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:

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

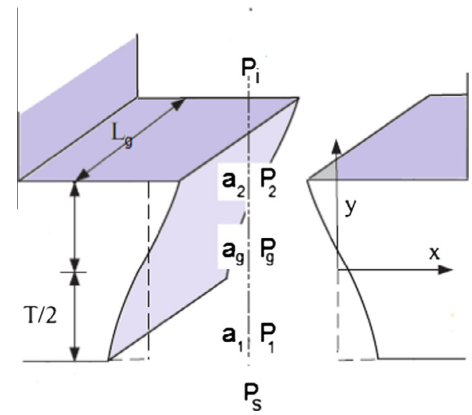

Fig. 1. Frontal section of the flapping model for the vocal folds.

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