forum acusticum 2026
8-12 September · Graz, Austria

AEROACOUSTIC ANALYSIS OF PRESCRIBED VOCAL FOLD MOTION IN THE SIMVOICE BENCHMARK USING WALL-MODELLED LARGE-EDDY SIMULATION

Thomas Brunner[1], Christoph Heigl[2], Stefan Kniesburges[3], Stefan Becker[2], Johannes Kreuzinger[4], Florian Schwertfirm[4], Stefan Schoder[1]

[1] *Institute of Fundamentals and Theory in Electrical Engineering, TU Graz*
[2] *Institute of Fluid Mechanics (LSTM), FAU Erlangen-Nürnberg*
[3] *HNO Klinik, Abteilung f. Phoniatrie und Pädaudiologie, Universitätsklinikum Erlangen*
[4] *KM Turbulenz GmbH.*

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Abstract

Voice disorders affect millions of people across Europe, with significant implications for quality of life. Understanding the physical mechanisms underlying phonation and their variation under disturbances remains a central challenge. While the fundamental principles of voice production are well established, accurately modelling the complex fluid–structure–acoustic interactions between the vocal folds and the laryngeal airflow remains challenging [1]. To foster reproducibility and cross-validation within the phonation research community, the SimVoice benchmark was established as a shared reference for both numerical and experimental investigations. In this study, complementary numerical and experimental analyses of the benchmark configuration are performed to investigate its aeroacoustic characteristics and further support methodological development. A computationally efficient framework based on the MGLET solver is employed to conduct Large-Eddy Simulations (LES) with prescribed vocal fold motion, permitting parameter studies of subglottal pressure variations and deformation patterns. Additional measurements using both rigid (non-oscillating) and deformable vocal folds provide a robust extension of the dataset for model calibration and parameter tuning. Exploring multiple operating points enhances the benchmark's characterization, advancing the broader goal of reproducible, physically grounded voice simulation. Our findings give new insights into the physics of voice production and contribute to the continued validation of the SimVoice benchmark.

Corresponding author: *thomas.brunner@tugraz.at.*

Keywords: *computational aeroacoustics, computational fluid dynamics, large eddy simulation, phonation*

1. Introduction

Voice production arises from the fluid-structure-acoustic interaction of airflow, deformable vocal folds, and the resulting acoustic field. Even small changes in tissue properties, geometry, or subglottal pressure can noticeably alter the generated sound, which shows the sensitivity of the underlying flow and acoustic mechanisms. While the basic principles of phonation are well understood, quantitatively predicting the aeroacoustic sources and the radiated sound remains challenging. Experimental studies with simplified laryngeal models and high-fidelity computational fluid dynamics (CFD), have improved the understanding of glottal jet dynamics and sound generation. However, differences in geometry, boundary conditions, and operating parameters often make direct comparison between studies difficult, limiting reproducibility and systematic validation.

To address this issue, the SimVoice benchmark was introduced as a shared and well-defined reference configuration for phonation research [2]. It provides a precisely specified geometry, operating conditions, and reference data, allowing comparison between numerical and experimental approaches. The benchmark was originally investigated in combined experimental and numerical studies using deformable vocal fold models. A key challenge remains the consistent separation of flow-induced sound generation from fluid–structure coupling effects. Conducting additional reference measurements using rigid models allows efficient preemptive aeroacoustic calibration of numerical models without the computational burden inflicted by full simulation of fluid-structure interaction. Further systematic characterization of the studied configura-

tion, particularly with respect to aeroacoustic source mechanisms, parameter sensitivity and methodological robustness is thus essential to further exploit the benchmarks potential.
In the present work, we build upon the SimVoice benchmark [2] by performing complementary numerical and experimental investigations of its aeroacoustic characteristics. A computationally efficient framework based on LES with optional prescribed vocal fold motion is presented to study the flow field and its acoustic emissions under controlled variations of subglottal pressure and deformation. The configurations using rigid vocal fold models serve as a controlled reference for fluid-dynamic and acoustic model tuning, isolating flow-induced sound generation from fluid–structure coupling effects, allowing more computationally efficient parameter tuning using non-deforming simulations.

2. Methodology

2.1. Experimental details

The experimental setup (shown in Figure 1) is based on a rectangular channel consisting of aluminium (sub- and supraglottal) and acrylic glass (glottal) sections with cross-sectional dimensions of (15×18mm). For further information on the geometry we want to reference to the benchmark definition [2]. At defined positions a vocal fold and a ventricular fold pair is glued into the channel. The models are based on the M5 model. Two version of vocal fold models are investigated. For the studies investigating rigid vocal folds an adjusted M5 model with an enlarged glottal gap of 1mm is used to reduce viscous blocking, therefore allowing higher flow speeds and stronger aeroacoustic sources, leading to enhanced acoustic excitation. The rigid models were 3D-printed using stereolithography (SLA). The deformable silicon vocal fold models are based on the classical M5 model and were manufactured using a 3D-printed mold.
As shown in Figure 1a.) total of 8 (7 supraglottal Kulite XCS-093 and 1 subglottal Kulite XCQ-093) wall pressure probes were installed centred along the side wall of the channel. The wall pressure probe locations (indexed from 0-7, see Figure 1a.)) are placed at distances of $10, 30, 50, 70, 110, 150, 190, -50\,\mathrm{mm}$ from the trailing edge of the vocal fold model. In the far-field three microphones (Brüel and Kjaer 4189-L-001 1/2") were installed at a fixed distance of 1m to the end of the supraglottal channel at azimuthal angles of 30°, 60°, and 90°. Additionally the flow rate was measured at a upstream position. The experimental sensor data consists of 60s measuring intervals at a sampling frequency $f_\mathrm{s} = 48$kHz for each operating point.

2.2. Aeroacoustic modelling

The low Mach number $\mathrm{Ma} < 0.3$ allows treating the laryngeal air flow as an incompressible constant density flow. Thus the fluid can be described by the incompressible Navier-Stokes equations. The air is

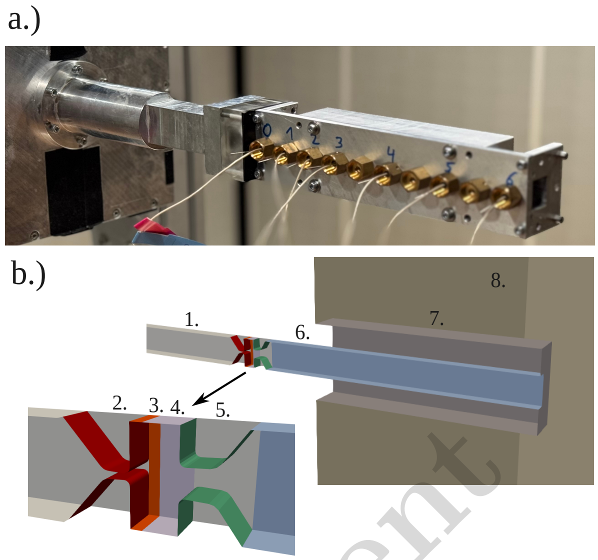

Figure 1: a.) the experimental setup with the marked wall pressure probe locations. b.) the computational setup coloured by the different regions referenced in the paper.

modelled using a constant density $\rho = 1.2\frac{\mathrm{kg}}{\mathrm{m}^3}$ and a kinematic viscosity of $1.53 \cdot 10^{-5}\frac{\mathrm{m}^2}{\mathrm{s}}$ To resolve the largest scales of turbulence inside the flow directly we use LES. The governing equations that need to be solved in LES are based on a spatial filtering operation which separates the flow into resolved $\bar{(\cdot)}$ and sub-grid scale (SGS) contributions. The filtered incompressible Navier-Stokes equations can be written as

$$\nabla \cdot \bar{\boldsymbol{v}} = 0\,, \tag{1}$$

$$\frac{\partial \bar{\boldsymbol{v}}}{\partial t} + (\bar{\boldsymbol{v}} \cdot \nabla)\bar{\boldsymbol{v}} = -\frac{1}{\rho}\nabla \bar{p}^{ic} + \nu \nabla^2 \bar{\boldsymbol{v}} - \nabla \cdot \boldsymbol{\tau}^{sgs}\,. \tag{2}$$

The system is closed by the assumed eddy viscosity ansatz for the SGS stress tensor $\boldsymbol{\tau}^{sgs}$ and the employed Smagorinsky model for the eddy viscosity.

To account for the aeroacoustic sources and acoustic propagation during voice production a hybrid aeroacoustic method is employed, namely a variant of the acoustic perturbation equations (APE). The used variant APE1 system [3] solves for the total compressible pressure $p^c := p' = \bar{p}^{ic} + p^a$ in the using the scalar wave equation

$$\frac{1}{c^2}\frac{\partial^2 p'}{\partial t^2} - \nabla^2 p' = -\nabla^2 \bar{p}^{ic}\,, \tag{3}$$

where the aeroacoustic source term $-\nabla^2 \bar{p}^{ic}$, that is based on the incompressible flow simulation, appears on the right side. This workflow represents a one-way coupling of flow to acoustics, which thus excludes acoustic feedback on the flow. The adopted hybrid formulation allows a clear distinction between hydrodynamic source generation and acoustic propagation.

2.3. Numerical treatment

Both the incompressible flow and the acoustic wave equation are solved using the commercial solver MGLET [4]. The solver requires an underlying Cartesian grid with octree-based refinement. The grid structure allows the efficient use of a geometric multigrid method. The simulation domain boundaries are treated by a cut-cell immersed boundary approach. The flow and acoustic simulation are temporally discretised by an explicit three and four stage Runge-Kutta scheme respectively, with several acoustic timesteps per flow timesteps due to the characteristic velocities ($\bar{v} \ll c_0$).

To allow for efficient computations we rely on a Werner-Wengle [5] wall model to treat the near-wall region of the flow. The use of wall modelling is motivated by the need to resolve aeroacoustically relevant flow structures at realistic Reynolds numbers within feasible computational cost. The flow was generated using pressure in- and outlet boundary conditions, which were chosen according to the pressure drop observed in the experiment. The walls of the subglottal (Number 1. in 1), supraglottal (6.), and exterior (7.) channels are modelled as rigid walls with a frequency independent acoustic absorption rate of 2%. The front wall of the supraglottal channel which houses the wall pressure probes has been modelled by a slightly higher (5%) absorption rate to incorporate the dissipative effects due to clearances. These assumptions are intended to represent unresolved dissipative effects of material properties and geometric tolerances in the experimental setup. The rigid vocal fold pair was modelled using a rough wall function with a equivalent sand roughness of $k_s = 3 \cdot 10^{-4}$m, which was adjusted to match the observed flow resistance. The rough wall model is introduced to represent the effective momentum loss associated with the unresolved interaction of the rough surface with the modeled hydrodynamic boundary layer. The roughness value also matches the real surface finish of the SLA-printed part. The effect of the rough surface of the vocal- and ventricular fold models on the acoustic damping was modelled by an absorption rate of 5%. The acrylic glass wall (2.) also has an absorption rate of 5%. The black rubber material (3.) used to seal the separate channels was explicitly modelled with an absorption rate of 50%, which was found to help reduce a unphysical artificial oscillation in the simulation. The pressure outlet boundary condition was prescribed in the far-field at the box shaped boundary (8.) with a total size of (0.438×0.384×0.384m). To avoid unphysical artificial sources the aeroacoustic source term were damped near the external domain boundaries using a sponge formulation. Radiation of acoustic waves was treated by non-reflecting boundary conditions at both the pressure inlet and in the far-field.

In the simulation probe points were placed at the positions of the installed wall pressure sensors. Three probe points were included at the corresponding exterior microphone positions at a reduced distance of 0.08m at the same azimuthal angles. To compare the corresponding pressure signals with the ones obtained at the experimental measurement distance the pressure levels were scaled assuming ideal spherical wave propagation $\mathrm{PSD}(r_{\mathrm{meas}}) = (r_{\mathrm{meas}}/r_{\mathrm{sim}})^2\, \mathrm{PSD}(r_{\mathrm{sim}})$. This approximation is expected to introduce some discrepancies between the experimental and numerical setup.

The generated numerical grid consists of 7 octree levels leading to a minimal cell size of $h_{\min} = 125\mu$m in the region of the vocal folds. In the channel a mesh size of $h_{\mathrm{channel}} = 0.5$mm was chosen to resolve large turbulent fluctuations and high frequency acoustics. An acoustic wavelength of a 10kHz wave is thus resolved by 68 second-order cells. The simulations with rigid vocal folds lead to a grid consisting of $\sim$ 20M cells. The deforming simulations lead to grids of $\sim$ 22M cells.

Table 1: An overview of the conducted experiments (Exp.) and simulations (Sim.) using both rigid and moving vocal fold models at their respective operating point consisting of the volume flux $\dot{V}$ in $\frac{1}{\min}$ and the mean subglottal pressure $\overline{P_{\mathrm{sub}}}$ in Pa.

Abbrev.	Type	VFs	$\dot{V}$	$\overline{P_{\mathrm{sub}}}$
SR50	Sim.	rigid	52.2	2000
SR65	Sim.	rigid	63.8	3000
SR75	Sim.	rigid	73.5	4000
SM90	Sim.	moving	90	2636
ER40	Exp.	rigid	40	1113
ER45	Exp.	rigid	45	1420
ER50	Exp.	rigid	50	1785
ER55	Exp.	rigid	55	2168
ER60	Exp.	rigid	60	2570
ER65	Exp.	rigid	65	3057
ER70	Exp.	rigid	70	3529
ER75	Exp.	rigid	75	4050
ER80	Exp.	rigid	80	4681
ER85	Exp.	rigid	85	5270
ER90	Exp.	rigid	90	6108
ER95	Exp.	rigid	95	6237
ER100	Exp.	rigid	100	6398
ER105	Exp.	rigid	105	7234
ER110	Exp.	rigid	110	8029
EM90(1)	Exp.	moving	90	3437
EM90(2)	Exp.	moving	90	3374
EM100	Exp.	moving	100	3368

2.4. Prescribed deformation

The prescribed deformation of the vocal fold in the deforming simulations is based on the function presented in the benchmark. An oscillation frequency of 148Hz was chosen to agree with the experimentally determined oscillation frequency. The addition of a tilting movement present in the benchmark was omitted due to restrictions of the prescribable motion in the solver. The time depended displacement vector $\boldsymbol{d}(t)$ is described by

$$\boldsymbol{d}(t) = \boldsymbol{x}(t) - \boldsymbol{X} = f(\tilde{Y}) \begin{bmatrix} -\tilde{X} \\ 0 \\ \tilde{Z} \end{bmatrix} g(t)\,, \qquad (4)$$

where $f(\tilde{Y}) = \frac{1}{2}(cos(2\pi\frac{\tilde{Y}}{L}) + 1)$ and the time function $g(t)$ from the benchmark [2]. This satisfies $\nabla \cdot \boldsymbol{d}(t) = 0$, which is the first order linearized approximation of the general geometrically nonlinear volume preservation condition $J = \det \mathbf{F} = 1$.

2.5. Overview of the conducted studies

In Table 1 we present an overview of all conducted numerical and experimental studies. These consist of three simulations investigating the rigid vocal fold models, with the corresponding experimental studies with differing volume fluxes $\dot{V} = 40 - 110\frac{\mathrm{l}}{\mathrm{min}}$. For the deforming vocal fold model three additional experimental measurements were conducted, where two studies at $90\frac{\mathrm{l}}{\mathrm{min}}$ showed mechanical contact and one study at $100\frac{\mathrm{l}}{\mathrm{min}}$ which did not show mechanical contact. The non-contacting study allows the analysis of the effect of contact on the acoustic excitation.

3. Results

3.1. Rigid models

The experimental data shows the characteristic radical dependence of the volume flux $\dot{V}$ on the subglottal pressure $\overline{P_{\mathrm{sub}}}$, as can be inferred by the data listed in Table 1. The self-adjusted volume flux of the incompressible simulations agree well with the experimental operating points. Field plots of the instantaneous velocity magnitude u extracted from the rigid SR65 simulation are shown in Figure 2. The flow is seen to transition from a laminar subglottal state to turbulence in the inter-glottal region. A jet is formed by the high speed flow that is ejected out of the vocal fold pair. The generated turbulent eddies are convected further downstream by the mean flow, where the small scales are observed to dissipate. In the SR65 simulation, maximum instantaneous flow velocities of up to (100m/s) occur only within localized turbulent fluctuations. The mean velocity reaches approximately (70m/s) in the constriction and jet region, while the mean velocity in the channel is 10m/s. Since the high-speed region near the vocal folds is acoustically compact, neglecting the convective term may locally affect the acoustic source region and the near-field propagation. However, acoustic propagation through the channel is dominated by the much slower mean flow. Therefore, the effect of neglecting convection in the acoustic propagation operator is expected to be small in the channel.

Investigating the spectra of the experiment at an example wall pressure sensor location in Figure 3 reveals a sudden change of fluid mechanical and acoustic behaviour above 95 l/min. A low frequency resonant effect at 300Hz, as well as extreme whistling at 10kHz are found. This behaviour suggests the presence of a coupled flow–acoustic feedback mechanism at higher

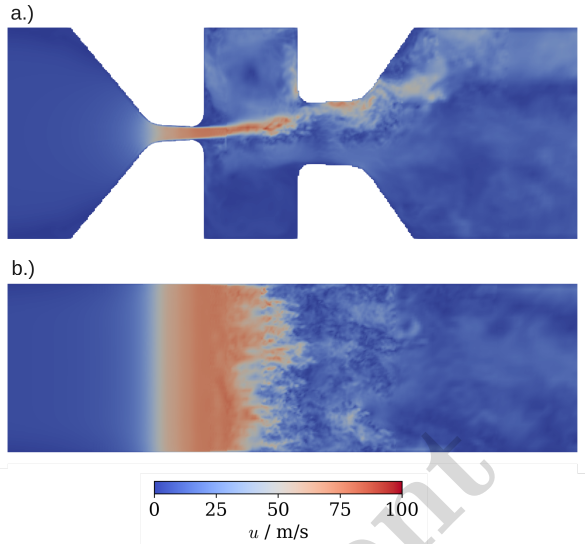

Figure 2: Field plots of the velocity magnitude u in the SR65 simulation, where a.) shows the centre slice in the x-z plane and b.) shows the centre slice in the x-y plane.

flow rates, which can not be modeled using the used numerical framework. Therefore we restrict our rigid comparison simulations to flow rates < 80l/min. The geometrical modes of the supraglottal channel, which correspond to formants in phonation research, are well defined in the experimental data.

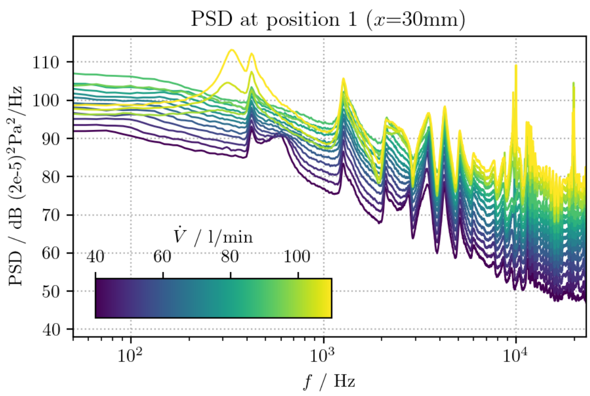

Figure 3: Overview of the measured PSD obtained at an example channel wall pressure probe location (1) for the different operating points ER40-110 of the experiment using the rigid models.

Comparing example spectra of a rigid simulation with the corresponding experimental operating point shows reasonable agreement, as can be observed in Figure 4. At frequencies below 20kHz the simulation over-predicts pressure levels. The structure of the geometric acoustic modes is well defined in the simulation data. Due to short simulation times, little windowing is possible without losing too much of the low frequency behaviour. We expect that longer simulation times would further improve the agreement of the low frequency spectral structures. Small deviations in the

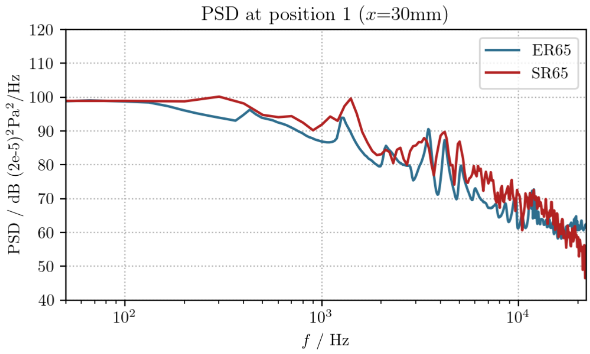

Figure 4: Comparison of the PSD obtained from the SR65 simulation and the corresponding experiment ER65 at an example channel wall pressure probe location (1).

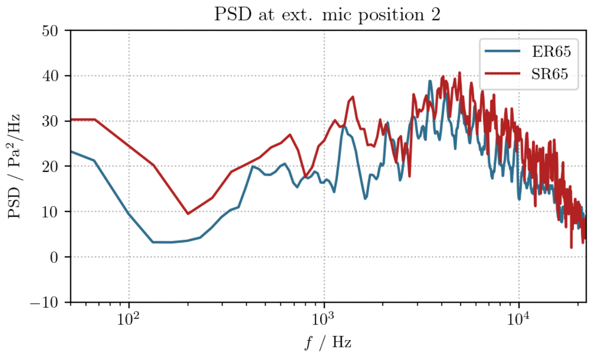

Figure 5: Comparison of the PSD of the radiated acoustics obtained from the SR65 simulation and the corresponding experiment ER65 at the external microphone position (2).

frequency values of the modes are apparent, which are likely caused by differences in the occurring acoustic boundary condition of the supraglottal channel at the ventricular folds. The numerical results therefore indicate that the dominant spectral structure is primarily influenced by the acoustic characteristics of the channel rather than by detailed turbulence modelling. The structure of the spectra obtained by the far-field probes, exemplarily shown in Figure 5, agree well. The pressure levels are again over-predicted in the simulations, which were sampled at a distance of 80mm and adjusted using the radial distance scaling of a spherical wave.

3.2. Deforming models

The use of the deforming silicon vocal fold model leads to the characteristic spectra related to the fluid-structure-acoustic interaction of the laryngeal airflow with the oscillating structure. The main dynamics of the deformation occurring at the oscillation frequency and its higher harmonics lead to tonal peaks, as can be seen in an example pressure probe spectrum in Figure 6. The opening and closing of the flow path leads to an oscillating volume flux, which influences both the time-dependent separation behaviour, as well as the glottal jet itself. The main tonal peaks are attributed

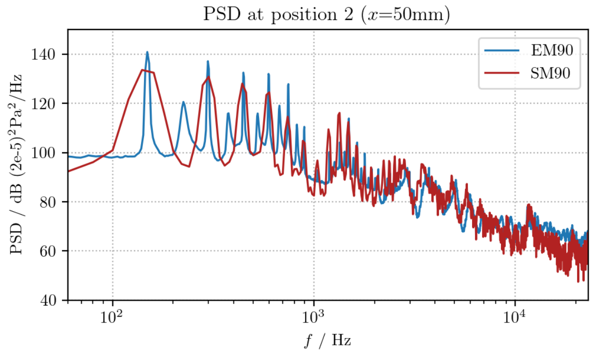

Figure 6: Comparison of the PSD obtained from the deforming SM90 simulation and the corresponding experiment EM90 at an example channel wall pressure prove position (2).

to this mechanism. Multiple geometric channel modes are superimposed on the underlying broadband aeroacoustic noise sources, which are caused by the turbulent fluctuations, and attenuate/amplify the spectrum close to their resonant frequency. A major difference in the spectrum is the emergence of odd harmonics $3/2, 5/2, \ldots \cdot f_{\text{osci}}$ of the first subharmonic of the oscillation frequency, which is caused by the more complex real deformation observed in the experiment. The deformation thus includes dynamics with differing periods of that of the main oscillation frequency, which are not represented in the prescribed-motion model and thus do not appear in its spectrum.

4. Discussion

The presented results show that the SimVoice benchmark provides an effective configuration for studying the interaction of hydrodynamic source mechanisms and the acoustic sound generation. The rigid vocal fold cases allow the investigation of flow-induced sound generation without the additional complexity of fluid-structure interaction. In these cases the measured and simulated spectra show that the spectral structure is strongly influenced by the acoustic properties of the supraglottal channel. The agreement of the modal structure between experiment and simulation indicates that the employed hybrid aeroacoustic formulation is able to capture the aeroacoustic mechanisms of the setup, as long as feedback effects (such as the encountered whistling at high flow rates) do not dominate the spectrum.

The observed discrepancies in pressure level show the limitations of the present modelling approach. In both the wall pressure and far-field spectra the numerical results tend to overpredict the measured amplitudes. This behaviour is likely caused by factors such as uncertainties in the acoustic boundary conditions, the simplified treatment of acoustic damping, and the extrapolation of the simulated far-field probe signals from the shorter radial distance in the simulation. A rough wall boundary condition was used to account for the momentum loss at the rigid vocal folds. The presented results show that these assumptions allow reproduction of the dominant spectral features.

The sudden change in the experimental rigid-model spectra at high volume fluxes further suggests that an additional coupled flow-acoustic feedback mechanism becomes relevant in this regime. The occurrence of strong whistling and low-frequency resonance above approximately 95 $\frac{1}{\text{min}}$ indicates that the flow is no longer governed only by one-way fluid-acoustic coupling. As these effects were not considered in the present numerical setup, the comparison was restricted to lower flow rate operating points, which are also used to imitate physiological flow conditions.

For the deforming vocal fold cases, the prescribed-motion simulations reproduce the main tonal peaks at the oscillation frequency and its harmonics, showing that the dominant periodic forcing of the flow is represented well by the chosen prescribed kinematics. However, the comparison with the experiment also reveals clear differences, such as the appearance of additional subharmonic-related peaks in the measured spectra. These point to more complex deformation dynamics in the experiment, which are not included in the motion prescription of the simulation.

Overall, the results support the use of rigid cases as an efficient first validation step before considering more complex deforming configurations.

5. Conclusion

A combined numerical and experimental investigation of the SimVoice benchmark has been conducted to assess the aeroacoustic behaviour of a simplified phonation model. Additional experimental measurements using rigid and deformable vocal fold models were performed across a range of operating conditions and compared with corresponding numerical simulations. The rigid vocal fold cases enabled a systematic analysis, isolating it from effects of fluid-structure interaction and provides a controlled reference for calibrating numerical models. The simulations were able to reproduce the dominant spectral characteristics of the measured wall-pressure and far-field acoustic signals, including the modal structure of the supraglottal duct. For the cases investigating deforming vocal fold models the comparison with experimental data shows that additional structural dynamics occur, which are not modeled in the simulation. These odd harmonics cause additional tonal peaks in the acoustic response. The present work extends the experimental and numerical reference data available for the SimVoice benchmark and illustrates the value of combining rigid and deforming configurations for the validation of aeroacoustic simulation methodologies. The findings provide practical guidance for the development of computational approaches to voice production and contribute to the broader objective of establishing reproducible and physically grounded benchmark cases in phonation aeroacoustics.

References

[1] M. Döllinger, Z. Zhang, S. Schoder, P. Šidlof, B. Tur, and S. Kniesburges, "Overview on state-of-the-art numerical modeling of the phonation process," *Acta Acustica*, vol. 7, p. 25, 2023.

[2] S. Schoder, S. Falk, A. Wurzinger, A. Lodermeyer, S. Becker, and S. Kniesburges, "A benchmark case for aeroacoustic simulations involving fluid-structure-acoustic interaction transferred from the process of human phonation," en, *Acta Acustica*, vol. 8, p. 13, 2024, ISSN: 2681-4617. DOI: 10.1051/aacus/2024005. Accessed: Oct. 22, 2025. [Online]. Available: https://acta-acustica.edpsciences.org/10.1051/aacus/2024005.

[3] 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, Jul. 2003, ISSN: 0021-9991. DOI: 10.1016/S0021-9991(03)00168-2. Accessed: Nov. 5, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999103001682.

[4] F. Kurz, J. Kreuzinger, F. Schwertfirm, and G. K. Dutzler, "Cut-Cell Immersed Boundary Methode für hybride CFD/CAA Simulationen mit bewegten Körpern: 40. Jahrestagung für Akustik, DAGA 2014," Mar. 2014.

[5] H. Werner and H. Wengle, "Large-Eddy Simulation of Turbulent Flow Over and Around a Cube in a Plate Channel," en, in *Turbulent Shear Flows 8*, F. Durst, R. Friedrich, B. E. Launder, F. W. Schmidt, U. Schumann, and J. H. Whitelaw, Eds., Berlin, Heidelberg: Springer, 1993, pp. 155–168, ISBN: 978-3-642-77674-8. DOI: 10.1007/978-3-642-77674-8_12.

6. Acknowledgments

The auhors would like to thank the Austrian Federal Ministry for Innovation, Mobility and Infrastructure and the Austrian Research Promotion Agency (FFG) for funding the bioCOMP4acoustics project under the TakeOff programme line (project number 913972) and the ECHODA project under the COMET programme line.