11th World Congress on Computational Mechanics (WCCM XI)
5th European Conference on Computational Mechanics (ECCM V)
6th European Conference on Computational Fluid Dynamics (ECFD VI)
E. Oñate, J. Oliver and A. Huerta (Eds)

3D NUMERICAL SIMULATIONS OF HUMAN PHONATION

Manfred Kaltenbacher[1], Stefan Zörner[1], Andreas Hüppe[1] and Petr Sidlof[2]

[1] Vienna University of Technology, 1040 Vienna, Austria, manfred.kaltenbacher@tuwien.ac.at

[2] Technical University of Liberec, 461 17 Liberec, Czech Republic, sidlof@it.cas.cz

Key words: Multiphysics Problems, Human Phonation

Abstract. The human phonation process is an aeroacoustic phenomenon, induced by air flowing through the larynx. Thereby, the air interacts with the vocal folds, forcing them to vibrate. To be able to capture the 3D flow field, we simplify the model by prescribing the motion of the vocal folds and then determine 3D flow field using the finite volume solver *OpenFoam*. In a next step the aeroacoustic source terms are calculated, with the help of our research code *CFS++*. Hereby, we compare the two hybrid methods Lighthill's analogy and a perturbation approach to determine their applicability to the problem at hand. Furthermore, a vocal tract and acoustic propagation is added to the acoustic domain. This enables us to include the impact of the additional volumes and compare it to measured data.

1 INTRODUCTION

The human phonation is a complex interaction of fluid mechanics, solid mechanics and acoustics. As the lungs compress, air flows through the larynx passing the vocal folds (VFs) which form a narrow constriction, the glottis. The air flow forces the VFs to vibrate which in turn create a pulsating air stream, which is the main sound generating mechanism for phonation. Hence, our modeling approach is to resolve, within the larynx and adjacent regions, the physical details of the phonation process in space and time by means of partial differential equations (PDEs).

Due to limitations in computer resources and current numerical methods, full coupling between all three fields for realistic 3D geometries is currently not feasible (see, e.g., [1, 2, 3]). Therefore, in a first step we restrict to 2D fully coupled simulations and compare to CFD simulation with prescribed VFs oscillation. As a main result, we can demonstrate that in case of driven vocal fold models (pure fluid simulation with prescribed change of the fluid domain) a pressure boundary condition is necessary to capture the physical effects. In a second step, we then apply this finding to 3D flow computations and turn our focus to the generation of sound. Thereby, we utilize the program *OpenFoam* as

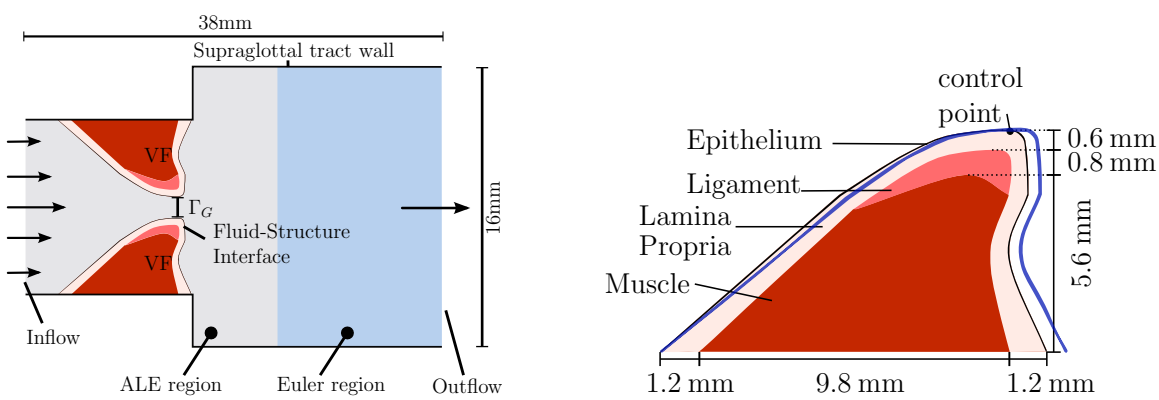
well as our in-house research code *CFS++* for solving the 3D incompressible Navier-Stokes equations, and the acoustic sources as well as sound propagation. In detail, we investigate Lighthill's analogy and its application towards the human phonation. It is a well-known fact that not all sources given by the Lighthill tensor cause propagating acoustic waves, audible in the far field. Therefore, we compare the method to a perturbation approach, which does not have these limitations.

2 PRESCRIBED VF MOVEMENT

A fully coupled 3D fluid-structure simulation is not feasible regarding computation costs, since several iterations are necessary in each time step to reach a state of equilibrium for the two physical fields. Therefore, we first analyze the effect of applying prescribed motions of the VFs instead of a full coupling by a 2D model, which is presented in the next section.

2.1 2D Geometry

The geometric model consists of a simple channel with two elastic bodies inside representing the VFs (see Fig. 1a). The VF model, from now on referred to as *VF-1*, has



(a) Fluid regions and boundary conditions. The glottis divides the fluid region into sub- and supraglottal area.

(b) *VF-1* Geometry and material model, consisting of four different regions. Blue line indicates the outline of *VF-2*.

different layers, the muscle, the ligament, the lamina propria and the epithelium. The muscle builds the base and supports the ligament, and both are covered by the lamina propria. The lamina propria itself is covered by the very thin tissue layer (0.05 mm), the so-called epithelium, represented by a thin line in Fig. 1b. Material parameters for these different materials are given in Tab. 1. A value of 0.45 is taken as Poisson's ratio for all four tissue types.

A second VF model named *VF-2* is introduced (see Fig. 1b), which has a slightly different geometry. Material parameters are identical to those of *VF-1*, as well as the

Table 1: Vocal fold material parameters for the 4 layer model.

Material	Elasticity modulus (kPa)
muscle	30
ligament	25
lamina propria	20
epithelium	50

topology of the layers.

2.2 CFD model

Since the maximal flow velocity during human phonation is usually below Ma = 0.2, the assumption of incompressibility can be assumed. Therefore, the airflow through the larynx is modeled by the incompressible Navier-Stokes equations, given by[1]

$$\frac{\partial \mathbf{u}^{\mathrm{ic}}}{\partial t} + \left(\mathbf{u}^{\mathrm{ic}} \cdot \nabla\right) \mathbf{u}^{\mathrm{ic}} + \nabla p^{\mathrm{ic}} - \nu \Delta \mathbf{u}^{\mathrm{ic}} = \mathbf{0}, \tag{1}$$

$$\nabla \cdot \mathbf{u}^{\mathrm{ic}} = 0. \tag{2}$$

In Eqs. (1) and (2) $\mathbf{u}^{\mathrm{ic}}$ denotes the fluid velocity, p^{ic} the kinematic pressure ($p^{\mathrm{ic}} = \frac{P^{\mathrm{ic}}}{\rho_{\mathrm{f}}}$), and ν ($\nu = \frac{\mu}{\rho_{\mathrm{f}}}$) the kinematic viscosity. The boundary conditions of the fluid velocity are set to $\partial \mathbf{u}^{\mathrm{ic}}_{\mathrm{in}} / \partial \mathbf{n} = 0$ at the inlet and no-slip condition $\mathbf{u}^{\mathrm{ic}} = 0$ on the fixed walls.

To accurately capture the change in the fluid domain due to structural movement, the Arbitrary-Lagrangian-Eulerian (ALE) approach is used. Numerical instabilities, arising by convection dominants, are remedied by the streamline upwind Petrov-Galerkin (SUPG) method. A detailed description on the numerical schemes solving the Naviers-Stokes equations, can be found in [5], and we refer to [1] for a discussion on the implementation in our code.

2.3 Solid mechanics

The vibrations of the vocal folds are modeled by Navier's equation

$$\nabla \cdot \sigma_{\mathrm{s}} = \rho_{\mathrm{s}} \frac{\partial^2}{\partial t^2} \mathbf{d} \ , \tag{3}$$

where σ_{s} denotes the Cauchy stress tensor, ρ_{s} the density of the solid and $\mathbf{d}$ the mechanical displacement. Introducing the tensor of elasticity $[c]$ and tensor of linear strain $[S]$, allows us to express Hook's law by

$$\sigma_{\mathrm{s}} = [c][S] \tag{4}$$

[1]The superscript ic indicates that the physical quantities are incompressible.

and the linear strain-displacement by

$$[S] = \nabla^{\mathrm{sym}}\mathbf{d} \; . \tag{5}$$

Substituting (4) and (5) into (3) results in the final PDE for linear elasticity

$$\mathcal{B}^T[c]\mathcal{B}\mathbf{d} = \rho_{\mathrm{s}}\frac{\partial^2}{\partial t^2}\mathbf{d} \tag{6}$$

with the differential operator $\mathcal{B}$ which computes in the 2d plane case as

$$\mathcal{B} = \begin{pmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{pmatrix} .$$

2.4 Fluid-structure coupling

For a correct representation of the fluid-structure interaction, two conditions must be met at the common interface Γ_{fs} between fluid and solid. First, the fluid velocity and structural velocity must be equal

$$\mathbf{u}^{\mathrm{ic}} = \frac{\partial}{\partial t}\mathbf{d} \qquad \text{on } \Gamma_{\mathrm{fs}} \; , \tag{7}$$

which indicates that the fluid adheres to the structure. For a fixed wall, this corresponds to "no-slip" condition.
The second condition is the continuity of stress in normal direction along the interface, indicating that fluid stress σ_{f} and solid stress σ_{s} must coincide, which is enforced by

$$[\sigma_{\mathrm{s}}] \cdot \mathbf{n} = [\sigma_{\mathrm{f}}] \cdot \mathbf{n} \qquad \text{on } \Gamma_{\mathrm{fs}}. \tag{8}$$

Thereby, the acting fluid forces can be split up into a pressure and a viscous component

$$\begin{aligned} \mathbf{f}_{\mathrm{fs}} =& [\sigma_{\mathrm{f}}] \cdot \mathbf{n} \\ =& \underbrace{\rho_{\mathrm{f}} \int\limits_{\Gamma_{\mathrm{fs}}} -p^{\mathrm{ic}} I \cdot \mathbf{n} \, \mathrm{d}x}_{\text{pressure force}} + \underbrace{\int\limits_{\Gamma_{\mathrm{fs}}} \mu \left(\nabla \mathbf{u}^{\mathrm{ic}} + \left(\nabla \mathbf{u}^{\mathrm{ic}}\right)^T \right) \cdot \mathbf{n} \, \mathrm{d}x}_{\text{viscous force}} \; . \end{aligned} \tag{9}$$

2.5 Results of prescribed movement

In a first fully coupled simulation the *VF-1* model is used and the flow is driven by a fixed pressure of 1 kPa at the inlet and 0 kPa at the outlet. Then, using the computed vibration of the vocal folds from the fully coupled simulation on *VF-1*-model, two separate simulations are performed: The first one, called *VF-2-V*, is a velocity driven flow with the inlet set to the velocity of *VF-1*. The second simulation, called *VF-2-P*, is a pressure

driven flow, which is set to 1 kPa at the inlet and 0 kPa at the outlet, as in the *VF-1* case. For both simulations we use the slightly different geometry (named *VF-2* in Fig. 1b).

These three simulations are then compared, using the volume flux through the glottis as a measure. The change of the geometry should have an impact on the volume flow, when the computations with the prescribed movements captures physics. Figure 1 shows a time excerpt of the flux, where the flow is fully developed. It shows that prescribed motion together with a velocity profile as inflow condition (*VF-2-V*) shows no difference to the original fully coupled *VF-1* case. The phase and amplitudes of the flux are identical, which is un-physically. However, when using a pressure driven flow, as in the case of *VF-2-P*, the flow field does react on the changed geometry.

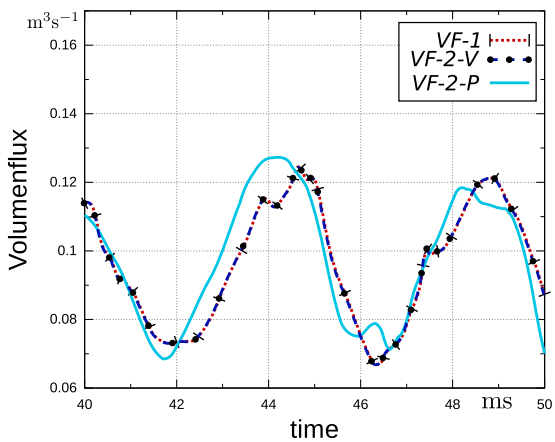


Figure 1: Geometry and mesh for the simulation.

Due to this study it is feasible to use a prescribed motion of the VFs, as long as the flow is driven by a pressure gradient.

3 AEROACOUSTIC MODEL OF THE HUMAN PHONATION

3.1 Geometric model

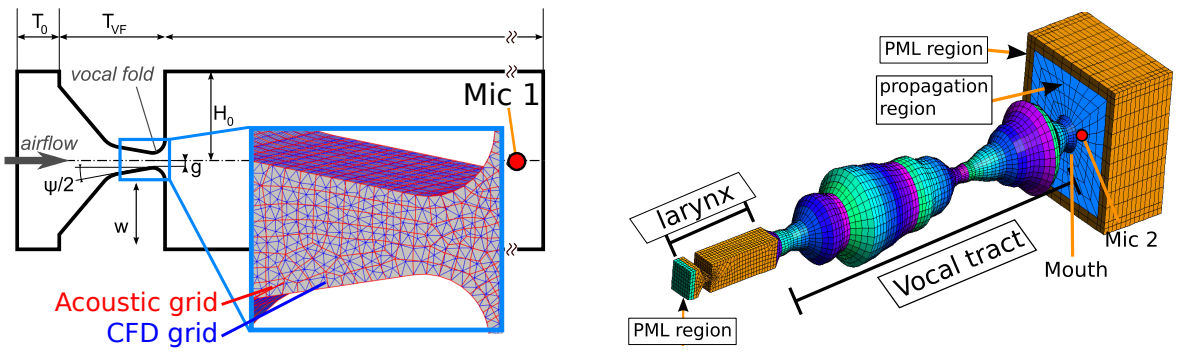
The 3D model used for the simulations consists of three domains, the larynx, the vocal tract and an acoustic propagation region in which the radiated sound is monitored (see Fig. 2b). The flow field simulation is constricted to the larynx as this is the region in which the sound is generated. The coronal section of the CFD domain is given in Fig. 2a, and consists of a short subglottal region of length $T_0 = 2.8$ mm, the vocal folds of length $T_{VF} = 7.2$ mm, and a supraglottal region of length $T_1 = 40$ mm. In depth and height the laryngeal channel has the same dimension of $2H_0 = 12.0$ mm – the shape does not change along the anterior-posterior axis. The geometry of the vocal folds are modeled according to the "M5" model proposed by [6] with a medial surface convergence angle of $\psi = -20°$. To simulate the vocal folds oscillation a sinusoidal motion is prescribed in inferior and

superior direction, given by

$$w = A \sin(2\pi f). \tag{10}$$

Thereby, a frequency of $f = 100$ Hz and a oscillation amplitude of $A = 4$ mm is used, ensuring that the minimal gap between the two vocal folds of $g = 0.2$ mm is kept.

For the acoustics, the vocal tract model is attached to the larynx and consists of multiple frustums concatenated one after another. The number of frustums and their radius determines the resulting sound radiating from the artificial mouth. In this case the vocal tract models the sound /u/ ("who"), according to the works of [7], who acquired 3D vocal tract shapes using magnetic resonance imaging (MRI).



(a) Geometric model of the human larynx in coronal section. Comparison of the fine CFD grid and coarse acoustic grid

(b) Larynx, vocal tract, propagation region, perfectly matched layer (PML) regions

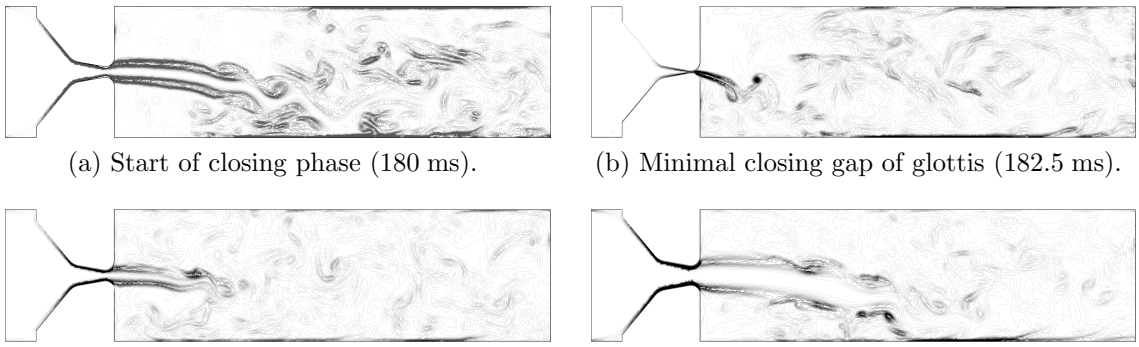
Figure 2: Geometry and mesh for the simulation.

3.2 Mathematical model

3.3 CFD model

The flow simulation is according to the 2D model in Sec. 2.2 with the difference of being a 3D simulation and using prescribed motion of the VFs. Other alterations include the pressure at the inlet is set to $p_{\text{in}} = 300$ Pa (i.e. lung pressure), $p_{\text{out}} = 0$ Pa at the outlet and $\partial p/\partial \mathbf{n} = 0$ for the channel walls. The boundary conditions of the fluid velocity are set to $\partial \mathbf{u}_{\text{in}}/\partial \mathbf{n} = 0$ at the inlet and no-slip condition $\mathbf{u} = 0$ on the fixed walls. On the vocal fold surfaces, the flow velocity is set to the velocity of the moving vocal fold.

In Fig. 3 four time instants of a full phonation cycle are shown. The flow field is dominated by a pulsating jet, which is gradually cut off by the VFs during the closing stage (see Fig. 3a and Fig. 3b). As the VFs open again (Fig. 3c) the jet starts to impinge into the supraglottal region and is fully developed at the maximal opening phase (Fig. 3d). The deflection of the jet is caused by interacting with a large scale vortex downstream. This redirection of the jet repeats itself from cycle to cycle, since the closing phase is too short for the large scale vortex to completely disperse.



(a) Start of closing phase (180 ms). (b) Minimal closing gap of glottis (182.5 ms).

(c) Opening phase of glottis (185 ms). (d) Maximal opening of glottis (187.5 ms).

Figure 3: Flow field at the cross section along glottis mid-line. Vorticity contours at four time instants during 18th (180–190 ms) cycle of vocal fold vibration.

3.4 CAA model I: Lighthill's analogy

Lighthill's approach [8] assumes that the compressible field variables, pressure p^{c}, velocity $\mathbf{u}^{\mathrm{c}}$ and density ρ^{c} may be split into a mean and a fluctuating part [2]

$$p^c = \overline{p} + p' \quad ; \quad \mathbf{u}^c = \overline{\mathbf{u}} + \mathbf{u}' \ . \tag{11}$$

It is essential for the validity of this ansatz that the perturbation quantities are much lower than the mean quantities (i.e. $p' << \overline{p}$). Thereby, the compressible Navier-Stokes equations can be reformulated into an inhomogeneous wave equation

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}p' - \frac{\partial^2}{\partial x_i^2}p' = \frac{\partial^2}{\partial x_i \partial x_j}T_{ij} \tag{12}$$

with p' the fluctuating pressure (superposition of flow pressure p^{ic} and acoustic pressure p^{a}), c the speed of sound and Lighthill tensor $\mathbf{T}$

$$T_{ij} = \underbrace{\rho u_i u_i}_{\text{Reynolds stress}} - \underbrace{\tau_{ij}}_{\text{Viscous stress}} + \underbrace{\left(p' - c^2\rho'\right)\delta_{ij}}_{\text{Heat conduction}}, \tag{13}$$

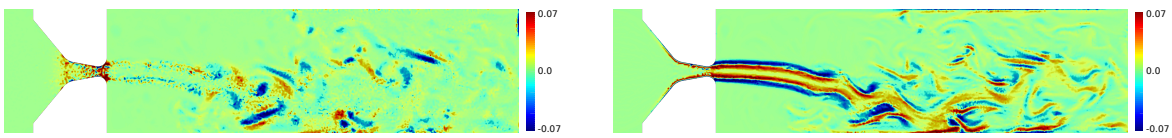
which models the aeroacoustic source term. In (13), τ_{ij} is the stress tensor, ρ' the fluctuating density and δ_{ij} the Kronecker delta. It is assumed that no heat conduction takes place, the viscous stress may be neglected (see [8]) and the flow is incompressible. Therefore an approximation of $\mathbf{T}$ can be given as

$$T_{ij} \approx \rho u_i^{\mathrm{ic}} u_j^{\mathrm{ic}} \ . \tag{14}$$

Here, we want to note that the solution of Lighthill's inhomogeneous wave equation is a fluctuating pressure p', which just approaches the acoustic pressure p^{a} in the far field (outside the flow region).

[2] The superscript c denotes that the physical quantities are compressible.

It can be shown that the divergence of the Lamb vector is the main acoustic source at low Mach numbers which radiates into the far field. Furthermore, we found that the second derivative of Lighthill's tensor leads to stronger noise in the acoustic source structures, as shown in Fig. 4. It needs to be pointed out, that the scale has been adjusted to make the sources in the turbulent region visible. Inside the glottis are the dominant source terms, which are more than an order higher.



(a) Second derivative in space of the Lighthill tensor (RHS of wave equation (12)).

(b) Divergence of the Lamb vector.

Figure 4: Visualization of different aeroacoustic source models on a cross section of the human larynx, at time step 180 ms. Results have been normed to their maximum value.

The wave equation (12) is solved using the Finite-Element Method (FEM) (see [9]). As a rule of thumb 20 linear finite elements should be used per wave length. The grid used for the flow simulation (characteristic length 0.15 mm) is unnecessarily small for computation of acoustic wave propagation. Therefore, the acoustic sources are conservatively interpolated (see [9]) onto a coarser grid. Since the maximum frequency encountered in this application is about 3.5 kHz a characteristic length of 1.3 mm is sufficient as it has a frequency resolution of up to 12 kHz. Once the acoustic sources are determined, the sound propagation is calculated.

We add so called perfectly matched layers (PML) at the inflow as well as at the propagation region to efficiently damp outgoing waves without reflections [10] (see Fig. 2b).

3.5 CAA model II: Perturbation equations

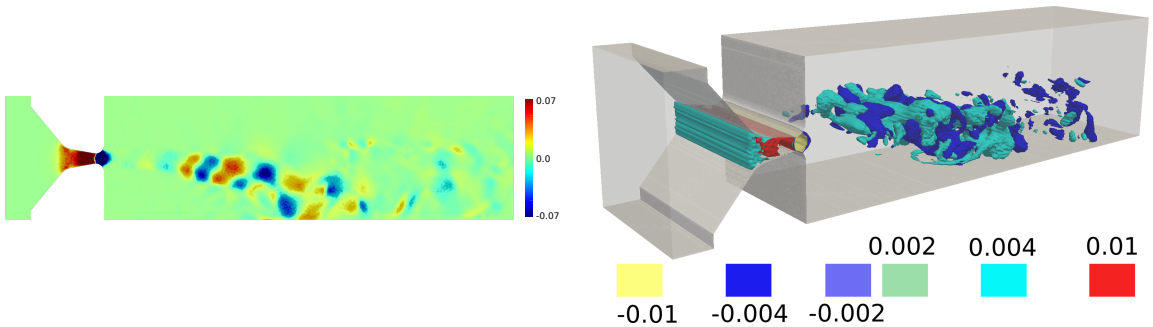
To correctly monitor the acoustics inside the flow field, we use a perturbation approach. The splitting of the field variable, as in (11), is thereby pushed one step further to separate fluctuating flow and acoustic quantities. For the velocity field, this splitting is based on the fact that the incompressible part of the velocity field is purely solenoidal $\nabla \cdot \mathbf{u}^{\mathrm{ic}} = 0$, whereas the acoustic components are irrotational $\nabla \times \mathbf{u}^{\mathrm{a}} = 0$. Thereby we obtain

$$\mathbf{u}^{c} = \overline{\mathbf{u}} + \mathbf{u}' = \overline{\mathbf{u}} + \mathbf{u}^{\mathrm{ic}} + \mathbf{u}^{\mathrm{a}} \ . \tag{15}$$

This is also analogously done for the pressure p^c and density ρ^c. Then based on the equation of mass, momentum and energy conservation, as well as the state equations of fluid, following perturbation equations (PE) can be derived

$$\begin{aligned} \frac{\partial}{\partial t} p^{\mathrm{a}} + \rho c^2 \nabla \cdot \mathbf{u}^{\mathrm{a}} + \nabla \cdot (p^{\mathrm{a}} \bar{\mathbf{u}}) &= -\frac{\partial}{\partial t} p^{\mathrm{ic}} - \bar{\mathbf{u}} \cdot \nabla p^{\mathrm{ic}}, \\ \frac{\partial}{\partial t} \mathbf{u}^{\mathrm{a}} + (\mathbf{u}^{\mathrm{a}} \cdot \nabla) \bar{\mathbf{u}} + (\bar{\mathbf{u}} \cdot \nabla) \mathbf{u}^{\mathrm{a}} + \frac{1}{\rho} \nabla p^{\mathrm{a}} &= 0, \end{aligned} \tag{16}$$

with the acoustic pressure p^{a} and acoustic particle velocity $\mathbf{u}^{\mathrm{a}}$. By also regarding the mean velocity $\overline{\mathbf{u}}$ convection and refraction effects are accounted for, which is not done for the wave equation (12). The right hand side, or source term, is determined with the help of the CFD results, which provides the incompressible variables p^{ic} and, by time averaging, the mean flow field $\overline{\mathbf{u}}$. Figure 5a presents a coronal section of the flow field, showing significantly less noise in the structures as Lighthill source term does (see Fig. 4). The 3D iso-surfaces in Fig 5b show the 3D structure, which correlates to the jet (see Fig. 3). Clearly visible are the significantly higher source term amplitudes inside the glottis.



(a) Coronal cross section of the human larynx.

(b) 3D iso-surfaces of acoustic sources.

Figure 5: Visualization of aeroacoustic sources of PE at time step 180 ms. Results have been normed to their maximal value.

For a detailed derivation of the system (16) we refer to [11] and [12], who points out the close relation of (16) to the *Acoustic Perturbation Equations* (APE), for which the splitting into perturbation quantities was initially introduced [13].

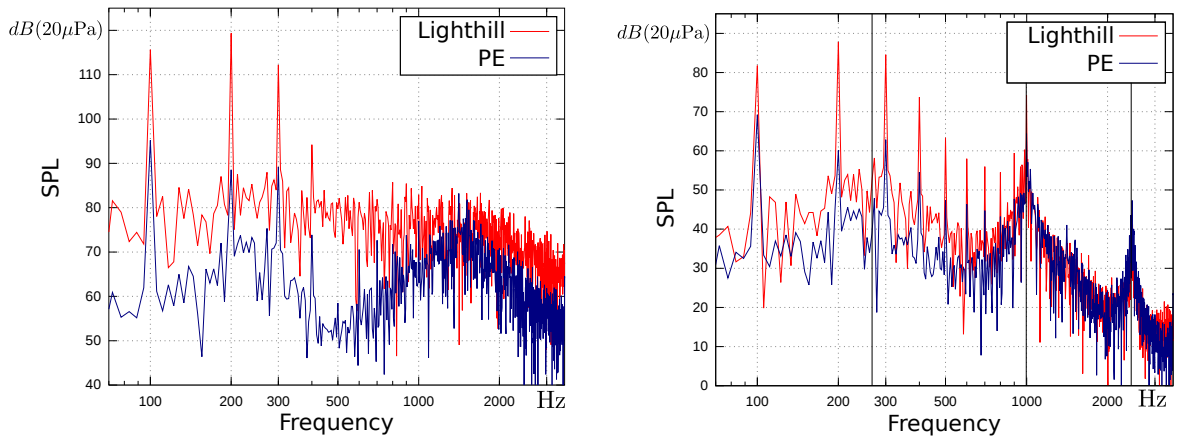
3.6 Aeroacoustic results and comparison

For the evaluation of the acoustic pressure and the comparison of the hybrid approaches, two monitoring points are considered. The first "MIC 1" is situated at the end of the larynx, inside the flow domain. "MIC 2", the second monitoring point, is 1 cm behind the vocal tract, in the acoustic propagation region. For these two monitoring points the acoustic sound pressure (SPL) is plotted in Fig. 6. As expected, Lighthill's approach using the fluctuation pressure p' has significantly stronger amplitudes inside the flow region, as evident in Fig. 6a. Comparing the splitting of the field quantities, between Lighthill's approach (11) and the PE approach (15), it is clear that the computed pressure for the wave equation case is superimposed by the incompressible flow pressure p^{ic}. However, this superimposition does not propagate into the far field for all frequencies, as Fig. 6b shows. Here, at "MIC 2", formants are aligned and are comparable to the formants recorded in natural speech, as given in Tab. 2. Both analogies, PE and Lighthill's approach are in very good agreement, except for the main frequency (100 Hz) and its harmonics. Here, Lighthill's wave equation again shows an over-prediction of the SPL for

Table 2: Recorded formants from natural speech "N", as given by [7], and simulated by our reproduced vocal tract "S" (Lighthill's analogy and PE are identical).

	$/\mathrm{u}/^N$	$/\mathrm{u}/^S$
F1	389 Hz	270 Hz
F2	987 Hz	1000 Hz
F3	2299 Hz	2484 Hz

these frequencies.



(a) Acoustic pressure at "Mic 1".

(b) Acoustic pressure at "Mic 2". Vertical black lines indicate the formant frequency.

Figure 6: Acoustic pressure for both monitoring points, for the wave equation and the PE.

3.7 Summary

Two aeroacoustic models of the human voice generation have been developed. CFD simulation were based on the unsteady incompressible Navier-Stokes equations in 3D utilizing the finite volume code OpenFoam. Vocal fold movement were simulated by prescribing a predefined vibration. Using the resulting fluid pressure and velocity, the acoustic source terms and wave propagation could be calculated based on Lighthill's analogy and the perturbation equations. Both showed a good agreement in the far field, between each other but also to measured data. However, Lighthill's analogy does over-predict the SPL of the fundamental frequency as well as its harmonics.

REFERENCES

[1] Link G., Kaltenbacher M., Breuer M., and Döllinger M. A 2D finite element scheme for fluid-solid- acoustic interactions and its application to human phonation. *Comput.*

Method. Appl. M. (2009).

[2] Alipour, F. et al. Mathematical models and numerical schemes for the simulation of human phonation. *Curr. Bioinform.* (2011).

[3] Seo, J. H. and Mittal, R. A high-order immersed boundary method for acoustic wave scattering and low-Mach number flow-induced sound in complex geometries. *J. Comput. Phys.* (2011).

[4] Mattheus W. and Brücker C. Asymmetric glottal jet deflection: Differences of two- and three-dimensional models. *J.Acoust. Soc. Am. Expr. Let.* (2011)

[5] Brooks A. N. and Hughes T. J. R. 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*, 32:199–259, (1982).

[6] Scherer, R. C., Shinwari, D. , Witt, K. J. D., Zhang, C. , Kucinschi, B. R. and Afjeh A. A. Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. *The Journal of the Acoustical Society of America*, 109(4):1616–1630, (2001).

[7] Story, B.H., Titze, I. and Hoffman, E.A. Vocal tract area functions from magnetic resonance imaging. *J Acoust Soc Am*, 100(1):537–554, (1996).

[8] Lighthill M. On sound generated aerodynamically I. General theory, *Proceedings of the Royal Society of London* 211:564–587, (1951).

[9] Kaltenbacher, M. *Numerical Simulation of Mechatronic Sensors and Actuators*, 2. edition (Springer, Berlin), (2007).

[10] Kaltenbacher, B., Kaltenbacher, M. and Sim I. 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. *J. Comput. Phys.*, 235(0):407–422, (2013).

[11] Hüppe, A. *Spectral finite elements for acoustic field computation*, (Measurement-, Actuator-, and Simulation-technology). PhD thesis, Shaker Verlag GmbH

[12] Hüppe, A. and Kaltenbacher, M. Spectral finite elements for computational aeroacoustics using acoustic perturbation equations. *J. Comput. Acoust.*, 20(02):1240,005 1–13, (2012).

[13] Ewert, R. and Schröder, W. Acoustic perturbation equations based on flow decomposition via source filtering. *J. Comput. Phys.*, 188(2):365–398, (2003).