Available online at www.sciencedirect.com
ScienceDirect
Speech Communication 66 (2015) 176–181
SPEECH COMMUNICATION
www.elsevier.com/locate/specom

# Relaxation to one-dimensional postglottal flow in a vocal fold model

D. Sciamarella [a,*], G. Artana [b]

[a] LIMSI-CNRS, BP 133, F 91403 ORSAY Cedex, France
[b] LFD, Facultad de Ingeniería, Universidad de Buenos Aires, CONICET, Argentina

Received 11 July 2014; received in revised form 5 September 2014; accepted 4 November 2014
Available online 13 November 2014

**Abstract**

Postglottal flow in low-order dynamical systems modeling vocal fold motion is customarily considered one-dimensional. A relaxation distance is however mandatory before the flow effectively complies with this approximation. A continuous vocal fold model is used to show that this relaxation distance can impact voice simulation through the coupling strength between source and tract. The degree of interaction raises if relaxation occurs closer to the glottis, introducing complexity in the response of the system.


*Keywords:* Voice production; Vocal fold models; Non-linear source-filter coupling

## 1. Introduction

This work addresses an issue that has plagued speech modeling for a number of years, that is, coupling a flow solution that requires a finite distance for the flow and acoustics to reach one-dimensionality, with commonly-employed acoustic solvers that assume that one-dimensionality occurs instantly at the glottis.

Voice production can be modeled with different degrees of complexity. The essentials of the fluid–structure-acoustics interaction process can be captured by simple ordinary differential equation systems (ODEs), where the folds are represented by a mass-spring system, the fluid is represented by a quasi-parallel (1D) flow, and the acoustic source is represented by a plane wave emitter at the glottis (Sciamarella and Artana, 2009). The mucosal-wave model (Titze, 1988) is an example of the low-order modeling approach, in which the flapping motion of the vocal folds is condensed in one second-order ODE. This model, initially conceived for small amplitude oscillations, was later extended to account for large amplitude oscillations (Laje et al., 2001). In the extended version, an ad hoc nonlinear damping term was added in the ODE to account for an ensemble of effects ranging from the formation of the glottal jet to the saturation mechanism responsible for stopping the folds and interrupting the flow during vocal fold collision. The extended model has the particular advantage of being continuous: the returning points of the oscillation are included without resorting to piecewise functions. The approach, shown to produce vocal fold oscillation with physiologically realistic values for the parameters (Lucero, 2005) and also applied to labial oscillation modeling in birdsong (Laje and Mindlin, 2008), was employed to study the effect of source-tract coupling in phonation, *i.e.* of delayed feedback on vocal fold dynamics.

Feedback arises when the glottal system is coupled to the vocal tract and pressure reverberations are allowed to perturb vocal fold motion after a time delay given by sound speed and vocal tract length. The inclusion of this delay transforms the single ODE system into a DDE system (delay differential equation), endowing the simple oscillator with a complexity that can lead to subharmonic and non-periodic solutions (Laje et al., 2001).

In an application of the DDE system to source-tract interaction in birdsong (Laje and Mindlin, 2008), the

* Corresponding author.

http://dx.doi.org/10.1016/j.specom.2014.11.002
0167-6393/

transition zone between the avian source and the base of the tract is modeled in terms of characteristic distances which are redefined in this work for application to the case of human voice. A transition or relaxation distance separates the glottal outlet from the region where postglottal flow can be effectively considered 1-D. This distance is incorporated into the continuous vocal fold model, leading to an expression for the pressure perturbations that depends on this length scale.

This study considers the role of the relaxation distance in human voice production. Unlike many of the parameters involved in low-order vocal fold models, the finite distance required for the flow and acoustics to reach one-dimensionality has a direct physical correlate in the development of the glottal jet. It corresponds to the distance it takes the flow exiting the glottis to regain a unidirectional profile across the vocal tract section. Different values of this parameter are to be expected depending on the spreading rate of the jet and on the geometry of the jet-developing region – epilarynx tube and vocal tract (Titze, 2008). The spreading rate of a jet is known to depend on numerous parameters (Gutmark and Grinstein, 1999), such as Reynolds number, nozzle geometry and aspect-ratio. The pulsating nature of the glottal jet makes the scenario still more complex, because most of these parameters are time-varying. Moreover, the elongated geometry of the glottal outlet leads to spreading rates with initially opposed tendencies in the coronal and sagittal planes, that result in axis switching (Sciamarella et al., 2012). Recent in vitro studies (Krebs et al., 2012) are addressing the quantification of the full flow field in the proximity of the glottis, and therefore on the problem on which this work focuses, with simple modeling tools. Correlations will be proposed in this work with experimental data, in order to show how the solution is affected by measured variations in the development length of the flow.

The paper is organized as follows. Section 2 presents the derivation of the equation system modeling human voice with the relaxation length as an additional parameter, together with an analysis of the involved scales. Section 3 contains numerical examples showing how the model produces qualitatively different behavior for different values of the parameter. It also shows how solutions are affected if the relaxation length is time dependent. Conclusions are provided in Section 4.

## 2. The relaxation length in the model

Titze's flapping model (Titze, 1988) is based on the geometrical sketch of the vocal folds presented in Fig. 1. The glottal areas at entry, mid-height and exit respectively are:

$$\begin{aligned} a_1 &= 2L_g(x_{01} + x + \tau x') \\ a_g &= 2L_g((x_{01} + x_{02})/2 + x) \\ a_2 &= 2L_g(x_{02} + x - \tau x') \end{aligned} \quad (1)$$

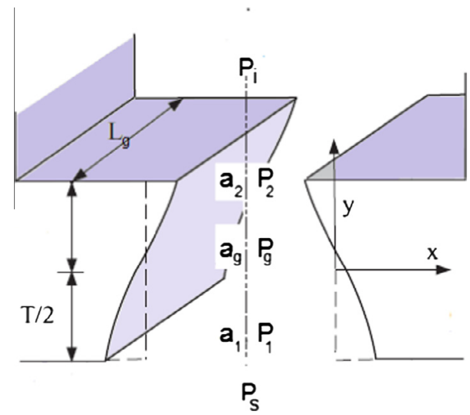



Fig. 1. Frontal section of the flapping model for the vocal folds.

where $x$ is the departure of the midpoint of the folds from the prephonatory profile, $L_g$ is the glottal length in the anteroposterior direction and $2\tau = T/c_w$ is the time it takes the surface wave to travel along the vocal fold body from bottom to top at speed $c_w$. The constants $x_{01}$ and $x_{02}$ correspond to the prephonatory positions, $\Delta x_0 \equiv x_{01} - x_{02}$. The equation describing the fluid–structure interaction is written by lumping the mechanical properties of the vocal fold tissue at the glottal midpoint:

$$Mx'' + Kx + f_d = P_i + (P_s - P_i)\frac{\Delta x_0 + 2\tau x'}{x_{01} + x + \tau x'} \quad (2)$$

where $M, K$ are the mass and stiffness (per unit area) of the vocal fold medial surface, $f_d$ is the dissipative force and pressures $P_s$ and $P_i$ stand respectively for the subglottal (lung) pressure, and for the input pressure at the vocal tract. The extension of the model by Laje et al. (2001) uses a nonlinear dissipative force that is quadratic in $(x - \bar{x})$, where $\bar{x}$ is the position of equilibrium. This allows for large amplitude oscillations since the squared term guarantees high dissipation every time the departure from the stationary position $\bar{x}$ is large.

$$f_d = B[1 + C(x - \bar{x})^2]x' \quad (3)$$

In this expression, $B$ is the damping per unit area and $C$ is a phenomenological coefficient. The effect of acoustic feedback is incorporated through the expression for $P_i = P_i(x, x')$. The pressure at the vocal tract input is composed of two parts: the forward-propagating perturbations generated by the time-varying flow injected by the vocal valve $s(t)$, and a backward-propagating sound wave $b(t)$ due to reflections occurring in the tract. As in Laje et al. (2001), the vocal tract is assumed to be a uniform tube of length $L$, so that:

$$P_i(t) = s(t) + RP_i(t - L/c_s) \quad (4)$$

where $c_s$ is the speed of sound and $R$ the reflection coefficient at the interface between the vocal tract end and the atmosphere. This simple boundary conditions assume that the wave propagating along the tract is a plane wave. Let us consider the region near the source, where the sound wave is not expected to be planar. Following Laje and Mindlin (2008), let us assume that the vocal valve is a local emitter of diverging spherical harmonic waves of the form:

$$s(r,t) = \frac{P_0}{r}\exp(i(\omega t - kr)) \tag{5}$$

As in Laje and Mindlin (2008), $P_0$ can be determined by the relationship between $s$ and $v$ (the particle velocity) for a spherical sound wave at a distance $d_1$ from the source:

$$\begin{aligned} s(d_1,t) &= Z(d_1)v(d_1,t) \\ Z(d_1) &= \frac{\rho((kd_1)^2 + i\omega d_1)}{(1+(kd_1)^2)} \end{aligned} \tag{6}$$

where $k = 2\pi f/c_s$ is the wave number and $f$ is the sound frequency. Combining Eq. (5) evaluated at $r = d_1$ and (6), we have $P_0 = \exp(-i(\omega t - kd_1))d_1 Z(d_1)v(d_1,t)$.

Let us now define $d_2$ as the relaxation distance where the one-dimensional flow approximation becomes valid, *i.e.* where pressure perturbations and particle velocity depend only on the streamwise coordinate, as illustrated in Fig. 2. Evaluating Eq. (5) at $r = d_2$, one obtains the expression that corresponds to $s(t)$ in Eq. (4). Since $v$ is also a harmonic function of time for a spherical harmonic sound disturbance, $v' = i\omega v$, and hence:

$$\begin{aligned} s(d_2,t) &= \frac{d_1}{d_2}Z(d_1)\exp(-ik(d_2-d_1)v(d_1,t)) \\ &= \mathcal{R}v(d_1,t) + \mathcal{I}v'(d_1,t) \end{aligned} \tag{7}$$

Let us now introduce some case-specific assumptions differing from those in Laje and Mindlin (2008). To apply the derivation to human voice, we will take as reference case the parameters used in Laje et al. (2001) to study coupling to the vocal tract. In this scenario, the resistive term $\mathcal{R}v$ is more than two orders of magnitude lower than $\mathcal{I}v'$. The same applies to the sine term (with respect to the cosine term) within $\mathcal{I}v'$, so that:

$$\mathcal{I}(d_1,d_2) \approx \frac{d_1^2\rho\cos[k(d_2-d_1)]}{d_2(1+(kd_1)^2)} \tag{8}$$

By flow conservation, $u(d_1,t) = A_s v(d_1,t)$ can be equated to $u_g(t) \approx a_g(t)v_g$ with $v_g^2 = 2P_s/\rho$:

$$v'(d_1,t) \approx \frac{a_g'(t)v_g}{A_s} = \frac{2L_g x'(t)\sqrt{2P_s/\rho}}{A_s} \tag{9}$$

This leads to an expression for the forward-traveling waves in terms of a coupling coefficient $\lambda$ depending on lengths $d_1$ and $d_2$:

$$s(d_2,t) = \lambda(d_1,d_2)x'(t) = \overline{\mathcal{I}}(d_1,d_2)2L_g\sqrt{2P_s/\rho}\,x'(t) \tag{10}$$

with $\overline{\mathcal{I}} = \mathcal{I}/A_s$. The insertion of $s(d_2,t)$ as $s(t)$ in (4) leads to a set of equations which explicitly include the distance it takes the flow to relax to a one-dimensional profile.

The values of the coupling factor $\overline{\mathcal{I}}(d_1,d_2)$ for the two length scales $d_1$ and $d_2$ are plotted in Fig. 3. As mentioned before, our default parameters are those of the extended model with coupling to vocal tract and feedback used in Laje et al. (2001). This includes a value of $\tau = 0.2$ ms, which corresponds to an oscillation regime with low tissue deformation. To evaluate $A_s$ in $\overline{\mathcal{I}}(d_1,d_2)$ we use $A_s = \pi(d_g/2)^2$ with $d_g = 1$ mm, the standard value for the mean glottal opening.

Normal speech values for the inertance are typically lower than 0.06 g/cm$^4$. Fig. 3 shows that for certain values of $d_1$ and $d_2$, the coupling factor $\overline{\mathcal{I}}$ attains values higher than ~0.08 g/cm$^4$, where tongues of subharmonic solutions are reported to occur (Laje et al., 2001). This indicates that $d_1$ and $d_2$ are potentially relevant scales in models implementing the simplest explanation for vocal fold oscillation with delayed feedback. Their values significantly affect the degree of coupling (or coupling strength) and can hence lead to qualitatively different dynamics, as will be shown in the next section.

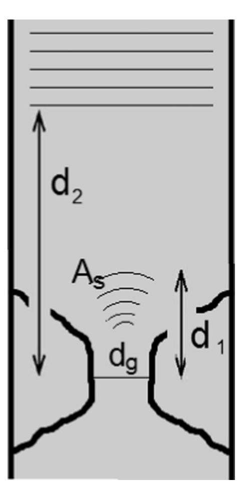



Fig. 2. Coupling between source and tract: $d_1$ indicates the length scale of spherical wave propagation and $d_2$ the length scale of the distance it takes the flow to become one-dimensional, $d_g$ stands for the mean glottal opening.

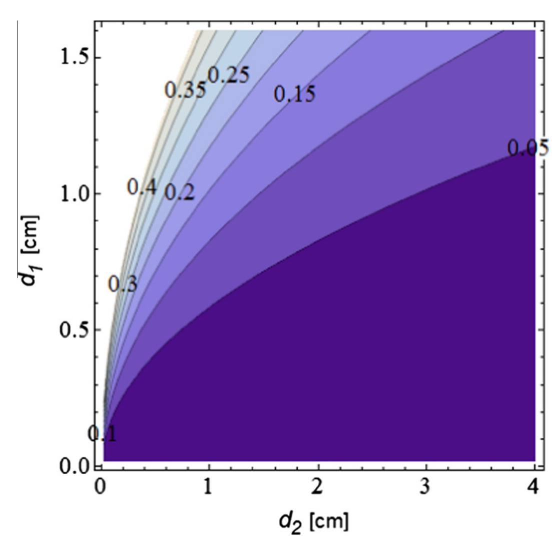
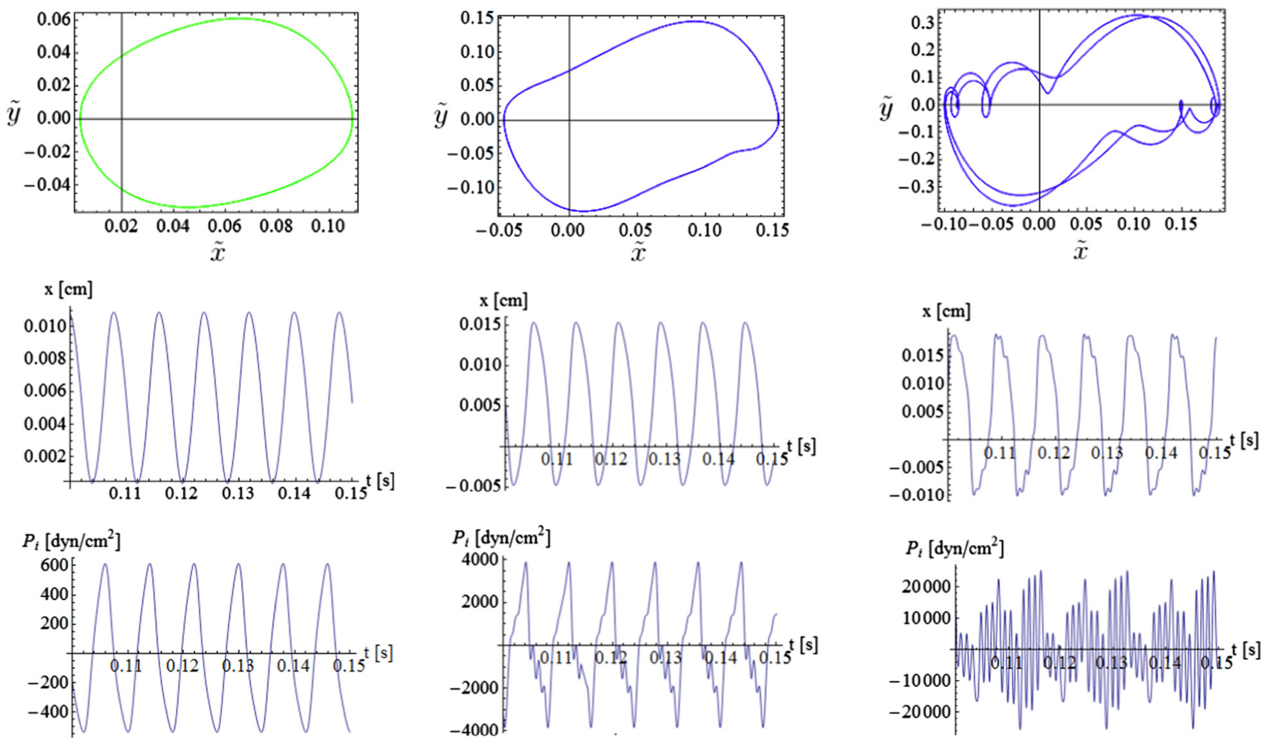



Fig. 3. Contour values of the coupling factor $\overline{\mathcal{I}}$ in (g/cm$^4$) as a function of $d_1$ and $d_2$ for the human voice reference case.




(a) $d_2 = 3\ cm$ (b) $d_2 = 1.2\ cm$ (c) $d_2 = 0.65\ cm$

Fig. 4. Solutions for different values of the relaxation distance $d_2$. The top pannel shows phase space $(\tilde{x}, \tilde{y})$, the middle pannel shows the time series for $x$, and the bottom pannel the time series for $P_i$.

## 3. Numerical examples

Let us now use the DDE system as a voice simulator for different values of the relaxation distance $d_2$. Variables are normalized using: $\tilde{x} = x/x_0 \equiv x/x_{01}$ and $\tilde{t} = tf \equiv t\sqrt{K/M}$ as in Lucero (2005). This yields:

$$\tilde{x}' = \tilde{y} \tag{11}$$

$$\tilde{y}' = -\alpha(1+\beta(\tilde{x}-\tilde{x}_0)^2)\tilde{y} - \tilde{x} + \widetilde{P}_i + (\widetilde{P_s} - \widetilde{P}_i)\frac{\Delta_0 + 2\delta\tilde{y}}{1+\tilde{x}+\delta\tilde{y}} \tag{12}$$

$$\tilde{P}_i(\tilde{t}) = \tilde{\lambda}(d_1, d_2)\tilde{y} + R\tilde{P}_i(\tilde{t} - \tau_1) \tag{13}$$

where the dimensionless control parameters are: $\alpha = B/f$, $\beta = Cx_0^2$, $\delta = \tau f$, $\Delta_0 = \Delta x_0/x_0$, $\tau_1 = L/(c_s f)$, $\widetilde{P}_{s,i} = P_{s,i}/(Kx_0)$, $\tilde{x_0} = \bar{x}/x_0$, and $\tilde{\lambda}(d_1, d_2)\sqrt{K\,M} = \lambda(d_1, d_2) = \overline{\mathcal{I}}(d_1, d_2)2L_g\sqrt{2P_s/\rho}$.

As mentioned in the previous section, our default parameter values correspond to the settings in Laje et al. (2001) to study delayed feedback: $K = 250$ kdyn/cm$^3$, $M = 0.45$ g/cm$^2$, $B = 100$ dyn/cm$^3$, $C = 10^5$ cm$^{-2}$, $x_{01} = 0.1$ cm, $x_{02} = 0.09$ cm, $P_s = 15$ kdyn/cm$^2$, $L_g = 1.4$ cm, $T = 0.4$ cm, $\tau = 0.2$ ms, $\rho = 0.00114$ g/cm$^3$, $R = -0.9$, $L = 17.5$ cm, $c_s = 3500$ cm/s, $d_1 = 1.5$ T.

Solutions corresponding to different values of $d_2$ are shown in Fig. 4. For each solution we display the oscillation in the $(\tilde{x}, \tilde{y})$ space, the time series for $x(t)$ and of the pressure fluctuations $P_i(t)$. As the relaxation length $d_2$ decreases, solutions become dynamically richer and $P_i$ is enhanced. The sequence shows three different solution types as $d_2$ is lowered: a quasi-sinusoidal oscillation, an asymmetric solution with skewing in the glottal pulse, and a period-2 solution with skewing and ripples in the glottal waveform.

Let us finally bring into consideration that $d_2$ is not a static parameter, since it depends on the dynamics of the glottal jet. Recent 3D experimental studies of the glottal jet (Krebs et al., 2012) show that the magnitudes quantifying its intra-cycle development are not constant. In consonance with these observations, let us conjecture that $d_2$ is correlated to the glottal flow pulsations given by $U(t) = U(x(t))$. This involves letting $d_2 = d_2(x(t))$.

Let us expressly assume that $d_2$ fluctuates as the flow rate increases and decreases. Since in this model $U(t) \propto x(t)$, let us pose $d_2(t) = d_{20} + fx_{01}\tilde{x}(t)$ with $f$ a phenomenological factor. To estimate this factor, we consider that $d_2$ advances (and recedes) as the jet axis-switching crossover position advances (and recedes) within the open phase of the cycle (Krebs et al., 2012). Let us set $f \sim 5$: this introduces an $\tilde{x}$ dependence in the delay equation for $\widetilde{P}_i(\tilde{t})$:

$$\widetilde{P}_i(\tilde{t}) = \tilde{\lambda}(d_2(\tilde{x}))\tilde{y} + R\tilde{P}_i(\tilde{t} - \tau_1) \tag{14}$$

Let us now numerically integrate the system formed by (11)–(14). Using $d_{20} = 0.65$ cm as in the last example of Fig. 4c, the solution that results retains some but not all of the features of the static 0.65 cm case. The rippled glottal

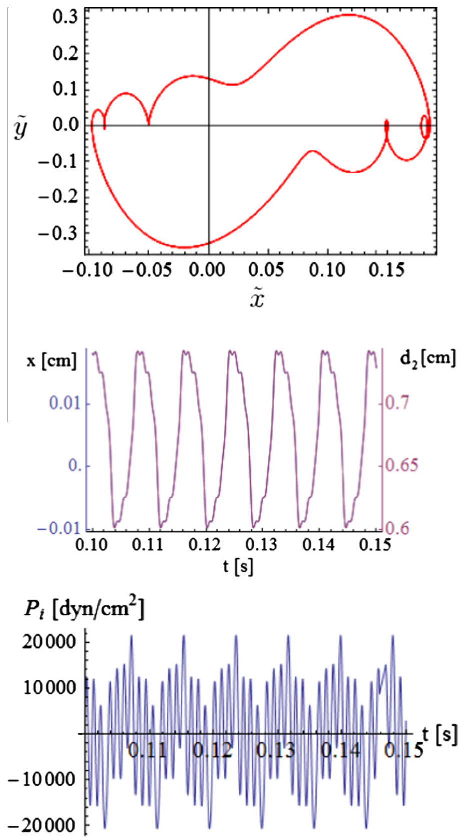



Fig. 5. Period-1 solution for $d_2 = 0.65\ \text{cm} + 5x(t)$. The top pannel shows phase space $(\tilde{x}, \tilde{y})$, the middle pannel shows the time series for $x$, and the bottom pannel the time series for $P_i$.

waveform is similar but the 2-period solution becomes a 1-period solution. The introduction in the model of a time-varying relaxation distance is thus found to help moderate the complexity of the solution (see Fig. 5).

## 4. Conclusions

Simple mathematical models are particularly useful to study in isolation the different sources of complexity that intervene in vocal fold behavior and hence in sound production. In studies devoted to the problem of source-tract (Laje and Mindlin, 2008) and source-source interaction (Laje et al., 2008) in oscine birds, coupling is treated with an approach that is more general than the traditional impedance approach. This rationale is applied back to speech communication in this work.

Following recent velocimetry measurements for the glottal jet (Krebs et al., 2012), it is postulated that glottal flow regains a quasi-one-dimensional profile across the tract at a finite distance from the source. Combining ingredients from human (Laje et al., 2001) and birdsong (Laje and Mindlin, 2008) toy models, a delayed differential equation system is obtained and used as a voice simulator to evaluate the influence of this relaxation distance. Numerical examples show that a significant sensitivity to this magnitude exists through the coupling strength.

The assumption of 1D postglottal flow is in fact an unphysical hypothesis. Our study shows that when feedback is taken into account, this common assumption is not deprived of consequences. For low values of the relaxation distance (*i.e.* comparable to the size of the vocal folds), the degree of coupling is shown to reach levels where subharmonic responses are possible. Delayed feedback is one of the possible mechanisms underlying subharmonicity, also reported to occur in the case of asymmetrical vocal fold motion (Steinecke and Herzel, 1995; Svec et al., 1996) or in the Kargyraa style of harmonic chant (Levin and Edgerton, 1999). Larger values of the relaxation distance, instead, are shown to undermine the enhancement of delayed feedback, moving the system back to normal vocal fold behavior – where subharmonic response is untypical. Moreover, an intra-cycle time-variation of the relaxation length (in phase with glottal flow, as indicated by experimental studies Krebs et al., 2012) can drive the response from the system back from a period-2 state to a period-1 solution, as shown by our numerical simulations.

In short, to the parameter set that controls the degree of source-tract interaction in phonation (including geometrical factors, such as the exact dimensions of the epilarynx), this work incorporates the role of commonly disregarded fluid dynamical factors, such as the non-planar characteristics and intra-cycle dynamics of postglottal flow: both are shown to have a moderating influence over coupling strength and therefore, over dynamical complexity. In other words, jet relaxation is found to provide a physical mechanism inducing a regularization of vocal fold behavior, there where the assumption of one-dimensionality occurring instantly at the glottis produces subharmonicity. There is still a paucity of information regarding the three-dimensional fluid dynamics of postglottal flow in phonation, which this work encourages to fulfill. An analysis of the manner in which the relaxation distance changes with subglottal pressure, stress–strain vocal fold characteristics, and other parameters controlling vocal fold motion would provide further insight on the effective role of this physical mechanism in real phonation.

## Acknowledgements

This research was supported by the 13STIC-08-P-MVP project of the SticAmSud program and by the LIA PMF-FMF (Franco-Argentinian Internacional Associated Laboratory in the Physics and Mechanics of Fluids).

## References

Sciamarella, D., Artana, G., 2009. A water hammer analysis of pressure and flow in the voice production system. Speech Commun. 51, 344–351.

Titze, I.R., 1988. The physics of small-amplitude oscillation of the vocal folds. J. Acoust. Soc. Am. 83 (4), 1536–1552.

Laje, R., Gardner, T., Mindlin, G.B., 2001. Continuous model for vocal fold oscillations to study the effect of feedback. Phys. Rev. E 64, 056201.

Lucero, J.C., 2005. Bifurcations and limit cycles in a model for a vocal-fold oscillator. Comm. Math. Sci. 3, 517–529.
Laje, R., Mindlin, G.B., 2008. Modeling source-source and source-filter acoustic interaction in birdsong. Phys. Rev. E 77, 011912.
Titze, I.R., 2008. Nonlinear source-filter coupling in phonation: theory. J. Acoust. Soc. Am. 123 (5), 2733–2749.
Gutmark, E.J., Grinstein, F.F., 1999. Flow control with noncircular jets. Ann. Rev. Fluid Mech. 31, 239–272.
Sciamarella, D., Silva, F., Artana, G., 2012. Similarity analysis of a glottal-like jet. Exp. Fluids 53 (3), 765–776.
Krebs, F., Silva, F., Sciamarella, D., Artana, G., 2012. A three-dimensional study of the glottal jet. Exp. Fluids 52, 1133–1147.
Laje, R., Sciamarella, D., Zanella, J., Mindlin, G.B., 2008. Bilateral source acoustic interaction in a syrinx model of an oscine bird. Phys. Rev. E 72, 036218.
Steinecke, I., Herzel, H., 1995. Bifurcations in an asymmetric vocal-fold model. J. Acoust. Soc. Am. 97 (3), 1874–1884.
Svec, J.G., Schutte, H.K., Miller, D.G., 1996. A subharmonic vibratory pattern in normal vocal folds. J. Speech Hear. Res 39, 135–143.
Levin, T., Edgerton, M., 1999. The throat singers of tuva. Sci. Am. 281 (3), 70–77.