Comments on single-mass models of vocal fold vibration

Richard S. McGowan
CReSS LLC, 1 Seaborn Place, Lexington, Massachusetts 02420

Michael S. Howe
College of Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Abstract: Proposed mechanisms for single-mass oscillation in the vocal tract are examined critically. There are two areas that distinguish single-mass models: in the sophistication of the air flow modeling near the oscillator and whether or not oscillation depends on acoustic feedback. Two recent models that do not depend on acoustic feedback are examined in detail. One model that depends on changing flow separation points is extended with approximate calculations.

© 2010 Acoustical Society of America

PACS numbers: 43.70.Gr [AL]
Date Received: February 17, 2010 **Date Accepted:** March 16, 2010

There has been substantial recent research activity on single-mass models of vocal fold vibration. Single-mass models could be more precisely termed single degree-of-freedom vibration models for the vocal folds, because each vocal fold is modeled with a single mass-spring system. Generally, it is assumed that the mass-spring systems modeling each vocal fold are identical, and the glottal flow is assumed to be the same on either side of the glottal centerline. These models are of particular interest theoretically because they must explain the net work done on the vocal folds by air flow in 1 cycle in terms of asymmetries between opening and closing phases in air flow conditions. That is, the flow asymmetries between the opening and closing phases of the vocal folds cannot be the result of asymmetries in the vocal fold geometry. This is in contrast to the two-mass model, which explains net energy input into the vocal folds using asymmetries in vocal fold geometry between opening and closing.

In this letter we will critically review the work done with single-mass models, which will reveal that, usually, acoustic loading effects above and below the vocal folds have been offered as mechanisms for net energy transfer from the air flow to the vocal folds for sustained single-mass oscillation. However, recent work offers other views on how single-mass oscillation could occur. Unless otherwise indicated, the models for vocal fold oscillation reviewed here assume that their motion is toward and away from the glottal and vocal tract centerline in a slab orthogonal to this centerline. A linear spring provides a restoring force to a vocal fold's displacement from its equilibrium position. The trap-door is an alternative geometry. In this configuration a vocal fold is assumed to oscillate about an axis at the wall of the vocal tract duct in rotational motion. A torsional spring is presumed to be attached that provides a restoring force if the trap-door is rotated from its rest position. The glottal area increases as the trap-door rotates downstream and decreases as it rotates upstream until it closes.

The first analysis of single-mass motion of vocal fold vibration known to us appears in the first volume of *The Journal of the Acoustical Society of America* by Wegel (1930). Wegel (1930), using a Lagrangian formulation, performs a linear stability analysis of flow through a general laryngeal system with each fold possessing one degree-of-freedom. "Gyrostatic" terms appear in the resulting equations, and they represent the modulating effect of the glottal aperture on the inertance of the glottal air flow. This permits a net energy flow from the air flow to the vocal folds over 1 cycle. Acoustic feedback is not necessary for oscillation; however, Wegel's (1930) Lagrangian formulation possesses an important flaw with the neglect of the energy source of the compression of air in the lungs.

Downloaded 13 May 2010 to 151.203.16.101. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

Flanagan and Landgraf (1968) use a one-mass model to excite a vocal tract computationally. They compute the static pressure on the masses representing the vocal folds to be the mean of the upstream and downstream pressures near the glottis. Essentially, this mean pressure is the subglottal pressure diminished by a constant multiple of $\rho_0 U_g^2/2A_g^2$, where ρ_0 is the density of air, U_g is the volume velocity of air through the glottis, and A_g is the cross-sectional area of the glottis. This diminishment of pressure in the glottal region is a model for total pressure loss at the glottal entrance and exit. Without a vocal tract coupled to this one-mass model, U_g depends on A_g only and there is no difference between opening and closing phases in the glottal pressure and vocal fold velocity as a function of glottal opening. This means that there can be no energy transfer between the air and the vocal folds over a single cycle of vocal fold oscillation.

On the other hand, the Flanagan and Landgraf (1968) one-mass model can sustain oscillation when a vocal tract is coupled to it and the fundamental frequency of vocal fold oscillation is less than the first resonance frequency of the vocal tract. In this case it is said that the glottal flow is driving an *inertive* load downstream of the glottis. Physically, this entails a delay in U_g with respect to A_g as the glottis opens than would otherwise be the case with no load. Further, as the glottal constriction closes, the flow decelerates more slowly than when no load is present (Titze, 2008). These phase effects can mean that there is a higher glottal static pressure for a given vocal fold opening speed than during the closing phase at the same vocal fold closing speed. Therefore there can be a net amount of work done on the vocal folds by the air flow over an oscillation cycle. Numerical simulations confirm this to be the case for this one-mass model (Flanagan and Landgraf, 1968).

There are some caveats to generalizing the results for this particular single-mass model. First, there must be a resistive component to the downstream load; otherwise the phase differences discussed above would collapse to zero. Further, although the inertive load downstream of the glottis appears to favor single-mass oscillation of many different types, including the trap-door configuration, the details of the flow model and its dependence on the motion of the oscillating structure cannot be ignored. There is no general theorem here that all types of oscillators will vibrate when an inertive load is placed downstream. For one thing, the magnitude of the work done on the structure must overcome the energy losses due to friction in the oscillator. The way that Flanagan and Landgraf (1968) calculate pressure in glottal region is somewhat *ad-hoc* in the sense that no vortex field is specified in the air flow calculations. It is possible to formulate a glottal pressure formula within the Flanagan and Landgraf (1968) model in which oscillation would not ensue.

Just as an inertive load tends to favor sustained single-mass oscillation when it is downstream of the glottis, the opposite phasing provided by capacitive loading upstream of the glottis tends to favor sustained oscillation (Titze, 2008). A capacitative load upstream of the glottis means that this region has its static pressure diminished more slowly as the glottis opens than would be the case with no load. Also, as the constriction closes, the region behind the glottis retains its lowered pressure as the capacitive load recharges with air. Again, a resistive component must also be part of the upstream load to ensure that the recharge time is not zero.

There have been models for single-mass oscillation with capacitive upstream loading offered in the speech production literature. In their linear analysis Gupta *et al.* (1973) show that a capacitive load enables oscillation as long as the mean flow velocity is large enough. McGowan (1992) argues that tongue-tip trills are made possible by a capacitive load supplied by the vocal tract wall. For the trap-door model of tongue-tip trill the result is a higher static pressure on the upstream side of the tongue-tip during opening than during closing for the same constriction area change speed. The air flow separates from the tongue tip in this model, so that the pressure on the downstream side of the tongue tip is always zero, in reference to atmospheric pressure. Thus the air performs net work on the tongue-tip in 1 cycle. The Gupta *et al.* (1973) and McGowan (1992) models, while specifying vortex fields, use crude approximations to calculate the air flow field near the oscillators.

Recent interest in one-mass vocal fold oscillation without acoustic loading effects has resulted from the experimental work of Park and Mongeau (2007). They examine the time variation of the glottal jet discharge coefficient C_d in two physical models of the vocal folds. The jet

Downloaded 13 May 2010 to 151.203.16.101. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

discharge coefficient is the product of two quantities $C_d = \sigma C_L$, where σ is the contraction ratio (the ratio of the downstream jet width to glottal width) and C_L is the loss coefficient that accounts for losses in volume flow due to viscous fluid effects, such as momentum diffusion. The experiments with the diverging glottal channel exhibit hysteretic effects in the discharge coefficient C_d (Park and Mongeau, 2007, Fig. 16). The discharge coefficient is smaller during the opening phase than during the closing phase when plotted against a Reynolds number that is based on the square root of glottal area and centerline glottal velocity.

An analytic model by Zañartu *et al.* (2007) for single-mass vocal fold oscillation grows out of the Park and Mongeau (2007) observations. The vocal folds are pictured as shutters that are rounded on the upstream side and squared-off on the downstream side. The flow is presumed to remain attached throughout the glottal channel bounded by the rounded vocal folds, and then to separate at the glottal exit where there is an abrupt change of area. These separation points do not change position during the vocal fold vibration. On the other hand, the discharge coefficient C_d is permitted to change value during the glottal cycle, so that it has the value 0.85 for the opening phase and 1.34 during the closing phase, based on the values measured by Park and Mongeau (2007). These discharge coefficients values are used as multiplicative factors to adjust the centerline air velocity at the glottal exit, which is calculated according to mass and momentum conservation. Under the quasi-steady approximation and along the centerline of the glottis, these conservation laws are expressed by the steady Bernoulli relation.

Zañartu *et al.* (2007) consider the case when both the subglottal and supraglottal ducts are semi-infinite, so that there are no reflected waves to interact with the vocal folds. In this case, for a constant transglottal pressure, vocal fold oscillation is sustained because of the time variable discharge coefficient. Without wave reflection, it can be assumed that the total pressure head is constant upstream to and into the glottis. Since the discharge coefficient is smaller during opening phase than during closing phase the particle velocity at any centerline position in the glottis is smaller during opening phase than during closing phase at any value of glottal area. Using an approximate calculation for streamline orientation this property is shown to hold for the fluid particle velocity at the vocal fold surfaces. Therefore, because total pressure head remains constant, the static pressure on the vocal folds must be greater during opening than during closing for all the degrees of glottal opening. Under these circumstances there will be net energy exchange from the air flow to the vocal folds. In this model the vortex field is specified, but approximations are employed to calculate air flow near the oscillator.

Howe and McGowan (2010) propose that the discharge coefficient $C_d = \sigma C_L$ differs in opening and closing phases in the experiments of Park and Mongeau (2007) because the expansion coefficient σ differs. The mechanism by which the expansion coefficient differs between opening and closing phases is through differences in the jet separation point positions between these two phases. While the folds are opening the separation points are on the upstream edge of the vocal folds. During the closing phase the flow remains attached until reaching downstream surface of each fold, where it does separate. In their model, each vocal fold is pictured as a rectangular shutter that is very thin in the direction of flow, so each fold can be idealized as an infinitely thin shutter for calculation purposes. When free streamline theory is employed to calculate the width of the jet emanating from the glottis, the contraction ratio σ is about 0.62 during the opening phase and is about 1.15 during the closing phase. The Howe and McGowan (2010) model does not account for the loss factor C_L in the discharge coefficient C_d. On the other hand, C_L does not appear to account for the hysteresis in C_d observed by Park and Mongeau (2007, Fig. 19b).

Forces tending to bring the shutters together are termed positive suction forces. In the Howe and McGowan (2010) model, when the shutters are opening and the separation points are on the upstream shutter surfaces, the pressure at the edges is atmospheric. Because the pressure is finite and the shutters are infinitely thin there is zero suction force on the shutters during the glottal opening phase. During the closing phase, the flow remains attached until it has turned the corner to the downstream side of each vocal fold. For the infinitely thin shutter, the pressure is infinitely negative at the edges of the shutters. But because the edges are infinitely thin, finite positive suction forces are generated on these edges. There must be such a forces on the flow

Downloaded 13 May 2010 to 151.203.16.101. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

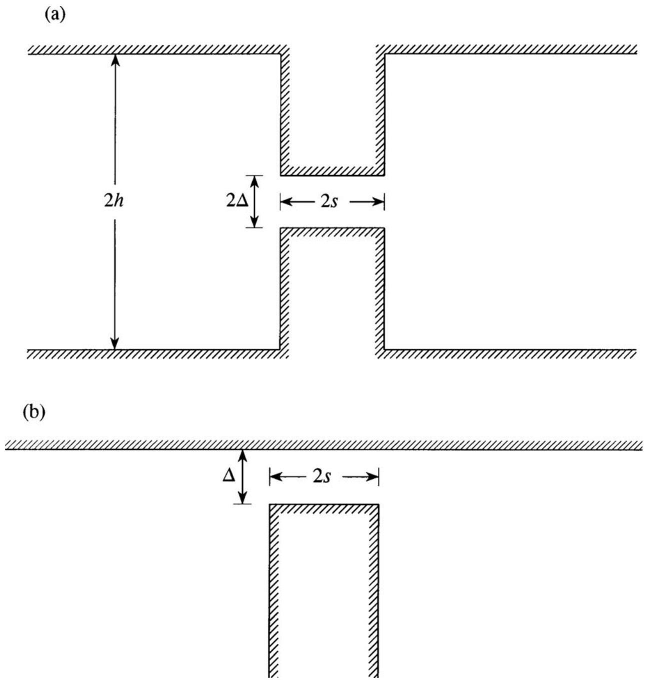

Fig. 1. (a) Geometry with shutters of finite thickness $2s$ in a duct of width $2h$. The shutters are separated by a distance of 2Δ, and (b) simplified geometry for suction force calculation on one shutter.

around the edge of each shutter; i.e., there are centripetal forces acting on the air and there are equal positive suction forces on the shutters (Howe, 2007, pp. 19, 118, and 119). This kind of suction force is part of the explanation for the Coanda effect in which a fluid stream tends to remain attached to a nozzle wall at an exit. Such forces are associated with the bending of flow streamlines around solid objects. With the vortex field specified, the flow around the oscillators is calculated using potential theory in the Howe and McGowan (2010) model.

The results in Howe and McGowan (2010) are for infinitely thin shutters, but here we offer some discussion of the effect of allowing the shutters to have finite thickness [Fig. 1(a)]. In the following we are interested in the magnitude of the suction force exerted on one of the shutters. As before, when the shutters are opening and the flow separates on the upstream sides, the suction force on one of the shutters is negative and equal to minus atmospheric pressure times the thickness of the shutter $2s$ times their depth ℓ_3. When the shutters are closing the separation points are assumed to move to the downstream faces of the shutters.

Instead of using free streamline theory to calculate the flow and suction force on one of the shutters during the closing phase, the calculation will be done for unseparated flow. The amount of suction force is overestimated doing this, because the curvatures of the flow streamlines is overestimated compared to their curvatures in separated flow. However, the complicated calculation using free streamline theory is not warranted for the conceptual points that are drawn in this letter. These simplified calculations provide an upper bound on the suction force that would be expected from flow that separates from the downstream face of the vocal folds.

In the case that $\Delta \ll h$ and $s \ll h$, for the purposes of the calculation of the suction force on one of the shutters, the geometry of Fig. 1(a) can be further simplified to an infinitely long shutter of finite thickness $2s$ perpendicular to and forming a gap of width Δ with a wall (i.e., glottal centerline) shown in Fig. 1(b). The potential flow field can be calculated using a conformal transformation from the upper half-plane with a source and a sink to the geometry of Fig. 1(b). In the following the flow will be presumed to be from upstream on the left to downstream on the right with velocity U in the duct. Specifically, a Schwartz–Christoffel mapping from the

Downloaded 13 May 2010 to 151.203.16.101. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

upper half ζ-plane to the z-plane of the shutter and wall is given by the derivative

$$\frac{dz}{d\zeta} = C\frac{\sqrt{\zeta^2 - 1}}{(\zeta^2 - \alpha^2)^{3/2}}, \quad (1)$$

where C and $\alpha > 1$ are constants. With $\beta = 1/\sqrt{\alpha}$ it can be shown that

$$\frac{\Delta}{s} = \frac{\int_0^{\pi/2} \sqrt{\beta + (1-\beta)\sin^2\theta}\, d\theta}{\beta \int_0^{\pi/2} \frac{\cos^2 \theta d\theta}{(1 - \beta \sin^2\theta)^{3/2}} d\theta}, \quad (2)$$

where $\beta \to 0$ corresponds to an infinitely thin shutter.

The suction force (above that with atmospheric pressure) on a shutter is

$$F = -\rho_0 \ell_3 \int_s^{-s} \frac{\partial w^*}{\partial t} dz^* - \frac{\rho_0}{2}\ell_3 \int_s^{-s} \left(\frac{dw}{dz}\right)^2 dz - 2s\ell_3 P, \quad (3)$$

where ρ_0 is the density of air, w is the complex velocity potential, and P is the total pressure (above atmospheric) in the duct (Howe, 2007, p 116). From Eqs. (1) and (3),

$$F = \frac{4}{\pi^2}\frac{(hU)^2}{\Delta}\mathcal{F}(\beta) - 2s\ell_3 P, \quad (4)$$

where $\mathcal{F}(\beta) = K(\beta)\int_0^{\pi/2}\sqrt{\beta + (1-\beta)\sin^2\theta}\, d\theta$ and K is the complete elliptic integral of the first kind. The first term of Eq. (3) is zero because of the symmetry of the flow. This term is related to the time derivative of the Rayleigh length of the glottis and the inertial term in the equation of fluid motion for the glottal jet. Even in the case where this term is nonzero it can be expected to be small, except, perhaps, during opening and closing of the glottis, because the Strouhal number of the oscillation is small.

The first term in Eq. (4) represents the suction force on a vocal fold caused by local flow effects [i.e., without the total pressure term, the second term in Eq. (4)], including streamline bending near and within the glottal channel. The suction force caused by local flow effects is often estimated by assuming straight flow through the glottal channel, without accounting for the bending of the streamlines at the ends of the channel. Figure 2(a) shows the ratio of the magnitude of the first term in Eq. (4) to the quantity $\rho_0(h/\Delta)^2 U^2 s\ell_3$, which is the corresponding expression for straight flow through the glottal channel. As might be expected, the bending of the streamlines is very important for glottal channels short compared to the glottal width (i.e., for $s \ll \Delta$), and it becomes insignificant in the other extreme (i.e., for $\Delta \ll s$). In all cases, as Fig. 2(a) shows, the bending of the streamlines increases the suction force on a shutter above what would be expected for straight flow through the glottis. Of course the suction force could be negative if there is a broad shutter because of the second term involving total pressure in Eq. (4). The effect total pressure is discussed below.

The fact that the calculation of suction force is performed for no flow separation means that the suction force has been overestimated compared to what would occur for flow that separates somewhere on the downstream side of the shutter. The results for $s/\Delta \to 0$ here can be compared to the Howe and McGowan (2010) calculation that uses free streamline theory for an infinitely thin shutter. The ratio of the suction force calculated by Howe and McGowan (2010) to that without flow separation is $D/3^{3/2}$ where $0.5 \lessapprox D \lessapprox 1$ for $\Delta \ll h$.

To make matters more explicit, an expression for the total pressure in the duct in terms of the duct flow velocity can be derived assuming that the flow through the glottis is established by constant volume velocity source far upstream of the shutters and that a sufficient time has passed so that transient effects have disappeared. For flow velocities corresponding to 10 cm H_2O lung pressure, the acoustic approximation can be made as follows, with c_0 = sound speed:

Downloaded 13 May 2010 to 151.203.16.101. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

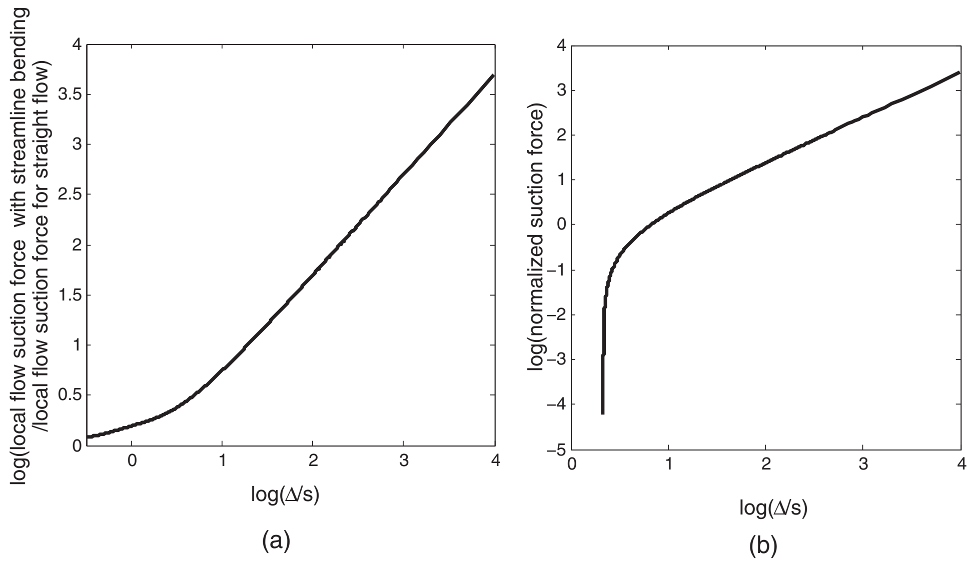

Fig. 2. (a) Logarithm of the ratio of suction force due to local flow with bending streamlines to suction force due to local flow with straight streamlines through the glottis as a function of $\log(\Delta/s)$, and (b) logarithm of the suction force above that caused by atmospheric pressure as a function of $\log(\Delta/s)$.

$$P = \rho_0 c_0 U\left(1 + \frac{U}{2c_0}\right) \approx \rho_0 c_0 U. \quad (5)$$

Using this expression and dividing Eq. (4) by $2s\ell_3 P$,

$$\frac{F}{2s\ell_3 P} = \frac{2}{\pi^2}\frac{\Delta}{s}\left(\frac{h}{\Delta}\right)^2\frac{U}{c_0}\mathcal{F}(\beta) - 1. \quad (6)$$

This normalized suction force as a function of Δ/s is shown in Fig. 2(b) with $h/\Delta=10$. There is positive suction force for $\Delta/s \gtrapprox 2.3$. This positive suction force increases as the shutters get thinner in relation to their distance to the glottal centerline. For more complex shutters with rounded ends, the *effective* ratio of Δ/s is greater than the ratio of measured minimum distance between the folds to the thickness of the folds, which will mean the suction force could remain positive for relatively thick folds, as long as the ends are rounded. On the other hand, the suction force has been overestimated in relation to what would be expected for separated flow in these calculations.

In summary, until recently, most single-mass models that have been offered involve acoustic feedback from resonators of the vocal system. There is an important caveat to statements that some types of acoustic loading enables single-mass oscillation: Each model must consider the details of the air flow near the vibrator. Two new models are now offered based on the experimental results of Park and Mongeau (2007), which measured hysteresis in the glottal jet discharge coefficient. These models are by Zañartu *et al.* (2007) and by Howe and McGowan (2010). The former model proposes differences between opening and closing phases of the glottis in glottal jet centerline velocities. The latter model proposes differences in the points of jet separation between the opening and closing phases. The effect of finite thickness of the vocal folds was also discussed in relation to the Howe and McGowan (2010) model.

Acknowledgment

This work was supported by subawards of Grant No. DC009229 from the National Institutes of Health to the University of California, Los Angeles.

Downloaded 13 May 2010 to 151.203.16.101. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

References and links

Flanagan, J. L., and Landgraf, L. L. (**1968**). "Self-oscillating source for vocal-tract synthesizers," IEEE Trans. Audio Electroacoust. **16**, 57–64.

Gupta, V., Wilson, T. A., and Beavers, G. S. (**1973**). "A model for vocal cord excitation," J. Acoust. Soc. Am. **54**, 1607–1617.

Howe, M. S. (**2007**). *Hydrodynamics and Sound* (Cambridge University Press, Cambridge).

Howe, M. S., and McGowan, R. S. (**2010**). "On the single-mass model of the vocal folds," Fluid Dyn. Res. **42**, 015001.

McGowan, R. S. (**1992**). "Tongue-tip trills and vocal-tract wall compliance," J. Acoust. Soc. Am. **91**, 2903–2910.

Park, B. P., and Mongeau, L. (**2007**). "Instantaneous orifice discharge coefficient of a physical model of the human larynx," J. Acoust. Soc. Am. **121**, 442–455.

Titze, I. R. (**2008**). "Non-linear source-filter coupling in phonation: Theory," J. Acoust. Soc. Am. **123**, 2733–2749.

Wegel, R. L. (**1930**). "Theory of vibration of the larynx," J. Acoust. Soc. Am. **1**, 1–21.

Zañartu, M., Mongeau, L., and Wodicka, G. R. (**2007**). "Influence of acoustic loading on an effective single mass model of the vocal folds," J. Acoust. Soc. Am. **121**, 1119–1129.

Downloaded 13 May 2010 to 151.203.16.101. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp