrt{\frac{2(P_s - P_r)}{k_t\rho}}A_2. \tag{11}$$

For simplicity we assume that the glottis is not coupled with the vocal tract. This implies that the atmospheric pressure may serve as the vocal tract input pressure, i.e. $P_r = 0$. We also assume that the supra-glottal area A_r is much greater than the glottal exit area A_2, i.e. $k_e \cong 0$. Then Eqs. (10) and (11) can be rewritten as

$$F_M = \begin{cases} S_M\left[\dfrac{P_s}{k_t}\left(1 - \dfrac{A_2}{A_1}\right)\right] & \text{if } A_1 > A_2, \\ 0 & \text{if } A_1 \le A_2, \end{cases} \tag{12}$$

$$U = \sqrt{\frac{2P_s}{k_t\rho}}A_2. \tag{13}$$

3. Numerical simulations

The mechanical model is summarized in Eq. (1) which is a second order ordinary differential equation (ODE). This may be integrated numerically by using some suitable ODE integrator. In the present work a fourth order Runge-Kutta method has been employed to find numerical solution of the symmetric vocal folds model represented by Eq. (1). This numerical scheme has been implemented by developing our own Matlab program. The sampling rate is chosen to be 22.05 kHz in all simulations. Values of other parameters have been listed in Table 1.

Table 1. List of parameters and their range of values.

Parameters	Values
M	0.00016 kg
B	0 N·s·m^{-1}
L	0.014 m
ρ	1.15 kg·m^{-3}
K	80 N·m^{-1}
P_s	300–3000 Pa (Pascal)
X_{01}	0.005–0.1 mm
X_{02}	0.005–0.1 mm
X_0	0.006–3 mm

4. Results and discussion

In the present section we demonstrate how the proposed model works with different parametric values. The role of the vertical and horizontal position of the fulcrum point has been studied for controlling the shape of the glottal flow. For simplicity we assume in all simulations that $X_{01} = X_{02}$ and keep the sub-glottal pressure constant at $P_s = 300$ Pa (3.059 cm H_2O). This constant value of the sub-glottal pressure falls in the range considered in (Drioli, 2002). First of all we present the simulation of the model for the case when the vertical position of the fulcrum point is in the middle of the glottis, i.e. $r = 1$. The horizontal position of the fulcrum point is taken as $X_0 = 0.04$ cm. The rest positions of the glottal entry and exit have been taken as $X_{01} = X_{02} = 0.0005$ cm. Further we assume that the rest position of the vocal folds is the convergent configuration of the glottis. Figure 3 shows the simulation results of the proposed model for the present case.

Figure 3a shows five cycles of the vocal fold vibration which may be regarded as an approximation of the real glottal flow. During the opening and closing phases of the glottis cycle, air is allowed to pass through the glottis while the vocal folds are in contact with each other in the closed phase and stop the air from

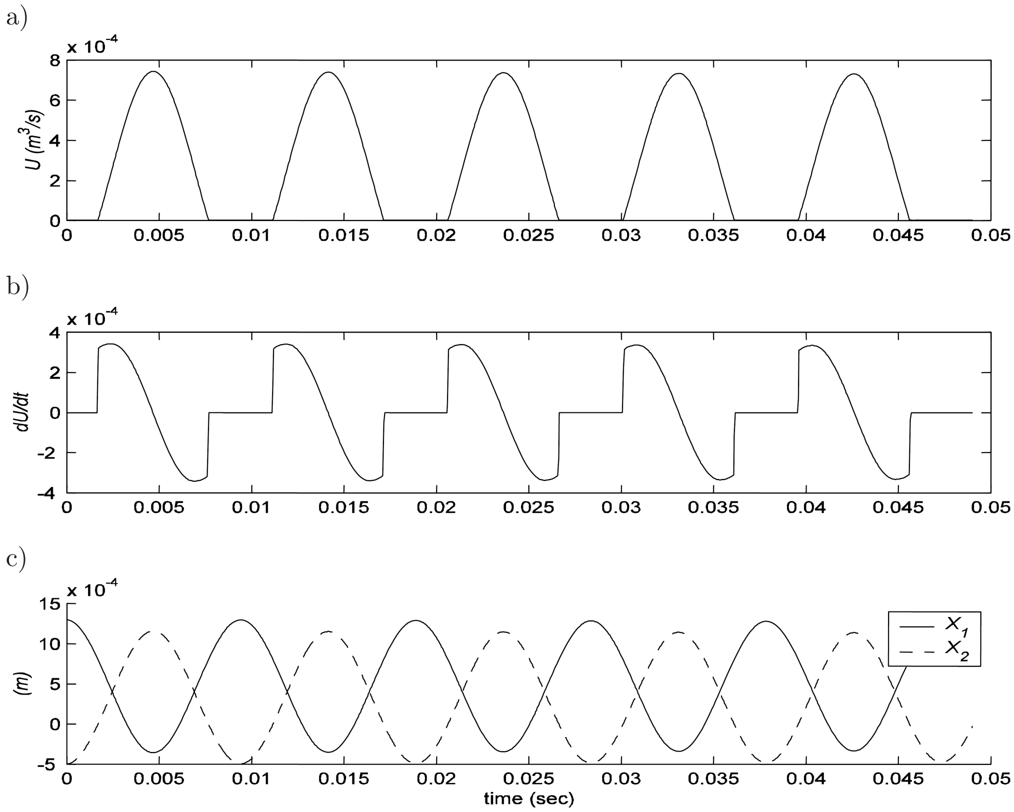

Fig. 3. Results of the simulation performed with the proposed model: a) the glottal volume flow, b) derivative of the glottal flow and c) displacements of the glottal entry and exit.

passing through the glottis. The derivative of the glottal flow is also presented in Fig. 3b. Figure 3c demonstrates the displacement of the glottal entry and exit during vibration, which is shown on the vertical axis with time presented at the horizontal axis. The negative displacement in the figure corresponds to collision of the vocal folds. We note that the glottal entry and exit displacements are in the phase difference of 180 degrees. Moreover, the glottal entry displacement is a little bit translated above the glottal exit displacement due to the influence of asymmetric force in the model.

Figure 4 demonstrates how the ratio r defined by the vertical position of the fulcrum point controls the amplitude of the glottal flow. With increment in the ratio r the amplitude of the displacement of the glottal exit increases as shown in Fig. 4c, which in turn increases the amplitude of the glottal flow as presented in Fig. 4a. We observe from Fig. 4b that the displacement of the glottal entry does not change during the variation of the ratio r because the displacement of the glottal entry is due to the displacement of the mass M which is attached to the spring K and the damper B as described earlier.

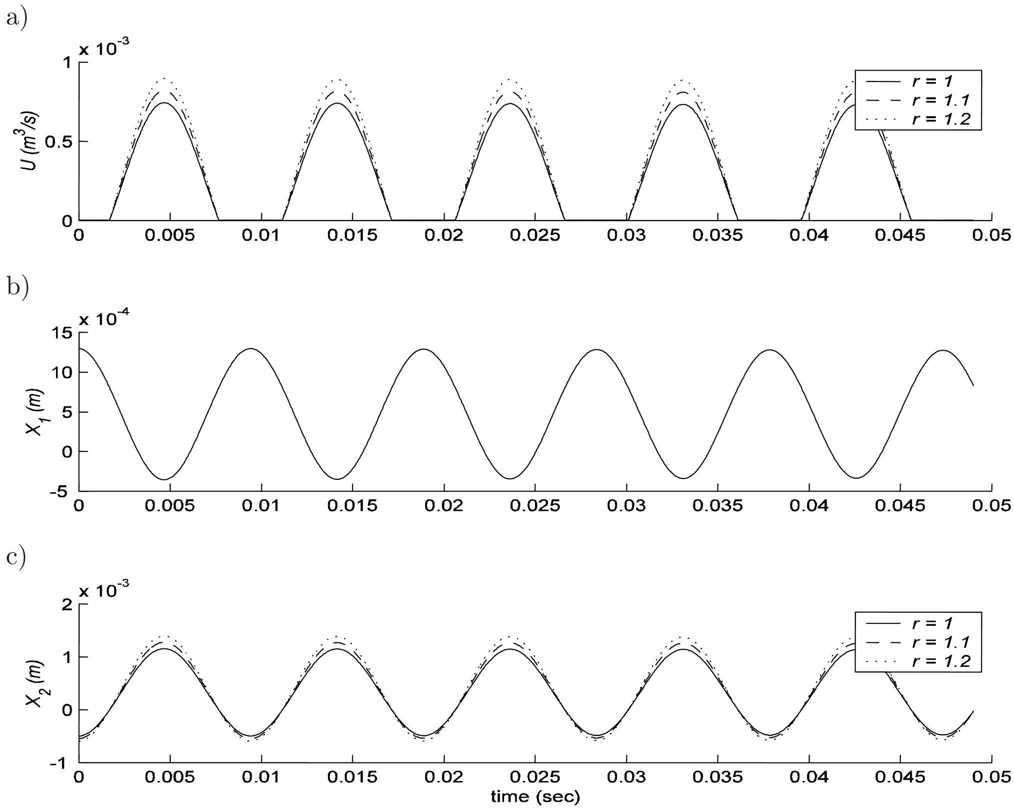

Fig. 4. Simulation of the model with variation in the ratio r: a) glottal volume flow with different values of r, b) displacement of the glottal entry and c) displacement of the glottal exit for different values of r.

On the other hand, the displacement of the glottal exit is due to the displacement of the glottal entry with respect to r. This case is well illustrated in Fig. 2.

Figure 5 represents the role of the horizontal position of the fulcrum point in generating different shapes of the glottal flow. When the fulcrum point is horizontally translated the displacements of the glottal entry and exit are also translated in the same direction as revealed by Figs. 5b and 5c. The translation of the fulcrum point in the horizontal direction has three implications in regard to the shape of the glottal flow: the first one is that it changes the amplitude of the glottal pulse; the second one is that it changes the opening and closing time periods of the glottal flow; and the third one is that it inversely changes the time period of the closed phase of the glottal flow. For example, if the horizontal position of the fulcrum point is increased, the amplitude, the opening and closing time periods of the glottal flow increase but the closed time period of the glottal flow decreases as shown in Fig. 5a. A very interesting and notable point is that

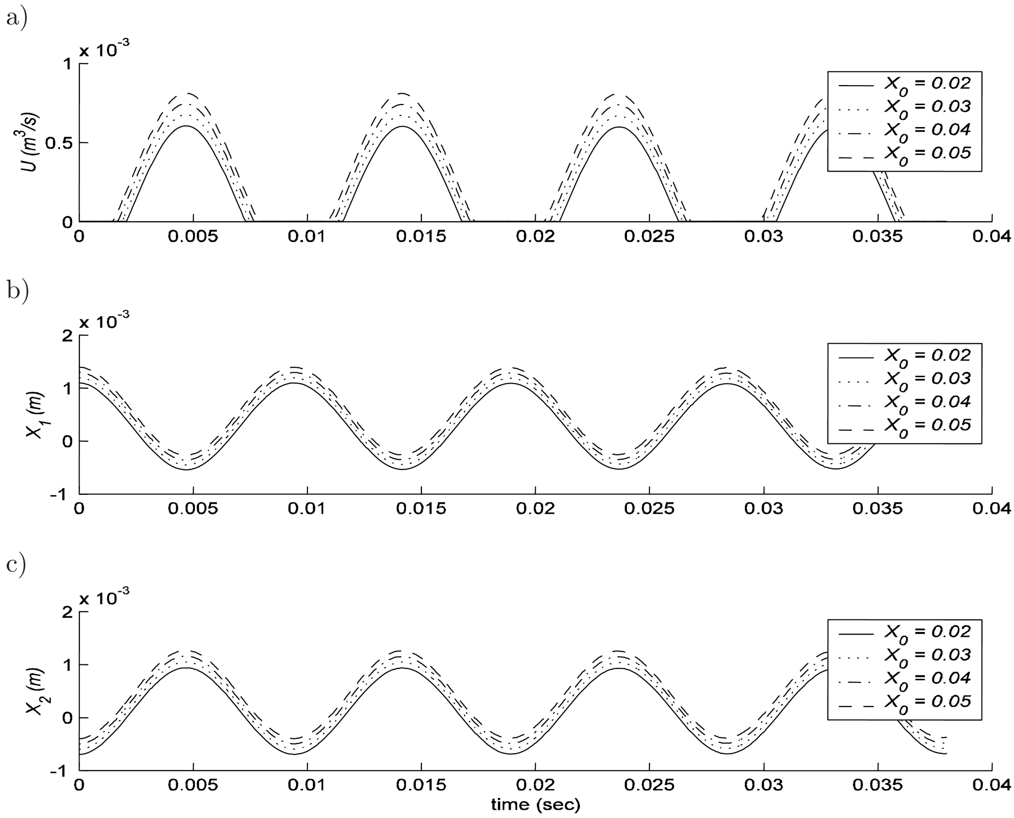

Fig. 5. Simulation of the model with variation in the horizontal position of the fulcrum point $P_0(X_0, Y_0)$: a) glottal volume flow for different values of X_0, b) displacement of the glottal entry for different values of X_0 and c) displacement of the glottal exit for different values of X_0.

the variations in the horizontal position of the fulcrum point bring about changes in the shape of the glottal flow but keep its pitch constant as shown in Fig. 5a.

We validate our proposed model by comparing it with the low-dimensional mucosal surface wave propagation model. In the proposed model the glottal exit displacement is expressed by the glottal entry displacement with phase difference of 180 degrees due to the oscillation of the seesaw-like glottis about its fulcrum point. We compare our results with the surface wave propagation method mentioned in (Titze, 1988; Drioli, 2002). In this approach the glottal exit displacement is represented by

$$X_2(t) = X_1(t) - \tau \dot{X}_1(t), \tag{14}$$

where τ is the time taken by the wave to propagate from the entrance to the exit of the glottis.

Figure 6 represents an excellent match of both the models. This demonstrates that the rotary motion of the seesaw-like glottis is well approximated by the

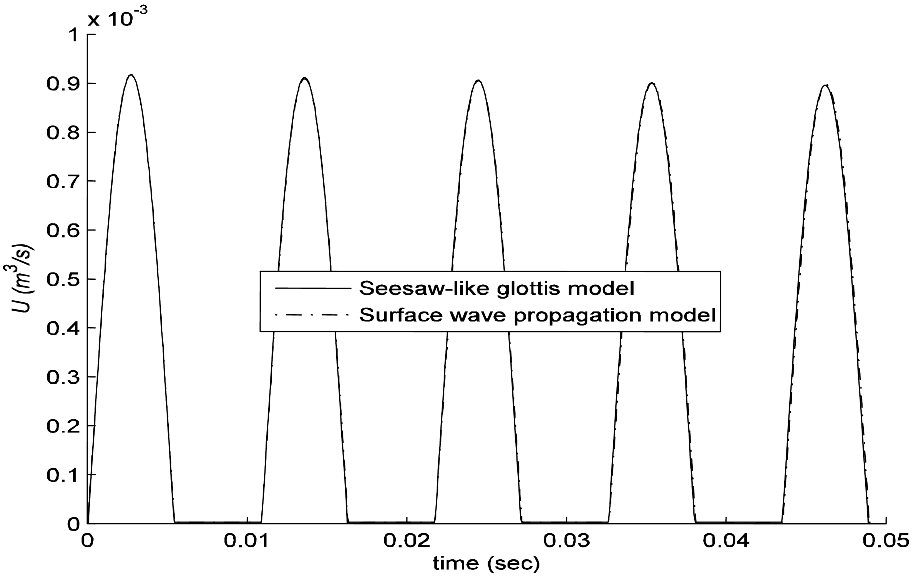

Fig. 6. A comparison of the proposed model with the surface wave propagation model.

surface wave propagation of the vocal folds in small amplitude oscillation. The parametric values taken by these two models have been mentioned in Table 2.

Table 2. Parametric values for seesaw-like and surface wave propagation vocal folds models.

Parameters	Seesaw-like glottis model	Surface wave propagation model	Units
M	0.00016	0.00016	kg
B	0	0.0065	$\mathrm{N{\cdot}s{\cdot}m^{-1}}$
L	0.014	0.014	m
ρ	1.15	1.15	$\mathrm{kg{\cdot}m^{-3}}$
K	86.8	80	$\mathrm{N{\cdot}m^{-1}}$
P_s	300	300	Pa (Pascal)
$X_{01} = X_{02}$	0.0005	0.0005	cm
X_0	0.006	–	cm
τ	–	0.00118	m/s
r	1.77	–	–

5. Conclusions

For small amplitude oscillations of the vocal fold a simplified version of the body-cover model has been developed. In this model mucosal surface wave propagation is approximated by the rotary motion of the seesaw-like glottis about its adjustable fulcrum point in both the horizontal and vertical directions. A spring-mass system is attached to the lower edge of the seesaw-like glottis to oscillate the displacement of the glottal entry which in turn oscillates the displacement of the glottal exit with a phase difference of 180 degrees. The mean pressure works as a driving force for the oscillation of the spring-mass system. With this proposed model we have investigated the effects of the horizontal and vertical positions of the fulcrum point on the shape of the glottal pulse. We observe that the vertical position of the fulcrum point affects the amplitude of the glottal exit displacement. That is the amplitude of the glottal exit displacement increases with an increase in the ratio in which the vertical position of the fulcrum point divides the glottis. A change in the horizontal position of the fulcrum point brings about a change in the shape of the glottal pulse. For example, an increment in the horizontal position of the fulcrum point leads to the following four changes in the glottal flow: it increases the amplitude of the glottal flow, the time period of the opening phase as well as the closing phase, and it decreases the time period of the closed phase. The present model has also been validated by comparing its results with a model based on the low-dimensional mucosal surface wave propagation for small amplitude oscillation of the vocal folds.

Our results lead to the following conclusions:

1. Seesaw-like design of the vocal fold with one spring-mass system can be successfully employed for glottal flow simulation.
2. Adjustable fulcrum point position provides a great design flexibility to govern glottal flow properties.
3. The extreme values of the minimum cross-sectional area can be controlled by adjusting of the horizontal position of the fulcrum point.
4. The proposed model has an advantage of controlling the entry and exit displacements with a phase difference of 180 degrees by only one degree of freedom.

Hopefully, this proposed approach may also trigger new series of research on vocal folds vibration in the form of seesaw-like vibration about its fulcrum point.

References

1. Avanzini F., Alku P., Karjalainen M. (2001), *One-delayed-mass model for efficient synthesis of glottal flow*, Proc. of Eurospeech Conf., pp. 51–54.

2. Deverge M., Pelorson X., Vilain C.E., Lagre'e P.-Y., Chentouf F., Willems J.F.H., Hirschberg A. (2003), *Influence of collision on the flow through in-vitro rigid models of the vocal folds*, Journal of the Acoustical Society of America, **114**, 3354–3362.

3. Drioli C. (2002), *A flow waveform adaptive mechanical glottal model*, Speech, Music and Hearing Quarterly Progress and Status Report, **43**, 69–79.

4. Flanagan J.L., Landgraf L.L. (1968), *Self-oscillating source for vocal tract synthesizers*, IEEE Transactions on Audio and Electroacoustics, **16**, 57–64.

5. Hirano M. (1974), *Morphological structure of the vocal cord as a vibrator and its variations*, Folia Phoniatrica, **26**, 89–94.

6. Ishizaka K., Flanagan J.L. (1972), *Synthesis of voiced sounds from a two mass model of the vocal cords*, Bell System Technical Journal, **51**, 1233–1268.

7. Ishizaka K., Flanagan J.L. (1977), *Acoustic properties of longitudinal displacement in vocal cord vibration*, Bell System Technical Journal, **56**, 889–918.

8. Ishizaka K., Matsudaira M. (1972), *Fluid mechanical considerations of vocal cord vibration*, *SCRL*-Monograph, Speech Communication Research Laboratory, Santa Barbara, Calif.

9. Koizumi T., Taniguchi S., Hiromitsu S. (1987), *Two-mass models of the vocal cords for natural sounding voice synthesis*, Journal of the Acoustical Society of America, **82**, 1179–1192.

10. Liljencrants J. (1991), *A translating and rotating mass model of the vocal folds*, Speech Transmission Laboratory-Quarterly Progress and Status Report, 1–18.

11. Lous N.J.C., Hofmans G.C.J., Veldhuis R.N.J., Hirschberg A. (1998), *A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design*, Acta Acustica (united with Acustica), **84**, 1135–1150.

12. Pelorson X., Hirschberg A., van Hassel R.R., Wijnands A.P.J. (1994), *Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model*, Journal of the Acoustical Society of America, **96**, 3416–3431.

13. Scherer R., Titze I.R. (1983), *Pressure flow relationships in a model of the laryngeal airway with a diverging glottis*, [in:] Vocal Fold Physiology: Current Research and Clinical Issues, D. Bless and J. Abbs [Eds.], pp. 179–193, College Hill, San Diego.

14. Sondhi M. (2002), *Articulatory modeling: a possible role in concatenative text-to-speech synthesis*, Proc. 2002 IEEE Workshop on Speech Synthesis, S. Monica (CA), 73–78.

15. Story B.H., Titze I.R. (1995), *Voice simulation with a body-cover mode of the vocal folds*, Journal of the Acoustical Society of America, **97**, 1249–1260.

16. Titze I.R. (1973), *The human vocal cords: A mathematical model, part I*, Phonetica, **28**, 129–170.

17. Titze I.R. (1988), *The physics of small amplitude oscillations of the vocal folds*, Journal of the Acoustical Society of America, **83**, 1536–1552.

18. Vilain C.E., Pelorson X., Fraysse C., Deverge M., Hirschberg A., Willems J.F.H. (2004), *Experimental validation of a quasi-steady theory for the flow through the glottis*, Journal of Sound and Vibration, **276**, 475–490.

19. de Vries M.P., Schutte H.K., Veldman A.E.P., Verkerke G.J. (2002), *Glottal flow through a two-mass model: Comparison of Navier–Stokes solutions with simplified models*, Journal of the Acoustical Society of America, **111**, 1847–1853.