INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

A unified numerical simulation of vowel production that comprises phonation and the emitted sound

Niyazi Cem Degirmenci[1], *Johan Jansson*[2,1], *Johan Hoffman*[1,2], *Marc Arnela*[3], *Patricia Sánchez-Martín*[3], *Oriol Guasch*[3], *Sten Ternström*[4]

[1]Department of Computational Science and Technology, School of Computer Science and Communication, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden
[2]BCAM- Basque Center for Applied Mathematics Mazarredo 14, 48009 Bilbao, Spain
[3]GTM Grup de recerca en Tecnologies Mèdia, La Salle, Universitat Ramon Llull, C/ Quatre Camins 30, Barcelona 08022, Catalonia, Spain
[4]Department of Speech, Music and Hearing, School of Computer Science and Communication, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden

ncde@kth.se, jjan@kth.se, jhoffman@kth.se, marnela@salle.url.edu, psanchez@salle.url.edu, oguasch@salle.url.edu, stern@kth.se

Abstract

A unified approach for the numerical simulation of vowels is presented, which accounts for the self-oscillations of the vocal folds including contact, the generation of acoustic waves and their propagation through the vocal tract, and the sound emission outwards the mouth. A monolithic incompressible fluid-structure interaction model is used to simulate the interaction between the glottal jet and the vocal folds, whereas the contact model is addressed by means of a level set application of the Eikonal equation. The coupling with acoustics is done through an acoustic analogy stemming from a simplification of the acoustic perturbation equations. This coupling is one-way in the sense that there is no feedback from the acoustics to the flow and mechanical fields.

All the involved equations are solved together at each time step and in a single computational run, using the finite element method (FEM). As an application, the production of vowel [i] has been addressed. Despite the complexity of all physical phenomena to be simulated simultaneously, which requires resorting to massively parallel computing, the formant locations of vowel [i] have been well recovered.

Index Terms: Numerical voice production, phonation, vocal tract acoustics, fluid-structure interaction, finite element method

1. Introduction

The physics of voice production is rather intricate and involves many different phenomena (see e.g., [1] for a review on the essentials and [2] for a comprehensive introduction to the topic). In a nutshell, the air emanating from the lungs impinges on the vocal folds (VF) and separates them apart until their elasticity takes over and the VF close again. The pressure at the glottis decreases due to the jet flow in the vocal tract (VT) by Bernoullis law and a self sustained oscillation is established. The flow dynamics and VF vibrations result in acoustic sources of monopolar, dipolar and quadrupolar character [3]. Acoustic waves are generated, propagate through the VT and become finally radiated outwards. In the case of vowels, the VT resonances (formants) get excited and this is what actually allows one to distinguish one vowel sound from another.

The numerical simulation of all the above phenomena poses a big challenge. Therefore, researchers have traditionally split the problem by either just focusing on the process of phonation, or solely addressing VT acoustics. Direct numerical simulations (DNS) have been achieved for the former (see e.g., [4, 5]) though at the price of artificially reducing the Reynolds number of the problem. An immersed boundary method is adopted in those works and contact is enforced by a kinematic constraint. With regard to VT acoustics, several works have been performed to date for static vowels sounds (see e.g., [6, 7, 8, 9, 10]), and not long ago for dynamic vowel sounds as well [11].

Recently, unified simulations of phonation and VT acoustics have also been attempted [12, 13, 14]. However, due to the high computational cost a two step hybrid approach is pursued in those works: the VF movement being obtained from a first FSI simulation, and then prescribed in the computation of the acoustic field.

In this work we endeavor a unified simulation of phonation and acoustics. At each time step of the computation an FSI problem is solved for the self oscillations and contact of the VF, following the strategy in [15]. Then a source term is computed and used as an inhomogeneous volume source for the wave equation in mixed form, in an arbitrary Lagrangian-Eulerian (ALE) frame of reference [11]. The implemented acoustic analogy may be viewed as a simplification of the acoustic perturbation equations for negligible convection velocity [16, 17, 18], which filters pseudo-sound to some extent. All the involved equations have been implemented in the open source finite element framework Unicorn [15] from the FEniCS-HPC project.

As a model problem we have chosen the production of vowel [i]. A correct reproduction of the formant locations is reported despite the complexity of all simulated phenomena and the simplifications introduced in our model. This is encouraging as there is room for its improvement considering multi-layer VFs, more detailed contact and acoustic analogy models, etc.

The outline of this report is as follows: in Section 2, we present the 3D problem to be solved. In Section 3 we describe the mathematical model, with subsections focused on the FSI model, the contact model, the acoustic coupling, the numerical implementation and some simulation details. Results are depicted in Section 4. They comprise the outputs from the FSI simulation with contact, the acoustic source term and finally, the generation of vowel [i]. Section 5 closes the paper with some conclusions and a discussion of future work.

http://dx.doi.org/10.21437/Interspeech.2017-1239

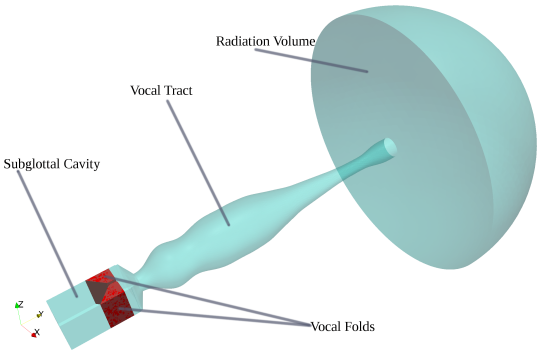

Figure 1: *Computational domain for the unified numerical production of vowel [i].*

2. Problem Statement

We consider the computational domain Ω shown in Figure 1, which we have made publicly available at [19]. The domain consists of a rectangular channel with dimensions $0.0194 \times 0.0384 \times 0.015$, which embeds the subglottal cavity and the vocal folds. This is connected to a circular vocal tract [20] built from the cross-sectional areas in [21]. The VT has length 0.155 and we have attached a half sphere of radius 0.072 to its end to account for free-field sound radiation outside the mouth. SI units are assumed throughout the paper.

A velocity is prescribed at the inlet of the subglottal cavity which triggers the VF folds self oscillations after some time steps. This results in the generation of acoustic waves that propagate inside the VT, get partially reflected, generating stationary waves, and get partially transmitted outside the mouth. The mathematical models used to account for all these phenomena will be next presented.

3. Mathematical Modelling

As described in the introduction, the mathematical modelling of the above physical phenomena essentially involves a full coupling between the mechanical and flow fields, and a one way coupling with the acoustic field by means of an acoustic analogy.

3.1. The self oscillations of the vocal folds

For robustness we have chosen a monolithic approach to FSI, which we derive from the basic conservation laws. We seek for a velocity $\boldsymbol{u}$, a density ρ and a phase function θ to track the fluid (glottal flow) and structure (VF) phases. We also introduce the unified Cauchy stress $\boldsymbol{\sigma}$ for all phases. The incompressible Unified Continuum fluid-structure model in ALE coordinates reads [22],

$$\rho(\partial_t \boldsymbol{u} + ((\boldsymbol{u}-\boldsymbol{\beta})\cdot\nabla)\boldsymbol{u}) + \nabla\cdot\boldsymbol{\sigma} = 0 \quad \text{in } Q, \tag{1}$$

$$\nabla\cdot\boldsymbol{u} = 0 \quad \text{in } Q, \tag{2}$$

$$\partial_t\theta + ((\boldsymbol{u}-\boldsymbol{\beta})\cdot\nabla)\theta + \theta^c = 0 \quad \text{in } Q, \tag{3}$$

where the phase function θ defines the solid and fluid domains

$$\begin{aligned}\Omega_s(t) &= \{\boldsymbol{x} : \boldsymbol{x}\in\Omega,\ \theta(\boldsymbol{x},t) = 0\},\\ \Omega_f(t) &= \{\boldsymbol{x} : \boldsymbol{x}\in\Omega,\ \theta(\boldsymbol{x},t) = 1\},\end{aligned} \tag{4}$$

and θ^c will be described in the next subsection.

Here $\Omega \subset \mathbb{R}^3$ is the spatial domain and $Q = \Omega \times I$ is a space-time domain with $I = [0,T]$ a time interval. Subtracting the pressure p from the Cauchy stress, and using the phase function θ, we can define separate constitutive laws for the two phases. Namely, a Newtonian fluid law for the stress $\bar{\boldsymbol{\sigma}}_f$ in the fluid phase and a Neo-Hookean law in rate form, for the stress $\bar{\boldsymbol{\sigma}}_s$ in the solid phase. We obtain,

$$\rho = \theta\rho_f + (1-\theta)\rho_s, \tag{5}$$

$$\boldsymbol{\sigma} = \bar{\boldsymbol{\sigma}} - p\mathbb{1}, \tag{6}$$

$$\bar{\boldsymbol{\sigma}} = \theta\bar{\boldsymbol{\sigma}}_f + (1-\theta)\bar{\boldsymbol{\sigma}}_s, \tag{7}$$

$$\partial_t\bar{\boldsymbol{\sigma}}_s = 2\mu_s\epsilon + \nabla\boldsymbol{u}\bar{\boldsymbol{\sigma}}_s + \bar{\boldsymbol{\sigma}}_s\nabla\boldsymbol{u}^\top, \tag{8}$$

$$\bar{\boldsymbol{\sigma}}_f = 2\mu_f\boldsymbol{\epsilon}, \tag{9}$$

where μ_s is the shear modulus (solid) and μ_f the dynamic viscosity (fluid). $\mathbb{1}$ is the identity tensor and $\boldsymbol{\epsilon} = \frac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^\top)$. $\boldsymbol{\beta}(\boldsymbol{x},t)$ in (1)-(3) stands for the velocity of the ALE coordinate system relative to the Eulerian coordinate system, which equals $\boldsymbol{u}$ in the structure domain. $\boldsymbol{\beta}$ in the fluid domain is arbitrary and used for improving the mesh quality while running the numerical simulation.

The equations (1)-(3) for the FSI problem are supplemented with the following boundary and initial conditions for $\boldsymbol{x} = (x,y,z)$ resulting in a Reynolds number $Re \sim 1200$,

$$\begin{aligned}&\boldsymbol{u}(\boldsymbol{x},t) = (0,0.5,0) && \boldsymbol{x}\in\{\boldsymbol{x} : \boldsymbol{x}\in\partial\Omega, y = 0\},\\ &\boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{0} && \boldsymbol{x}\in\{\boldsymbol{x} : \boldsymbol{x}\in\partial\Omega, y > 0, y < 0.2\},\\ &p(\boldsymbol{x},t) = 0 && \boldsymbol{x}\in\{\boldsymbol{x} : \boldsymbol{x}\in\partial\Omega, y \geq 0.2\},\\ &\boldsymbol{u}(\boldsymbol{x},0) = \boldsymbol{0},\\ &\bar{\boldsymbol{\sigma}}_s(\boldsymbol{x},0) = \boldsymbol{0}.\end{aligned} \tag{10}$$

3.2. The contact model

To resolve the contact phenomenon, the term θ^c is introduced in (3) to switch the continuum between fluid and solid, according to the results of the Eikonal equation.

Assuming an initial state without contact, we define

$$\Omega_{f0}(t) := \{\boldsymbol{x} : \boldsymbol{x}\in\Omega,\ \theta(\boldsymbol{x},0) = 1\}. \tag{11}$$

For each $t\in[0,T]$, let $D_b^t : \Omega_{f0}(t) \to \mathbb{R}$ denote the distance from the boundaries,

$$|\nabla D_b^t(\boldsymbol{x})| = 1, \quad \boldsymbol{x}\in\Omega_{f0}(t) \tag{12}$$

$$D_b^t = 0, \quad \boldsymbol{x}\in\partial\Omega_{f0}, \tag{13}$$

and let $M^t : \Omega_{f0}(t) \to \mathbb{R}$ be the medial axis,

$$M^t(\boldsymbol{x}) := \begin{cases}1 & \text{if } |\nabla D_b^t(\boldsymbol{x})| < \alpha, D_b^t(\boldsymbol{x}) < \beta,\\ 0 & \text{otherwise .}\end{cases} \tag{14}$$

We next introduce $D_m^t : \Omega_{f0}(t) \to \mathbb{R}$ as the distance from the medial axis,

$$|\nabla D_m^t(\boldsymbol{x})| = 1, \quad \boldsymbol{x}\in\Omega_{f0}(t), \tag{15}$$

$$D_m^t = 0, \quad \boldsymbol{x}\in\{\boldsymbol{x} : \boldsymbol{x}\in\Omega_{f0}(t),\ M^t(\boldsymbol{x}) = 1\}, \tag{16}$$

and finally attain an expression for the phase change $\theta^c : (\Omega,t) \to \mathbb{R}$ in (3), assuming a maximal oscillation frequency of $1/\delta$ ($\delta << 1$),

$$\theta^c(\boldsymbol{x},t) := \begin{cases}-1/\delta & \text{if } \boldsymbol{x}\in\Omega_{f0}(t),\ D_m^t < \gamma,\ \theta(\boldsymbol{x},t-\delta) = 1,\\ 1/\delta & \text{if } \boldsymbol{x}\in\Omega_{f0}(t),\ D_m^t \geq \gamma,\ \theta(\boldsymbol{x},t-\delta) = 0,\\ 0 & \text{otherwise,}\end{cases} \tag{17}$$

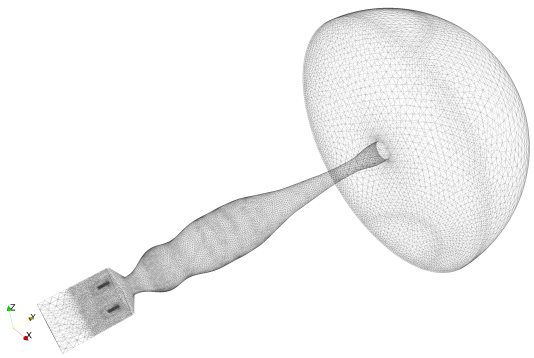

Figure 2: *Computational mesh for the unified simulations of vowel [i].*

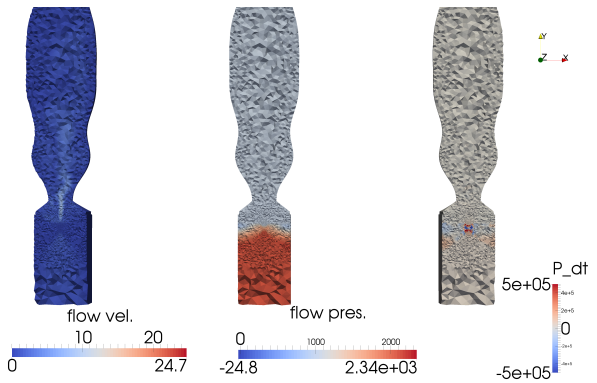

Figure 3: *Snapshot of the hydrodynamic velocity(m/s), hydrodynamic pressure(Pa) and its derivative(Pa/s) for unified simulations of vowel [i] at* $t = 0.055$.

with contact parameters $\alpha, \beta, \gamma, \delta$. The following values have been taken for the parameters in the model, $\rho_f = 1$, $\rho_s = 800$, $\mu_f = 17 \times 10^{-6}$, $\mu_s = 3 \times 10^3$, $\alpha = 0.05$, $\beta = 3h$ and $\gamma = 50h$, with the mesh size $h = 5 \times 10^{-5}$. The contact parameter δ is taken as the time step of the FSI solver (see subsection 3.5).

3.3. The acoustic model

The solution to (1)-(3) will provide the incompressible flow velocity $\boldsymbol{u}$ and pressure p, which can be used to construct the acoustic source terms for a wave operator, following an acoustic analogy approach. In an ALE setting, the standard wave equation becomes inappropriate for that purpose [11] and one has to resort to a mixed formulation [23] for the acoustic pressure, p_a, and acoustic particle velocity, $\boldsymbol{u}_a$, to account for the moving boundaries (in our case the VF). Let $Q_a = \Omega_f \times I$ stand for the space-time acoustic domain with $\boldsymbol{\beta}$ denoting again the ALE velocity. The equations to be solved are given by [11],

$$\frac{1}{\rho_0 c_0^2}\partial_t p_a - \frac{1}{\rho_0 c_0^2}\,\boldsymbol{\beta}\cdot\nabla p_a + \nabla\cdot\boldsymbol{u}_a + \alpha_a p_a = Q_s \qquad \text{in } Q_a, \tag{18a}$$

$$\rho_0\partial_t \boldsymbol{u}_a - \rho_0\,\boldsymbol{\beta}\cdot\nabla\boldsymbol{u}_a + \nabla p_a + \alpha_a^*\boldsymbol{u}_a = \boldsymbol{f}_s \qquad \text{in } Q_a. \tag{18b}$$

$Q_s(\boldsymbol{x}, t)$ represents a volume source distribution and $\boldsymbol{f}_s(\boldsymbol{x}, t)$ an external body force per unit volume. To perform the one-way coupling between FSI and the acoustics equations, the terms $Q_s = (\rho_0 c_0^2)^{-1}\partial_t p$ and $f_s = \mathbf{0}$ have been chosen. That corresponds to the mixed form of the analogy in [24], which as said, can be viewed as a simplification of the acoustic perturbation equations in [17, 18].

Equation (18) is to be supplemented with the following boundary and initial conditions

$$\begin{aligned}
&\boldsymbol{u}_a(\boldsymbol{x},t)\cdot\boldsymbol{n} = \gamma_a p_a \quad && \boldsymbol{x}\in\{\boldsymbol{x}:\boldsymbol{x}\in\partial\Omega_f,\\
&&& y > 0.012, y < 0.1854\},\\
&\boldsymbol{u}_a(\boldsymbol{x},t)\cdot\boldsymbol{n} = 0 && \boldsymbol{x}\in\{\boldsymbol{x}:\boldsymbol{x}\in\partial\Omega_f, y = 0.1854,\\
&&& \sqrt{x^2+(z-0.01045)^2}\le 0.06\},\\
&\boldsymbol{u}_a(\boldsymbol{x},0) = 0, &&
\end{aligned} \tag{19}$$

where we have taken $\gamma_a = 0.05/c_0\rho_0$, $c_0 = 350$, $\rho_0 = 1.225$.

As regards the hemisphere of the radiation domain, an absorbing condition has to be imposed on it to avoid reflections from impinging waves. A perfectly matched layer (PML) has been used to do so. The terms containing α_a and α_a^* have been included in (18) for that reason following the strategy in [7]. However, new formulations for α_a and α_a^* have been designed. These are

$$\alpha_a(\boldsymbol{x}) := \begin{cases} \alpha_{max}\dfrac{(y-b_0)^2}{(b_f-b_0)^2}(-2y+3b_f-b_0) \\ \qquad\qquad \text{if } y\le 0.012 \\ \alpha_{max}\dfrac{(r(\boldsymbol{x})-r_0)^2}{(r_f-r_0)^2}(-2r(\boldsymbol{x})+3r_f-r_0) \\ \qquad\qquad \text{if } y\ge 0.1854, r(\boldsymbol{x}) > r_0 \\ 0 \\ \qquad\qquad \text{otherwise} \end{cases} \tag{20}$$

with

$$\alpha_a^* := \rho_0^2 c_0^2 \alpha_a, \tag{21}$$

and

$$r(\boldsymbol{x}) := \sqrt{x^2+(y-0.1854)^2+(z-0.01045)^2}. \tag{22}$$

The values $\alpha_{max} = 0.4$, $b_0 = 0$, $b_f = 0.012$, $r_0 = 0.06$ and $r_f = 0.072$ have been chosen for the parameters in the model.

3.4. Numerical Implementation

The weak formulations of the PDE's in subsections 3.1, 3.2 and 3.3 are discretized in space using CG1 and DG0 elements on conforming tetrahedral meshes. For the time discretization an implicit Crank-Nicholson scheme is used with a CFL number of 0.5. The boundary conditions for the FSI-contact equations are applied strongly while they are imposed weakly for the wave equation in mixed form.

It is known that numerical instabilities may arise in finding numerical solutions to those PDE's for finite dimensional discretizations. A typical example for hyperbolic systems is the oscillations caused by convection dominated flows [25]. Another example for mixed problems is the violation of the Ladyshenskaya, Babuska, Brezzi (LBB) condition [26]. This prevents from using equal interpolations for the problem unknowns and demands resorting to stabilization strategies. For the FSI weak equations we have employed several stabilization strategies of the streamline diffusion type (see e.g., [27, 28, 25, 29, 30] whereas an algebraic subgrid scale (ASGS) has been applied to the weak formulation of the mixed wave equation [11].

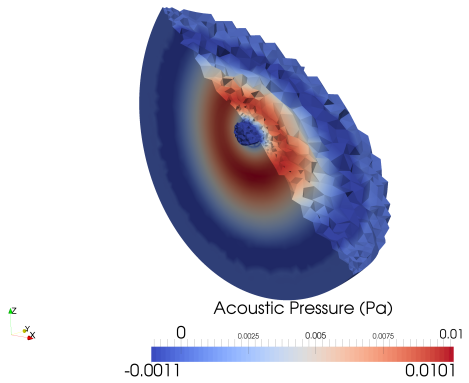

Figure 4: *Snapshot of the acoustic pressure for the unified simulation of vowel [i] at* $t = 0.055$.

3.5. Simulation Details

A computational mesh with 445K nodes and 2409K elements (shown in Figure 2) was created in ANSA [31], a computer-aided engineering tool for pre-processing. It should be noted that the smallest elements ($h \sim 5e{-}5$) are within the VF region, where resolving the boundary layer was necessary to trigger the VF self oscillations. The mesh for the acoustic simulation is a subset of the FSI mesh excluding the solid cells, and it is created in memory on the fly during the simulations. The latter are performed until $T \sim 0.121$. The time step for the acoustic solver is 1.25×10^{-5} while it is decided through a CFL condition for the FSI solver, being of order $\mathcal{O}(10^{-6})$.

The computations were performed on Beskow, a Cray XC40 system, where each node has two CPUs (Intel E5-2698v3) with 16 cores. 320 cores were allocated for 264 hours in total for the simulations.

4. Results

4.1. FSI simulations with contact

To compute the source terms for the acoustic wave equation in mixed form, the FSI-contact solver has been used together with two mesh smoothers, whose detailed description can be found in [22]. These simulations took roughly 90% of the total computational time for the given example.

An oscillation frequency of ~ 112 was obtained for the VF, which is in the natural range of human individuals. Snapshots of the flow velocity, flow pressure and its time derivative to be used for the acoustic source term $(\rho_0 c_0^2)^{-1}\partial_t p$ are provided in Figure 3. The deflecting flow jet and the flow pressure pattern due to the VF can be observed. It can also be appreciated how the acoustic source term concentrates close to the the VF contact region.

4.2. Acoustic Simulations

In Figure 4 we present a snapshot of the computed of the acoustic pressure. It can be clearly observed how spherical front waves emanate from the mouth. The acoustic pressure is essentially valid outside the mainstream of the jet flow leaving the mouth, though the implemented acoustic analogy is able to filter some pseudo-sound at very low convection velocities, which is the case.

In Figure 5 we show the pressure spectrum at a point located at $(-0.0036, 0.2480, 0.0151)$ in the radiation domain, far apart

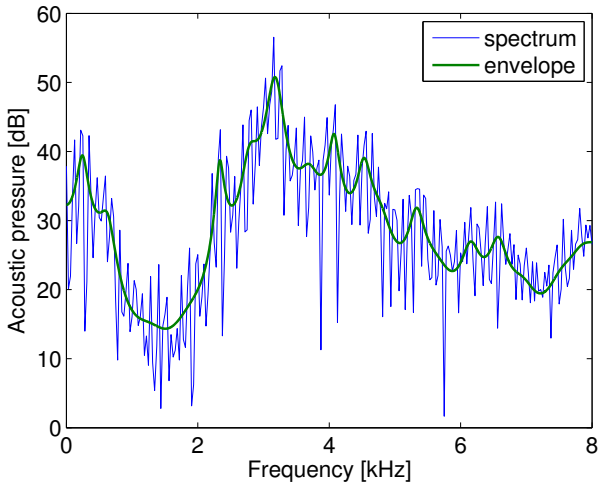

Figure 5: *Spectrum of the acoustic pressure at a point located at position* $(-0.0036, 0.2480, 0.0151)$ *outside the mouth. The formants of vowel [i] can be clearly appreciated.*

from the emanating jet flow. Moreover, an LPC (Linear Predictive Coding) analysis was also performed to extract the spectral envelope. As observed in the figure, the formants for vowel [i] are well reproduced. In particular, the following values have been obtained for them, $f_1 = 243$, $f_2 = 2338$ and $f_3 = 3175$. These values are in accordance with those reported in the literature.

5. Conclusions

In this work we have presented a unification strategy for the numerical production of static vowel sounds. The strategy involves performing a coupled fluid-structure-acoustic simulation that includes from the self oscillations of the vocal folds, including contact, to the radiated sound outside the mouth. The methodology has been applied to generate vowel [i] and the first three formant locations have been captured successfully.

Future areas of improvement encompass investigations on more complex vocal fold and contact models, the use of realistic geometries, the numerical production of dynamic vowel sounds like diphthongs, and the implementation of a compressible continuum approach to attain a two-way coupling between the FSI and the acoustics.

6. Acknowledgments

This work has been supported by EU-FET grant EUNISON 308874. The authors from the Universitat Ramon Llull also acknowledge the Agencia Estatal de Investigación (AEI) and FEDER, EU, through project GENIOVOX TEC2016-81107-P, and the grant 2014-SGR-0590 from the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (Generalitat de Catalunya). The fourth and sixth authors also respectively thank the support of grants 2016-URL-IR-013 and 2016-URL-IR-010 from the Generalitat de Catalunya and the Universitat Ramon Llull.

7. References

[1] R. Mittal, B. D. Erath, and M. W. Plesniak, "Fluid dynamics of human phonation and speech," *Annual Review of Fluid Mechan-*

ics, vol. 45, pp. 437–467, 2013.

[2] I. R. Titze and F. Alipour, *The myoelastic aerodynamic theory of phonation*. National Center for Voice and Speech, 2006.

[3] W. Zhao, C. Zhang, S. H. Frankel, and L. Mongeau, "Computational aeroacoustics of phonation, Part I: Computational methods and sound generation mechanisms," *Journal of the Acoustical Society of America*, vol. 112, no. 5, pp. 2134–2146, 2002.

[4] X. Zheng, R. Mittal, Q. Xue, and S. Bielamowicz, "Direct-numerical simulation of the glottal jet and vocal-fold dynamics in a three-dimensional laryngeal model," *Journal of the Acoustical Society of America*, vol. 130, no. 1, pp. 404–415, 2011.

[5] Q. Xue, X. Zheng, R. Mittal, and S. Bielamowicz, "Subject-specific computational modeling of human phonation," *Journal of the Acoustical Society of America*, vol. 135, no. 3, pp. 1445–1456, 2014.

[6] T. Vampola, J. Horáček, and J. G. Švec, "FE modeling of human vocal tract acoustics. part I: Production of Czech vowels," *Acta Acustica united with Acustica*, vol. 94, no. 3, pp. 433–447, 2008.

[7] H. Takemoto, P. Mokhtari, and T. Kitamura, "Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method," *Journal of the Acoustical Society of America*, vol. 128, no. 6, pp. 3724–3738, 2010.

[8] M. Arnela, O. Guasch, and F. Alías, "Effects of head geometry simplifications on acoustic radiation of vowel sounds based on time-domain finite-element simulations," *Journal of the Acoustical Society of America*, vol. 134, no. 4, pp. 2946–2954, 2013.

[9] M. Arnela, R. Blandin, S. Dabbaghchian, O. Guasch, F. Alías, X. Pelorson, A. Van Hirtum, and O. Engwall, "Influence of lips on the production of vowels based on finite element simulations and experiments," *Journal of the Acoustical Society of America*, vol. 139, no. 5, pp. 2852–2859, 2016.

[10] M. Arnela, S. Dabbaghchian, R. Blandin, O. Guasch, O. Engwall, A. Van Hirtum, and X. Pelorson, "Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds," *Journal of the Acoustical Society of America*, vol. 140, no. 3, pp. 1707–1718, 2016.

[11] O. Guasch, M. Arnela, R. Codina, and H. Espinoza, "A stabilized finite element method for the mixed wave equation in an ALE framework with application to diphthong production," *Acta Acustica united with Acustica*, vol. 102, no. 1, pp. 94–106, 2016.

[12] S. Zörner, M. Kaltenbacher, and M. Döllinger, "Investigation of prescribed movement in fluidstructure interaction simulation for the human phonation process," *Computers & Fluids*, vol. 86, pp. 133–140, 2013.

[13] P. Šidlof, S. Zörner, and A. Hüppe, "A hybrid approach to the computational aeroacoustics of human voice production," *Biomechanics and Modeling in Mechanobiology*, vol. 14, no. 3, pp. 473–488, 2015.

[14] S. Zörner, P. Šidlof, A. Hüppe, and M. Kaltenbacher, "Flow and acoustic effects in the larynx for varying geometries," *Acta Acustica united with Acustica*, vol. 102, no. 2, pp. 257–267, 2016.

[15] J. Hoffman, J. Jansson, R. V. de Abreu, N. C. Degirmenci, N. Jansson, K. Müller, M. Nazarov, and J. H. Spühler, "Unicorn: Parallel adaptive finite element simulation of turbulent flow and fluid–structure interaction for deforming domains and complex geometry," *Computers & Fluids*, vol. 80, pp. 310–319, 2013.

[16] R. Ewert and W. Schröder, "Acoustic perturbation equations based on flow decomposition via source filtering," *Journal of Computational Physics*, vol. 188, no. 2, pp. 365–398, 2003.

[17] A. Hüppe and M. Kaltenbacher, "Spectral finite elements for computational aeroacoustics using acoustic perturbation equations," *Journal of Computational Acoustics*, vol. 20, no. 02, p. 1240005, 2012.

[18] O. Guasch, P. Sánchez-Martín, A. Pont, J. Baiges, and R. Codina, "Residual-based stabilization of the finite element approximation to the acoustic perturbation equations for low mach number aeroacoustics," *International Journal for Numerical Methods in Fluids*, vol. 82, no. 12, pp. 839–857, 2016.

[19] "the EUNISON project," http://fp7eunison.com, accessed: 2017-02-28.

[20] M. Arnela and O. Guasch, "Finite element computation of elliptical vocal tract impedances using the two-microphone transfer function method," *Journal of the Acoustical Society of America*, vol. 133, no. 6, pp. 4197–4209, 2013.

[21] B. H. Story, "Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002," *Journal of the Acoustical Society of America*, vol. 123, no. 1, pp. 327–335, 2008.

[22] J. Hoffman, J. Jansson, and M. Stöckli, "Unified continuum modeling of fluid-structure interaction," *Mathematical Models and Methods in Applied Sciences*, vol. 21, no. 03, pp. 491–513, 2011.

[23] R. Codina, "Finite element approximation of the hyperbolic wave equation in mixed form," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, no. 13, pp. 1305–1322, 2008.

[24] M. Roger, "Aeroacoustics: Some theoretical background the acoustic analogy," *Computational Aeroacoustics*, pp. 2006–05, 2006.

[25] C. Johnson and J. Saranen, "Streamline diffusion methods for the incompressible Euler and Navier-Stokes equations," *Mathematics of Computation*, vol. 47, pp. 1–18, 1986.

[26] A. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, 1994.

[27] A. Brooks and T. Hughes, "Streamline upwind/petrov-galerkin formulations for convection dominated flows with particular emphasis on the incompressible navier-stokes equations," *Computer methods in applied mechanics and engineering*, vol. 32, pp. 199–259, 1982.

[28] C. Johnson, U. Nävert, and J. Pitkäranta, "Finite element methods for linear hyperbolic problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 45, pp. 285–312, 1984.

[29] C. Johnson and A. Szepessy, *Shock-capturing streamline diffusion finite element methods for nonlinear conservation laws*. Recent Developments in Computatioonal Fluid Dynamics, edited by T. J. R. Hughes and T. Tezduyar, AMD-Vol 95, 1988.

[30] C. Johnson, A. Szepessy, and P. Hansbo, "On the convergence of shock-capturing streamline diffusion finite element methods for hyperbolic conservation laws," *Mathematics of Computation*, vol. 54, no. 189, pp. 107–129, 1990.

[31] "ANSA Pre-processor V.15.2.0.linux.x64.2014. Produced by BETA CAE Systems S.A., Greece."