DOI: https://doi.org/10.14311/TPFM.2020.027

AEROACOUSTIC SIMULATION OF HUMAN PHONATION WITH THE WALE SUB-GRID SCALE MODEL

Petr Šidlof [1, 2], **Martin Lasota** [1, 3]

[1] **Technical University of Liberec, NTI FM, Studentská 2, 461 17 Liberec 1, Czech Republic**

[2] **Institute of Thermomechanics, The Czech Academy of Sciences, Dolejškova 5, 182 00 Prague 8, Czech Republic**

[3] **Faculty of Mechanical Engineering, Department of Technical Mathematics, Czech Technical University in Prague, Karlovo nám. 13, Prague, Czech Republic**

Abstract

The paper reports on an aeroacoustic model of voice generation in human larynx, based on Large Eddy Simulation with the Wall-Adapting Local Eddy-Viscosity (WALE) sub-grid scale (SGS) model. The simulation uses a three-step hybrid approach, with an incompressible finite volume CFD computation providing the filtered velocity and pressure, evaluation of the aeroacoustic sources, and simulation of the sound propagation by finite element discretization of the Acoustic Perturbation Equations. The WALE SGS model is used to overcome the limitation of the classical Smagorinski SGS model, which overpredicts the SGS viscosity in regions of high shear, especially within the boundary layer in the glottal constriction. Results of the 3D CFD simulation, location of the aeroacoustic sources and the spectra of the radiated sound for two vowels are presented.

Keywords: Large Eddy Simulation, turbulence modelling, voice generation, aeroacoustics

1 Introduction

Generation of human voice is a highly complex biophysical process, where the viscoelastic multi-layered tissues of the vocal folds interact with the airflow expired from lungs, start to self-oscillate and close the channel periodically. The vocal fold oscillation and glottal closure modulate the mass flux, creating pressure disturbances and complex turbulent structures in the supraglottal region. The acoustic resonances of the subglottal and supraglottal spaces do not only play the role of a passive linear acoustic filter modulating the source signal generated in the larynx, but do also have influence on the vocal fold oscillation [1]. Thus, voice generation is actually governed by three-way interaction between the structure, airflow and acoustics.

In computational fluid dynamics (CFD) of turbulent flows, three approaches are currently used. First, the Direct Numerical Simulations (DNS), i.e. straightforward discretization of the Navier-Stokes equations on a sufficiently fine computational mesh, where all turbulent scales up to smallest dissipating vortices are resolved. Even for moderate Reynolds numbers encountered in laryngeal airflow, this type of simulation is prohibitively expensive in terms of computational requirements. The second approach is the current industrial standard, Reynolds-Averaged Navier-Stokes equations (RANS). RANS completely gives up to resolve the turbulent fluctuations and dynamic evolution of vortical structures and aims to calculate the mean, time-averaged flow. The influence of turbulence on the mean flow is modeled using some of the plethora of more or less complex turbulence models. Clearly, RANS is unsuitable for aeroacoustic simulations of voice where the unsteady turbulent motions represent a crucial portion of the aeroacoustic sources.

With rigorous DNS unfeasible and Reynolds-Averaged Navier-Stokes equations inapplicable, numerical modelling of human phonation can use either the laminar simulations (such as in the works [2, 3, 4, 5, 6, 7]) or the third, arguably most promising approach: the Large Eddy Simulation (LES). The LES concept resolves the large, anisotropic energy-carrying fluid motions and models the effect of sub-grid scale, largely isotropic turbulent structures. Large eddy simulations are still computationally expensive, especially if the boundary layer is to be resolved properly. In numerical simulation of human phonation, the LES approach has been already used in recent years. One of the first was Suh et al. [8], who used compressible LES and Ffowcs Williams-Hawkings acoustic analogy in a static model of human glottis for far-field sound predictions. Mihaescu et al. [9, 10]

employed the LES capability of ANSYS Fluent to study the laryngeal airflow both during phonation and respiration. Recently, Sadeghi et al. [11, 12] simulated the laryngeal flow and the effect of ventricular folds using the LES feature of STAR-CCM+ and they quantified the computational requirements on parallel architectures.

A systematic comparison of the airflow and the sound generated in human larynx, obtained using a laminar and LES flow simulations, can help to assess the importance and the impact of turbulence modelling in aeroacoustic simulations of voice generation. This study builds on the previous publications [4, 13] and focuses on the impact of the numerical approach to turbulence modelling on the sound calculated by a hybrid aeroacoustic simulation. For this sake, the fluid-structure interaction is neglected and the vocal fold oscillation is prescribed. The current paper shortly reports on the first results obtained using the WALE (Wall-Adapting Local Eddy-viscosity) sub-grid scale (SGS) model. A follow-up full paper will compare and analyze in detail the results of computational aeroacoustic (CAA) simulations obtained using a laminar CFD simulation and two LES simulations with the Smagorinski and WALE SGS models.

2 CFD model of the incompressible flow in the larynx

2.1 Mathematical model

Large Eddy Simulation is a mathematical concept for modeling turbulent flows, which deals with flow structures carrying most kinetic energy, i.e. spatially organized large scales [14]. In the LES concept, any flow variable f may be decomposed as

$$f(\boldsymbol{x},t) = \overline{f}(\boldsymbol{x},t) + f'(\boldsymbol{x},t), \tag{1}$$

where $\overline{f}(\boldsymbol{x},t) = G * f(\boldsymbol{x},t) = \int G(\boldsymbol{r},\boldsymbol{x},\Delta) f(\boldsymbol{x}-\boldsymbol{r},t) d\boldsymbol{r}$ is the large-scale component, obtained by spatial filtering, and $f'(x,t)$ is the small, sub-grid scale contribution. The filter kernel G separates spatial scales $|\boldsymbol{r}| \leq \Delta/2$ from $|\boldsymbol{r}| > \Delta/2$, where Δ is the filter width. The effect of sub-grid scale (SGS) contributions on the large flow scales relies on the assumption of isotropic (non-directional) small-scale turbulence and is modeled.

The continuity and momentum equations for the incompressible fluid with applied LES filtering can be written as

$$\frac{\partial \overline{U}_i}{\partial x_i} = 0, \quad \frac{\partial \overline{U}_i}{\partial t} + \frac{\partial}{\partial x_j}(\overline{U}_i \overline{U}_j) = -\frac{1}{\rho}\frac{\partial \overline{P}}{\partial x_i} + \nu \frac{\partial^2 \overline{U}_i}{\partial x_j \partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j}, \tag{2}$$

where $\overline{U}_i$ and $\overline{P}$ are the filtered velocity vector and pressure, respectively. Except for the last righ-hand side term, equations (2) are equivalent to the standard Navier-Stokes equations. The term τ_{ij} is the subgrid scale tensor

$$\tau_{ij} = \overline{U_i U_j} - \overline{U}_i \overline{U}_j, \tag{3}$$

which needs to be modeled. The eddy-viscosity assumption of Smagorinski [15] states that

$$\tau_{ij} - \frac{1}{3}\tau_{kk}\delta_{ij} = -2\nu_t \overline{S}_{ij} = \nu_t \left(\frac{\partial \overline{U}_i}{\partial x_j} + \frac{\partial \overline{U}_j}{\partial x_i} \right), \tag{4}$$

where the $\overline{S}_{ij}$ is the symmetric part of the velocity gradient $g_{ij} = \partial \overline{U}_i / \partial x_j$ and

$$\nu_t = (C_S\, \Delta)^2 \sqrt{\overline{S}_{ij} \overline{S}_{ij}} \tag{5}$$

is the kinematic sub-grid scale viscosity. The current study uses the Wall-Adapting Local Eddy-viscosity (WALE) SGS model by Nicoud [16]. The WALE model overcomes the deficiency of the classical Smagorinski model, which overestimates the SGS viscosity near the walls (in the regions of high shear). In the WALE model, the SGS viscosity is modeled as

$$\nu_t = (C_m \Delta)^2 \frac{(s^d_{ij} s^d_{ij})^{3/2}}{(\overline{S}_{ij}\, \overline{S}_{ij})^{5/2} + (s^d_{ij} s^d_{ij})^{5/4}}, \tag{6}$$

where s_{ij}^d is the traceless symmetric part of the square of the velocity gradient tensor

$$s_{ij}^d = \frac{1}{2}\left(\frac{\partial \overline{U}_i}{\partial x_k}\frac{\partial \overline{U}_k}{\partial x_j} + \frac{\partial \overline{U}_j}{\partial x_k}\frac{\partial \overline{U}_k}{\partial x_i}\right) - \frac{1}{3}\delta_{ij}\frac{\partial \overline{U}_k}{\partial x_k}\frac{\partial \overline{U}_k}{\partial x_k} = \frac{1}{2}(\overline{g}_{ij}^2 + \overline{g}_{ji}^2) - \frac{1}{3}\delta_{ij}\overline{g}_{kk}^2 \tag{7}$$

and $\overline{g}_{ij}^2 = \overline{g}_{ik}\,\overline{g}_{kj}$. C_m is the model constant ranging between 0.55 - 0.6.

The Navier-Stokes equations in the LES concept are solved in a 3D computational domain modeling the larynx. The geometry of the domain is shown in Fig. 1, the boundary conditions are listed in Tab. 1. The vocal folds undergo sinusoidal prescribed oscillations with two degrees of freedom, with the amplitude $A = 0.3 \cdot 10^{-3}$ m, frequency $f = 100$ Hz and phase difference between the inferior and superior margin $\eta = \pi/2$ (for details see [4]).

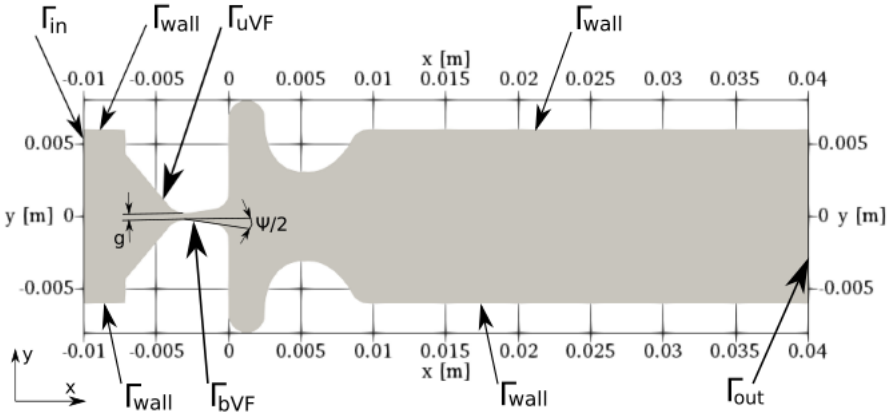

Figure 1: Computational domain in coronal section with vocal folds and ventricular folds. The front and back boundaries of the domain are denoted Γ_{front} and Γ_{back}.

Table 1: Boundary conditions for the mean velocity $\boldsymbol{U}$ and pressure P. The vector $\boldsymbol{n}$ is the unit outer normal and $h(\boldsymbol{x}, t)$ is the prescribed sinusoidal displacement of the vocal folds in medial-lateral direction.

Boundary	$\boldsymbol{U}$ [ms^{-1}]	P [Pa]
Γ_{in}	from flux if $\boldsymbol{U}\cdot\boldsymbol{n} < 0$ 0 if $\boldsymbol{U}\cdot\boldsymbol{n} > 0$	307.5
Γ_{out}	$\nabla(\boldsymbol{U})\cdot\boldsymbol{n} = 0$ if $\boldsymbol{U}\cdot\boldsymbol{n} > 0$ $\boldsymbol{U} = 0$ if $\boldsymbol{U}\cdot\boldsymbol{n} < 0$	0
Γ_{bVF}, Γ_{uVF}	$\boldsymbol{U} = \dot{h}(\boldsymbol{x}, t)$	$\nabla(P)\cdot\boldsymbol{n} = 0$
Γ_{wall}	$\boldsymbol{U} = 0$	$\nabla(P)\cdot\boldsymbol{n} = 0$
Γ_{front},Γ_{back}	$\boldsymbol{U} = 0$	$\nabla(P)\cdot\boldsymbol{n} = 0$

2.2 Numerical solution

The CFD model is implemented in the OpenFOAM CFD library. The equations are solved numerically using the cell-centered finite volume method with central differencing scheme for the spatial discretization of the diffusion term. The upwind-based schemes bring high numerical diffusion and hence they are not recommended in LES studies. The spatial discretization of the convective term is realized using the total variation diminishing scheme, combining benefits from first order schemes to keep stability and second order schemes for high accuracy. The temporal discretization is realized using the two-step second-order backward scheme. The simulation was run in parallel on 16 processors of a computational cluster for twenty periods of vocal fold vibration, i.e. $T = 0.2$ s, and took about 480 hours of walltime. The timestep was adjusted by an automatic algorithm to keep the Courant number below a predefined limit, here $Co < 1$.

The numerical mesh is hexahedral block-structured, consisting of 1.9 million elements. The numerical simulation is performed on a moving mesh, which is deformed due to vocal fold oscillation. The grid point coordinates are updated each timestep by solving an auxiliary Laplace equation. The distance of the first gridpoints in wall units varies according to the phase of oscillation of the vocal folds. At time instant $t = 0.19$ s, the average and maximum values on the vocal fold surfaces are $y^+_{avg} = 1.43$ and $y^+_{max} = 8.39$.

3 CAA model of sound generation and propagation

3.1 Mathematical model

The computational aeroacoustic model is based on the Acoustic Perturbation Equations (APE) for low Mach-number flows [17]. The approach splits the fluid variables p and $\boldsymbol{U}$ into mean, fluctuating solenoidal non-acoustic (incompressible) and irrotational acoustic components, respectively:

$$\boldsymbol{U} = \overline{\boldsymbol{U}} + \boldsymbol{U}^{ic} + \boldsymbol{U}^{a}, \quad p = \overline{p} + p^{ic} + p^{a} \ . \tag{8}$$

Under the assumption of isentropic incompressible flow, the perturbation equations are

$$\frac{Dp^a}{Dt} + \rho_0 c_0^2 \nabla \cdot \boldsymbol{U}^a = -\frac{Dp^{ic}}{Dt}, \tag{9}$$

$$\rho_0 \frac{\partial \boldsymbol{U}^a}{\partial t} + \rho_0 \nabla(\overline{\boldsymbol{U}} \cdot \boldsymbol{U}^a) + \nabla p^a = 0, \tag{10}$$

where the material derivative is $\frac{D}{Dt} = \frac{\partial}{\partial t} + (\overline{\boldsymbol{U}} \cdot \nabla)$ and the non-acoustic parts are obtained from CFD results. This workflow is called the hybrid aeroacoustics approach and involves three steps: first, unsteady CFD computation; second, computation of the acoustic sources – right-hand side of equation (9); third, simulation of the acoustic field by numerical solution of equations (9),(10). The unknown variables $\boldsymbol{U}^a$ and p^a are related to the acoustic scalar potential ψ^a as

$$U_i^a = -\frac{\partial \psi^a}{\partial x_i}, \ p^a = \rho_0 \frac{\partial \psi^a}{\partial t} + \rho_0 \overline{U_i} \frac{\partial \psi^a}{\partial x_i} \, . \tag{11}$$

Thus, the set of APEs can be rewritten into a scalar-valued Perturbated Convective Wave Equation [18]

$$\frac{1}{c_0^2} \frac{D^2 \psi^a}{Dt^2} - \sum_{i=1}^{3} \frac{\partial^2 \psi^a}{\partial x_i^2} = -\frac{1}{\rho_0 c_0^2} \frac{Dp_{ic}}{Dt} \, . \tag{12}$$

3.2 Numerical solution

The numerical solution of the acoustic field is realized using the finite element code CFS++ [19]. The computational domain, shown in Fig. 2, consists of the larynx (red), vocal tract (purple), radiation zone (blue) and Perfectly Matched Layers (PML, green) at inlet and outlet with a damping function implemented to suppress reflection of acoustic waves on this artificial boundary [20]. The vocal tract was modeled according to the data published by Story et al. [21].

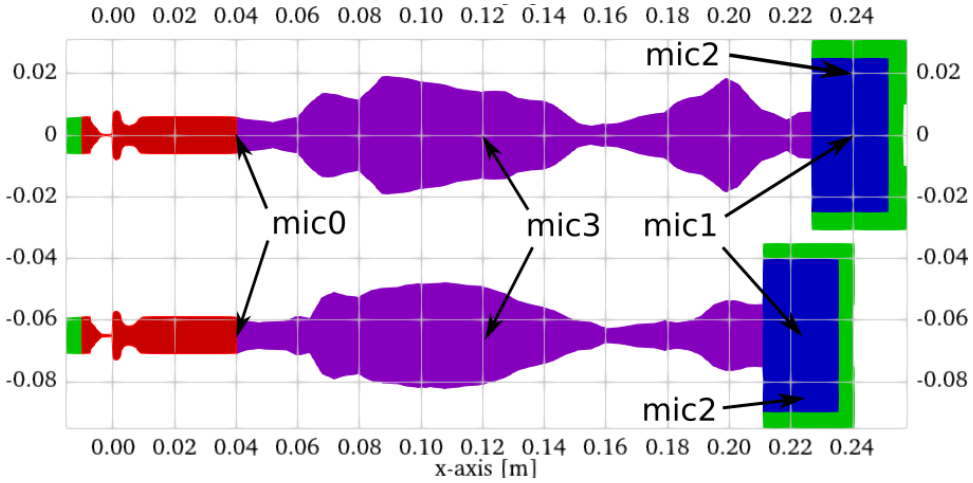

Figure 2: Schematic of the aeroacoustic computational domains for vowels [u:] (top) and [i:] (bottom) with locations of the acoustic probes.

Since about ten elements per wavelength are sufficient for the acoustic simulation, the CAA mesh can be significantly coarser than the CFD mesh. In the second step of the hybrid scheme, the aeroacoustic sources calculated from the CFD results are conservatively interpolated onto the CAA grid in the laryngeal part of the computational domain (for details see [13]).

4 Results

Fig. 3 shows the vorticity contours in the mid-coronal plane at time instant $t = 0.19$ s, i.e. after 19 periods of vocal fold oscillation. The time instant is chosen in the divergent phase of the vocal fold oscillation, close to the time where the glottal gap is narrowest. In this phase, the jet formed by the vocal fold is almost interrupted and attached to the lower vocal fold. The subglottal velocity field is laminar, in the supraglottal volume residual turbulent structures interacting with the jet can be seen.

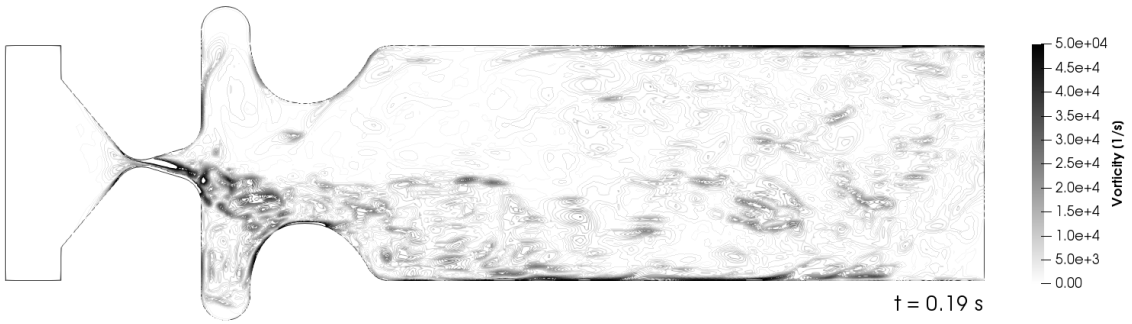

Figure 3: Vorticity contours in the subglottal and supraglottal region of the larynx. Beginning of the 20$^{\text{th}}$ period of vibration, divergent phase of the vocal fold motion close to the minimum glottal opening.

The location of the dominant aeroacoustic sources at the fundamental frequency $F_0 = 100$ Hz and first harmonic frequency $f_1 = 200$ Hz, calculated as a Fourier transform of the right-hand side of equation (9), is demonstrated in Fig. 4. It can be seen that the most important sources are located within the glottal gap and further on the surface of the lower ventricular fold, where the glottal jet hits the wall.

The acoustic spectrum evaluated at the acoustic probe (microphone 1), located in the radiation region 1 cm downstream of the mouth (see Fig. 2), is shown in Fig. 5. Two cases are demonstrated

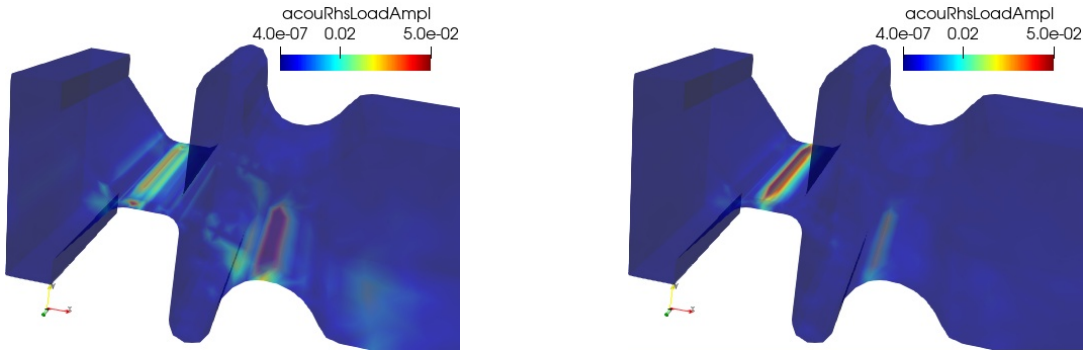

Figure 4: Aeroacoustic sources in frequency domain. Left: 100 Hz, right: 200 Hz

– simulation B13 realized on a vocal tract geometry modeling the vowel [u:] and simulation C13 for the vocal tract [i:], together with the formant frequencies (resonance frequencies of the vocal tract) taken from [22]. It can be seen that the simulated formants match well the theoretical ranges especially in the case of F_2 and F_3 for the vowel [u:], and F_2 for the vowel [i:].

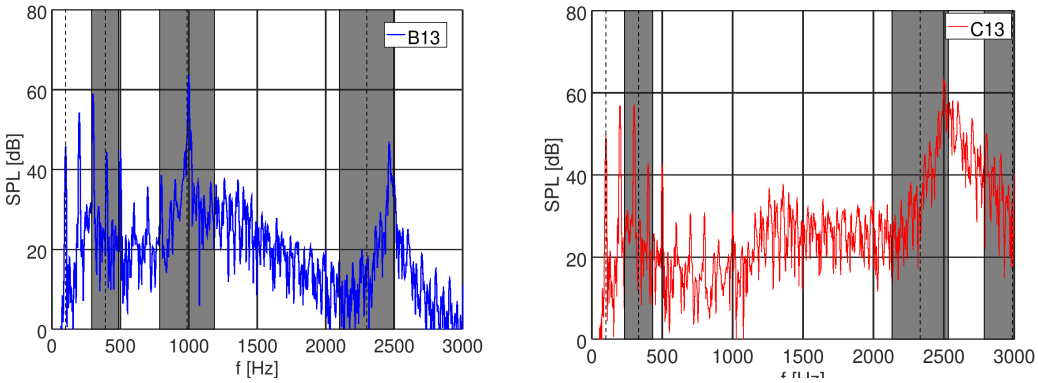

Figure 5: Acoustic spectra evaluated at the probe "mic1". Left: vocal tract for vowel [u:], Right: vocal tract for vowel [i:]. The gray strips represent the formant frequency ranges according to [22].

5 Conclusions

In the current paper, a hybrid aeroacoustic model of human phonation based on a 3D LES simulation of airflow in human larynx with oscillating vocal folds was presented. The WALE SGS model, which should correctly predict the SGS viscosity even in the regions of high shear (near the walls and in the shear layer of the glottal jet), was employed. The spectra of the radiated sound for the two simulated vowels [u:] and [i:] match well the formant frequencies known from real speech, though the predicted sound pressure levels are quite low.

In a follow-up full paper, the results obtained using the LES simulation with the WALE model will be analyzed in detail and compared to the simulation using a classical Smagorinski SGS model, and a "laminar" CFD simulation (with no turbulence modelling). As many research groups worldwide dealing with numerical simulation of voice generation still use laminar or underresolved LES flow simulations, the goal of the current study is to assess what is the impact of the SGS model on the aeroacoustic sources and on the emitted acoustic signal.

Acknowledgments

The research was supported by the Czech Science Foundation, project 19-04477S *Modelling and measurements of fluid-structure-acoustic interactions in biomechanics of human voice production.* The aeroacous-

tic simulations were realized in the finite element software CFS++ developed by the group of prof. M. Kaltenbacher at TU Vienna.

References

[1] I. R. Titze: Nonlinear source filter coupling in phonation: Theory, *The Journal of the Acoustical Society of America*, vol. 123, pp. 2733–2749, May 2008.

[2] M. de Oliviera Rosa, J. Pereira, M. Grellet, and A. Alwan: A contribution to simulating a three-dimensional larynx model using the finite element method, *Journal of the Acoustical Society of America*, vol. 114, pp. 2893–2905, Nov. 2003.

[3] S. Thomson:, Investigating coupled flow-structure-acoustic interactions of human vocal fold flow-induced vibration, *The Journal of the Acoustical Society of America*, vol. 138, pp. 1776–1776, sep 2015.

[4] P. Šidlof, S. Zörner, and A. Hüppe: A hybrid approach to computational aeroacoustics of human voice production, *Biomechanics and Modeling in Mechanobiology*, vol. 14, pp. 473–488, June 2015.

[5] W. Jiang, X. Zheng, and Q. Xue: Computational modeling of fluid–structure–acoustics interaction during voice production, *Frontiers in Bioengineering and Biotechnology*, vol. 5, feb 2017.

[6] P. Sváček and J. Horáček: Finite element approximation of flow induced vibrations of human vocal folds model: Effects of inflow boundary conditions and the length of subglottal and supraglottal channel on phonation onset, *Applied Mathematics and Computation*, vol. 319, pp. 178–194, 2018.

[7] J. Valášek, M. Kaltenbacher, and P. Sváček: On the application of acoustic analogies in the numerical simulation of human phonation process, *Flow, Turbulence and Combustion*, vol. 102, pp. 129–143, apr 2019.

[8] J. Suh and S. Frankel, Numerical simulation of turbulence transition and sound radiation for flow through a rigid glottal model, *Journal of the Acoustical Society of America*, vol. 121, pp. 3728–3739, June 2007.

[9] M. Mihaescu, S. M. Khosla, S. Murugappan, and E. J. Gutmark: Unsteady laryngeal airflow simulations of the intra-glottal vortical structures, *Journal of the Acoustical Society of America*, vol. 127, no. 1, pp. 435–444, 2010.

[10] M. Mihaescu, G. Mylavarapu, E. J. Gutmark, and N. B. Powell: Large eddy simulation of the pharyngeal airflow associated with obstructive sleep apnea syndrome at pre and post-surgical treatment, *Journal of Biomechanics*, vol. 44, pp. 2221–2228, Aug. 2011.

[11] H. Sadeghi, M. Döllinger, M. Kaltenbacher, and S. Kniesburges, Aerodynamic impact of the ventricular folds in computational larynx models, *The Journal of the Acoustical Society of America*, vol. 145, pp. 2376–2387, apr 2019.

[12] H. Sadeghi, S. Kniesburges, S. Falk, M. Kaltenbacher, A. Schützenberger, and M. Döllinger: Towards a clinically applicable computational larynx model, *Applied Sciences*, vol. 9, p. 2288, jun 2019.

[13] 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.

[14] M. Lesieur, O. Métais, and P. Comte: *Large-eddy simulations of turbulence*. Cambridge University Press, 2005.

[15] J. Smagorinsky: General circulation experiments with the primitive equations: I. the basic experiment, *Monthly weather review*, vol. 91, no. 3, pp. 99–164, 1963.

[16] F. Nicoud and F. Ducros: Subgrid-scale stress modelling based on the square of the velocity gradient tensor, *Flow, turbulence and Combustion*, vol. 62, no. 3, pp. 183–200, 1999.

[17] 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.

[18] A. Hüppe, J. Grabinger, M. Kaltenbacher, A. Reppenhagen, G. Dutzler, and W. Kühnel: A Non-Conforming Finite Element Method for Computational Aeroacoustics in Rotating Systems, in *20th AIAA/CEAS Aeroacoustics Conference*, 2014-2739, 2014.

[19] M. Kaltenbacher, *Numerical simulation of mechatronic sensors and actuators*, vol. 2. Springer, 2007.

[20] B. Kaltenbacher, M. Kaltenbacher, and I. Sim: A modified and stable version of a perfectly matched layer technique for the 3-d second order wave equation in time domain with an application to aeroacoustics, *Journal of Computational Physics*, vol. 235, no. 0, pp. 407 – 422, 2013.

[21] B. H. Story and I. R. Titze: Parameterization of vocal tract area functions by empirical orthogonal modes, *Journal of Phonetics*, vol. 26, no. 3, pp. 223–260, 1998.

[22] B. H. Story, I. R. Titze, and E. A. Hoffman: Vocal tract area functions from magnetic resonance imaging, *The Journal of the Acoustical Society of America*, vol. 100, no. 1, pp. 537–554, 1996.