TECHNISCHE UNIVERSITÄT WIEN

Efficient FEM Model for Human Phonation

Master's Thesis

Paul Maurerlehner

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

TECHNISCHE UNIVERSITÄT WIEN

DIPLOMARBEIT

Efficient FEM Model for Human Phonation

ausgeführt zum Zwecke der Erlangung
des akademischen Grades eines Diplom-Ingenieurs
unter der Leitung von

Dipl.-Ing. Dr.techn. Stefan Schoder
Univ.-Prof. Dipl.-Ing. Dr.techn. Manfred Kaltenbacher
Institut für Mechanik und Mechatronik, E325-A3

eingereicht an der Technischen Universität Wien
Fakultät für Maschinenwesen und Betriebswissenschaften
von

Paul Maurerlehner
01427390
Biberbachstraße 2
3353 Seitenstetten

Wien, am 20.12.2019

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich meine Diplomarbeit mit dem Titel *Efficient FEM Model for Human Phonation* selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe und dass ich alle Stellen, die ich wörtlich oder sinngemäß aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe. Die Arbeit hat bisher in gleicher oder ähnlicher Form oder auszugsweise noch keiner Prüfungsbehörde vorgelegen.

Wien, am 25.02.2020

Paul Maurerlehner

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Danksagung

An dieser Stelle möchte ich all jenen danken, die durch ihre fachliche und persönliche Unterstützung zum Gelingen dieser Diplomarbeit beigetragen haben. Allen voran möchte ich meinen Eltern Brigitte und Robert Maurerlehner danken, die mir mein Studium ermöglicht haben. Ein großer Dank gilt vor allem auch meiner Freundin Magdalena Wallner für ihre bedingungslose Unterstützung während meines gesamten Studiums.

Ein besonderer Dank gebührt Dipl.-Ing. Dr.techn. Stefan Schoder für die äußerst engagierte Betreuung. Darüberhinaus bedanke ich mich bei Univ.Prof. Dipl.-Ing. Dr.techn. Manfred Kaltenbacher für die Möglichkeit in seiner Arbeitsgruppe arbeiten und auch diese Arbeit verfassen zu können. Großer Dank gilt auch unserem Projektpartner Universitätsklinikum Erlangen (Abteilung für Phoniatrie und Pädaudiologie), insbesondere M.Sc. Sebastian Falk, für die gute Kooperation. Ebenso möchte ich mich bei all meinen Kollegen des Instituts für Mechanik und Mechatronik für die fachliche Unterstützung und das angenehme Arbeitsklima bedanken. Abschließend danke ich Dipl.-Ing. Michael Weitz für die Unterstützung beim Projekt und der Abschlussarbeit auch über seine Anstellung am Institut hinaus.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

Voiced speech is the basic and most important human communication tool. Thus, voice disorders drastically restrict the affected individual and, moreover, have an enormous economic impact on our society. Therefore, the project *Numerical computation of the human voice source*, to which the present thesis contributes to, aims to develop a numerical model that supports clinicians on treating voice problems by predicting the voice based on given laryngeal geometry and characteristics.

In order to ensure the computational efficiency necessary for a clinical application, the model is based on the hybrid aeroacoustic approach, which separates the task of sound computation into a flow (fluid dynamics) and an acoustics part. Thereby, the present thesis focuses on the simulation of the acoustic propagation including the required pre-processing (e.g., source term computation and conservative interpolation).

After providing a brief overview of the general development and the state of the art in the field of numerical models of human phonation, preliminary work related to the model of the present thesis is presented. Having depicted the applied methodology and the theoretical background, the numerical model and the corresponding workflow are established in a next step. Thereby, several aspects (e.g., boundary condition and conservative interpolation scheme) of the model related to the numerical efficiency and accuracy are investigated. As the presented model corresponds to an experimental setup with synthetic vocal folds and non-realistic vocal tract, it is subsequently adapted to real human phonation before it is applied to two case studies. The first case study investigates the impact of abnormal vocal fold characteristics (glottal gap and asymmetric vocal fold motion) on the acoustic sources and the resulting voice, whereas the second utilizes clinical data of patients in order to evaluate the agreement with real phonation. In the first study, it was found that voice quality decreases with rising glottal gap due to insufficiently developed acoustic sources and that asymmetric vocal fold motion can compensate for glottal insufficiency. The simulated configurations of the second case study only partly agreed with the clinical data. Deviations in the higher frequency range suggest to more extensively assess the accordance of the vocal tract shape used for the acoustic simulation. Furthermore, the amplitudes at lower frequencies are underestimated, which indicates that the neglected convective part of the acoustic source term might play a role in this frequency range.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Kurzfassung

Gesprochene Sprache ist die wichtigste und grundlegenste Art der menschlichen Kommunikation. Deshalb sind Menschen, die an einer Stimmerkrankung leiden, in ihrem Alltag und ihrer Lebensqualität stark eingeschränkt. Daher wird im Rahmen des Projektes *Numerical computation of the human voice source*, wozu diese Arbeit beiträgt, ein numerisches Modell entwickelt, welches die Behandlung von Stimmerkrankungen unterstützen soll, indem es in der Lage ist, die Stimme aufgrund der Kehlkopfgeometrie und -charakteristik zu berechnen.

Die numerische Effizienz, welche für einen klinischen Einsatz erforderlich ist, wird durch den Einsatz des hybriden aeroakustischen Ansatzes erzielt. Dieser unterteilt die Berechnung der Schallentstehung und -ausbreitung in eine Strömungs- und Akustiksimulation, wobei die für die zweitere erforderlichen akustischen Quellterme aufgrund der Strömungsergebnisse berechnet werden. Der Fokus dieser Arbeit liegt auf der Akustiksimulation inklusive dem notwendigen Pre-Processing, wie die Quelltermberechnung und die konservative Interpolation.

Um einen Überblick zu verschaffen, wird in dieser Arbeit eingangs die allgemeine Entwicklung und der Stand der Technik in der Modellierung der menschlichen Stimme vorgestellt. Danach werden die Arbeiten, auf denen das in dieser Arbeit verwendete Modell basiert, präsentiert. Nach der Erläuterung der angewandten Methodik wird dann näher auf das verwendete numerische Modell eingegangen, welches vorerst einen experimentellen Versuch mit synthetischen Stimmlippen nachbildet. Bevor das Modell in einem nächsten Schritt an die reale menschliche Stimmbildung angepasst wird, werden einige Aspekte im Zusammenhang mit der Effizienz und Simulationsgenauigkeit näher untersucht. Das adaptierte Modell wird dann in zwei Studien angewandt. In der ersten Studie wird der Einfluss anormaler Stimmlippencharakteristik, wie gestörtes Schließverhalten oder asymmetrische Schwingung, auf die entstehenden akustischen Quellen und die resultierende Stimme analysiert. Die Ergebnisse zeigen, dass die Stimmqualität mit steigender initialer Stimmlippenöffnung abnimmt, wobei dabei asymmetrische Stimmlippenbewegung die Stimme positiv beeinflussen kann. In der zweiten Studie zeigen Simulation und Messung nur teilweise gute Übereinstimmung. Große Abweichungen im hohen Frequenzbereich lassen eine in der Simulation unpassende verwendete Geometrie des Vokaltraktes vermuten, während die unterschätzten Amplituden im niedrigen Frequenzbereich auf das Fehlen des vernachlässigten konvektiven Anteils des Quellterms hinweisen könnten.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Used Symbols

E^{a}	Finite element of the acoustic grid
E^{f}	Finite element of the flow grid
$\mathrm{Err}_{\mathrm{rel}}{}^{\mathrm{L}_2}$	Relative error based on L2-norm
F	Acoustic source
$L_{P\mathrm{a}}$	Sound pressure level
M^{f}	Set of flow cells
Ma	Mach number
N	Finite element basis function, number of elements
U_∞	Free stream velocity
V	Volume, vertex velocity of finite element
X	Vertex coordinate of finite element
c	Speed of sound
d	Diameter
$\boldsymbol{d}$	Mechanical displacement
f	Frequency, acoustic source density
$\boldsymbol{f}$	Force density
l	Characteristic element length
$\hat{n}$	Surface normal of finite element
p	Pressure
$\vec{p}$	Arbitrary Euclidean vector
q	Source density
r	Distance
t	Time
$\boldsymbol{v}$	Fluid velocity
$\boldsymbol{x}$	Cartesian coordinate
Γ	Mesh ratio
γ	Ratio of specific heats
ν	Kinematic viscosity
ϕ	Field variable of finite element method
$\boldsymbol{\xi}$	Local coordinate in a finite element
λ	Wave length
ρ	Mass density
ψ	Velocity potential, test function
$\boldsymbol{\sigma}$	Stress tensor

Subscripts

a	Acoustic part
f	Flow part
c	Compressible part
ic	Incompressible part

BI	Boundary intercept
GC	Ghost cell
IP	Image point

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Operators

$(\star)'$	Fluctuation or perturbation operator
$\overline{(\star)}$	Mean operator
∇	Nabla operator with respect to the Cartesian coordinate system $\boldsymbol{x}$
$\frac{\mathrm{D}\star}{\mathrm{D}t}$	Substantial derivative with respect to parameter t
$\frac{\partial\star}{\partial t}$	Partial derivative with respect to parameter t
$\cdot$	Scalar product

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Used Acronyms

2D	two-dimensional
3D	three-dimensional
ABC	absorbing boundary condition
ALE	arbitrary lagrangian–eulerian
AMG	algebraic multigrid
APE	acoustic perturbation equation
ASD	amplitude spectral density
BC	boundary condition
BI	boundary intercect
CAA	computational aeroacoustic
CFD	computational fluid dynamics
CT	computed tomography
DOF	degree of freedom
fVF	false vocal fold
FE	finite element
FEM	finite element method
FFT	fast Fourier transformation
FSAI	fluid-structure-acoustic interacion
FSI	fluid-structure interacion
FV	finite volume
FVM	finite volume method
FWF	Fonds zur Förderung der wissenschaftlichen Forschung (Austrian Research Council)
GAW	glottal area waveform
GC	ghost cell
GCA	ghost cell algorithm
HBC	homogeneous boundary condition
HFE	high frequency energy
IBM	immersed boundary method
IP	image point
LPCE	linearized perturbed compressible equations
LES	large eddy simulation
NSC	no-slip condition
MRI	magnete resonance imaging
PCWE	perturbed convective wave equation
PISO	pressure-implicit with splitting of operators
PDE	partial differnential equation
PML	perfectly matched layer
RHS	right hand side
SPL	sound pressure level
VF	vocal fold
WALE	wall-adapting local eddy-viscosity

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

1 Introduction **3**
- 1.1 Human Phonation . . . 3
- 1.2 Aim of the Thesis . . . 4

2 State of the Art for Simulation of Human Phonation **5**
- 2.1 General Development of Numerical Models of Human Phonation . . . 5
- 2.2 Development of the Applied Model . . . 13

3 Methodology **19**
- 3.1 Hybrid Aeroacoustic Approach . . . 19
- 3.2 Perturbed Convective Wave Equation . . . 20
- 3.3 Conservative Interpolation . . . 21

4 Introduction of the Numerical Model **25**
- 4.1 Workflow . . . 25
- 4.2 CFD Simulation Setup . . . 28
- 4.3 CAA Simulation Setup . . . 30
  - 4.3.1 Spatial Discretization . . . 31
  - 4.3.2 Modifications and Investigations . . . 33

5 Application of the Numerical Model **41**
- 5.1 Adaption to Real Phonation . . . 41
- 5.2 Case Study 1: Primary Voice Source . . . 43
  - 5.2.1 Configurations . . . 44
  - 5.2.2 Analysis of Acoustic Source Terms . . . 45
  - 5.2.3 Acoustic Spectra . . . 50
  - 5.2.4 Discussion of Results . . . 51
- 5.3 Case Study 2: Validation of the Physiological Correctness . . . 52
  - 5.3.1 Configurations . . . 52
  - 5.3.2 Analysis of Acoustic Source Terms . . . 53
  - 5.3.3 Acoustic Spectra . . . 54
  - 5.3.4 Discussion of Results . . . 55

6 Conclusion **57**

Bibliography . . . 59

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Chapter 1

Introduction

Voice disorders are an underestimated issue in our society. The capability of speaking is often taken for granted, but once we catch a cold and get hoarse, we realize how dependent we are on a healthy voice since voice production is the basic and most important communication tool for us humans. However, voice problems do not only restrict the quality of life, lead to isolation and depression [1–3], but also endanger the livelihood of people suffering from chronical voice disorders. According to Ruben, people with severe speech disabilities are more often found to be unemployed or in a lower economic class than people with a healthy voice [4]. Besides the consequences for the individual, voice disorders also have a huge economic impact. In [4], the costs arising from communication disorders are estimated to be up to three percent of the Gross National Product of the United States, whose economy has become dependent on communication-based employment. Furthermore, it is reported that this trend will increase and thus communication disorders will be a major public health challenge for the 21st century.

1.1 Human Phonation

Although we use our voice continuously and take it for granted, human voice production (phonation) is a complex process consisting of the interaction between the tracheal airflow and the vocal folds (see Fig. 1.1) and is only working properly under certain conditions. Due to self-excitation caused by the expiration air stream, the two elastic vocal folds are excited to oscillations, which periodically interrupt the airflow forming the primary acoustic voice signal. Depending on the air stream and the dynamic properties of the vocal folds, the frequency of the primary acoustic signal, which is the fundamental frequency f_0 of the acoustic signal, normally lies between 100 Hz and 300 Hz.[1] By pharynx constrictions (movement of mouth and jaws), tongue motion, and oral-nasal stream division, the primary voice signal is modulated and then emitted from the mouth, resulting in the audible voice. It has to be noted that the false vocal folds (fVF) depicted in Fig 1.1 do not oscillate and can thus be assumed as rigid.

A prevailing assumption is that for a good primary acoustic voice signal, the vocal fold vibrations need to fulfill three conditions:

1. Symmetrical oscillation of the vocal folds.
2. Periodical oscillation of the vocal folds.
3. Total closure of the glottal gap during every oscillation cycle. This requires full contact of the vocal folds (see Fig. 1.1). Otherwise, this kind of phoniatric malfunction is called the glottal gap or glottis closure insufficiency.

[1] The fundamental frequency is depending on gender. In general, the fundamental frequency of female is higher.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

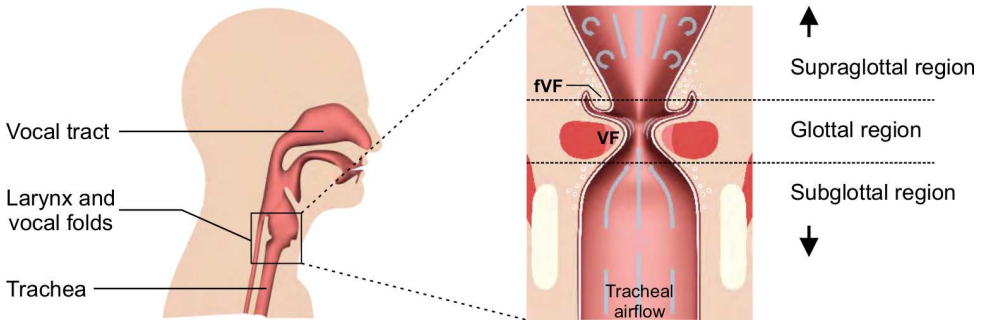


Figure 1.1: Parts involved in phonation process (left). Detail of the larynx with vocal folds (VF), false vocal folds (fVF) and indicated airflow (right). [5]

1.2 Aim of the Thesis

In former years, voice research mainly has been done experimentally by applying physical replicas of the vocal folds that can either be synthetic or real animal and human larynx (*ex-vivo* or *in-vivo*). In general, experiments have the drawback of a lack of spatial resolution compared to numerical simulations. The spatial resolution of measurements of physical quantities, such as pressure or particle velocity, is not only restricted by the space and the cost of sensors but can also affect the measured physical field. In contrast, numerical tools can provide physical quantities in every location of the computational domain. With the development of computational hardware power and efficient numerical methods, aeroacoustic simulation lately has become a powerful and valuable tool in voice research. In other fields, for example for cardiovascular systems [6–8], numerical tools have been recently used to successfully support post-surgical planning, whereas numerical simulation of phonation has not yet been applied clinically. Therefore, Universitätsklinikum Erlangen (Department for Phoniatrics and Pedaudiology) and TU Wien initiated the joint project *Numerical computation of the human voice source*, which is funded by the German Research Foundation (DFG) and the Austrian Research Council (FWF). The objective of the project is the development of a numerical tool called *simVoice*, which is capable of predicting the voice, depending on various parameters e.g. the incoming airflow, the prescribed motion of the vocal folds and the geometry of the rigid larynx and vocal tract. Considering a prospective clinical application, the project is focused on numerical efficiency in order to limit the required temporal and computational resources. Therefore, a hybrid approach is used to separate fluid dynamics and acoustic simulation in order to reduce the spacial resolution for the acoustic propagation and thus save computational time. Within this thesis, parts of working package 2 (WP2) and WP3 as a whole are accomplished. The objective of WP2 is to increase the numerical efficiency of the present aeroacoustic validation setup, whereas in WP3 the physiologic correctness is assessed by comparing simulation results of real geometries with clinical data. Additionally, a variety of configurations, which don't fulfill all of the three conditions for a proper primary voice signal (see Sec. 1.1), is simulated and the impact is investigated. It should thereby be noted that this work only contributes to the computational aeroacoustic part (CAA) of the hybrid workflow, which is based on the perturbed convective wave equation (PCWE), as the computational fluid dynamics (CFD) data is provided by Universitätsklinikum Erlangen.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Chapter 2

State of the Art for Simulation of Human Phonation

2.1 General Development of Numerical Models of Human Phonation

This chapter provides an overview of work done so far in the field of simulation of human phonation as well as the state of the art. According to Fig. 2.1, models of human phonation can be distinguished between lumped mass models and partial differential equation (PDE) based models, which can be either two or three-dimensional. These models are mostly based on the finite volume method (FVM) and the finite element method (FEM) and can further be separated depending on the motion of the vocal folds, which can be static, externally driven (prescribed motion) or self-excited due to fluid-structure-acoustic interaction (FSAI). The latter approach is the most realistic one but also the computationally most expensive and complex one as the mechanic field of the tissue of the vocal folds has to be solved in addition to the flow field, which requires knowledge of the material properties. This can be avoided by prescribing the movement of the vocal folds which has to be provided by experiments, though. Thereby, the mesh adaption due to the moving boundary is still a challenge and has some numerical issues. Therefore, and because of the simplicity of the model, also the least realistic model with static vocal folds (mostly in combination with a prescribed waveform of the inflow velocity profile) is often applied. Regarding computational effort and complexity of the mathematical system of phonation models in general, the number of degrees of freedom (DOFs) can reach from one (one-mass model) to several millions (FE or FV models).

Numerical simulations have first been applied in the field of human phonation at the end of the 1960s [9]. Thereby, the vocal folds were modeled as lumped masses (with one-dimensional dampers and springs, see Fig. 2.2a) driven by a function of the subglottal pressure and the Bernoulli pressure, assuming a quasi-steady and incompressible flow. Furthermore, the models of the vocal folds were integrated in acoustic circuits that also considered a vocal tract by assuming plane wave propagation, in order to compute the resulting voice [9, 10].

Whereas in the past lumped mass models were used for principal investigations of the excitation mechanism of the vocal folds, they are still applied nowadays but in a three-dimensional arrangement (multi-mass models) to investigate material properties of the vocal folds. Yang *et al.* [11] fitted three-dimensional vocal fold dynamics, extracted from *in vitro* experiments by a 3D multi-mass-model using an optimization procedure (see Fig. 2.2b). Therewith, they optimized the parameters of the 25 mass

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

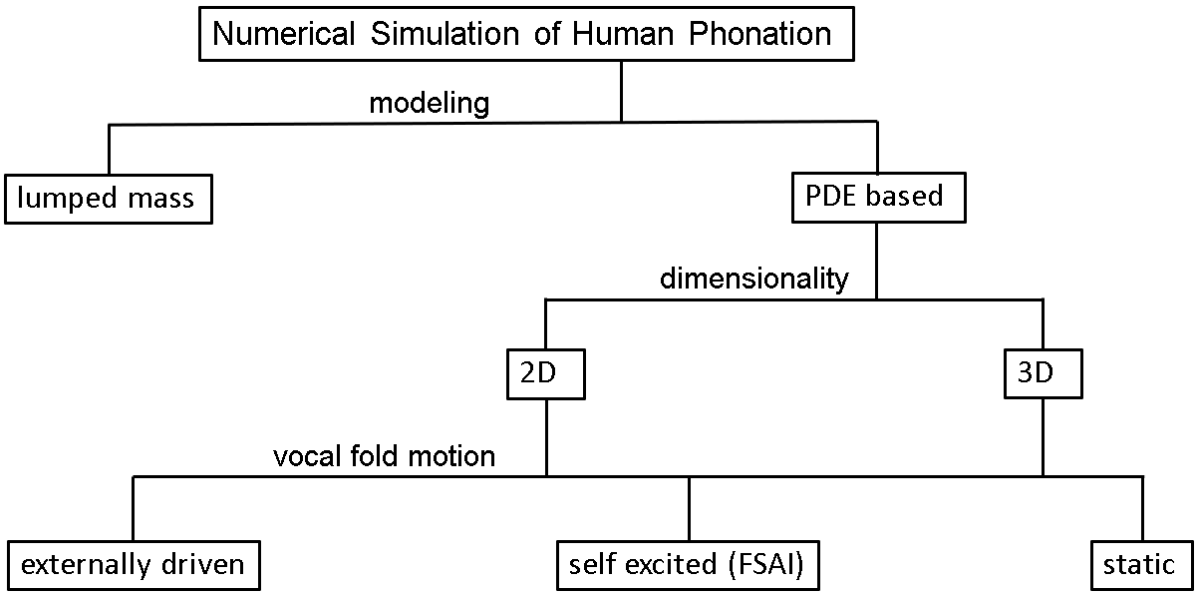
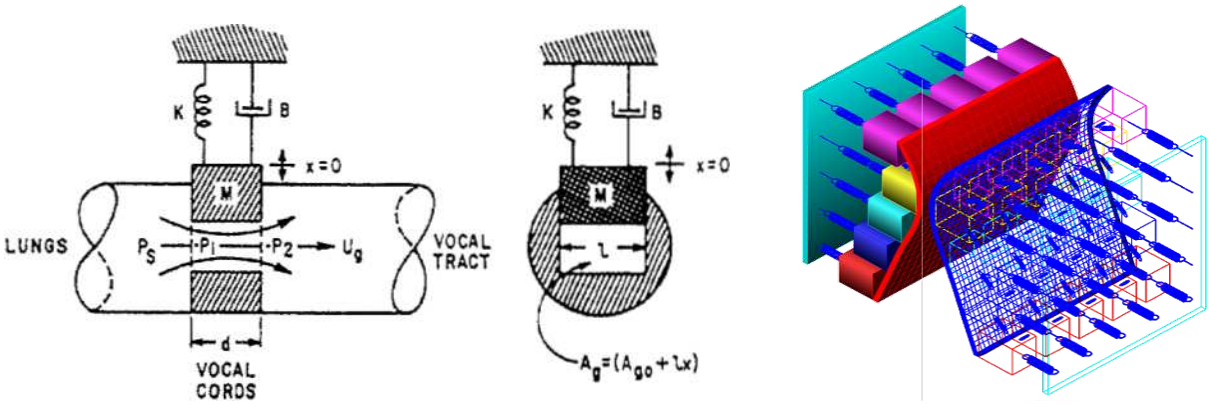


Figure 2.1: Overview of simulation types applied in research of human phonation.



(a) Second-order single mass model [9].

(b) 3D multi-mass model [13].

Figure 2.2: Lumped mass models of the vocal folds with different levels of complexity.

elements of each of the two vocal folds, such as effective mass, lateral, vertical and longitudinal stiffness in order to match the vocal fold motion obtained from hemilarynx studies[1]. Thus, the spatial distributions of local effective mass and stiffness of the vocal folds were yielded. As material properties of the tissue of the vocal folds can't be measured directly during *in-vivo* human phonation, this physiologically relevant information can be captured by using the described 3D-modeling technique [11, 12].

Although lumped mass models are able to simulate the pulsating airflow through the larynx, they don't represent all physical aspects due to the strong simplifications. However, the development of computational power and numerical methods made PDE-based tools like CFD (computational fluid dynamics) and CAA (computational aerodynamics) available, which aim to resolve all physical details and are therefore very promising for principal research in human phonation, as well as for clinical applications.

In order to capture the entire physics of phonation, the flow field (airflow), the acoustic field (acoustic wave propagation) and the field of structural mechanics (tissue) including their interactions (couplings)

[1] Hemilarynx studies are capable of tracking the three-dimensional motion of the vocal folds on an excised human larynx.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

have to be considered. As displayed in Fig. 2.3, fluid forces thereby act on the bordering vocal folds and deform them. Due to the deformation, the fluid domain changes and the mesh has to be updated during the simulation which requires special treatment in order to fulfill the boundary conditions. The fluid-acoustics interaction is described by aeroacoustic analogies or by applying acoustic perturbation equations (APE) whereas the solid-acoustic coupling is realized by equating the velocity of the solid and the acoustic particle velocity on the bounding surface. However, due to the complexity and computational effort, current PDE-based approaches have to simplify the described causalities and e.g. neglect couplings that have minor effects on voice production [14].

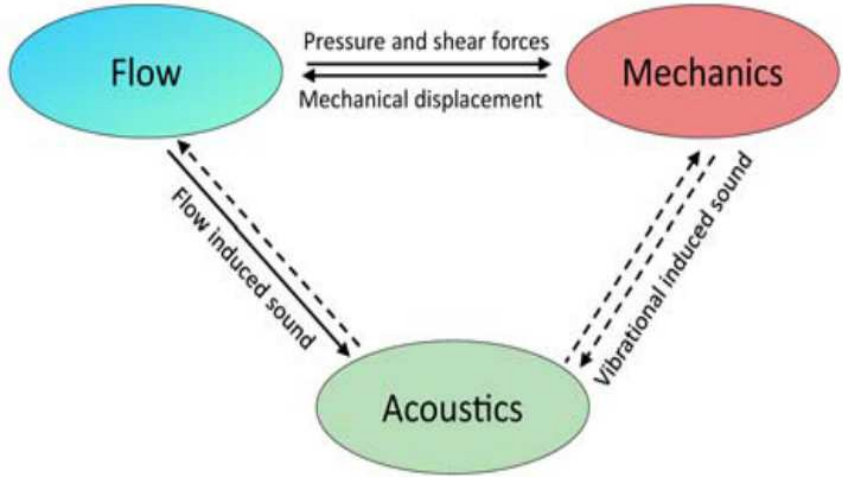


Figure 2.3: Fluid-Solid-Acoustic interactions in human phonation. The dashed lines define couplings which can be neglected in human phonation as they only play a minor role [15].

In [16], a first numerical model considering fluid-structure interaction was introduced by Alipour *et al.* starting from a 3D continuum model (FEM) of the vocal folds to analyze their dynamic properties by investigating the eigenmodes. Thereby, they assumed linear elasticity of the transversely isotropic tissue consisting of three layers (see Fig. 2.4) and that the vibration only takes place in a single plane. Furthermore, the effect of grid motion during FE space integration was neglected assuming fixed control volume for integration. At every time step of the transient simulation, the global matrices first were calculated, and then the pressure distribution from the airflow was obtained and used to compute the forcing vector. When the vocal folds touch each other, the contact nodes lose one degree of freedom. Finally, the mesh can be updated by the solved displacement field for the next time step. Despite the coarse grid, they were able to capture the self-excited oscillation of the vocal folds with this methodology, but as they assumed steady and laminar incompressible flow, they were not able to capture aerodynamic effects. In [17], Rosa *et al.* simulated the fluid-structure interaction with a similar approach but with a tubular geometry and advanced contact procedure and also considered the ventricular folds. They found that the ventricular folds have considerable influence on the airflow and therefore affect the glottal wave form.

The described approach for treating fluid-structure interaction suffers from poor element quality due to the deformation of the mesh. Especially the fluid elements inside the glottal gap are distorted drastically as the vocal folds theoretically close every oscillation cycle. Therefore, alternative techniques, such as the *immersed boundary method* were developed, which don't have this drawback. Due to the mathematical complexity and computational cost, only 2D models could be used in the beginning [18]. With the 2D model and the *immersed boundary method*, Zheng *et al.* were able to show the influence of the false vocal

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

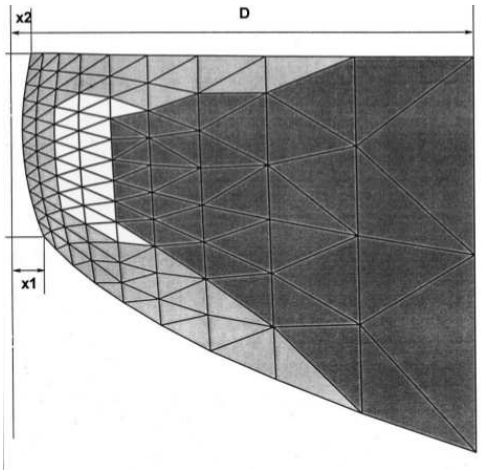


Figure 2.4: Discretization of the vocal fold consisting of body (dark grey) ligament (white) and cover (light grey). The glottal half widths x1 and x2 define the glottal adduction [16].

folds on phonation. Their simulations indicated that the presence of the false vocal folds reduces the translaryngeal flow impedance and increases the mean flow rate through the glottis, which results in a larger amplitude of vibration in the true vocal folds. Therefore, the false vocal folds tend to enhance the sound intensity for a given effort [19].

However, Mattheus and Brücker showed that the behavior of the laryngeal airflow shows significant differences depending on the dimensionality of the model [20]. In their work, they compared the airflow of a 2D geometry and a 3D geometry with a lens-like and a slit-like orifice of the glottis by performing an incompressible flow simulation with a prescribed motion of the vocal folds (see Fig. 2.5). The results show that there is a significant deflection of the glottal jet for the 2D case, whereas for the slit-like there is only little and for the lens-like orifice there is no significant jet deflection. For the 3D lens-like shape, the vorticity is redistributed from the spanwise component into the cross-wise and streamwise direction whereas for the slit-like orifice this effect is weaker. For the 2D case, there is no redistribution of vorticity at all, as vortices can only occur in the spanwise direction. Consequently, the spanwise vorticity of the 2D model is stronger which causes drastic jet deflection resulting in clearly different flow conditions. These results emphasize the importance of dimensionality for flow computation.

Figure 2.6 shows a typical glottal jet attached to one side of the supraglottal duct, obtained from a 2D flow simulation. The phenomenon of the deflection of the glottal jet is called *Coanda effect* and is likely to induce asymmetric airflow as well as different aerodynamic forces acting on the two vocal folds [21].

Due to the changing fluid and structural domain, a major challenge of simulations of human phonation is fluid-structure interaction of the vocal folds. In [22], Mittal *et al.* presented the *sharp interface immersed boundary method* (sharp interface IBM) for incompressible flows with complex boundaries, which makes it possible to simulate the 3D fluid-structure interaction and thus drastically increases the fidelity of the numerical model. The method is capable of satisfying the boundary condition (BC) on the immersed boundary without the need for a conform interface between fluid and structural domain allowing arbitrary movement and deformation of the structural domain. The IBM can be divided up into two categories [23]. The first category are methods that employ "continuous forcing" wherein a forcing term is added to the continuous Navier-Stokes equations before they are discretized, whereas in the second category the forcing is applied to the discretized Navier-Stokes equations. The disadvantage of the first category of methods is that a "diffuse" boundary is established which means that the boundary condition

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

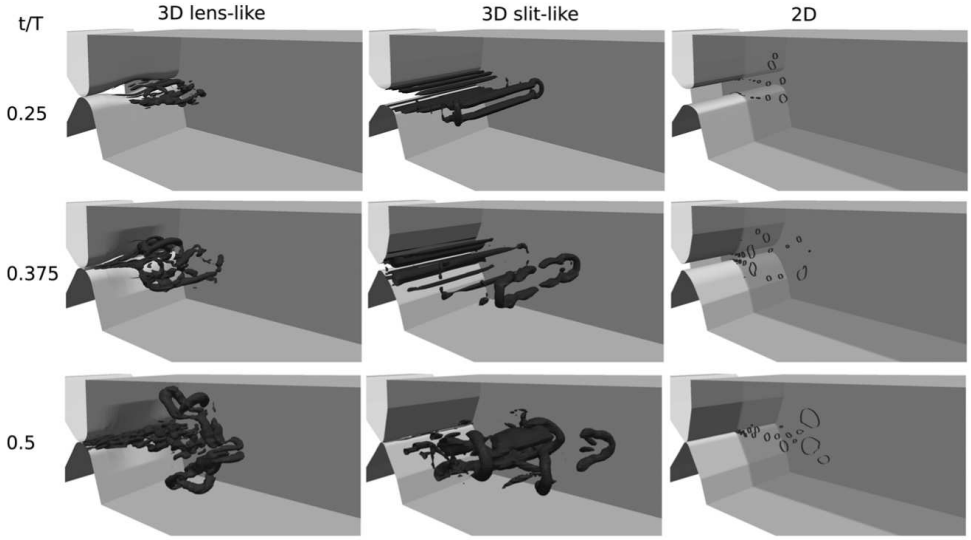


Figure 2.5: Vortical structures visualized at certain times t/T during cycle for different orifices (3D lens-like, 3D slit-like, 2D) [20].

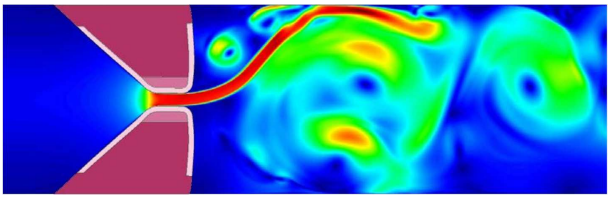


Figure 2.6: Deflected glottal jet attaching to the subglottal surface (*Coanda effect*) [14].

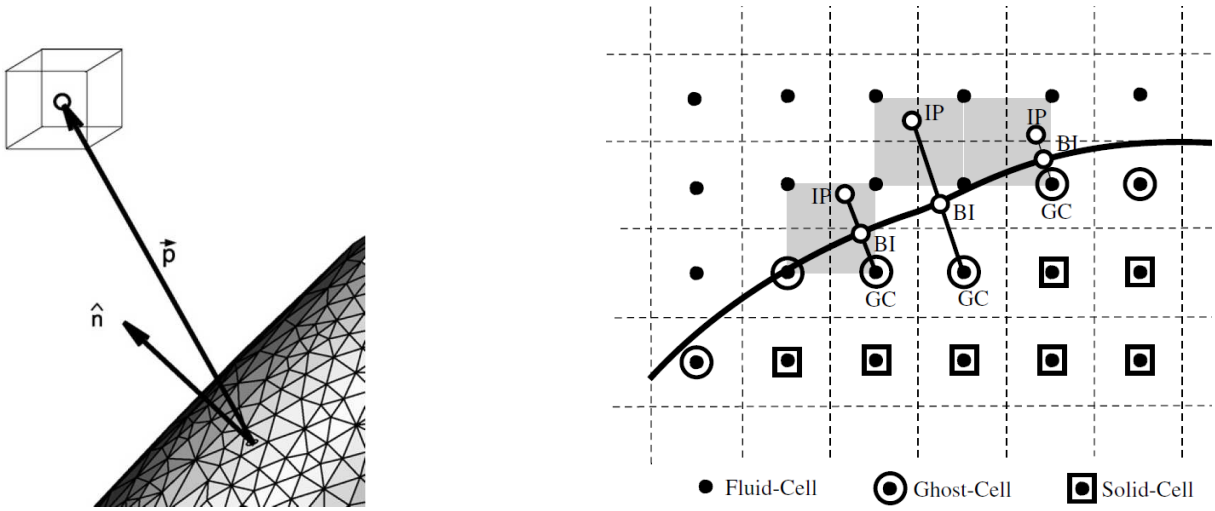
is not satisfied exactly on the immersed boundary but within a localized region around it. The second category doesn't have this drawback and is therefore called the *sharp immmersed boundary method*. The key advantage of the non-sharp methods is that they are hardly dependent on the spatial discretization and can be implemented into an existing Navier-Stokes solver relatively easily [22].

In the described method, the surface of the solid, which is called the immersed boundary, is represented by triangular elements, while the flow is computed on a non-uniform cartesian (hexagonal) grid with cell-centered representation of the solution variables. Thereby, the immersed boundary can be located arbitrarily within the cartesian volume grid as indicated in Fig. 2.7b. In order to be able to satisfy the BCs precisely, the solution quantity has to be (at least implicitly) known on the immersed boundary. Therefore, the authors of [22] developed the so called *ghost cell algorithm* (GCA), which in a first step, identifies cells whose nodes are inside the solid boundary (termed "solid cells") and cells that are outside the body (termed "fluid cells"). This is done by taking a dot-product of the vector $\vec{p}$ extending from the surface element closest to a given element to the node of the element with the surface normal of the surface element $\hat{n}$ (see Fig. 2.7). A positive value of the dot-product $(\vec{p} \cdot \hat{n})$ indicates that the node is outside the body (fluid cell) and a negative value means that the node is inside (solid cell). Solid cells with a neighboring fluid cell are declared as ghost cells (see Fig. 2.7b). This determination needs to be done for every time step for moving boundaries and is therefore computationally relatively demanding. However, assuming that the immersed boundary only travels a distance of the order of the cartesian grid size per time step, the number of fluid cell nodes that need to be identified can be reduced drastically by using the declaration of the previous time step. In the next step a line segment from the node of the ghost cell (GC) towards a node in the fluid cell (image point, IP) normally intersecting the immersed boundary such that

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

the boundary intercept (BI) is midway between point GC and IP. Once the BI and the corresponding IP are identified, the value of the solution variable ϕ is interpolated three-dimensionally with respect to the eight surrounding nodes to obtain ϕ_{IP}. For the ghost cell, the variable ϕ_{GC} is known and ϕ_{BI} can be obtained using the linear approximation $\phi_{BI} = \frac{1}{2}(\phi_{IP} + \phi_{GC})$. Having found an approximation of ϕ_{BI}, it can be plugged in for the definition of the boundary condition that is considered for solving the PDEs of the coupled fields. Boundary motion is accomplished by moving the nodes of the surface triangles with the vertex velocity V_i, which is either prescribed or can be computed from a dynamical equation if the body motion is coupled to the flow field. Equation (2.1) describes the mesh update from time step n to $n+1$ [22].

$$\frac{X_i^{n+1} - X_i^n}{\Delta t} = V_i^{n+\frac{1}{2}} \tag{2.1}$$



(a) Procedure for cell identification.

(b) Definition of fluid, solid and ghost cells. The black line indicates the immersed boundary.

Figure 2.7: Schematics indicating the procedure of the ghost cell algorithm [22].

In [24,25] Zheng *et al.* applied IBM to an idealized shaped vocal fold model with reduced vocal tract to take the interaction of the glottal airflow and the vocal folds into account without investigating acoustics. The computed features of the vocal fold oscillation, such as glottal mean and peak flow rates, and the open and skewness quotients, are within the typical range of experiments. A major issue of fluid-structure interaction is to find an appropriate material model for the vocal fold tissue. In the respective study, transverse isotropy of the three layers of the vocal folds was assumed. Thereby, the longitudinal stiffness incorporates the effect of muscle activation but for better fidelity, this effect would have to be modeled. According to [24], a result of this simplification is a shifted oscillation base frequency. Later, the model of Zheng *et al.* was modified by Jiang *et al.* and a simplified vocal tract was added in order to also account for acoustic propagation [26]. Thereby, the glottal airflow was governed by the incompressible Navier–Stokes equations

$$\nabla \cdot \boldsymbol{v}^{\mathrm{ic}} = 0 \tag{2.2}$$

$$\frac{\partial \boldsymbol{v}^{\mathrm{ic}}}{\partial t} + (\boldsymbol{v}^{\mathrm{ic}} \cdot \nabla)\boldsymbol{v}^{\mathrm{ic}} = -\frac{1}{\rho_0}\nabla p^{\mathrm{ic}} + \nu_0 \nabla^2 \boldsymbol{v}^{\mathrm{ic}} \tag{2.3}$$

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

with $\boldsymbol{v}^{\mathrm{ic}}$, ρ_0, p^{ic}, ν_0 being the incompressible flow velocity, density, pressure, and kinematic viscosity. The governing equation for structural dynamics of the vocal folds was the Navier equation

$$\rho_{\mathrm{tiss}} \frac{\partial^2 \boldsymbol{d}}{\partial t^2} = \nabla \cdot \boldsymbol{\sigma} + \rho_{\mathrm{tiss}} \boldsymbol{f} \tag{2.4}$$

with a linear stress–strain relationship where ρ_{tiss} is the tissue density, $\boldsymbol{d}$ is the displacement, σ is the stress tensor, and $\boldsymbol{f}$ is the body force.

Regarding acoustics, a hybrid approach is applied that divides the task into an incompressible flow simulation and an acoustic simulation governed by the linearized perturbed compressible equations (LPCE)

$$\frac{\partial \rho\prime}{\partial t} + (\boldsymbol{v}^{\mathrm{ic}} \cdot \nabla)\rho\prime + \rho_0(\nabla \cdot \boldsymbol{v}\prime) = 0 \tag{2.5}$$

$$\frac{\partial \boldsymbol{v}\prime}{\partial t} + \nabla(\boldsymbol{v}\prime \cdot \boldsymbol{v}^{\mathrm{ic}}) + \frac{1}{\rho_0}\nabla p\prime = 0 \tag{2.6}$$

$$\frac{\partial p\prime}{\partial t} + (\boldsymbol{v}^{\mathrm{ic}} \cdot \nabla)p\prime + \gamma p^{\mathrm{ic}}(\nabla \boldsymbol{v}\prime) + (\boldsymbol{v}\prime \cdot \nabla)p^{\mathrm{ic}} = -\frac{Dp^{\mathrm{ic}}}{Dt} \tag{2.7}$$

wherein the right hand side of (2.7) is the aeroacoustic source term provided by the flow solver. Moreover, $\rho\prime$, $\boldsymbol{v}\prime$, $p\prime$ are the acoustic perturbed flow density, velocity, and pressure, respectively, and γ is the ratio of the specific heats.

The procedure for explicit coupling of the solvers for the transient simulation is the following:

1. Flow solver: get the incompressible flow pressure (p^{ic}) and velocity ($\boldsymbol{v}^{\mathrm{ic}}$)
2. Acoustics solver: get the compressible flow perturbed pressure ($p\prime$) and velocity ($\boldsymbol{v}\prime$)
3. Update traction on vocal folds using p^{ic}, $\boldsymbol{v}^{\mathrm{ic}}$, $p\prime$, and $\boldsymbol{v}\prime$
4. Solid solver: get the tissue displacement ($\boldsymbol{d}$) and velocity of the bounding surface ($\boldsymbol{v}_{\mathrm{tiss}}$)
5. Deform computational domain and update boundary condition
6. Go to Step 1

The model and details of the vocal folds are depicted in Fig. 2.8. Therein, the geometry of the larynx is obtained from a thin-slice computed tomography (CT) scan and the simplified geometry of the vocal tract is based on [27]. The model also accounts for the collision of the vocal folds by using a hard wall contact at two planes that are located very close to the midplane[2]. This approach is simple but has several drawbacks. On the one hand, a small leakage flow is caused due to the finite distance of the contact planes, and on the other hand, the location of the contact is prescribed to be at the midplane which is not the case for asymmetric oscillation of the vocal folds. This aspect is critical, as asymmetric vibration (e.g. due to a polyp on one vocal fold) is a common pathology of phonation but can't be captured with such a contact model. Thus, a more general model would be necessary, which allows the collision of the vocal folds at an arbitrary position and the exchange of energy between them as this is an important mechanism that helps to synchronize the oscillation of the two vocal folds.

However, the described numerical model of human phonation is one of the most exact in literature which considers most of the important physical aspects of human phonation. But due to the complexity of the model, the computational expenses are enormous. The total wall time is reported to be 900

[2]Due to numerical reasons the contact planes can't be congruent on the midplane.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

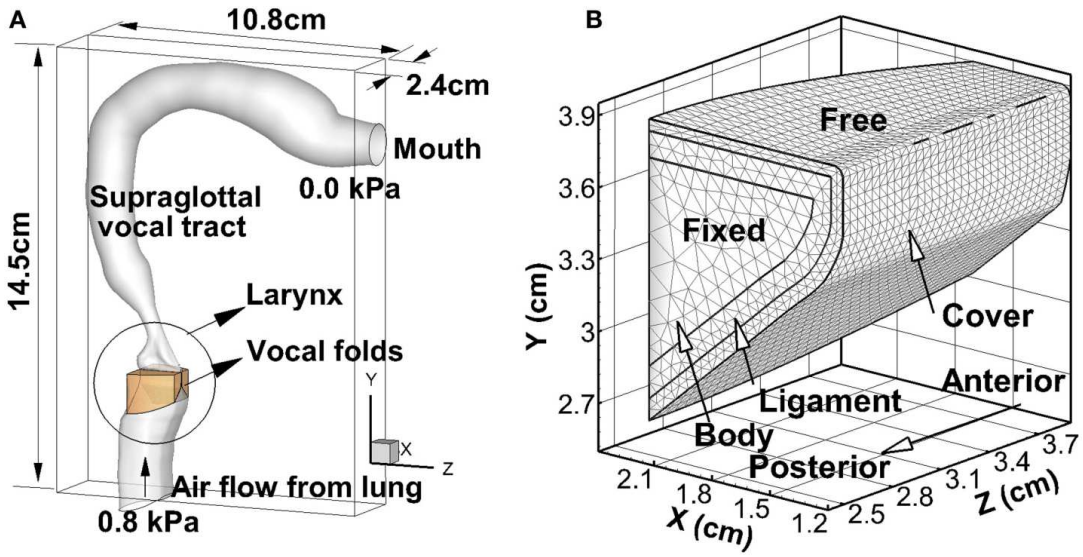


Figure 2.8: Model of the larynx and vocal tract (A) and detail of the vocal folds (B) [26].

hours (37.5 days) using 256 cores to simulate 15 oscillation cycles of the vocal folds. The huge computational effort demands both massive computational resources as well as computational time, which is why the model won't be applicable as a clinical tool without a ground-breaking development in computing hardware. Therefore, in this thesis, another approach is followed by prescribing the movement of the vocal folds instead of modeling fluid-structure interaction. Consequently, the challenge of material definition, including the longitudinal muscle activation of the vocal folds is omitted as well as the costly fluid-structure interaction. The development of the model used in this thesis is presented in the following chapter.

Besides interaction of the different physical fields, contact of the vocal folds and computational efficiency, another challenge is to obtain realistic geometries of the human airways that are necessary for a clinical application. This challenge is about to be accomplished, as recent advances of magnetic resonance imaging (MRI) make it possible to provide exact geometries of the human airways [28–30]. In [30], MRI is applied to capture the complex geometry of the vocal tract for different vowels (see Fig. 2.9). Using the obtained geometries, a compressible flow simulation was performed assuming non-moving boundaries (static vocal folds). Thereby, a typical waveform for the glottal volume flow from literature is used to model the pulsating glottal flow. Furthermore, the fluctuating pressure is assumed to be the acoustic pressure obtained from the compressible CFD simulation, which is applicable for a quiescent fluid (e.g. outside the mouth) but not inside the airways where the pressure fluctuation also contains a fluid dynamic part. The resulting sound pressure level (SPL) for each of the two formant frequencies of each vowel is illustrated in Fig. 2.9.

Summed up, a major challenge is to model the oscillation of the vocal folds[3], which can be accomplished in three ways. The most general approach is to consider fluid-structure interaction but thereby, the issue of material model and its parameters (spatially varying mass/stiffness and muscle activation) emerges. A less complex approach is to prescribe the motion of the vocal folds, which requires the exact knowledge of the movement. Lastly, a less accurate approach is static vocal folds with a prescribed waveform of the flow velocity at the glottis inlet, which is the simplest, as no grid adaption is necessary. Another challenge is the treatment of the collision of the vocal folds. This is a crucial aspect, as it is a

[3]The remaining boundaries can be assumed as static.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

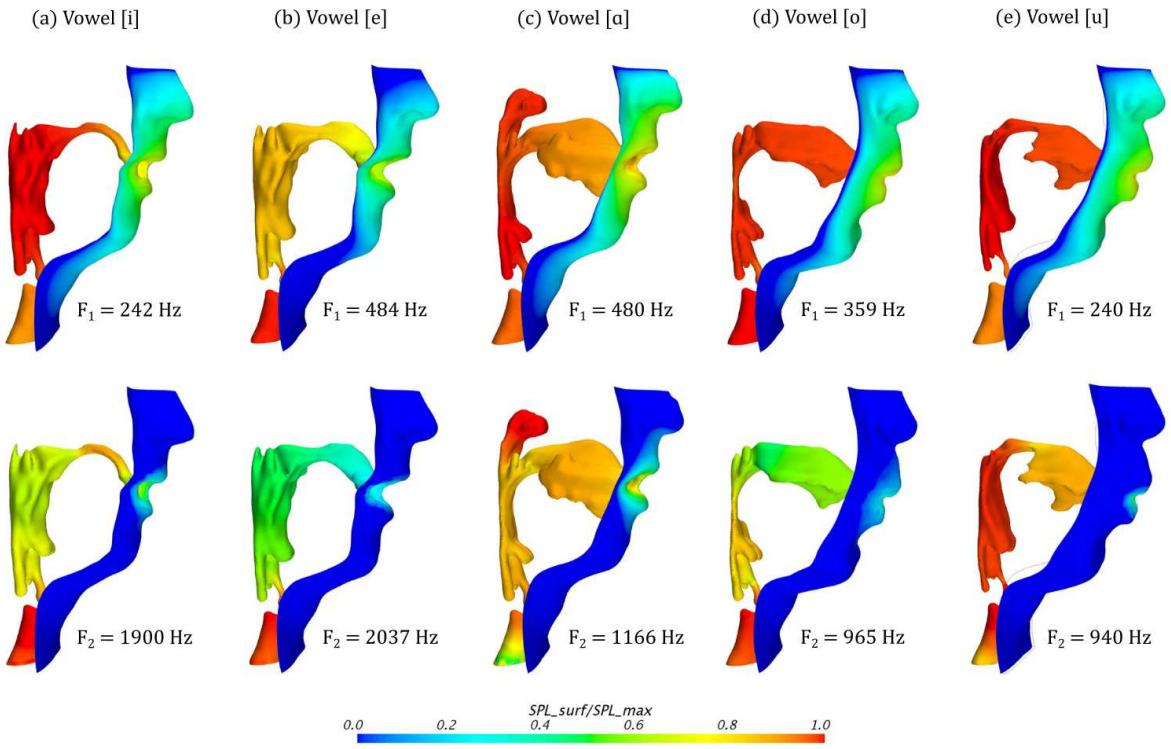


Figure 2.9: Normalized surface SPLs of the near-field acoustic pressure for the formant frequencies F_1 and F_2 of the vowels [i], [e], [a], [o], and [u] in (a)–(e). The SPL at the surface of the airway walls has been normalized by its maximum for each vowel. [30]

mechanism to exchange energy and synchronize the oscillation of the vocal folds. Finally, the challenge of realistic geometries of the airways of patients is about to be solved, but with complex geometries and complexity of the models (e.g. FSI, contact etc.) in general, the computational expenses increase dramatically. Regarding a prospective clinical application, the arising challenge of computational effort can either be solved by advances in computing hardware or by using different simplified models that are tailored to certain pathologies. However, the latter solution requires greater user knowledge and experience for the choice of the appropriate model.

2.2 Development of the Applied Model

In this section preliminary work, related to the head of our research group [31] Univ.Prof. Dipl.-Ing. Dr.techn. Manfred Kaltenbacher, who has been extensively researching in the field of modeling and numerical simulation of human phonation, is reported [32]. Thereby, the development of the model applied in this thesis is explained.

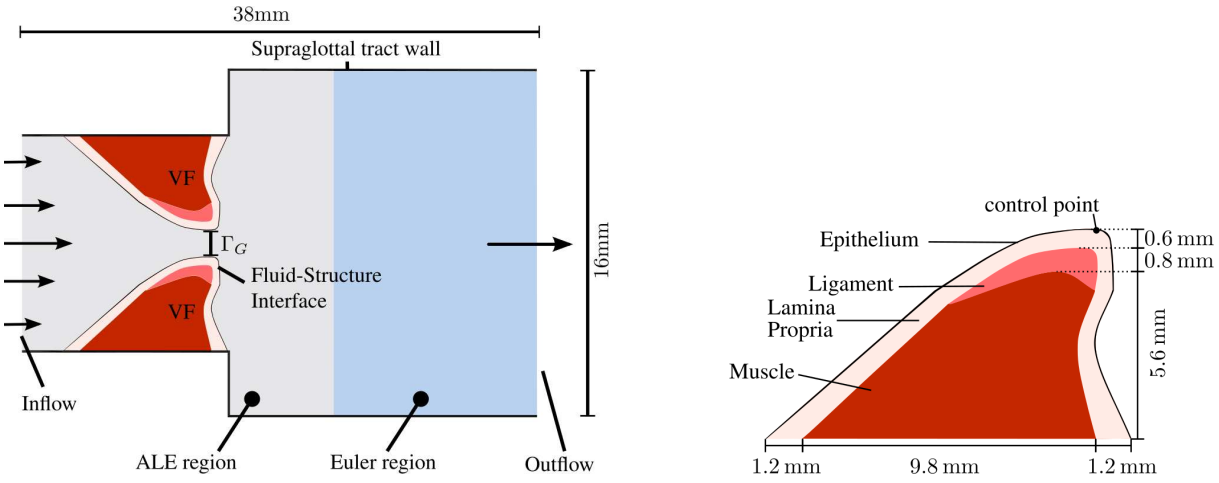
In 2009, Link *et al.* published a first 2D model of human phonation considering the interaction of all three involved physical fields (FSAI) [33] as previous investigations were either limited to fluid-structure or fluid-acoustic coupling. Therein, the interaction of the Eulerian fluid field with the Lagrangian mechanical field was based on the Arbitrary-Lagrangian–Eulerian (ALE) formulation. Moreover, Lighthill's analogy was applied for fluid-acoustic interaction. However, this analogy is not applicable for the computation of the near field sound which is why the correct physical computation of the sound within the entire region requires a perturbation ansatz as demonstrated in [34].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Therefore, Kaltenbacher *et al.* established a 2D FSAI-model which takes fluid-acoustic coupling into account by deploying acoustic perturbation equations within a hybrid aeroacoustic approach that only considers a forward coupling from the flow to the acoustic field (the considered couplings are indicated in Fig. 2.3 by solid lines). The fluid-solid interaction is realized via a strong coupling scheme that iterates within each time step between the two physical fields until an equilibrium is reached. Additionally, also weak coupling without iteration was tested but showed poor performance regarding numerical stability which makes it not applicable for long term computations as necessary in human phonation to achieve a good resolution of the whole frequency spectrum [15].

Due to the substantial computational demand of fluid-structure interaction, Zörner *et al.* replaced the fully-coupled interface condition by a one-way coupling [35]. Therewith, the costly coupled simulation was reduced to a pure flow simulation with prescribed movement of the vocal folds and appropriate boundary conditions.

In the study, first a fully-coupled simulation with the model of [15] was performed in order to gain the vocal fold motion which is then used as input data for the simplified (one-way coupled) simulation. Figure 2.10 illustrates the model of the fully-coupled simulation with the three-layered structure of the vocal folds. Apart from the prescribed motion of the vocal folds (instead of FSI), all parameters (including the geometry) are the same for the pure one-way coupled flow simulation. Thereby, the influence of the interface condition, the geometry and the inflow-BC were investigated within four case studies. The effect of the different aspects was then assessed by comparing features of the resulting flow (e.g. particle velocity and volume flux) with the flow features obtained from the fully-coupled simulation.



(a) Computational domain.

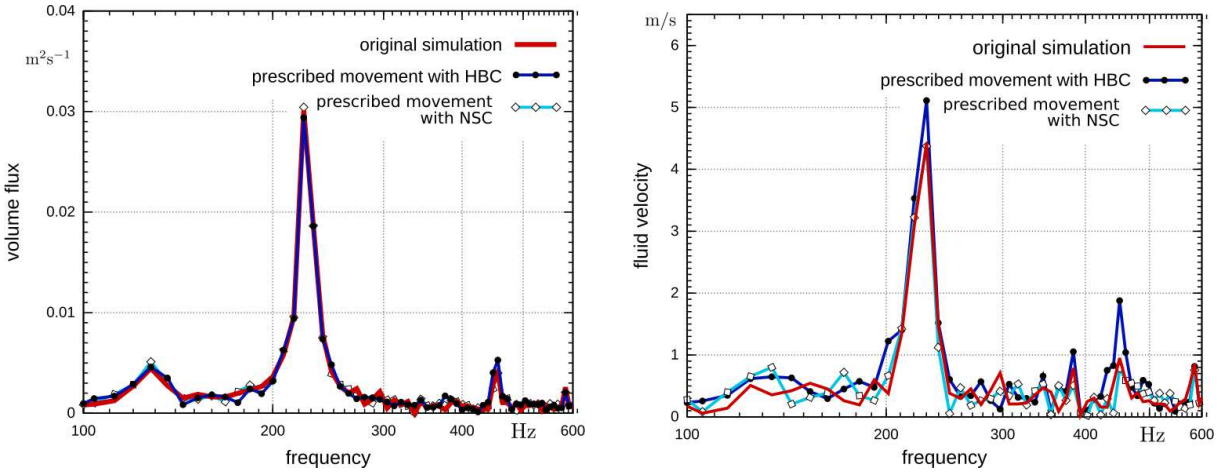
(b) Detail of structural model of the VF.

Figure 2.10: Model of fully-coupled simulation [35].

The influence of the interface condition (velocity) was assessed in the first case study. Thereby, all parameters of the model were identical to the original, fully-coupled model (i.e., geometry, boundary condition apart from the fluid–structure interface). On the one hand a simulation with a homogeneous boundary condition (HBC) with the velocity set to $\boldsymbol{v} = \mathbf{0}$ at the interface boundary, and on the other hand with a no-slip condition (NSC) with the fluid velocity matched to the vocal fold velocity at the interface, was performed. The results of the case study (see Fig. 2.11) show that the volume flux of HBC and NSC match the original model quite well, while for the velocity HBC differs from the other two models in amplitude of the characteristic frequencies. In particular, the amplitude of the second harmonic is twice as high. These differences will likely have an impact on sound generation. On the other hand, the simulation with NSC is in strong agreement and will result in a similar acoustic field as the

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

fully-coupled simulation.



(a) Volume flux through glottis.

(b) Particle velocity at point inside the glottis.

Figure 2.11: Spectral results of case study of the interface condition [35].

In the second case study, the effect of the size of the inlet pressure is investigated, as measuring the exact pressure in humans is difficult due to restricted access to the larynx, which is why measurement errors arise. Therefore, the inlet pressure of 1 kPa was reduced to 0.8 kPa and a pure flow simulation with prescribed movement and NSC was conducted. Additionally, also a fully-coupled simulation with the reduced inlet pressure of 0.8 kPa was performed. The results in Fig. 2.12 show that the inlet pressure has a significant impact, especially on the amplitude of the first harmonic, which illustrates that special effort should be put into determining the exact pressure.

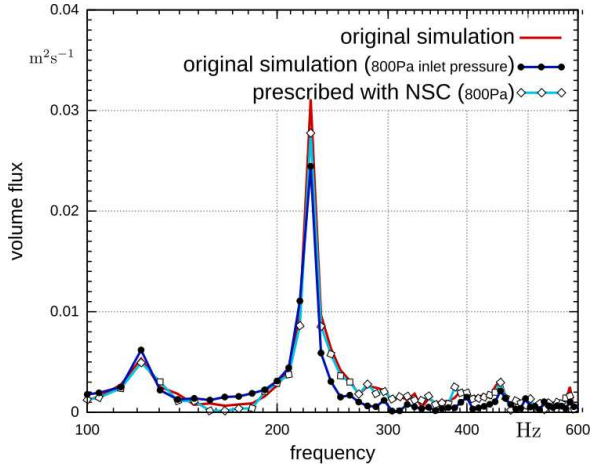
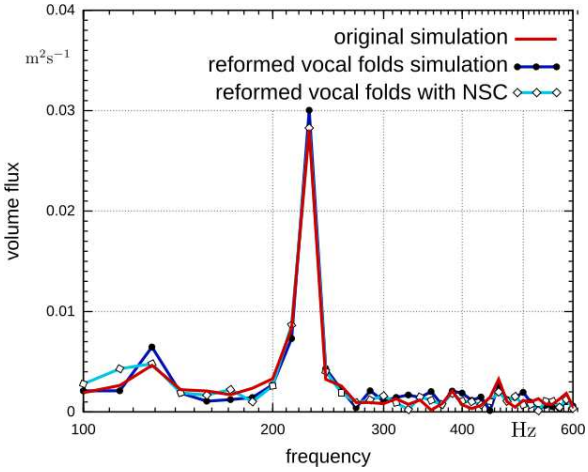


Figure 2.12: Spectral volume flux through glottis for case study of the inlet pressure [35].

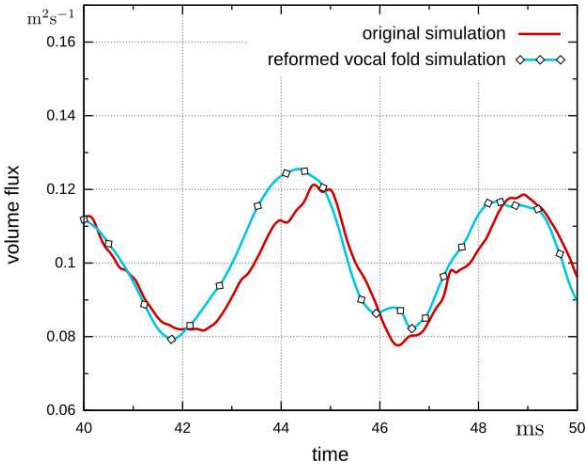
Besides a full-coupled simulation, 3D measurements (hemilarynx studies) can be used to obtain the motion of the vocal folds. Thereby, simplified models are acquired since realistic patient specific vocal fold models are not yet state of the art. Therefore, the third case study investigated the effect of incorrect models or occurring measurement inaccuracies by using a slightly reformed shape of the vocal folds. To obtain a prescribed movement, the displacements from the fully-coupled simulations with the original vocal folds was projected onto the deformed shape. A fully-coupled and a pure flow simulation with prescribed motion were then executed in addition to the original simulation. Although the geometry change is minor, there is an effect on the volume flux for the fully-coupled simulation (see Fig. 2.13a).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

However, the fact that the prescribed movement has the same characteristic peaks as the original simulation, implicates that small measurement errors in the vocal fold shape do not have a strong impact on the volume flux and that a fluid simulation with prescribed movement is feasible.



(a) Spectrum of volume flux through glottis.



(b) Volume flux at inflow in time domain.

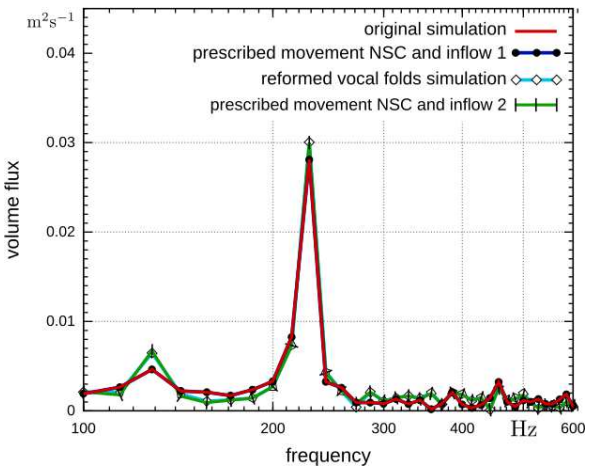
Figure 2.13: Results of case study for the fully-coupled simulations with original and reformed vocal folds [35].

In the last case study, the influence of the inflow-BC was analyzed. Therefore, the obtained velocity profiles at the inlet from the fully-coupled original case and from the fully-coupled case of the reformed geometry were used as an inlet-BC for the pure flow simulation with prescribed motion instead of the pressure BC at the inlet. The first inflow condition is referred to as "inflow 1", while the inflow condition gained from the simulation with reformed vocal folds is referred to as "inflow 2". The resulting volume flux of the two inflow conditions can be seen in Fig. 2.13b. Next, simulations with prescribed movement obtained from the original simulation and projected onto the reformed vocal folds were performed, once with "inflow 1" and once with "inflow 2". The results show that the simulations with the same inflow have an identical volume flux through the glottis, regardless if it is prescribed or fully-coupled (see Fig. 2.14). According to the authors, even if the prescribed movement from the original simulation is applied, but "inflow 2" condition is used, this results in the same volume flux as the fully-coupled simulation with reformed vocal folds although, the vocal fold motion is different. Consequently, the influence of geometry and motion of the vocal folds on the resulting volume flux through the glottis is not as strong as the velocity profile of the inflow-BC (if a velocity inflow-BC is applied). This fact also explains why models with steady vocal folds and prescribed waveform of the inflow-velocity provide proper results.

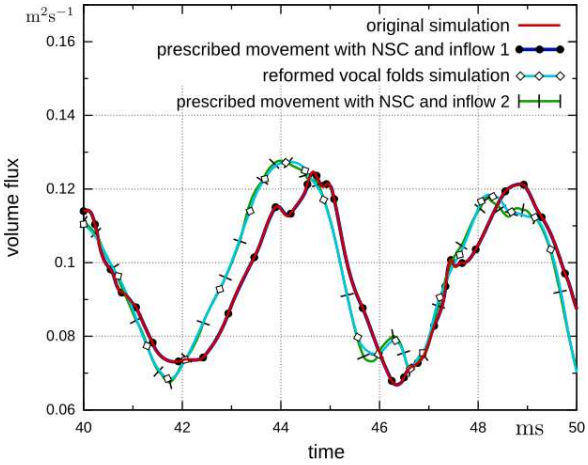
In general, the presented results showed that a pure CFD-simulation with a prescribed structural movement can substitute the fully-coupled approach. However, certain conditions have to be fulfilled in order to obtain good results. First, a no-slip boundary condition has to be applied to the interface of the moving vocal folds. Furthermore, the approach is sensitive to the driving inlet pressure, which therefore should be known as exactly as possible. Finally, small deviations of the geometry have minor effects on the resulting volume flux through the glottis, whereas the velocity profile used for the inlet-BC does have a strong impact (when it is used) and can compensate for inexact geometry and motion of the vocal folds [35].

In this work, the approach of pure flow simulation with a prescribed movement of the vocal folds is deployed (one-way coupling). As the applied model is three-dimensional, no results of fully-coupled simulations are available (inlet velocity profile and vocal fold motion) due to the computational effort.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Spectrum of volume flux through glottis.



(b) Volume flux at inflow in time domain.

Figure 2.14: Results of case study of the inflow BC [35].

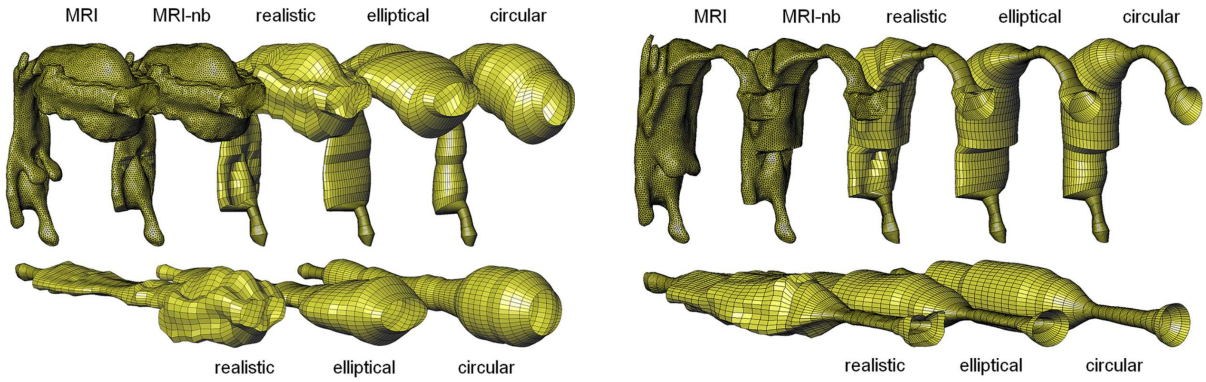
Hence, a pressure BC is used at the inlet instead of prescribing a velocity profile and vocal fold motion is provided by measurements instead of fully-coupled simulations.

Whereas for fluid dynamic simulations of the glottal airflow a model of the larynx and the lower part of the vocal tract is sufficient to obtain realistic flow conditions and to capture the principal acoustic sources, for acoustic simulations the entire vocal tract is essential to correctly predict the voice. Several studies show that only 3D vocal tract models entail higher order modes that typically appear in the higher frequency range above 5 kHz, whereas simplified, straight vocal tracts of a circular cross-sectional area or 1D models that assume plane wave propagation are not able to yield higher modes [29, 36–38]. For phonation types with significant high frequency energy (HFE) content, such higher order modes may be relevant to obtain a more natural voice and may be important for voice quality, speech localization, speaker recognition, and intelligibility [37, 39].

However, with regard to a prospective clinical application of the model in this thesis, we are willing to neglect the HFE content to preserve computational efficiency. Thus, for the model of real phonation, a simplified vocal tract is employed instead of employing a real geometry of the human airways which would demand a complex and fine mesh (for FEM and FVM) in order to properly resolve the shape. Therefore a simplified vocal tract according to Arnela *et al.* is applied [38]. In their work, they investigated the acoustic properties of vocal tracts for vowels /*a*/, /i/, and /u/ for the real MRI-based geometry and different degrees of simplification. Thereby, six cases were considered consisting of realistic, elliptical, and circular cross-section interpolated through the bent (original) or straight (simplified) midline as indicated for the vowels /*a*/ and /i/ in Fig. 2.15. For the case of the realistic cross-section, the original shape of each cross-section was preserved, while for the elliptical and circular case the cross-sections were approximated by an ellipse and a circle of equivalent area as the original one. Then, the straight configurations were obtained by straightening the curved centerline with the corresponding perpendicular cross-sections. In order to evaluate the acoustic characteristics of the various configurations, the respective transfer functions were computed by the fraction of the Fourier transform of the acoustic input and output signal. For doing so, transient acoustic simulations were performed by employing FEM and prescribing the acoustic particle velocity of a Gaussian pulse at the cross-section of the vocal tract inlet. Additionally, PML regions at the vocal tract inlet and outlet ensured free field radiation.

By comparing the transfer functions, it was found that for frequencies below $4 - 5$ kHz, the impact of bending and cross-sectional shape is weak, while above these values bent vocal tracts with realistic

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Vowel /a/. (b) Vowel /i/.

Figure 2.15: Realistic vocal tract obtained by MRI and geometries with different degree of simplification corresponding to the vowel /a/ and /i/. [38]

cross-sections are necessary to correctly emulate higher-order mode propagation. These findings were reproduced experimentally in [40] by applying 3D printed replicas of the vocal tracts with realistic and simplified geometries.

As in this thesis, the acoustic pressure is evaluated up to a frequency of 5 kHz, so that the application of a simplified vocal tract is valid and providing an effective way to reduce the complexity of the numerical model while preserving accuracy.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Chapter 3

Methodology

3.1 Hybrid Aeroacoustic Approach

The hybrid aeroacoustic approach is a powerful tool that decomposes the task of flow-induced sound prediction into a computational fluid dynamics (CFD) part and a computational aeroacoustic (CAA) part. Thereby, the solvers are highly specialized for their respective task allowing an efficient computation of the two physical fields. Additionally, the computational domains and the meshes for CFD and CAA can be tailored to their specific requirements. While the flow field usually has to be spatially resolved relatively fine[1], a coarser uniform mesh is required for the acoustic simulation[2]. Furthermore, the computational domains can be adapted to the respective application. In the case of human phonation, e.g. the modeling of the propagation region (see Sec. 4.3), which is essential for the acoustic propagation simulation, is not necessary for the CFD simulation as it has no significant impact on the resulting sources of sound. Therewith, the number of DOFs and thus, the computational expenses can be reduced considerably.

Figure 3.1 illustrates the workflow of the hybrid aeroacoustic approach. First, the flow field has to be provided, which is usually obtained by CFD simulations but in general, also measurements with sufficient spatial resolution can be utilized. For low Mach number flow (Ma < 0.3), compressible effects can be neglected and an incompressible CFD simulation can be performed which drastically reduces the computational cost. In the second step, the flow data has to be processed for the subsequent acoustic simulation which consists of the computation of the acoustic sources. When using separate meshes for CFD and CAA, a conservative interpolation from the flow to the acoustic grid (see Sec. 3.3) is necessary. Then, the two post-processing tasks can be carried out in an arbitrary order according to the Leibnitz integral rule, which allows the interchange of the integral and partial differential operators (see Sec. 3.3), and the optimal order of the two post-processing tasks depends on the application. On the one hand, performing the interpolation of the flow data required for source term interpolation (fluid-dynamic pressure and/or velocity) first, reduces the overall computational cost as the source term computation can be executed on the coarser CAA grid (fewer DOFs). On the other hand, conducting the source term computation first (on the CFD grid), in general increases the accuracy of the source term because spatial derivatives are computed more accurately due to the finer spatial resolution of the CFD grid.

Generally, the source term depends on the acoustic equation that is used for the CAA simulation and is basically the right hand side (RHS) of the respective equation. Within a hybrid aeroacoustic approach, either aeroacoustic analogies (e.g. Lighthill analogy) or acoustic perturbation equations (APE) [41] can be applied as governing equations for the generation and propagation of flow-induced sound. For the

[1] All relevant turbulent scales need to be resolved.
[2] The acoustic wave length needs to be resolved.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

simulations in this work, the perturbed convective wave equation (PCWE) is chosen [42]. Compared to other APEs, the PCWE is very efficient since it only needs to be solved for one unknown (the acoustic velocity potential ψ^{a}, see Sec. 3.2). Another advantage of the PCWE is that the convective part of the acoustic source term can be neglected for low velocities of the grid and the mean flow resulting in a source term only consisting of a temporal derivative. Thus, the disadvantage of inexact spatial derivatives on coarse grids is eliminated allowing to perform the source term computation more efficiently on the CAA mesh.

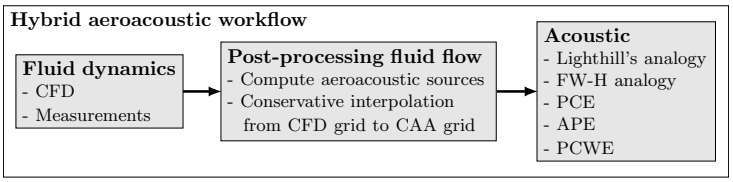


Figure 3.1: Workflow of the hybrid aeroacoustic approach.

3.2 Perturbed Convective Wave Equation

Hardin and Pope [43] introduced the acoustic/viscous splitting technique for the prediction of flow-induced sound. Afterwards, many scientists applied this idea and derived linear and non-linear wave equations [41, 44–46]. The essence of all these methods is that the flow field quantities are split into compressible and incompressible parts

$$p = \bar{p} + p^{\mathrm{ic}} + p^{\mathrm{c}} = \bar{p} + p^{\mathrm{ic}} + p^{\mathrm{a}} \tag{3.1}$$

$$\boldsymbol{v} = \bar{\boldsymbol{v}} + \boldsymbol{v}^{\mathrm{ic}} + \boldsymbol{v}^{\mathrm{c}} = \bar{\boldsymbol{v}} + \boldsymbol{v}^{\mathrm{ic}} + \boldsymbol{v}^{\mathrm{a}} \tag{3.2}$$

$$\rho = \bar{\rho} + \rho_1 + \rho^{\mathrm{a}}\,. \tag{3.3}$$

In this sense, the field variables are decomposed into mean ($\bar{p}$, $\bar{\boldsymbol{v}}$, $\bar{\rho}$) and fluctuating parts. Moreover, the fluctuating parts are further split into acoustic (p^{a}, $\boldsymbol{v}^{\mathrm{a}}$, ρ^{a}) and flow components (p^{ic}, $\boldsymbol{v}^{\mathrm{ic}}$). Finally, a density correction ρ_1 is built according to (3.3). Introducing an arbitrary Lagrangian-Eulerian description for the operators

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \left(\overline{\boldsymbol{v}} - \boldsymbol{v}_{\mathrm{rg}}\right) \cdot \nabla\,, \tag{3.4}$$

where $\boldsymbol{v}_{\mathrm{rg}}$ is the relative velocity of the grid, we arrive at the PCWE for moving meshes [42]

$$\frac{1}{c^2}\,\frac{D^2\psi^{\mathrm{a}}}{Dt^2} - \Delta\psi^{\mathrm{a}} = -\frac{1}{\bar{\rho}c^2}\,\frac{Dp^{\mathrm{ic}}}{Dt}\,. \tag{3.5}$$

This scalar convective wave equation is computationally efficient and describes aeroacoustic sources generated by incompressible flow structures and wave propagation through moving media. This formulation reduces the number of unknowns (acoustic pressure $p^{\mathrm{a}} = \bar{\rho}\frac{D\psi^{\mathrm{a}}}{Dt}$ and particle velocity $\boldsymbol{v}^{\mathrm{a}} = -\nabla\psi^{\mathrm{a}}$) to just one scalar unknown, the acoustic velocity potential ψ^{a}. The material derivative (3.4) in (3.5) is approximated for low Mach number flows and low grid velocities simply by the partial time derivative, which is applicable to human phonation. The incompressible pressure p^{ic} of the source term (RHS) is provided by the flow simulation (see Sec. 4.2).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

3.3 Conservative Interpolation

In this section, the two conservative interpolation schemes provided by the in-house pre-processing tool *CFSdat* are introduced. As these schemes are developed for the conservative interpolation of the acoustic source terms, the methodology is explained by means of the source term interpolation. However, in this work, not the source terms, but the flow data (needed for acoustic source term computation) is interpolated from the flow to the acoustic grid[3].

The presented hybrid aeroacoustic approach requires a conservative transformation of the acoustic sources f, which is the RHS of (3.5)

$$f := -\frac{1}{\bar{\rho}c^2}\frac{Dp^{\mathrm{ic}}}{Dt} \tag{3.6}$$

from the flow grid (superscript f) to the acoustic grid (superscript a). Based on this acoustic source density, the finite element right hand side

$$\int_{\Omega} \psi f \mathrm{d}\xi\,, \tag{3.7}$$

with $\psi \in H^1(\Omega)$ being the test function, is computed by the cell-centered approach

$$\int_{E^{\mathrm{a}}} N_i^{\mathrm{a}}(\boldsymbol{\xi}) f^{\mathrm{a}} \mathrm{d}\xi = \sum_{k \in M^{\mathrm{f}}} N^{\mathrm{a_i}}(\boldsymbol{\xi}_{E_k^{\mathrm{f}}}) F_k^{\mathrm{f}}\,, \tag{3.8}$$

where E_k^{f} denotes the cell of the flow grid, $\boldsymbol{\xi}_{E_k^{\mathrm{f}}}$ the local coordinate, M^{f} the set of flow cells, and the subscript i the node number on the CAA grid.

The algorithmic procedure of the cell-centered approach starts with the domain of source influence of an acoustic element. First, the volumetric loads are assembled on the CFD mesh. To reduce the computational complexity for fine flow grids, we simplify the integration over the source volume to a first order volume weighting

$$F_k^{\mathrm{f}} = V_{E_k^{\mathrm{f}}} f_k^{\mathrm{f}}(\boldsymbol{\xi}_{E_k^{\mathrm{f}}})\,. \tag{3.9}$$

Second, all loads F_k^{f} on the CFD grid that are inside the acoustic finite element E^{a} are collected by the set M^{f}. As illustrated in Fig. 3.2, the algorithm has to find the source location $\boldsymbol{x}_k$ inside an acoustic element (set M^{f}).

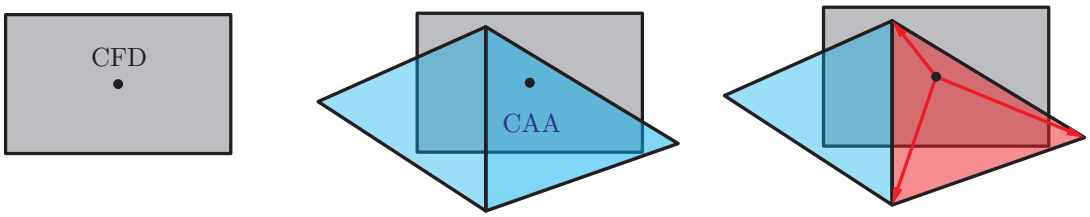


Figure 3.2: Steps for interpolation based on cell-centered approach. Based on a CFD centroid, the CAA element is determined and the sources are interpolated.

Third, the local position $\boldsymbol{\xi}_{E_k^{\mathrm{f}}}$ inside the reference finite element ($\boldsymbol{x}_k \multimap \boldsymbol{\xi}_{E_k^{\mathrm{f}}}$), corresponding to the global position $\boldsymbol{x}_k$ of the loads F_k^{f}, is determined. Finally, the loads F_k^{f} are interpolated to the nodes of the acoustic mesh by using the finite element basis functions $N^{\mathrm{a_i}}$ [47].

[3]The source terms are computed on the acoustic grid after the interpolation of the flow data.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

This cell-centered approach conserves the energy globally, whereas the energy is redistributed locally, which causes grid dependent local energy errors. In order to preserve the acoustic energy globally and locally, the nodal loads F_k^{f} of the flow grid have to be interpolated to a coarser acoustic grid (see Fig. 3.3). However, the mesh size of flow computations inside the glottis varies from fine meshes between the vocal folds to coarse meshes towards the subglottal and supraglottal channel.

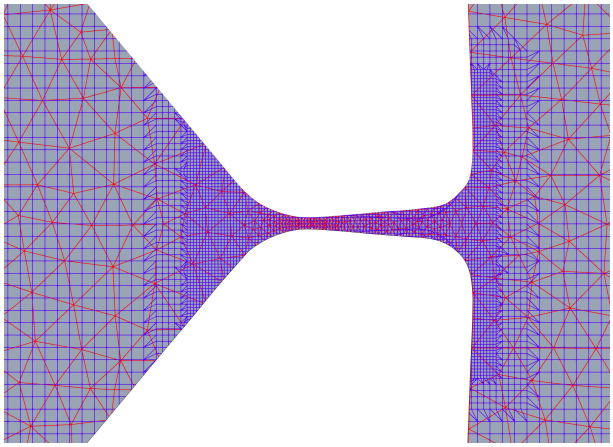


Figure 3.3: CFD grid (blue) and CAA grid (red).

In cases where the flow discretization is larger than the acoustic discretization, the standard approach fails locally because it neglects the contributions weighted by the cell-volume. Since the discretizations in the glottal region are in the same order of magnitude, both methods are compared during the validation of the computational setup, and the most appropriate method with respect to computational efficiency and accuracy is selected (see Sec. 4.3.2).

The improved approach (see Fig. 3.4) determines for each acoustic volume element E_l^{a} the set of intersecting flow cells E^{f}, which defines the intersection set $E_l^{\mathrm{f}\cap\mathrm{a}}$

$$E_l^{\mathrm{f}\cap\mathrm{a}} = E_l^{\mathrm{a}} \cap E^{\mathrm{f}} \,. \tag{3.10}$$

In this sense, the energy is conserved globally as well as locally for different mesh sizes. The finite element right hand side is calculated by

$$\int_{E^{\mathrm{a}}} N_i^{\mathrm{a}}(\boldsymbol{\xi}) f^{\mathrm{a}} \mathrm{d}\xi = \sum_{k \in M^{\mathrm{f}\cap\mathrm{a}}} N^{\mathrm{a_i}}(\boldsymbol{\xi}_k) F_k^{\mathrm{f}\cap\mathrm{a}} \,. \tag{3.11}$$

The nodal loads $F_k^{\mathrm{f}\cap\mathrm{a}}$ are volume-weighted based on the intersection. For a constant aeroacoustic source density f_l^{f} over the fluid cell, the integral simplifies to a multiplication of the intersection volume $V_{\mathrm{c}} \multimap V_{E_l^{\mathrm{f}\cap\mathrm{a}}}$ with the source density f_l^f at the volumetric centroid $\boldsymbol{x}_{\mathrm{c}} \multimap \boldsymbol{\xi}_{E_l^{\mathrm{f}\cap\mathrm{a}}}$ of the intersection polyhedron. These loads at the volumetric centroid $\boldsymbol{x}_{\mathrm{c}} \multimap \boldsymbol{\xi}_{E_l^{\mathrm{f}\cap\mathrm{a}}}$ of the intersection polyhedron

$$F_k^{\mathrm{f}\cap\mathrm{a}} = V_{E_k^{\mathrm{f}\cap\mathrm{a}}} f_k^{\mathrm{f}}(\boldsymbol{\xi}_{E_k^{\mathrm{f}\cap\mathrm{a}}}) \tag{3.12}$$

are now treated like nodal loads in the standard approach.

In contrast to the cell-centered approach, source contributions are added based on the intersection volume. Thereby, one-way aeroacoustic coupling is covered in which acoustic elements are entirely embedded within a single flow cell. Table 3.1 compares both methods as to how a varying mesh ratio Γ changes the relative energy content e_{r} in the actual wavelength compared to the total energy content of the signal. A mesh ratio of $\Gamma = 1$ means that the CFD cells and the acoustic elements are of the same

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

size; if $\Gamma < 1$ then the CFD cells are finer than the acoustic elements; if $\Gamma > 1$ then the CFD cells are larger than the acoustic elements.

Table 3.1: Ratio of the energy e_r corresponding to the actual wavelength and the total energy for the cell-centered procedure and the cut-volume-cell procedure, respectively. [48]

Γ	Cell-centered e_r	Cut-volume-cell e_r
0.5	99.59 %	99.34 %
1	99.51 %	99.14 %
2	87.50 %	98.76 %
5	55.00 %	98.94 %
10	31.23 %	98.96 %
20	15.85 %	99.06 %
30	10.60 %	99.10 %

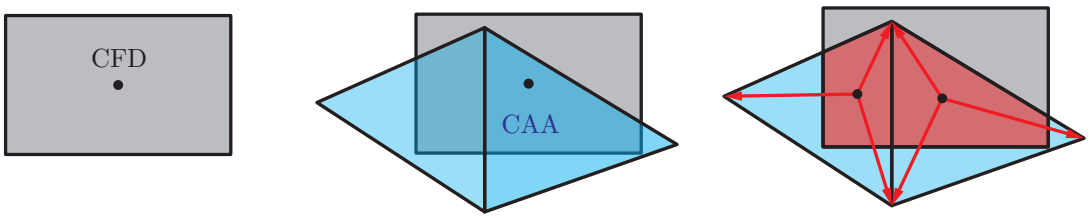


Figure 3.4: Steps for interpolation based on cut-volume-cell approach. The intersection elements are determined, then intersected, and the forces are interpolated based on this intersections. [48]

The more sophisticated conservative cut-volume-cell interpolation outperforms the cell-centered procedure, especially for large mesh ratios. In this regime, the cell-centered interpolation transfers energy of the aeroacoustic sources to unphysical wave lengths. The cut-volume cell approach conserves the energy for all investigated mesh ratios. For both methods, the computational demand increases linearly with the number of elements, but the more accurate cut-volume-cell intersection takes longer since the flow and the acoustic grid are intersected.

Based on the Leibnitz integral rule, the spatial integral and the partial time derivative can be interchanged:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\chi(t)}^{\varphi(t)} f(x,t)\mathrm{d}x = \int_{\chi(t)}^{\varphi(t)} \frac{\partial}{\partial t} f(x,t)\mathrm{d}x + f(\varphi(t),t)\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) - f(\chi(t),t)\frac{\mathrm{d}}{\mathrm{d}t}\chi(t)\,. \tag{3.13}$$

In the case of fixed integration limits, the terms considering the lower and upper integration limits are zero and the rule simplifies to

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\chi}^{\varphi} f(x,t)\mathrm{d}x = \int_{\chi}^{\varphi} \frac{\partial}{\partial t} f(x,t)\mathrm{d}x\,. \tag{3.14}$$

In this work, the geometry (mesh) of the glottis is not adapted during the CAA simulation - in contrast to the CFD simulation with prescribed movement of the boundary - but the condition of the maximum opening is used (i.e. CFD domain at maximum opening of VFs) in order to capture all acoustic sources. Therefore, the boundaries for the spatial integration are fixed allowing the interchange of the conservative interpolation and the partial time derivative for the source term computation as indicated in (3.6).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Chapter 4

Introduction of the Numerical Model

In this chapter, the numerical model for the validation of the applied methodology is introduced. For the purpose of validation, this model has a simplified geometry and corresponds to the experimental setup (including the BCs) that provides the data for validation. In contrast, the model applied for the case studies of Chap. 5 represents real human phonation and thus, employs more complex geometries and different boundary conditions. Besides the introduction of the model, investigations and modifications of the existing validation model with regard to the computational effort and accuracy are presented in Sec. 4.3.2 where the model is adapted towards high efficiency. Thereby, it has to be noted that the CFD and CAA model have already been validated in an earlier stage of the project.

The experimental setup consisted of a mass flow generator, a subglottal channel with synthetic vocal folds (silicone) and a supraglottal channel (acrylic glass) with stiff ventricular folds (see Fig. 4.1). For validation purposes, the supraglottal channel, which leads into the farfield and represents the vocal tract, has a simplified rectangular shape. The glottal area waveform (GAW), which is required for prescribing the vocal fold motion in the CFD simulation, was extracted from high-speed video footage (with a sample rate of 4000 frames per second). Moreover, the measurements were conducted in an anechoic chamber with a cutoff frequency of about 300 Hz. Four microphones measured the sound radiation at 1 m distance from the end of the supraglottal channel and an angle of 45° from the center line (see Fig. 4.16a and 4.2). Further details of the experimental setup can be found in [49]. The geometry of the flow and acoustic model derived from the experiment follows the rules of the hybrid aeroacoustic approach. The CFD domain is depicted by the bluish region in Fig. 4.1. It includes the rectangular subglottal channel, the vocal folds, the ventricular folds, and the rectangular supraglottal channel, that is, all parts of the test rig where the flow speed is sufficiently high. Additionally to the CFD domain, the CAA geometry (reddish region) models the sound radiation to the microphone positions in the farfield, and moreover avoids reflections at boundaries, where the domain is truncated.

4.1 Workflow

The applied workflow (Fig. 4.3) is defined by the hybrid aeroacoustic workflow. Therein, the results of the CFD simulation, required for source term computation (incompressible pressure p^{ic}), are provided by our cooperation partner, Universitätsklinikum Erlangen (Department for Phoniatrics and Pedaudiology), for every time step of the transient CAA simulation. In order to preserve the possibility to also consider the convective term within the source term computation or the application of a different acoustic equation (e.g. Lighthill), the fluid velocity is transferred as well. Furthermore, the mesh-file is provided for every

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

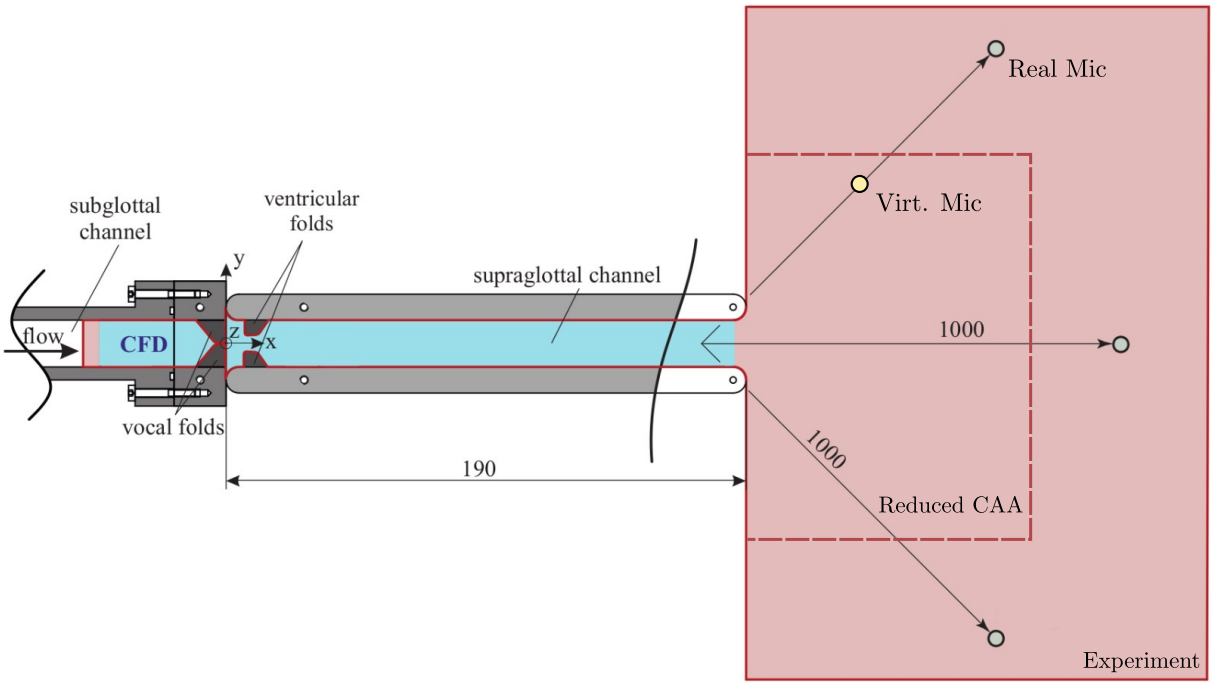


Figure 4.1: Schematic of the synthetic larynx model adapted to the experimental setup with indicated CFD (bluish) and CAA (reddish) domain . The acoustic pressure is transformed from the virtual to the real microphone position. All dimensions in mm. [49]

time step as it is occasionally needed during the interpolation when the connectivity between CFD data and cell index gets lost[1].

In a first step, the incompressible pressure p^{ic} is conservatively interpolated from the flow to the acoustic grid via the in-house pre-processing tool *CFSdat*. Besides the CFD simulation itself, this is the computationally most expensive step due to the fine CFD mesh ($\approx$ 1.64 million cells). Having accomplished the interpolation, the following computations become less demanding because of the coarser resolution of the acoustic mesh ($\approx$ 280 000 cells), which proves the capability of the hybrid aeroacoustic approach.

Due to numerical issues within the CFD simulation that mainly emerge because of the prescribed movement of the vocal folds, the pressure results of the individual fluid cells contain non-physical peaks (see Fig. 4.4, blue graph). Therefore, the (interpolated) pressure of each acoustic node is filtered in a next step (see Fig. 4.4, red graph) employing *Butterworth* low-pass filter (10th order) in *MATLAB*© with a cutoff frequency of 5 kHz. A continuous pressure signal is of great importance, as the subsequent source term computation consists of the partial time derivative of the incompressible pressure $\frac{\partial p^{\mathrm{ic}}}{\partial t}$, which is executed by a second order smoothed noise differentiator [50] that is also implemented in the in-house pre-processing tool *CFSdat*. Without filtering, the artifacts of the raw signal would lead to non-physically high acoustic source terms. For the source term computation, the convective part of the source term (see Sec. 3.2 and 3.3) is neglected due to the relatively low mean velocity of the flow. The admissibility of this assumption is examined in Sec. 4.3.2.

Once the acoustic sources are computed, the transient acoustic simulation is executed. Therein, the partial differential equations are solved with the in-house research FE code *CFS++* [42]. Therewith, the evolution of the acoustic pressure is obtained within the entire CAA domain. The propagation region of the computational domain of the acoustic setup does not contain the microphone positions of the

[1] Normally, the mesh-file of the first time step can be used for the entire interpolation (all time steps). However, the CFD software occasionally changes the cell numbering when exporting the results. When this is the case, the mesh-file is updated during the interpolation and used until the next change of numbering.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

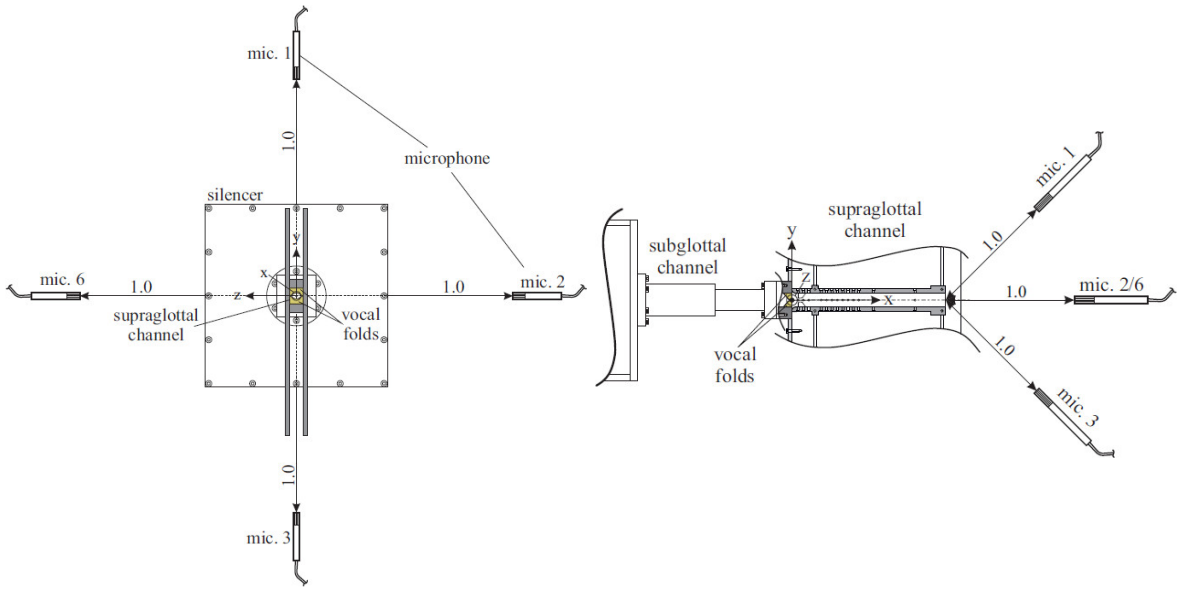


Figure 4.2: Front and side view of the experimental setup for sound measurements at the indicated microphone positions. All dimensions in m. [49]

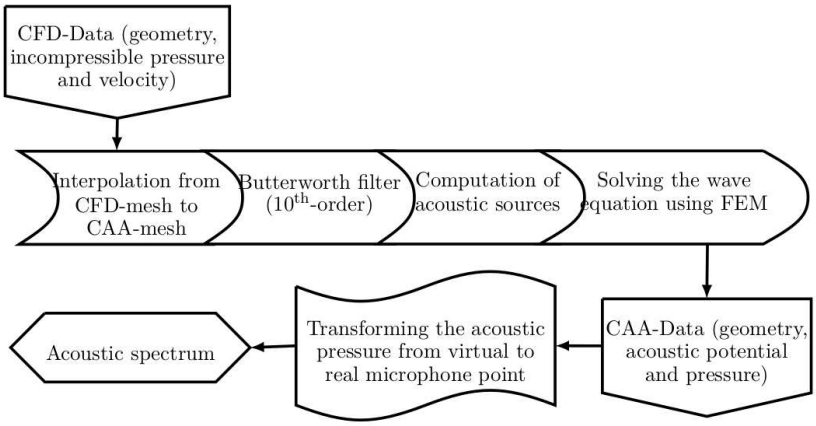


Figure 4.3: Workflow of the aeroacoustic simulation, including pre- and post-processing.

experiment (see Fig. 4.1) as including it would cause an unnecessary large domain and thus, unnecessary DOFs and computation time. Instead, the acoustic pressure at the virtual microphone point (simulation) is transformed to the location of the observation point of the experiment. According to the Fraunhofer approximation, the sound in the farfield decays with $1/r$ with r being the distance from the source position. Therefore, the acoustic pressure at the virtual microphone point with a distance of $r_{virt} = 80\,\text{mm}$ from the outlet of the vocal tract is transformed to the position of the real microphone in the experiment with a distance of $r_{real} = 1000\,\text{mm}$ by

$$p^{\text{a}}_{\text{real}} = p^{\text{a}}_{\text{virt}} \frac{r_{\text{virt}}}{r_{\text{real}}} = p^{\text{a}}_{\text{virt}} \frac{8}{100}. \tag{4.1}$$

Finally, depending on the purpose of the simulation the spectrum of the sound pressure level (SPL) or the amplitude spectral density (ASD) is computed with *MATLAB*©. Due to the long measuring time of the experiment, the ASD is used for the validation model, which is the square root of the power spectral

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

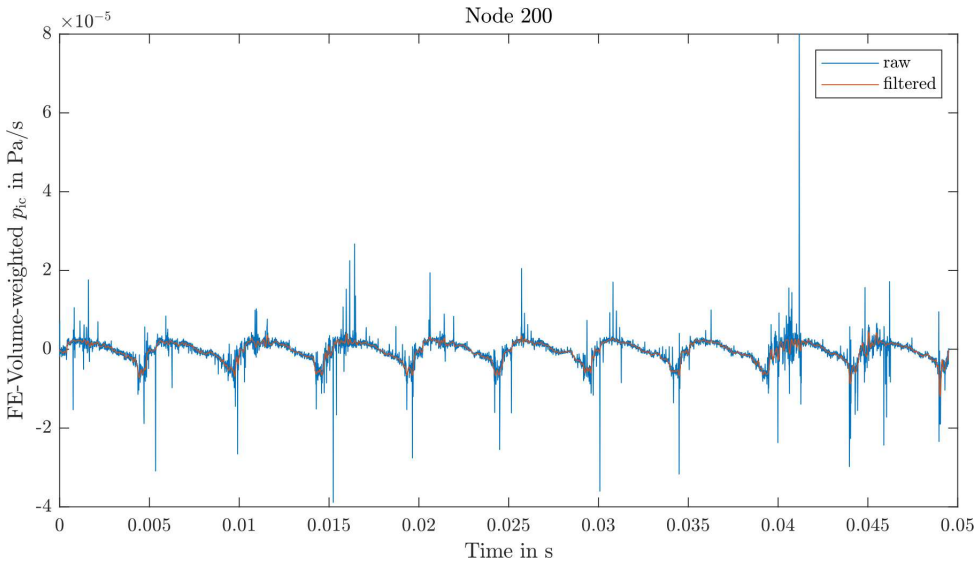


Figure 4.4: Raw and filtered signal of the interpolated incompressible pressure for a characteristic node inside the glottis.

density (PSD) computed via the function *pwelch* in combination with a hamming-window. Otherwise, (for the model of real phonation, Chap. 5) the SPL is used for evaluation as it is a more common quantity in acoustics. The sound pressure level L_{Pa} is defined by

$$L_{Pa} = 20 \log_{10} \frac{p_a}{p_{a,ref}} \tag{4.2}$$

where the threshold of hearing at 1 Hz $p_{a,ref} = 20\,\mu$Pa, is used as reference pressure.

In order to validate the transformation of the acoustic pressure from the virtual to the real microphone point (Eq. (4.1)), the transformation was carried out for a real microphone point located inside the propagation region. Consequently, the pressure signal of the real microphone point is known and can be compared with the transformed signal. Figure 4.5 shows the results of the transformation of the acoustic pressure signal from a microphone point at 50 mm in front of the vocal tract outlet (virtual mic.) to 80 mm (real mic.). First, the deviation between the virtual and real signal thereby is $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.61$, but reduces after transformation to a negligible error of $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.039$.

4.2 CFD Simulation Setup

This chapter presents the setup and parameters applied for the pure flow simulation with prescribed motion of the vocal folds. The incompressible CFD simulations were performed by our cooperation partner Universitätsklinikum Erlangen using the commercial CFD software *Star-CCM+*. Details of the 3D model are shown in Fig. 4.6. Thereby, it has to be noted that the investigations of [51], where the figure is extracted from, did not consider the ventricular folds, whereas in the models (CFD and CAA) in this thesis, they are included (see Fig. 4.8). Apart from not considering the ventricular folds, the setup including the boundary conditions of [51] and of this thesis are the same. A thorough description of the CFD setup including its validation can be found in [51]. Additionally, in [52], the computational expenses are presented.

A large eddy simulation (LES) on a finite-volume cell-centered non-staggered grid with a wall-adapting local eddy-viscosity (WALE) subgrid scale model was performed to model the turbulent flow in the larynx.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

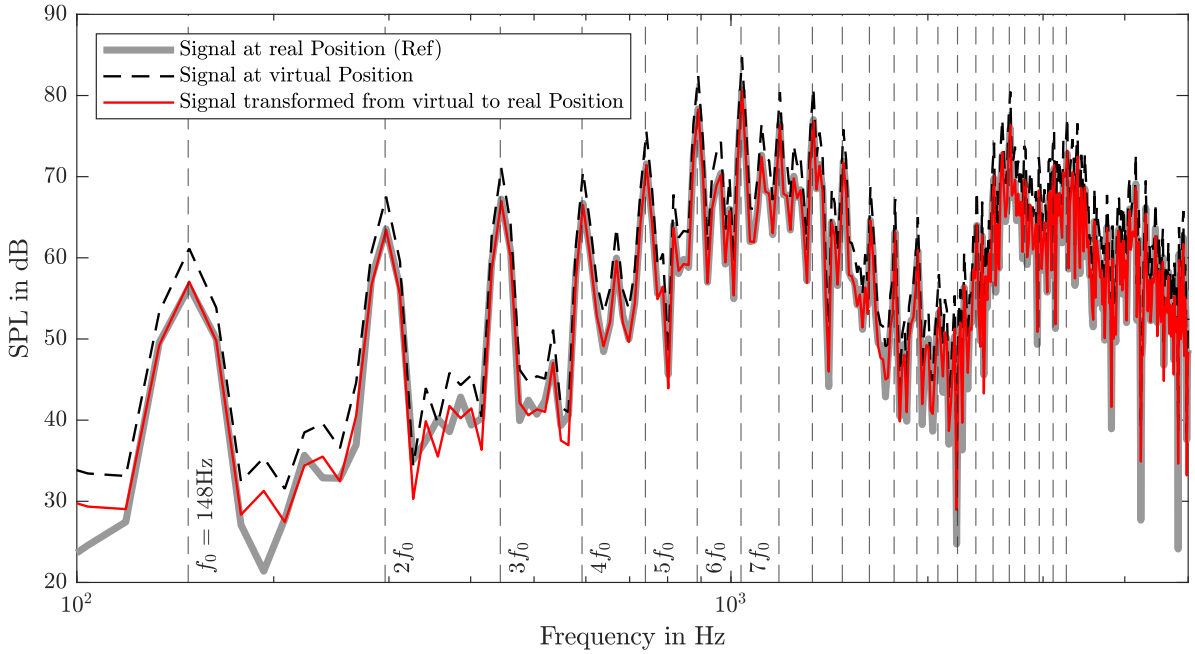


Figure 4.5: ASD of virtual, real and transformed pressure signal. $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.039$.

Thereby, a central difference scheme with second-order accuracy was used to discretize the convective and diffusive terms of the Navier-Stokes equations. These pressure-velocity linked equations were solved non-iteratively within a PISO (Pressure-Implicit with Splitting of Operators) algorithm. Furthermore, an algebraic multigrid (AMG) method with a Gauss-Seidel relaxation scheme and biconjugate gradient accelerator was applied to solve the final linear system of equations. The density and kinematic viscosity of air were considered to be $\rho = 1.18\,\frac{\mathrm{kg}}{\mathrm{m}^3}$ and $\nu = 1.57 \cdot 10^{-5}\,\frac{\mathrm{m}^2}{\mathrm{s}}$, respectively. As boundary conditions for the validation setup, the pressure at the inlet was set to be constant and equal to the mean value of the experimental subglottal pressure, while the outlet pressure reproduced the temporal evolution determined in the experiments. For the model of real phonation (Chap. 5), no measurements of the outlet pressure were available. Therefore, the outlet pressure was set to zero. At all walls (including the moving walls of the VFs), no-slip no-injection boundary conditions were employed.

The mesh consisted of hexahedral elements and was refined in the glottal area (glottis refinement) and near the supraglottal region. Figure 4.7 shows the finest mesh used for the validation of the model consisting of 2.4 million cells and a mesh size reaching from 0.063 mm to 0.5 mm. The overset mesh approach provided by *Star-CCM+* was applied to treat the mesh motion for moving boundaries. However, this methodology doesn't allow a complete closure of the glottis. Therefore, a minimum glottal gap of 0.2 mm was forced during the glottal closure, which caused a negligible leakage flow.

For the temporal discretization, the time step was chosen to be in the magnitude of 1 µs depending on the oscillation frequency of the vocal folds, which was 148 Hz for the validation case and between 80 Hz and 320 Hz for the case studies of Chap. 5. In order to reduce the transfer of data, every tenth time step is exported and used for the source term computation. Regarding the duration of the transient simulations, experience of previous simulations, as well as literature, advise to capture at least ten oscillation cycles of the vocal folds in order to obtain well developed spectra of characteristic flow and acoustic quantities. Additionally, the impact of the simulation duration on the spectra is investigated in Sec. 4.3.2.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

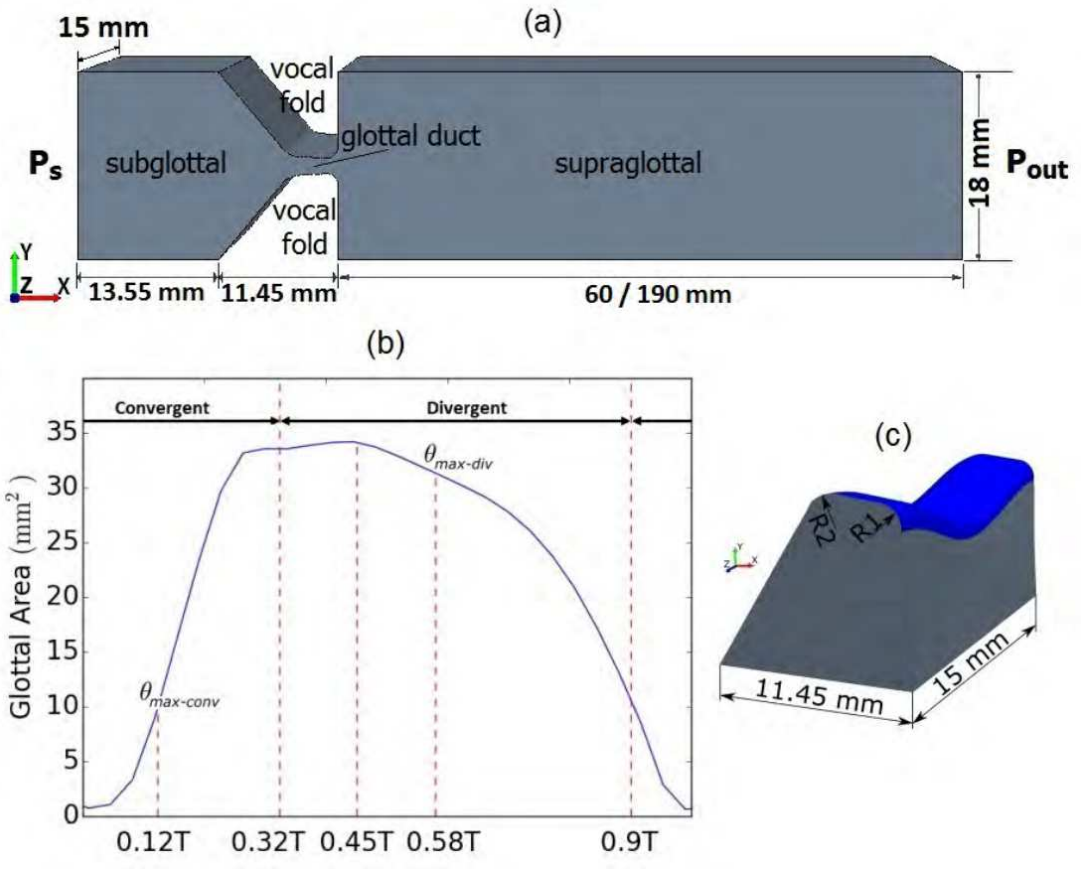


Figure 4.6: (a) Schematic of the numerical larynx model without consideration of the ventricular folds. (b) GAW obtained from the experimental model of flow-induced motion of the VFs. (c) Geometry of the vocal fold at the middle of the cycle [51].

4.3 CAA Simulation Setup

Figure 4.8 shows the computational setup for the acoustic simulation of the validation case. In addition to the CFD domain (larynx and vocal tract), the acoustic setup contains a propagation region which is attached downstream of the vocal tract and is necessary to calculate the sound radiation to the microphone points. The dimension of the propagation domain in x-, y- and z-direction is 85 mm × 120 mm × 120 mm. Additionally, two perfectly matched layer (PML) regions ensure free radiation and avoid reflections of the sound at the boundaries where the model is truncated. A thorough description of PMLs can be found in [53]. Apart from the bottom of the supraglottal propagation region, which is modeled sound-hard (corresponding to the experimental setup), one PML is attached adjacent to the propagation region. As PML must have orthogonal bounding surfaces in the current implementation of the in-house acoustic solver *CFS++*, the purpose of the other propagation domain at the inlet of the larynx is just to simplify the meshing procedure and circumvent a non-conforming interface on the subglottal PML region. The meshes of the larynx and the vocal tract are generated together in one step in order to avoid a non-conforming interface in the region with major sources of sound. In contrast, at the glottis inlet and vocal tract outlet where only minor sources occur, the non-conforming interfaces allow efficient mesh generation with high element quality. The treatment of non-conforming interfaces is implemented in *CFS++*. Furthermore, all channel walls are modeled as fully reflecting sound-hard walls, which is justified by a sufficiently high acoustic impedance jump between air and tissue. The fluid density of the air is set to $\bar{\rho} = 1.18\,\mathrm{kg/m^3}$ corresponding to the CFD simulation and the speed of sound is given by $c = 343.4\,\frac{\mathrm{m}}{\mathrm{s}}$.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

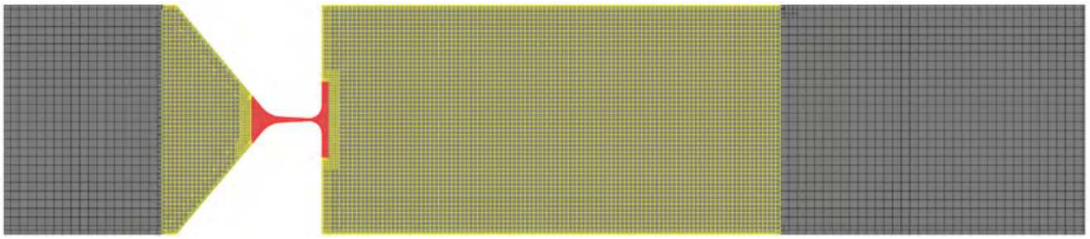


Figure 4.7: Sectional view of the finest CFD mesh for validation. Gray: cells with base size. Yellow: Cells with basic refinement. Red: Cells with glottis refinement [51].

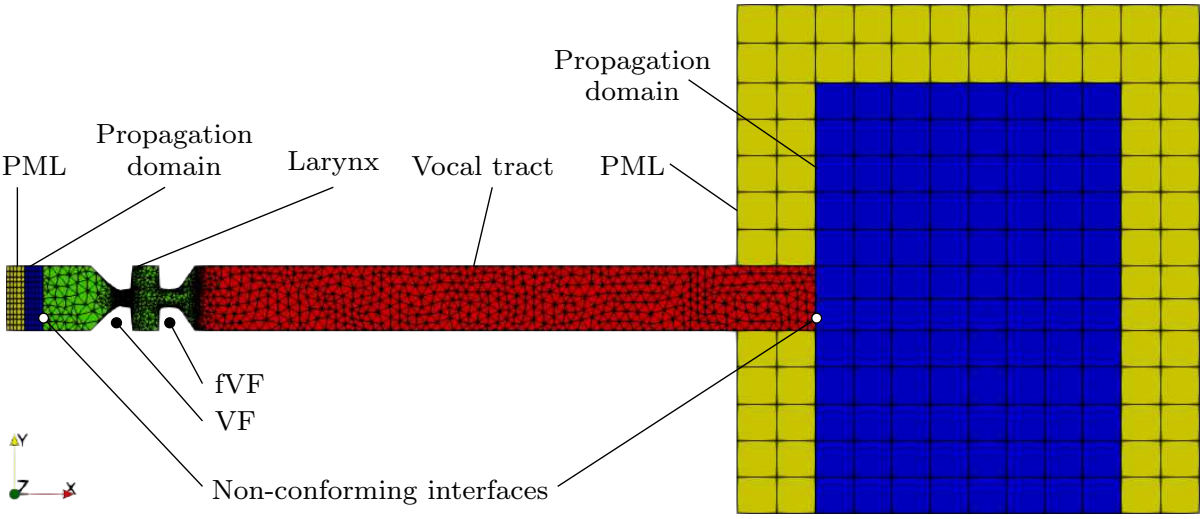


Figure 4.8: The validation setup for the acoustic simulation with a - in the midplane - sliced view of the supraglottal PML region.

The larynx and the vocal tract together match the CFD domain and contain the aeroacoustic sources. However, inside the middle and rear part of the vocal tract, only minor acoustic sources occur (see Fig. 4.9). Therefore, these regions could be neglected using a blending function. But since this doesn't increase the efficiency of the simulation, all source terms are considered [54].

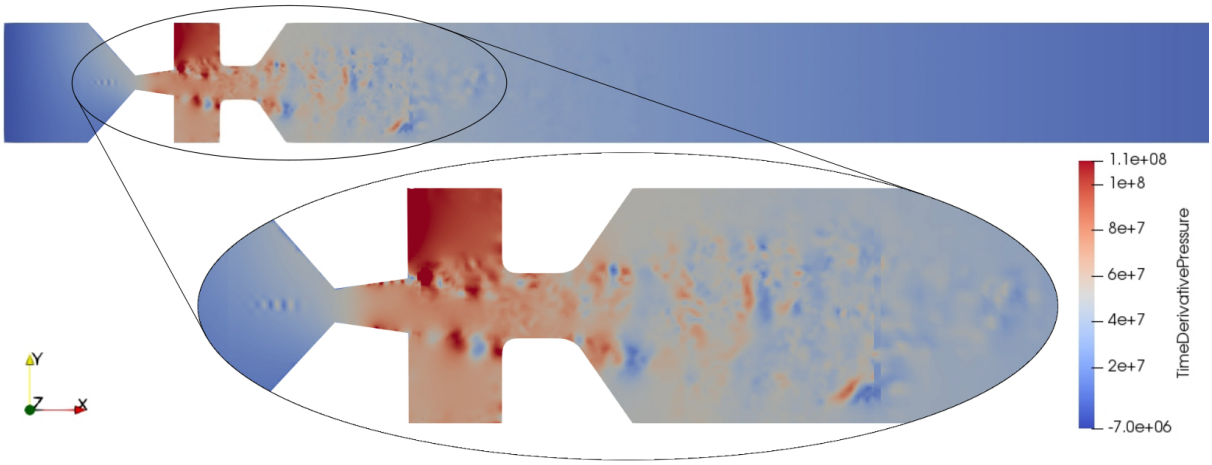
In contrast to the flow simulation, the computational domain is not adapted during the simulation, but a steady geometry representing the maximum opening of the vocal folds is deployed in order to capture all acoustic sources in every time step. The grid motion caused by the vocal folds has a Mach number below $\mathrm{Ma}_{\mathrm{gr}} < 0.015$ and the mean convective Mach number inside the glottis is $\overline{\mathrm{Ma}} < 0.1$ since the averaged flow speed is relatively low [51]. Therefore, the movement of the vocal folds as well as the convective term of the aeroacoustic source can be neglected in the acoustic simulation, which drastically reduces the complexity of the model, the data transferred, and the computational cost. The results of the flow simulation, which are required for the source term computation, are exported into EnSight Gold-files. Thereby, every tenth CFD time step is written to the files, resulting in a temporal resolution of the source term of $\Delta t \approx 10^{-5}\,\mathrm{s}$ used for the acoustic simulation.[2]

4.3.1 Spatial Discretization

The mesh for the acoustic simulation is generated with *ANSYS ICEM CFD*. As mentioned above, the domain of the larynx and the vocal tract are meshed together in one step in order to ensure a conform

[2] The exact size of the time step depends on the cycle duration of the oscillation.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



low

Figure 4.9: Qualitative visualization of the acoustic source terms in the region of the larynx and vocal tract for a characteristic time step by means of the partial time derivative of the acoustic pressure.

mesh. Thereby, the curvature based refinement was enabled to sufficiently resolve the three-dimensionally curved geometry of the vocal folds. Despite the simple rectangular shape of the vocal tract, tetrahedral elements are used within the automatic meshing procedure in order to allow the use of more complex geometries of the vocal tract after the validation of the numerical model. The meshes of the propagation regions and the PML regions consist of hexahedron elements. Table 4.1 summarizes the number of elements and the maximum element size in the different regions.

Table 4.1: Details of the CAA mesh.

Region	Elements	max. Element size
Larynx	46 400	4.8 mm
Simple vocal tract	45 100	4.5 mm
Propagation domain	1 150	10.6 mm
Prop. domain subglottal	600	1.5 mm
PML subglottal	600	1.5 mm
PML propagation	1 530	10.6 mm
Total	95 800	10.6 mm

The existing capabilities of the simulation software allow well suited meshes for the subdomains. Inside the larynx, a mesh consisting of tetrahedral elements approximates the continuous geometry by 46 400 elements with a maximum element size of 4.8 mm. The discretization is chosen considerably different to the CFD grid, but the presented conservative interpolation strategy allows a CFD grid independent meshing at least for the cut volume cell approach (see Sec. 3.3). However, inside the glottis region, the characteristic length of the acoustic mesh is in the range of the flow discretization. Hence, the application of the faster cell-centered interpolation approach is investigated in Sec. 4.3.2. Based on the simple estimation formula for the characteristic acoustic element length Δl_a

$$\Delta l_\mathrm{a} = \frac{\lambda}{20} = \frac{c}{20 f_\mathrm{max}} = 3.4\,\mathrm{mm}\,, \tag{4.3}$$

a maximum frequency of $f_\mathrm{max} = 5\,\mathrm{kHz}$ (cutoff frequency of low-pass filter) is resolved with 20 linear finite elements per acoustic wave length λ.

The existing mesh is the result of a grid study, which was part of a previous working package, with a trade-off between accuracy and computational effort [55]. Thereby, the model was altered towards

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

increasing efficiency by reducing the number of elements within the thickness of the supraglottal PML region and by mesh coarsening in general. Although the maximum grid size (see Tab. 4.1) exceeds the characteristic acoustic element length of $\Delta l_{\mathrm{a}} = 3.4\,\mathrm{mm}$, the results show negligible deviations from models with finer spatial discretization and thus the respective mesh was chosen for further investigations due to computational efficiency.

4.3.2 Modifications and Investigations

In this section, various aspects of the existing CAA model for the validation case apart from the grid resolution are investigated in order to find further potentials for an increase in efficiency and accuracy. Thereby, the influence of

- the approach for the source term interpolation,
- the modeling of the boundary at the larynx inlet,
- the simulation time (number of VF oscillation cycles),
- the consideration of the convective term within the source term computation, and
- obstacles in the experimental setup

is assessed by means of a transfer function or the acoustic spectrum at the microphone point 2 (see Fig. 4.2)). In order to quantitatively evaluate the relative error of the acoustic spectrum compared to a reference, the L_2-norm error

$$\mathrm{Err_{rel}}^{\mathrm{L_2}} = \sqrt{\frac{\sum_{f_l=10Hz}^{f_u=5kHz} (\mathrm{ASD_{ref}}(f) - \mathrm{ASD}(f))^2}{\sum_{f_l=10Hz}^{f_u=5kHz} \mathrm{ASD_{ref}}^2(f)}}. \tag{4.4}$$

is introduced, which is computed within the range of $f_l = 10\,\mathrm{Hz}$ to $f_u = 5\,\mathrm{kHz}$. The workflow and model are then adapted depending on the results of the investigations before applying them for the case studies of Chap. 5.

Approach for the Source Term Interpolation

As described in Sec. 4.3.2, the cell-centered interpolation procedure is faster than the improved cut-volume-cell approach because the intersection volumes don't have to be computed. However, its application is limited to low mesh ratios $\Gamma = l_{\mathrm{CFD}}/l_{\mathrm{CAA}} < 1$ in order to locally conserve the energy. The existing mesh ratio reaches from $\Gamma_{\mathrm{min}} = 0.05$ to $\Gamma_{\mathrm{max}} = 1$. According to Table 3.1, this would allow the application of the cell-centered interpolation scheme and thus save computational time.

The resulting spectra at the microphone point (Fig. 4.10) for acoustic simulations carried out with source terms interpolated with both algorithms validates this assumption. Thereby, the reduction of computational time for the interpolation procedure in the illustrated case (20 oscillation cycles of the vocal folds) is about five percent, but depends on the number of elements. The relative error of the faster cell-centered procedure compared to the more sophisticated cut-volume-cell interpolation scheme is $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.0007$. Therefore, the faster cell-centered approach can be applied for the present meshes without a significant loss of accuracy.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

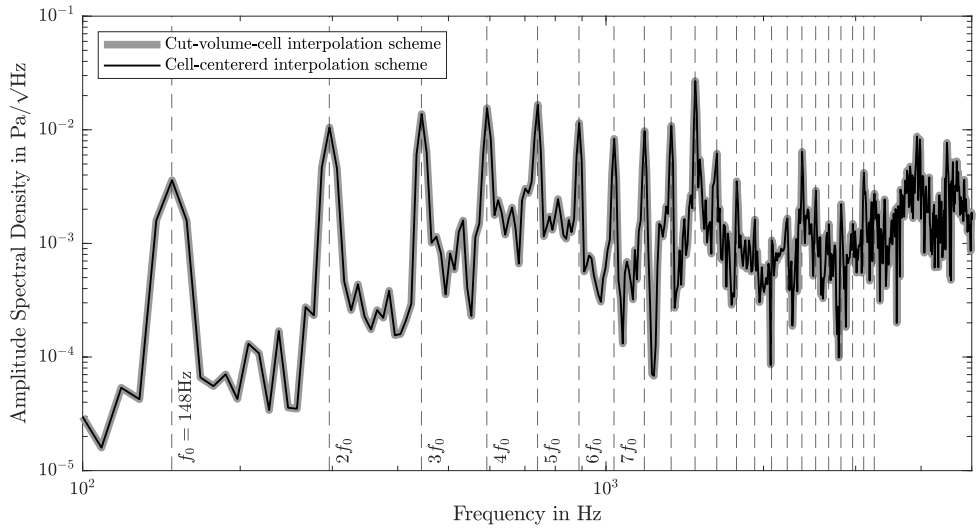


Figure 4.10: Resulting ASD at microphone point 2 for cut-volume-cell and cell-centered interpolation scheme. $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.0007$.

Modeling of the Boundary at the Larynx Inlet

In channels with a dimension d of the cross-section that is small compared to the acoustic wave length λ, the wave can be assumed as a plane wave and a so called absorbing boundary condition (ABC) can be used (instead of a PML region) to model free radiation where the channel is truncated. With a maximum diameter of the glottis cross-section of $d = 15\,\mathrm{mm}$, the condition

$$f_{\max} = 5\,\mathrm{kHz} < \frac{c_0}{2d} = 11.5\,\mathrm{kHz} \tag{4.5}$$

is fulfilled clearly and thus the plane wave approximation is applicable.

By substituting the subglottal PML region for an ABC, the subglottal propagation region including the non-conforming interface can be omitted, as the ABC is applied directly on the surface elements on the boundary at the glottis inlet. Due to this substitution, the number of DOFs is decreased and thus, computational time reduced. Thereby, reducing the number of PML elements is especially effective, as one PML node possesses five DOFs in contrast to a conventional node with only one DOF. The resulting acoustic spectrum for the modified setup compared to the existing setup in Fig. 4.11 shows that the simplification is permissible, as there are no significant deviations ($\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.001$). Therefore, this modification is incorporated into the model for further applications.

Simulation Time

As described before, the sampling duration of the simulation is much shorter than the duration of measurements in the experiment. This leads to the question of how much of an impact the number of simulated cycles (of the phonation process) has on the resulting spectrum. One cycle corresponds to the duration of one oscillation of the vocal folds, which determines the base frequency of the resulting voice. The above discussed spectra were obtained from simulations with an output data that includes 0.135 01 s, which is equivalent to 20 oscillation cycles. In Fig. 4.12, the resulting acoustic spectrum of the simulation of 20, 18, 15, and 10 cycles is shown. Especially the spectra of 18 and 15 cycles are still similar to 20 cycles ($\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.15$, resp. $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.34$), but this already results in a reduction of computational

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

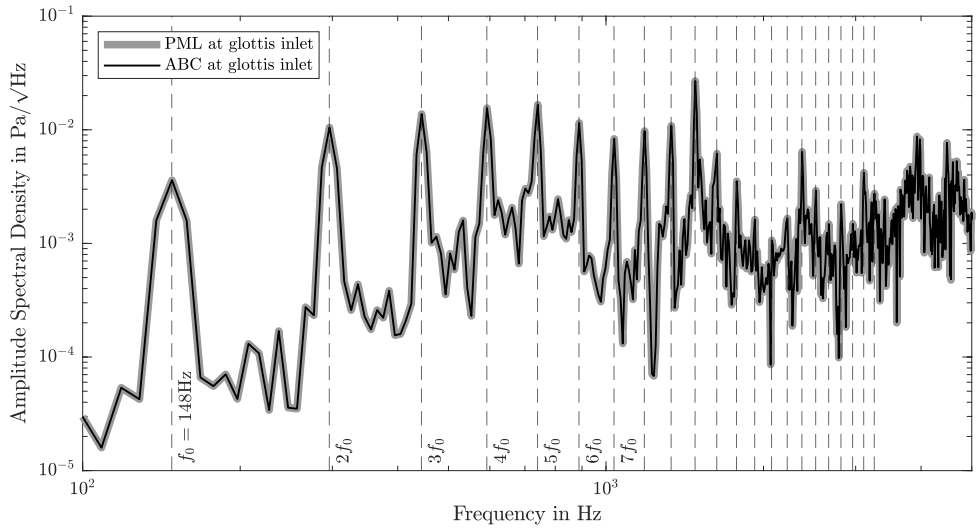


Figure 4.11: Resulting ASD at microphone point 2 for PML and for absorbing boundary condition (ABC) at larynx inlet. $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.001$.

time of 10 % or 25 %, respectively. By considering only half of the initial CFD simulation data and time (10 cycles), the reduction of the computational cost is 50 %, with obvious deviations, especially at the peaks, which are the main reason for the significant error of $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.60$. The reason for the deviations is that the truncation effects the finite signal (at the beginning and the end of the simulation) becomes more dominant for shorter signals. Furthermore, it has to be noted that the logarithmic scale of the amplitude has to be considered when comparing the graphs as the deviation of large amplitudes seems smaller than it actually is. However, a relative comparison of different larynx geometries and vocal fold motions, as it will be investigated in the case studies of Chap. 5, will be still valid also for a short simulation time.

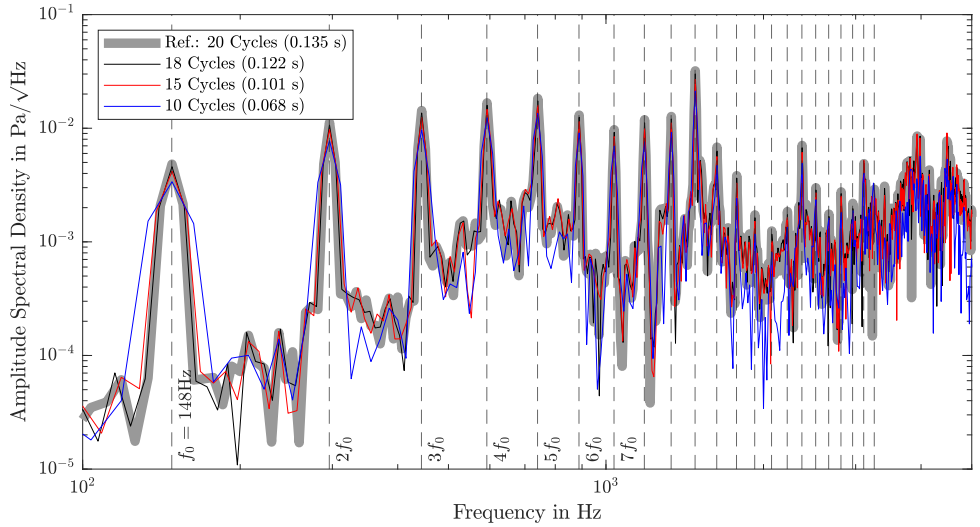


Figure 4.12: Resulting ASD at microphone point for different number of simulated cycles. 18 cycl.: $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.15$; 15 cycl.: $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.34$; 10 cycl.: $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.60$.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Consideration of the Convective Term

The convective term $\left(\overline{\boldsymbol{v}} - \boldsymbol{v}_{\mathrm{rg}}\right) \cdot \nabla p^{\mathrm{ic}}$ of the (full) source term $\frac{\partial p^{\mathrm{ic}}}{\partial t} + \left(\overline{\boldsymbol{v}} - \boldsymbol{v}_{\mathrm{rg}}\right) \cdot \nabla p^{\mathrm{ic}}$ of the PCWE (3.5) has been neglected so far because of the partial time derivative $\frac{\partial p^{\mathrm{ic}}}{\partial t}$ is dominating due to relatively low flow and grid velocities. The velocity at the inlet of the glottis is $U_\infty \approx 0.8\,\frac{\mathrm{m}}{\mathrm{s}}$ and the maximum flow velocity is $u_{\max} \approx 110\,\frac{\mathrm{m}}{\mathrm{s}}$. Furthermore, the grid velocity v_{rg} is zero because of neglecting the vocal fold motion in the acoustic model. In order to investigate its impact, the convective part is considered in the source term computation and then the spectrum of the acoustic simulation is compared with the spectrum obtained with the simplified source term. Therefore, the workflow for source term computation has to be adapted as indicated in Fig. 4.13. In addition to the incompressible pressure, the convective term also has to be provided by the CFD simulation[3] and is added to the simplified source term, which is the partial time derivative of the incompressible pressure $\frac{\partial p^{\mathrm{ic}}}{\partial t}$. As with the incompressible pressure, the convective term also needs to be filtered after the conservative interpolation due to the procedure used in the flow simulation for grid motion.

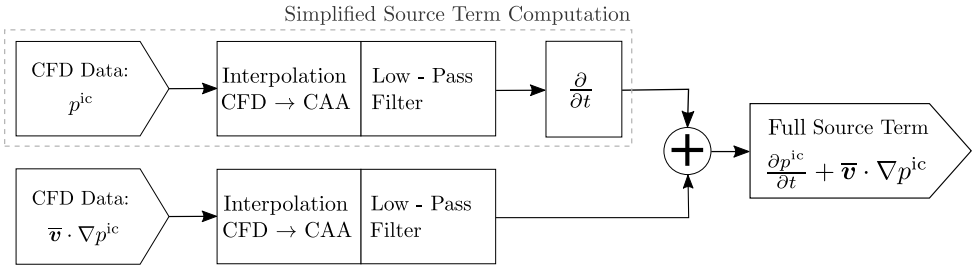


Figure 4.13: Adapted procedure for the computation of the full source term.

However, for this investigation, CFD-data (including the convective term) of only four oscillation cycles was available. Thus, the peaks of the resulting spectrum (Fig. 4.14) are rather wide as explained above. Despite the wide peaks, the negligible influence of the convective term can be seen ($\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.023$). However, for this first attempt of computation of the convective term, the flow velocity was taken instead of utilizing the appropriate quantity, which is the velocity of the mean flow, being aware that this is a substantial difference. Therefore, this aspect has to be investigated more extensively in the future and a method of computing the mean flow velocity $\overline{\boldsymbol{v}}$ needs to be developed. The mathematical background can be found in [41]. Nevertheless, the worflow for the consideration of the convective term within the source term computation was established with the present investigation.

Adaption to Experimental Setup

Figure 4.15 was obtained as part of the validation carried out in an earlier stage of the project and shows the ASD of the simulation and the experiment. Additionally, a tolerance band is illustrated that is generated from ASDs computed from chunks with the same duration as the simulation (0.135 s) obtained by dividing the total measurement signal of 60 s. By doing so, ASDs of 444 chunks were generated and their minimum and maximum values represent the lower and upper boundary of the tolerance band. The overall characteristic of the spectra of the simulation and measurement agree well, but especially in the lower frequency range, the simulation underestimates the amplitude.

[3] Another option is to export and conservatively interpolate the flow velocity (as with the incompressible pressure) and compute the convective term on the acoustic grid which, however, is less accurate since the spatial derivative has to be computed on the coarser acoustic grid.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

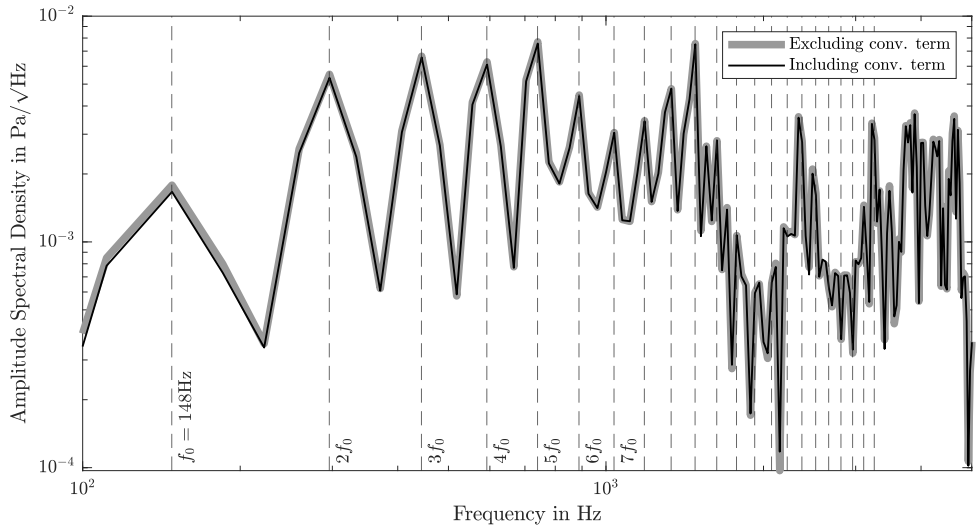


Figure 4.14: Resulting ASD at microphone point for source term with and without consideration of the convective term. $\mathrm{Err_{rel}}^{\mathrm{L_2}} = 0.023$.

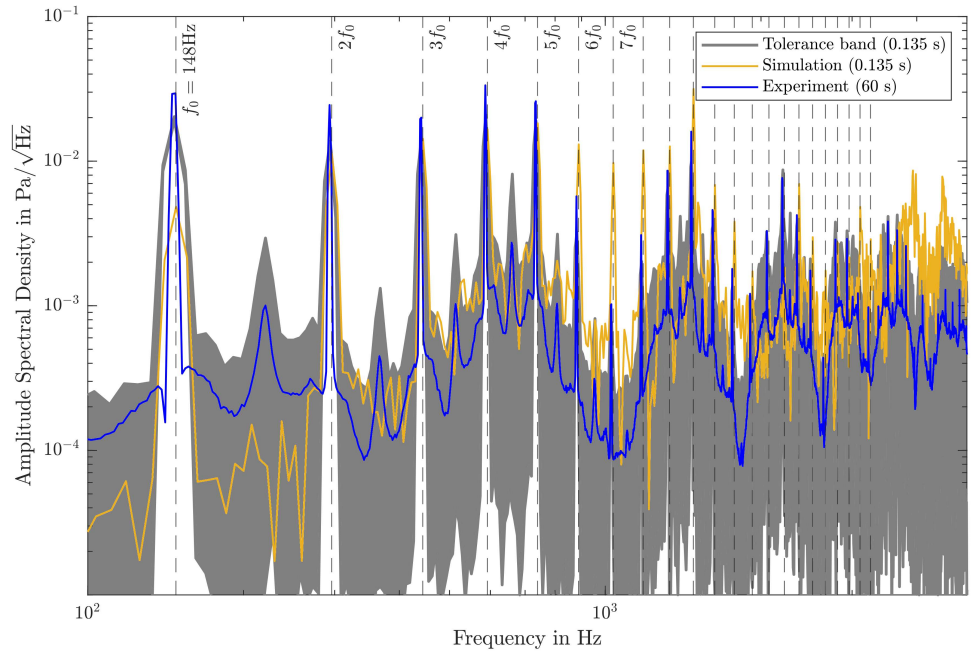
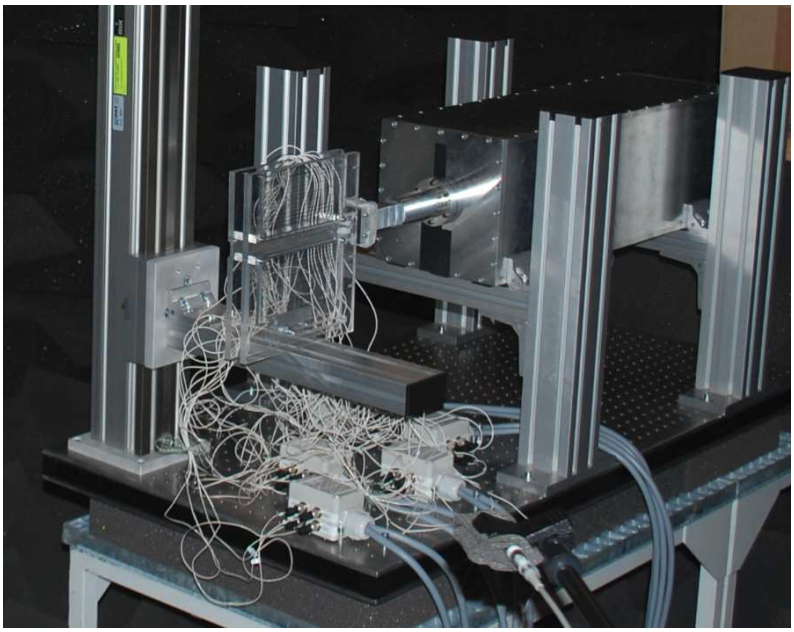


Figure 4.15: Comparison of ASD for experiment and simulation.

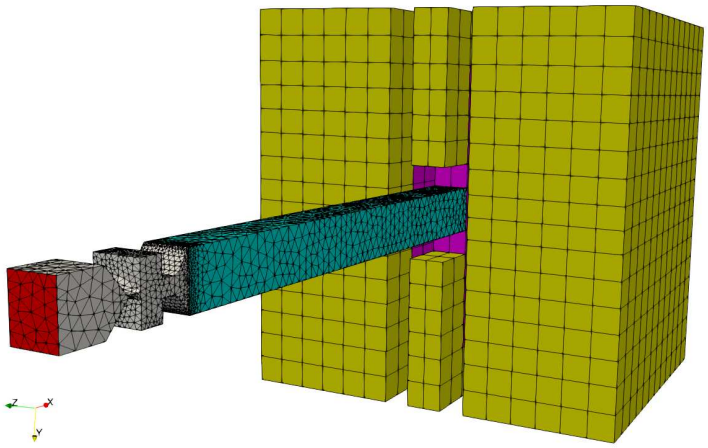
A reason for the deviations could be obstacles within the propagation domain in the experiment that were not considered in the numerical model so far. One of these obstacles that are most likely to have an impact on the sound propagation are the walls enclosing the vocal tract (see Fig. 4.16a) as they are in the immediate vicinity of the outlet. The top and bottom walls are aluminum plates, whereas the side walls are plastic plates which do not terminate at the vocal tract outlet, but extend about 20 mm into the propagation domain. Furthermore, they also exceed the contour of the vocal tract vertically by about 100 mm.

In order to assess their the impact on sound propagation, the enclosing walls are considered as sound-

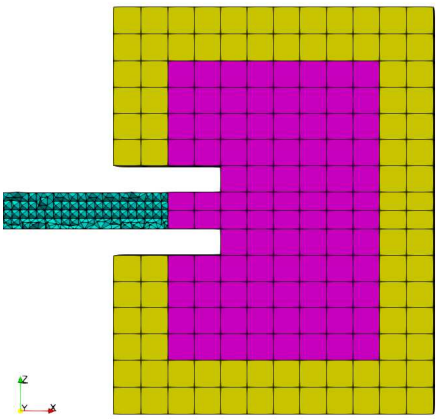
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Experimental setup for fluid dynamic and acoustic measurements.



(b) 3D view of CAA setup.



(c) Horizontal section of CAA setup.

Figure 4.16: Experimental and corresponding numerical setup of validation configuration.

hard walls for a harmonic acoustic simulation (10 Hz to 5 kHz) with a point source of magnitude 1 located on the center line of the larynx between the false vocal folds. The simulation setup is shown in Fig. 4.16. Therewith, the transmission behavior to the evaluation points can be obtained by simply evaluating the acoustic pressure with respect to the frequency at the microphone point which represents a transfer function. This transfer function is illustrated in Fig. 4.17 for the original and the adapted model. The transfer functions do not differ significantly in the lower frequency range indicating that the walls have only a minor influence, and thus the consideration would not lead to a better agreement of the acoustic spectra of experiment and simulation.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

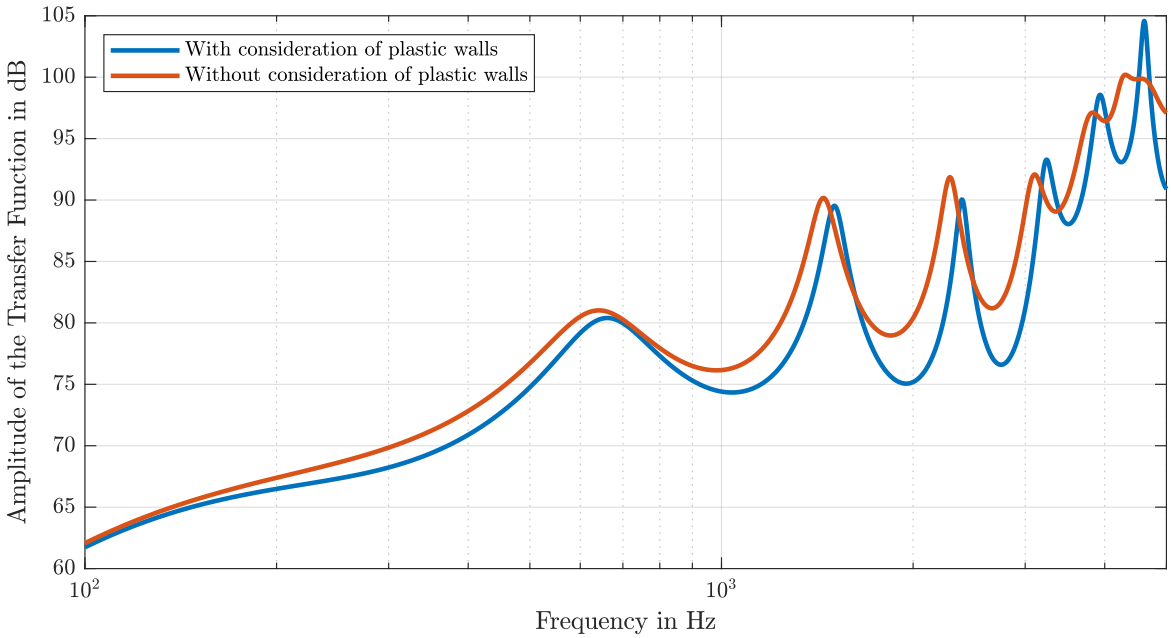


Figure 4.17: Transfer function for point source inside the glottis to microphone position 2.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Chapter 5

Application of the Numerical Model

In this chapter, the introduced numerical model (model of validation experiment with synthetic VFs) is applied for investigations of various aspects of real human phonation. For doing so, the above introduced model needs to be adapted towards real phonation, which is described in Sec. 5.1. Then, in Section 5.2, a case study of the vocal fold motion is presented that aims to asses the impact of certain anomalies of the glottis (e.g. asymmetric oscillation of VFs, glottal gap). The final part (Sec. 5.3) is another case study with the purpose of the validation of the physiological correctness of the numerical model by utilizing real data of patients and is part of WP3 of the project *Numerical computation of the human voice source*. Besides evaluating the predicted voice, the acoustic source terms are analyzed to gain a better understanding of how certain glottal malfunctions affect the preliminary voice signal (before being filtered by the vocal tract) and voice production in general.

5.1 Adaption to Real Phonation

The above introduced numerical model is representing the experimental setup with synthetic vocal folds with the purpose of validation and thus, does not represent real phonation. Therefore, the existing numerical model needs to be adapted to be able to investigate aspects of real phonation. For the validation model, the bottom of the propagation region was modeled sound-hard corresponding to the experimental setup, whereas for real phonation we assume the sound emitted from the mouth is exclusively reflected by the face. Thus, only the side of the propagation domain with the interface to the vocal tract (representing the human face) is modeled fully reflective, whereas the other sides have adjacent PML regions (see Fig. 5.1 and Fig. 5.2b for corresponding 3D view) acounting for free field radiation. Thereby, the dimensions of the propagation domain were kept equal. Furthermore, a realistic vocal tract representing the vowel /*a*/ is considered instead of the simple rectangular duct of the validation model. In this context, it has to be noted, that this vocal tract is still a simplification and does not represent real phonation in terms of its shape but exhibits the acoustic characteristics of a real vocal tract within the considered frequency range ($100\,\mathrm{Hz} < f < 5000\,\mathrm{Hz}$). The geometry is rotationally symmetric and is derived from a real geometry of a vocal tract by preserving the area of the cross sections and straightening the midline according to [38] (see Sec. 2.2). The conform mesh of the larynx and vocal tract is created separately for each configuration of the case studies as they differ slightly in the shape of the maximum VF opening (which is used for the CAA grid) due to a differing prescribed VF motion.

The details of the mesh of the adapted CAA setup shown in Tab. 5.1 are similar to the setup of the validation model (see Tab. 4.1). Only the spatial resolution in the rear part of the vocal tract (where only minor source terms occur) is reduced based on the experience of previous investigations which revealed

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

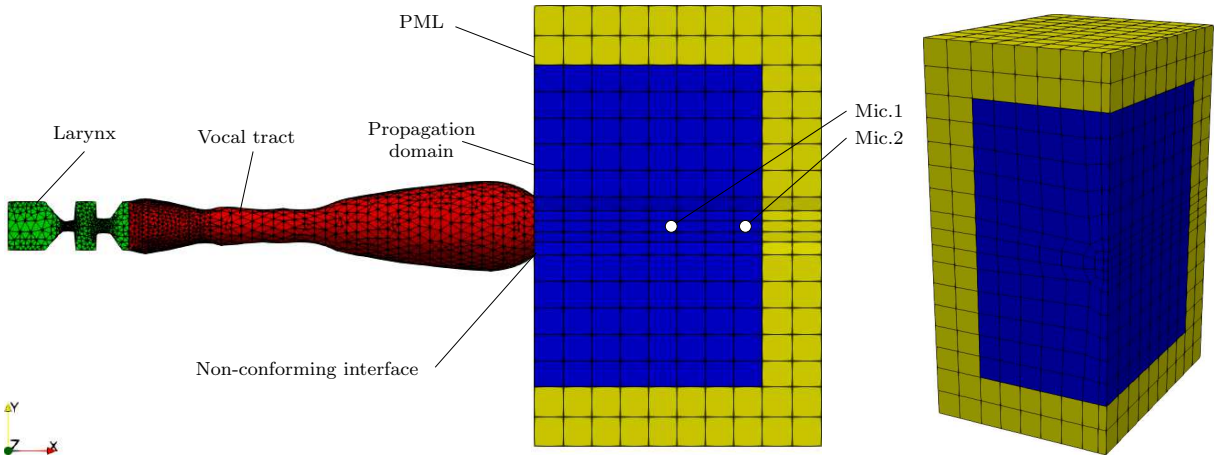


Figure 5.1: The adapted CAA setup for real phonation with a detail of the mesh of the PML and propagation domain depicted in a clipped view.

Table 5.1: Details of the mesh applied for real human phonation.

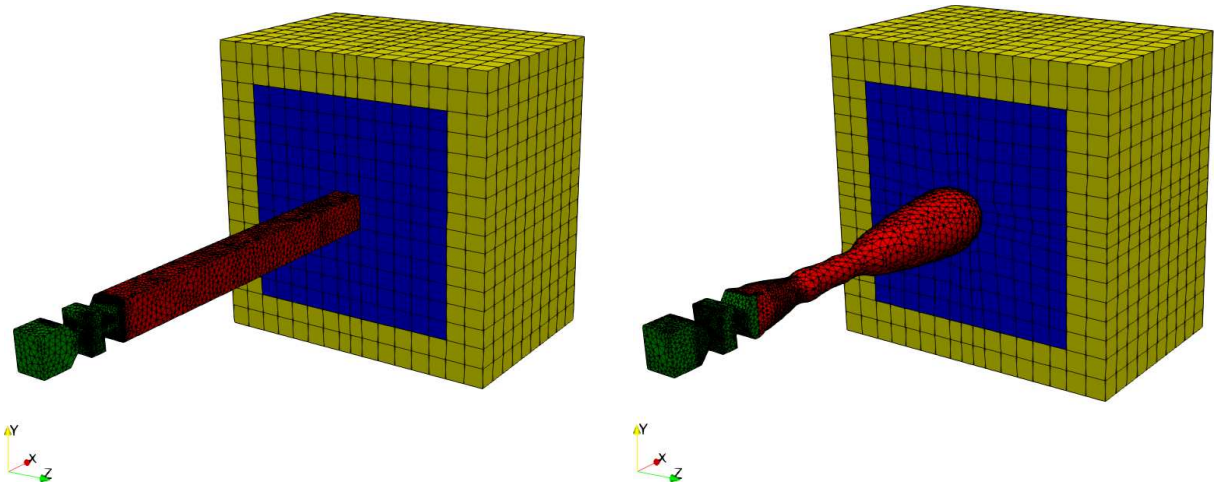
Region	No. elements	Max. element size
Larynx	48 750	5.0 mm
Vocal tract	23 530	6.3 mm
Propagation domain	1 450	10.6 mm
PML region	1 560	11.0 mm
Total	93 200	11.0 mm

that fine resolution is only necessary within the region of the major acoustic sources (see Fig. 4.9). The mesh of the larynx and the vocal tract is again generated in one step to ensure a conform mesh within the source region and tetrahedral elements are used to resolve the three-dimensionally curved geometry. Due to the circular (non-conforming) interface between the vocal tract and the propagation domain the so called *O-grid* in *Ansys ICEM CFD* is applied for the propagation domain (see detail in Fig.5.1). The geometry for the larynx at the maximum opening of each configuration of the case study is derived from the transferred CFD data (mesh) and the geometry of the realistic vocal tract is also provided by Universitätsklinikum Erlangen, but is equal for all cases.

In order to compare the transmission behavior of the simple and the realistic vocal tract independently from the VF vibration, the transfer functions for a point source between the VFs were computed by performing harmonic simulations (as described in Sec. 4.3.2) for both vocal tracts within the adapted setup (equal BCs, see Fig. 5.2). The resulting transfer functions illustrated in Fig. 5.3 reveal significant differences in the characteristics of the vocal tract. Especially, the first formant deviates clearly in frequency and amplitude. The first formant frequency of the simplified and real vocal tract is 0.7 kHz and 1.015 kHz, respectively and the amplitude of the first formant of the realistic vocal tract is about 10 dB larger.

In addition to geometry and BCs, the position of the observation points for the evaluation of the resulting voice is adapted to match the microphone positions of the voice recordings conducted on real patients. The adapted microphone positions of (*Mic.1* and *Mic.2*) are located on the center line 50 mm and 80 mm from the mouth (end of the vocal tract) as indicated in Fig. 5.1. Since the propagation domain captures both microphone points, no transformation of the pressure signal is necessary.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Simple vocal tract (validation case). (b) Realistic vocal tract for vowel /a/ (real phonation).

Figure 5.2: Setups for the comparison of the transmission behavior of the vocal tracts.

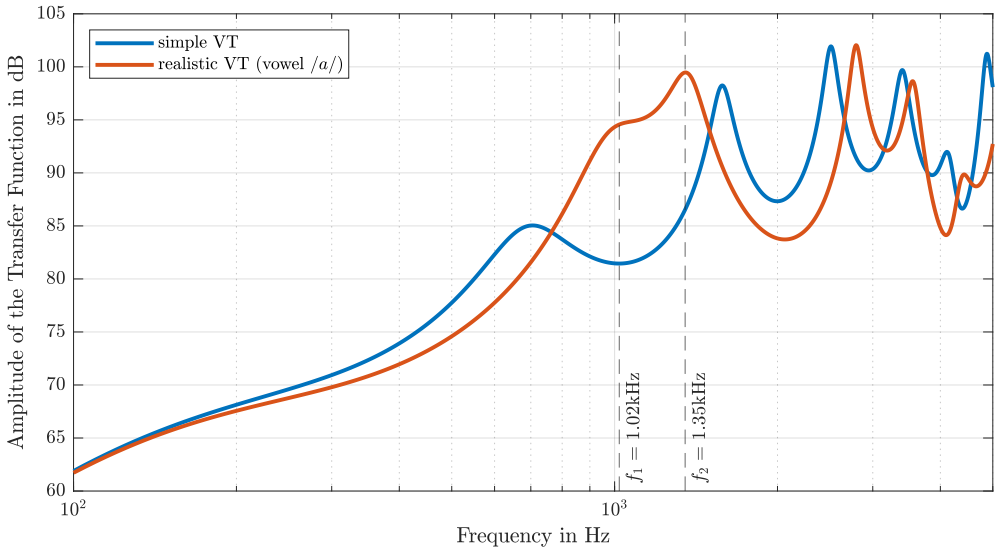


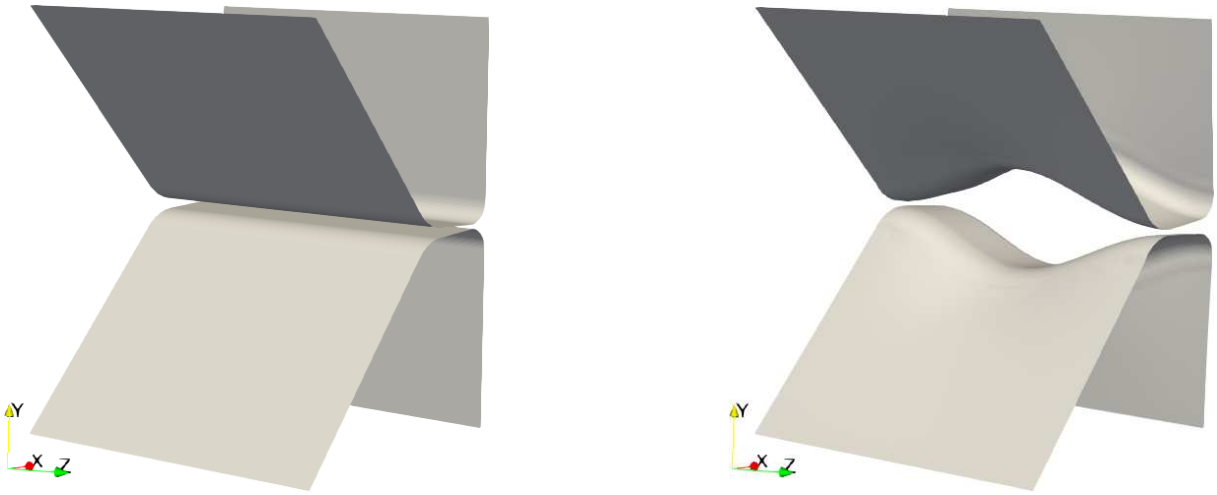
Figure 5.3: Comparison of the transfer characteristics of the vocal tracts with indicated first two formant frequencies of for the realistic vocal tract.

5.2 Case Study 1: Primary Voice Source

The purpose of the herein presented case study is to evaluate how certain anomalies of VF vibration affect voice production. For doing so, the prescribed VF motion of a healthy voice (reference) is varied in the CFD simulation in terms of symmetry and glottal closure of the VF vibration. Then, the workflow described in Sec. 4.1 consisting of conservative interpolation from CFD to CAA grid, low-pass filtering, source term computation, and the actual acoustic simulation is carried out. The shape of the closed (min. amplitude) and opened (max. amplitude) condition of the VFs is illustrated in Fig. 5.4. As can be seen on Fig. 5.4a, the VFs do not contact and thus, the glottis is not fully closed at minimum oscillation amplitude due to numerical issues of the CFD simulation, which causes a negligible leakage flow as described above. Nevertheless, in the following the condition of minimum oscillation amplitude is denoted as the closed condition.

The fundamental frequency of the oscillation is $f_0 = 148\,\mathrm{Hz}$ as for the validation case and one oscilla-

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Closed condition (min. amplitude). (b) Opened condition (max. amplitude).

Figure 5.4: Geometry of the vocal folds in closed and opened condition (reference configuration).

tion cycle is resolved by 495 steps resulting in a time step size of $\Delta t = 13.6\,\text{ms}$ for the acoustic simulation. Ten oscillation cycles are considered in the simulations which are sufficient for a relative comparison of the investigated configurations.

5.2.1 Configurations

Table 5.2 provides an overview of the ten configurations that are evaluated in the case study. Therein, four different glottis configurations (GC), with the respective initial (closed) condition, exist, namely GC1, GC2, GC3, and GC4. These glottis configurations initially have a diverging shape in spanwise direction (z direction) of 0, 40, 70, and 100%, respectively of the VF length as illustrated in Fig. 5.5 (first column). For GC1, additionally, a closed condition with a constant initial opening is introduced, namely the half open configuration (case 02 and 07). Furthermore, for each of the five initial conditions a symmetric and asymmetric motion of the vocal folds is considered resulting in ten configurations that are considered in the case study. In the asymmetric case, the upper vocal fold oscillates with half of the amplitude of the lower vocal fold. The opened conditions (max. amplitude), for the symmetric and asymmetric case, are illustrated in the second and third column of Fig. 5.5 and belong to the initial conditions of the first column.

Table 5.2: Overview of case study 1 (C...constant, D...diverging initial opening shape).

Case	Name	GC	Initial opening	VF motion
01	01_SymClosed_GC1	GC1	0% C	sym
02	02_SymHalfopen_GC1	GC1	100% C	sym
03	03_Sym_GC2	GC2	40% D	sym
04	04_Sym_GC3	GC3	70% D	sym
05	05_Sym_GC4	GC4	100% D	sym
06	06_Asym_Closed_GC1	GC1	0% C	asym
07	07_Asym_Halfopen_GC1	GC1	100% C	asym
08	08_Asym_GC2	GC2	40% D	asym
09	09_Asym_GC3	GC3	70% D	asym
10	10_Asym_GC4	GC4	100% D	asym

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

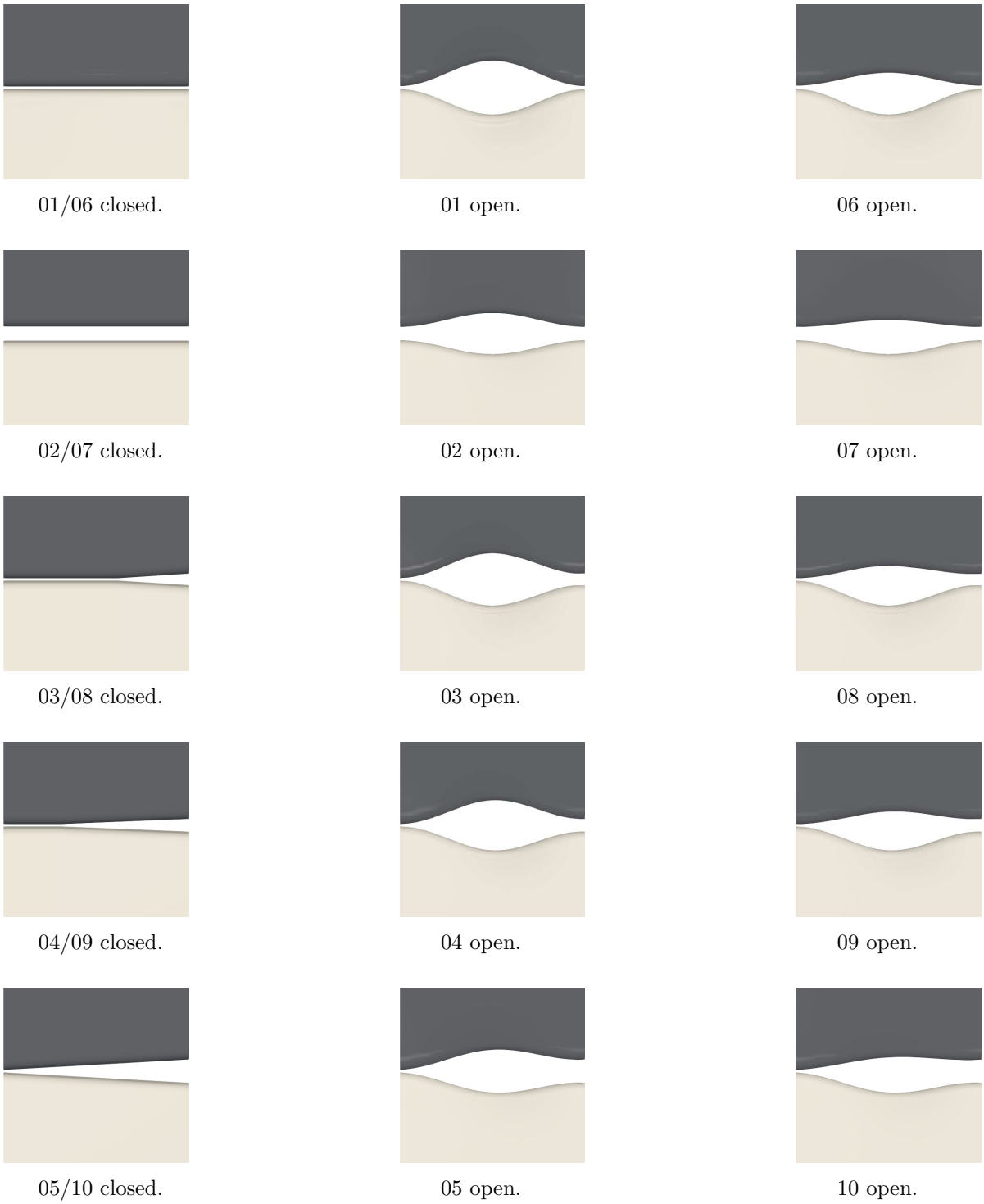


Figure 5.5: Glottal orifices of the case study in closed (first column) with corresponding opened conditions for symmetric (second column) and asymmetric (third column) VF motion illustrated in front view (y/z layer).

5.2.2 Analysis of Acoustic Source Terms

To get a better understanding of the sound generation mechanisms and the impact of certain glottal anomalies on the preliminary acoustic signal, the source terms are analyzed in the frequency domain. For proper visualization of the acoustic sources inside the source domain (larynx and vocal tract), a much

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

higher spatial resolution as for the actual acoustic simulation is necessary[1]. Therefore, the incompressible pressure from the CFD simulation is interpolated to two orthogonal meshes, consisting of only one layer of hexahedral elements with a uniform cell size of 0.2 mm × 0.2 mm and a thickness of the element layer of 0.1 mm, located on the symmetry planes of the model. Therewith, the data is reduced to the regions of interest (planes of symmetry) enabling efficient handling of the highly resolved data. Moreover, the cut-cell algorithm, which considers the intersection volumes, ensures that the pressure obtained on the two reduced meshes corresponds to the pressure which would be obtained from a mesh that captures the entire domain. Afterword, the incompressible pressure is low-pass filtered and subsequently, the source terms are computed following the usual workflow for source term computation. Thereby, it has to be noted that the scalar $-\frac{1}{\bar{\rho}c^2}$ of (3.5) is not included as it is applied just directly within the source term definition of the acoustic simulation. Since the analysis of the source terms is purely qualitative[2], the neglection of this scalar is justified. Having obtained the finely resolved source term, a field fast Fourier transformation (field FFT) is performed in *MATLAB*©, which is an FFT of the time signal of every nodal load on the computational grid. For the following plots, the real part of the Fourier transformed nodal loads is chosen to be visualized.

In Fig. 5.6, the result of the field FFT of the region with the principle acoustic source terms is depicted corresponding to the 30$^{\text{th}}$ harmonic of the VF oscillation. Furthermore, Fig. 5.7 provides the results of the x/y-symmetry plane for several selected additional frequencies. Therein, the amplitudes of five harmonic as well as one non-harmonic frequency, which serves as a random representative of the broadband noise spectrum, are visualized. For other frequencies of the broadband spectrum, which can not all be presented in this thesis, the results are comparable, concerning distribution and amplitude. For the present thesis, the field FFT on the fine mesh is carried out only for the reference configuration (*01_SymClosed_GC1*) but the analysis will be expanded soon to cover all configurations of the case study.

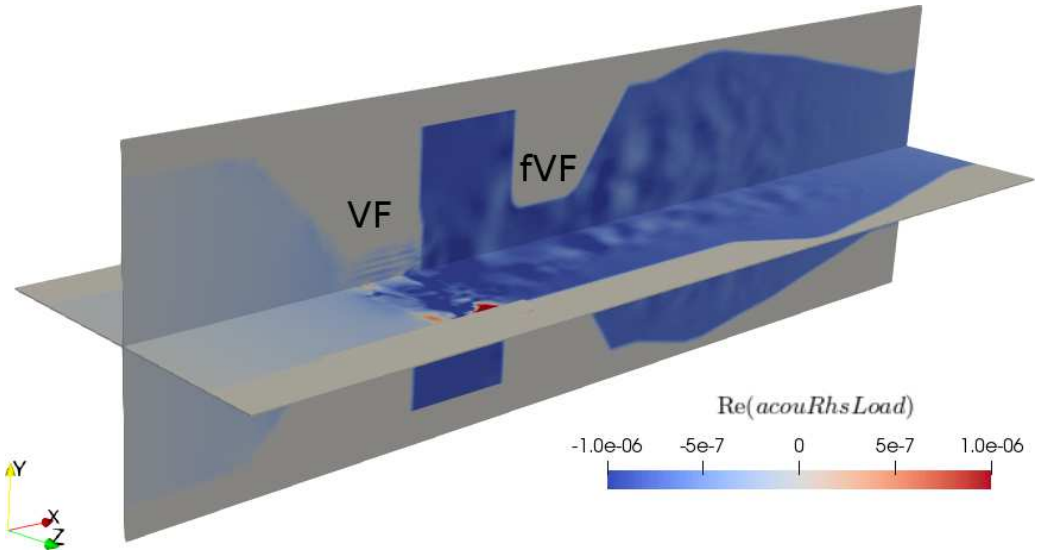


Figure 5.6: Visualization of the real part of the Fourier transformed acoustic source term (in $\frac{\text{Pa m}^3}{\text{s}}$) in the planes of symmetry for the 30$^{\text{th}}$ harmonic ($f = 30 f_o$).

Nevertheless, for assessment of the source terms of all configurations of the case study, the source terms are compared globally. For the global assessment, an averaged field FFT of the acoustic source

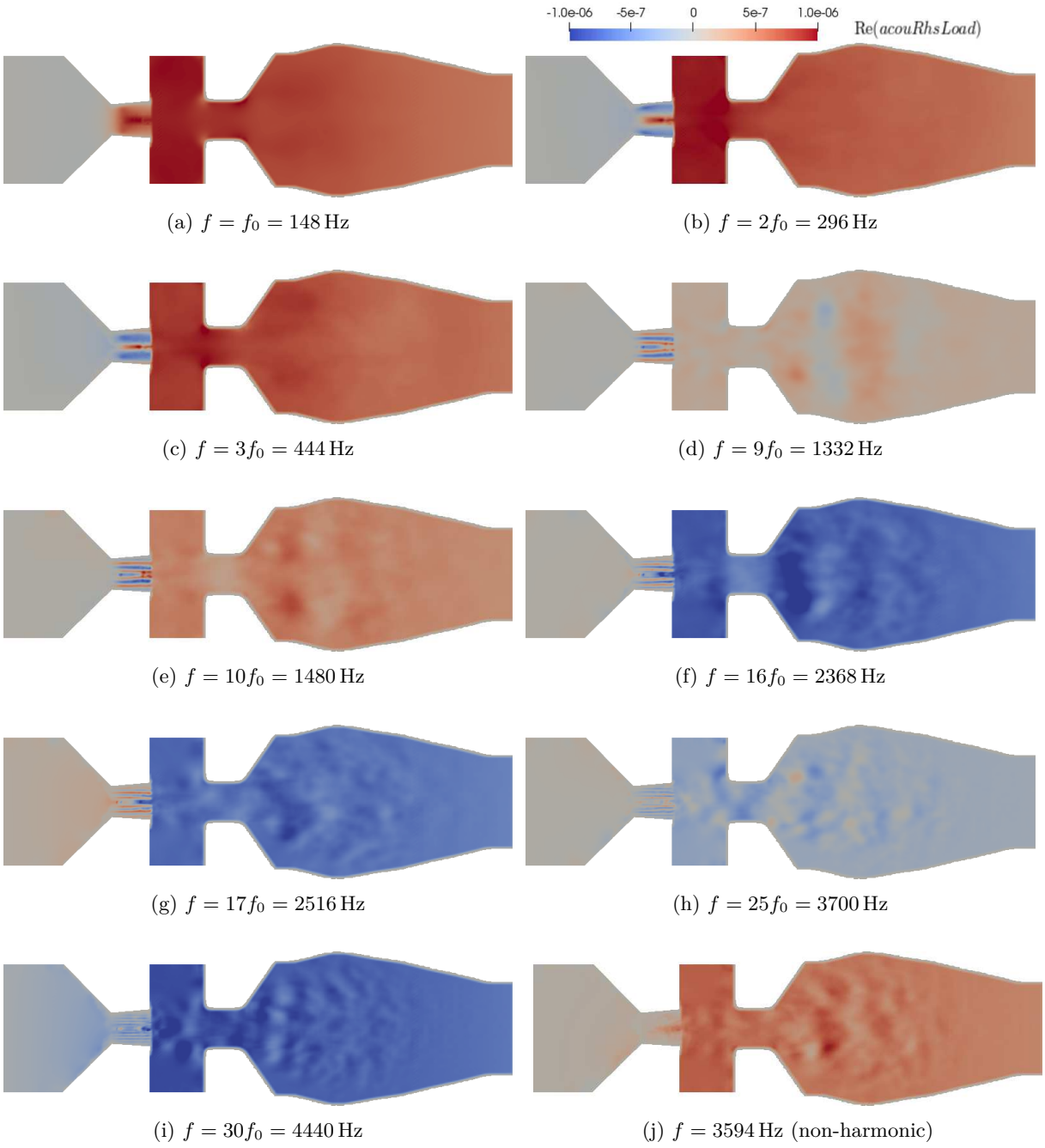
[1] Higher spatial resolution inside the source domain does only benefit in a minor improvement of accuracy of the acoustic results as shown in the convergence study in [54].

[2] The purpose of the source term analysis is on the one hand to find regions with the largest occurring source terms, and on the other hand to compare the source terms of the different configurations of the case study.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

term density f^{a} of the acoustic simulation (coarser spatial resolution) is performed, which computes the average amplitude $\overline{f}^{\mathrm{a}}(f) = 1/N \sum_{i=1}^{N} f_{\mathrm{i}}^{\mathrm{a}}(f)$ of all cells inside the source domain for each frequency f resolved by the field FFT with N being the number of cells. This is done by interpolating the source terms from nodes to cells (with *CFSdat*) and then dividing by the respective cell volume. In the next step, a field FFT is performed and for each frequency the average amplitude of the source density over all cells of the source domain is computed. Due to the mesh refinement in the glottal region, the influence on the average amplitude of these cells is stronger because of the higher number. But since the refinement is in the region of the major source terms this results in a more distinctive spectrum. Moreover, the refinement is equal for every configuration of the case study allowing an objective comparison. The resulting spectra of the averaged source density of the ten configurations are depicted in Fig. 5.8. Therein, the non-harmonic contribution to the source term is depicted by simple thin lines, whereas the harmonics of the VF vibration are indicated by dots. As mentioned above, the displayed quantity does not include the scalar $-\frac{1}{\bar{\rho}c^2}$ of (3.5).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) $f = f_0 = 148\,\text{Hz}$ (b) $f = 2f_0 = 296\,\text{Hz}$

(c) $f = 3f_0 = 444\,\text{Hz}$ (d) $f = 9f_0 = 1332\,\text{Hz}$

(e) $f = 10f_0 = 1480\,\text{Hz}$ (f) $f = 16f_0 = 2368\,\text{Hz}$

(g) $f = 17f_0 = 2516\,\text{Hz}$ (h) $f = 25f_0 = 3700\,\text{Hz}$

(i) $f = 30f_0 = 4440\,\text{Hz}$ (j) $f = 3594\,\text{Hz}$ (non-harmonic)

Figure 5.7: Distribution of the real part of the Fourier transformed acoustic source term (in $\frac{\text{Pa}\,\text{m}^3}{\text{s}}$) in x/y-symmetry plane for selected frequencies.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

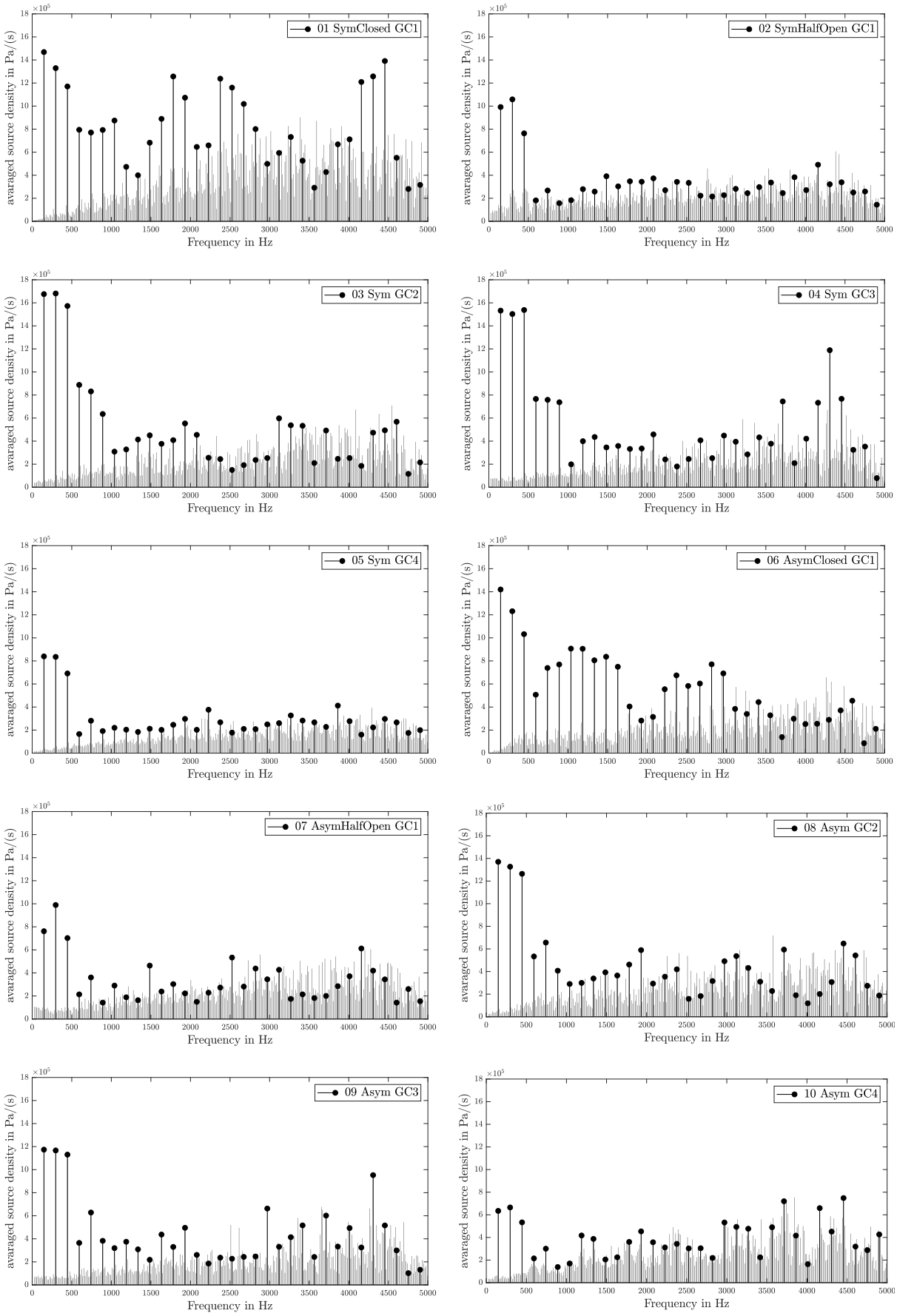


Figure 5.8: Averaged field FFT of the acoustic source density for the source region of case study 1.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

5.2.3 Acoustic Spectra

The spectra of the voice predicted by the numerical model of real phonation are depicted in Fig. 5.9 for each of the ten configurations of the case study.

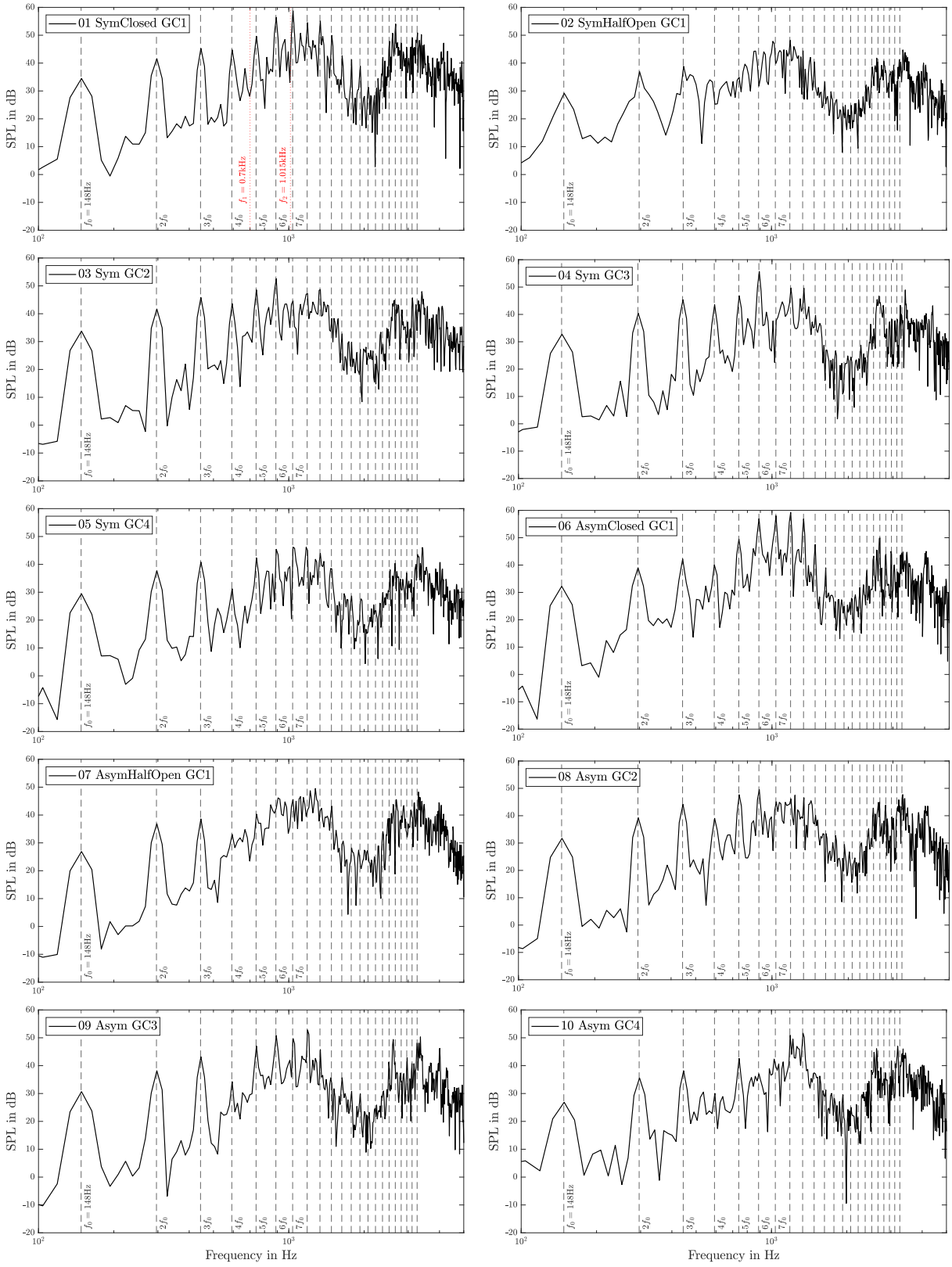


Figure 5.9: Spectral results of case study 1 at *Mic.2*.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

For the case 01 (reference case), the first two formant frequencies of the vocal tract are indicated. Additionally, the above introduced error $\mathrm{Err}_{\mathrm{rel}}{}^{\mathrm{L}_2}$ computed by (4.4) is provided in Tab. 5.3. However, herein the error is not used as a measure of inaccuracy of the simulation, but as a measure of deviation from the spectrum of the healthy voice (reference) in order to quantify the deviation with one value.

Table 5.3: Deviation from the reference voice by means of the error $\mathrm{Err}_{\mathrm{rel}}{}^{\mathrm{L}_2}$ (C...constant, D...diverging initial opening shape).

Case	01	02	03	04	05	06	07	08	09	10
Opening	0%	100%	40%	70%	100%	0%	100%	40%	70%	100%
Opening shape	C	C	D	D	D	C	C	D	D	D
Deviation	-	73%	61%	63%	72%	52%	67%	64%	63%	70%

5.2.4 Discussion of Results

The plots of the field FFT in Fig. 5.7 show that the distribution of the acoustic source term is relatively diffuse at low frequencies, whereas at higher frequencies more pronounced structures occur. A reason for the diffuse distribution could be the neglection of the convective term since its relative contribution to the total source term is greater at lower frequencies as the other term $\frac{\partial p^{\mathrm{ic}}}{\partial t}$ is smaller. The smaller total source term due to the neglection could also be an explanation for the underestimation of the first few harmonics of the sound simulated by the validation model (see Fig. 4.15). This assumption is supported by [56] since the authors also applied acoustic perturbation equations but considered the convective part within the acoustic source term and yielded more distinct structures of the source terms at lower frequencies. Moreover, they did not underestimate the amplitudes of the first harmonics of the predicted voice with their approach. Therefore, this aspect will be investigated in the near future. Interestingly, also the horizontal stripes between the VFs, which indicate their vertical motion, are not present in Fig. 5.7a, even though it corresponds to the first harmonic of the VF vibration. The plot in Fig. 5.7j doesn't contain the horizontal stripes either, but this is because it corresponds to a non-harmonic frequency. Due to the independence of the base frequency of the VF oscillation, the cause of these acoustic sources is mainly a turbulent phenomenon occuring predominantly in the supraglottal region due to vortex shedding as stated in [56]. In this context, it has to be mentioned that the applied low-pass filter also has an influence on the resulting acoustic sources for the higher frequency range near the cut-off frequency of 5 kHz. In general, the source terms with the largest magnitude occur in the region between the VFs at most of the frequency. However, at some frequencies, strong source terms also appear in the region between the VFs and fVFs (see Fig. 5.7i), at the impinging edge of the false VFs (see Fig. 5.7b) or shortly after the fVFs (see Fig. 5.7f).

Generally, the plots of the averaged field FFTs (see Fig. 5.8) show that in the lower frequency range the harmonic contribution to the source term is dominating, whereas at higher frequencies the appearance of broadband noise rises. Consequently, also the resulting acoustic spectra exhibit this characteristic. However, the vocal tract amplifies the frequencies near its formant frequencies $f_1 = 0.7\,\mathrm{kHz}$ and $f_2 = 1.015\,\mathrm{kHz}$ (indicated in Fig. 5.9 for the reference configuration), which is why the amplitudes of these frequencies become greater than the ones of the lower harmonics (e.g. first three harmonics), which are dominating in the averaged field FFT for most of the configurations. The overall intensities of the source terms in Fig. 5.7 correspond to the amplitudes of the spectra in Fig. 5.8 very well, which demonstrates the purpose of the averaged field FFT being the indication of the content of energy of acoustic source terms related to a certain frequency within the entire source domain.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Initial glottal opening obviously causes reduced harmonics of the source terms leading to a smaller ratio of the amplitude of the harmonics and broadband noise. Consequently, the acoustic spectrum is less developed with lower and broader peaks of the harmonics. In this context, the half open, non-diverging closed condition seems to produce the worst primary voice signal, which is feasible since it has the largest orifice in terms of the cross-sectional area. The error with respect to the reference case (normal VF motion) in Tab. 5.3 verifies the tendency of decreasing voice quality with an increasing degree of initial opening. Regarding voice quality, case 2 yields the worst results as the peaks of the harmonics are poorly developed compared to the other configurations. The reason can be found in the averaged FFT of the source terms where the amplitudes of the harmonics are smaller compared to the other configurations and also compared to the non-harmonic amplitudes (noise). Interestingly, the spectrum of configuration 7, which is the asymmetric version, reveals that the asymmetric motion actually improves the voice in this case. For the other cases, the asymmetric motion has a smaller impact on the source terms as well as on the acoustic spectrum. Especially, the averaged field FFT provides very similar characteristics for the symmetric and asymmetric cases. For example, the averaged spectra of the source terms of cases 3 and 8 and cases 4 and 9 match quite well.

In conclusion, the case study revealed, that a decreased degree of glottal closure causes worse voice quality. Furthermore, insufficient closure stronger affects the voice than asymmetric vibration. Moreover, asymmetric vibration can even compensate for glottal insufficiency as it restricts the glottal airflow more strongly causing a higher pressure fluctuation due to the periodic interruption of the flow and thus, leading to larger source terms. However, in this context, it has to be noted that the vocal fold motion in this case study is only asymmetric in terms of the amplitude of the oscillation. According to several studies, a generally asymmetric oscillation e.g., the difference in frequency, or the modal shape have a more severe impact on phonation [57–59].

5.3 Case Study 2: Validation of the Physiological Correctness

Having validated and adapted the numerical model to real human phonation, it is now applied to clinical data of patients with the goal to evaluate the physiologic correctness of the model, which is part of WP3 of the project *Numerical computation of the human voice source*. For doing so, the findings of the previous working packages are transferred to real laryngeal dynamics and acoustics being available from clinical in-vivo [60–63] and ex-vivo [64, 65] studies to investigate the qualitative and quantitative accordance of the simplified vocal fold shape and motion in combination without (in-vivo) and with (ex-vivo) the realistic vocal tract. As with the vocal tract, the laryngeal geometry and the VF motion are derived from real humans (measurements of patients) and simplified while preserving the aeroacoustic properties for the present range of application in this thesis. The clinical data used for the physiological validation contains female and male healthy voices (normal vocal fold vibrations) as well as hoarse voices due to common voice disorders, such as glottal gap, asymmetric or aperiodic VF vibrations.

It has to be noted, that the main part of the clinical validation of the acoustic results is carried out by the cooperation partner Universitätsklinikum Erlangen (Department for Phoniatrics and Pedaudiology) as they have the tools, experience, and clinical data required for the evaluation of the simulated voice.

5.3.1 Configurations

For the validation of the physiological correctness of the numerical model, in total eight data sets with different features of vocal fold characteristics are taken into account, which are divided into in-vivo (with a vocal tract) and ex-vivo (without a vocal tract) cases as follows:

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

- Ex-vivo vocal fold vibrations (without vocal tract)
  - ExVivo1: healthy female voice
  - ExVivo2: glottal gap, symmetric vibrations (hoarse female voice)
  - ExVivo3: glottal gap, asymmetric vibrations (hoarse female voice)
- In-vivo vocal fold vibrations (with vocal tract)
  - InVivo1: healthy female voice
  - InVivo2: healthy male voice
  - InVivo3: glottal gap (hoarse female voice)
  - InVivo4: asymmetric VF vibrations (hoarse female voice)
  - InVivo5: aperiodic VF vibrations (hoarse female voice)

Thereby, the configurations are chosen to cover a variety of potential vocal fold characteristics and are reproduced by the numerical model to be subsequently validated by means of previous experimental studies with real vocal folds. In this context, it is noted that the ex-vivo cases are only mentioned for the sake of completeness as the CFD data sets of these cases were not available until the completion of the thesis and therefore, the cases are not considered in the following.

Asymmetric vibration means that the upper VF oscillates with only half of the amplitude and aperiodic that the VFs oscillate at different frequencies. Details of the temporal resolution of the exported CFD data used for the CAA simulation can be found in Tab. 5.4. Therein, the number of time steps discretizing one oscillation cycle depends on the base frequency of the VF vibration f_0 since it was tried to maintain a time step size of $\Delta t \approx 13.5\,\mu s$. Furthermore, the number of oscillation cycles considered in the simulations differs slightly between 15 and 17.

Table 5.4: Overview of case study 2 and details about the temporal resolution.

Case	Name	VF characteristic	f_0	TS/Cycle	No. cycles	TS size Δt
1	InVivo1	healthy, female	279 Hz	266	16	13.48 µs
2	InVivo2	healthy, male	127 Hz	578	16	13.62 µs
3	InVivo3	glottal gap	276 Hz	266	17	13.62 µs
4	InVivo4	asymmetric vibration	217 Hz	338	16	13.63 µs
5	InVivo5	aperiodic	80/320 Hz	919	15	13.40 µs

5.3.2 Analysis of Acoustic Source Terms

As for the previous case study, the source terms are again assessed by means of an averaged field FFT, which is depicted in Fig. 5.10 for each configuration. For case *InVivo5*, two different harmonics are indicated, as it has an aperiodic VF motion meaning that the two VFs oscillate with a base frequency f_1 and f_2, respectively.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

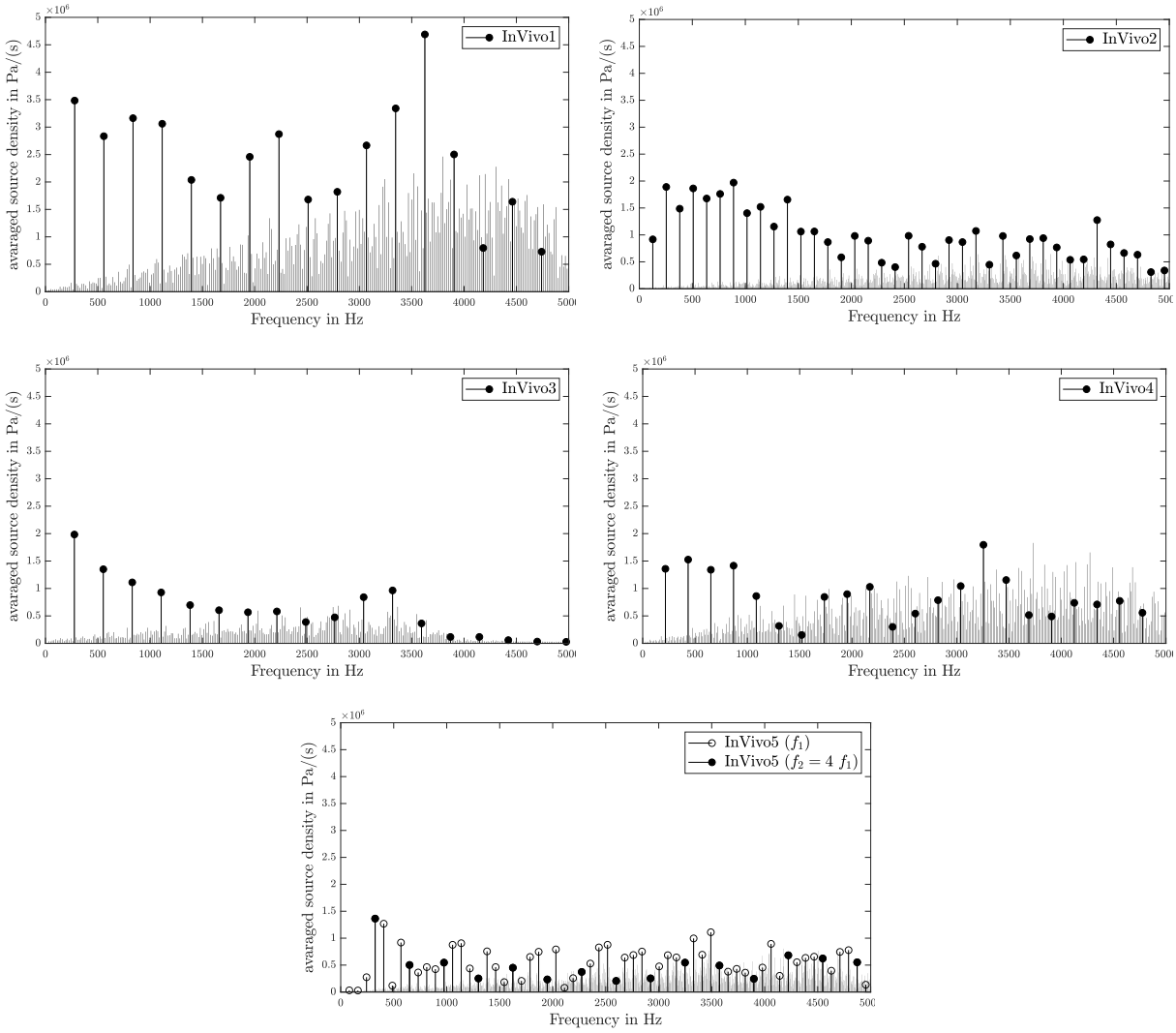


Figure 5.10: Averaged field FFT of the acoustic source density for the source region of case study 2.

5.3.3 Acoustic Spectra

Figure 5.11 displays the acoustic spectra of the five in-vivo cases of the simulation and the measurement obtained from a microphone point located at equal position (*Mic.2* from the simulation). Therein, the harmonics of the varying base frequency f_0 are indicated by dashed lines. For the aperiodic case (*InVivo5*), the harmonics of both frequencies are illustrated.

Again, the above introduced error $\mathrm{Err_{rel}}^{\mathrm{L_2}}$ defined by Eq. (4.4) is stated in Tab. 5.3 as a measure of deviation from the acoustic spectrum of the measurement of each case.

Table 5.5: Deviation from the measured voice by means of the error $\mathrm{Err_{rel}}^{\mathrm{L_2}}$ (f...female, m...male voice).

Case Characteristic	InVivo1 healthy (m)	InVivo2 healthy (f)	InVivo3 glottal gap (f)	InVivo4 asym. (f)	InVivo5 aper. (f)
Error ($100\,\mathrm{Hz} < f < 5\,\mathrm{kHz}$)	351%	307%	268%	282%	1210%
Error ($100\,\mathrm{Hz} < f < 800\,\mathrm{Hz}$)	90%	84%	86%	92%	200%

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

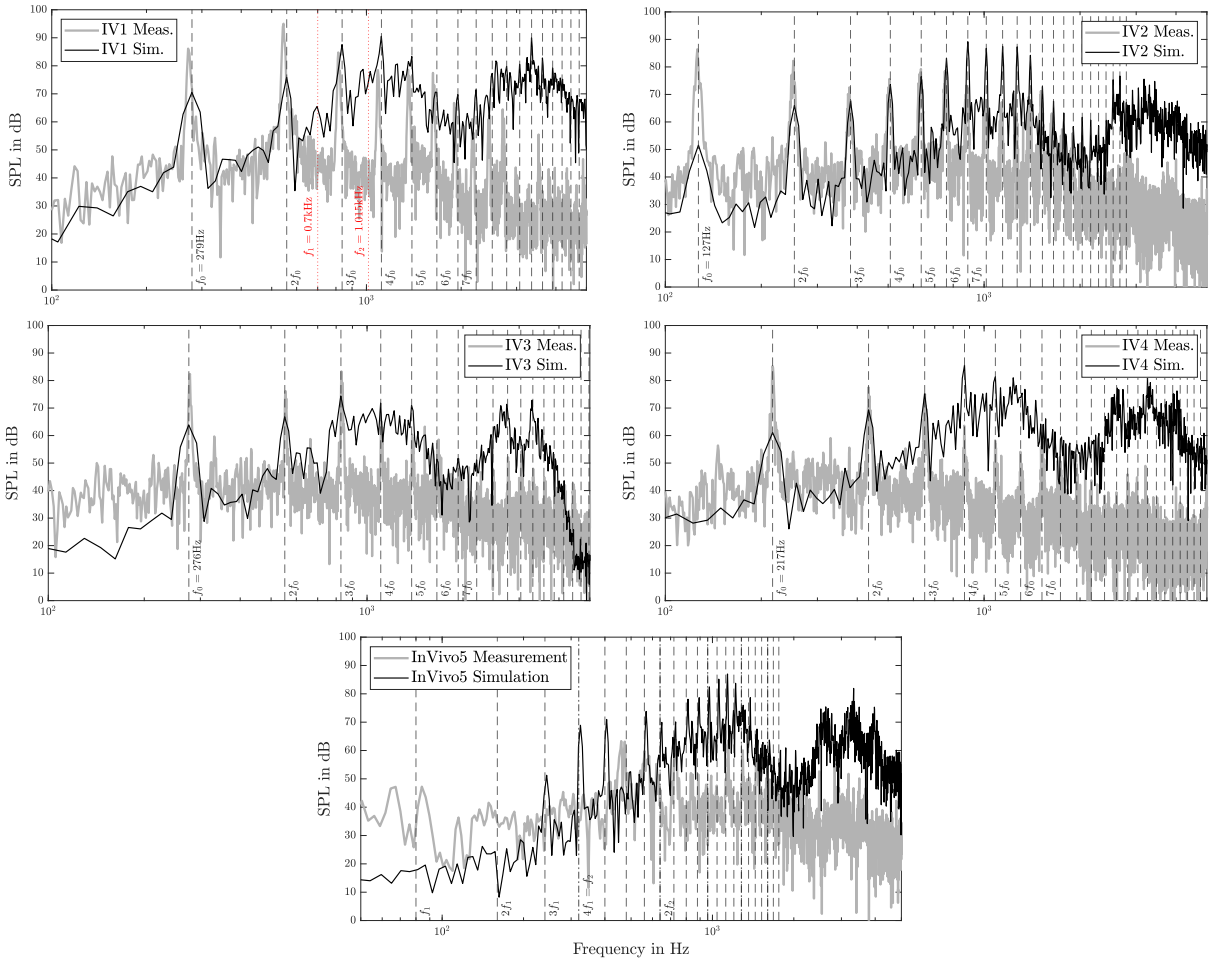


Figure 5.11: Spectral results of case study 2 at *Mic.2*. For the configuration *InVivo1*, the first two formant frequencies of the vocal tract are indicated.

5.3.4 Discussion of Results

The acoustic spectra of the computed voices of the in-vivo cases based on measurements on patients (VF motion) show quantitatively little agreement as indicated by the error in Tab. 5.5, which lies between 2.68 and 12.1. Thereby, the relatively large error is less obvious in the diagrams due to the logarithmic scale. However, the harmonics of the simulated voice are visible and especially in the low frequency range, the spectra partly match the measurements quite well. Especially for the configuration *InVivo2*, the spectrum from the second to the sixth harmonic agrees well with the measurement. As with the previous case study, the amplitude of the first and second harmonic are underestimated for most of the configurations. The neglection of the convective term could be a reason for that, as explained above. As indicated in Tab. 5.5, the overall error is mainly caused by the deviations in the higher frequency range as the errors for the reduced frequency range of $100\,\mathrm{Hz} < f < 800\,\mathrm{Hz}$ are drastically smaller. In the frequency range around the formant frequencies of the vocal tract (see Fig. 5.3), where the acoustic pressure is amplified, the amplitudes differ substantially, which mainly contributes to the large error. Moreover, the overall wavy characteristics of the spectrum in this frequency range can be seen in for every configuration. Consequently, it is assumed that the vocal tract used for the numerical model does not properly represent the vocal tract of the test persons in terms of the transfer characteristics. But since the ex-vivo configurations of the case study do not consider the vocal tract, these simulations will help to evaluate the impact of the vocal tract on the resulting voice. Furthermore, this aspect could be

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

investigated by capturing also the shape of the vocal tract (e.g. MRI) while measuring the VF motion and use the obtained shape for the acoustic simulation. Besides the vocal tract shape, the low-pass filter could contribute to the deviation of the spectra in the higher frequency range since its cut-off frequency (5 kHz) is located there. As the acoustic source term is computed by the time derivative of the filtered incompressible pressure, the parameters of the filter might significantly affect the resulting voice. For the aperiodic case *InVivo5* the largest deviation from the measured voice was achieved. However, the disorder of the VF motion of this case has the most severe impact on the resulting measured voice as hardly any distinctive peaks occur at the harmonic frequencies. Due to the aperiodic motion, the principle mechanism of voice production, being the periodic interruption of the air stream, is disturbed. In the simulation, the glottis closes every fourth oscillation cycle concerning $f_1 =$ Hz because of $f_2 = 4\,f_1$. If in reality, a slight phase shift occurs between the oscillation of the two VFs, the glottis will hardly ever close. Thus, the aperiodic case will be very sensitive to a phase shift and the exact value of the oscillation frequencies and consequently, an inexact reproduction of the VF motion could be the cause of the deviation from the measurement.

The averaged field FFTs of the source terms in Fig 5.10 show that the ratio of broadband sound to harmonic amplitudes is increased for the hoarse voices (*InVivo3* to *InVivo5*). Regarding the share of broadband sound, configuration *InVivo2* (female healthy voice) exhibits the best developed harmonics compared to the noise resulting in the clearest voice of the case study with the most distinct peaks at the harmonic frequencies in the acoustic spectrum. Generally, the source terms of the hoarse voices are smaller, which yields reduced SPLs for the respective computed voices (compared to the healthy voices).

Further analysis of the acoustic results will be carried out by the cooperation partner Universitätsklinikum Erlangen (Department for Phoniatrics and Pedaudiology) by applying typical tools for quality assessment of the voice.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Chapter 6

Conclusion

After providing a brief overview of the general development in the field of numerical models of human phonation, the development of the model, which is applied in the present thesis, was presented. Moreover, the methodology, including the governing equations on which the model is based on, was explained. Thereby, the hybrid aeroacoustic approach, which utilizes the PCWE, and the available schemes that are required for the conservative interpolation of the acoustic source terms from the flow to the acoustic grid were introduced. Then, the applied workflow was illustrated and the flow and acoustic model were established. Before applying the acoustic model, various modeling aspects were assessed. By incorporating several potentials which were found to yield improvements, the acoustic model was developed towards an efficient, clinically applicable tool.

Having derived an appropriate model representing the experimental setup for validation, the CAA model was adapted to agree with real human phonation in a next step. Subsequently, two case studies were conducted. The first study investigates the impact of abnormal vocal fold characteristics (glottal gap, asymmetric and aperiodic motion) on the acoustic sources and the resulting voice, whereas the second study utilizes clinical data of patients in order to evaluate the physiological correctness of the model. Thereby, in the first case study, it was found that voice quality decreases with a rising glottal gap due to insufficiently developed acoustic sources. Figure 6.1, which depicts the relationship between the degree of initial opening and the deviation from the normal voice by means of the error $\mathrm{Err}_{\mathrm{rel}}{}^{\mathrm{L}_2}$, supports this tendency. Interestingly, the case study revealed that asymmetric VF motion (in terms of oscillation amplitude) can compensate for glottal insufficiency. In the second case study, simulations based on healthy as well as disordered vocal fold motion of patients were performed and compared with acoustic measurements. The substantial quantitative and partly qualitative deviation from the measurements, especially in the higher frequency range, suggest to evaluate the accordance of the vocal tract applied in the numerical model with the vocal tract of the patients where measurements were conducted on. However, the results of the ex-vivo cases, which are not computed yet, will reveal first information regarding the influence of the vocal tract. Furthermore, the impact of the parameters of the low-pass filter, applied for filtering CFD data, on the acoustic results needs to be investigated more extensively as the cut-off frequency is located at the upper frequency range considered for evaluation.

The spatial source term distribution as well es the underestimation of the amplitudes of the lower harmonics suggest that the convective part might significantly contribute to the total source term at lower frequencies and thus, its consideration could reduce the deviation between the spectra of the simulated and measured sound. Therefore, this aspect will be investigated in the near future. For doing so, first, a methodology for obtaining the mean flow velocity, which is required for the computation of the convective

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

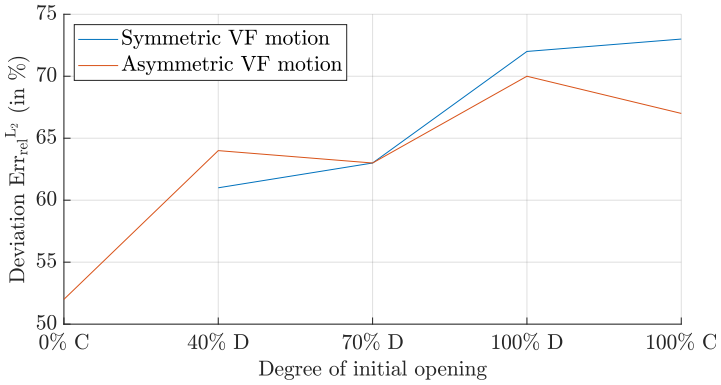


Figure 6.1: Relationship between initial glottal opening (with regard to the VF length) and the deviation from a normal voice by means of the error $Err_{rel}^{L_2}$ (D...diverging, C...constant initial opening).

term, has to be established first. Furthermore, the field FFT in the symmetry planes of the source region will be expanded in order to cover all configurations of case study 1 (variation of VF motion and glottal closure). Therewith, fundamental findings about the impact of glottal insufficiency, asymmetric and aperiodic VF oscillation on the spatial distribution of acoustic sources are expected.

Regarding the progress of the FWF project *Numerical computation of the human voice source*, WP3 will be accomplished when the three ex-vivo configurations (case study 2) are simulated. This will be done as soon as the flow data for source term computation is available. Having simulated and analyzed the remaining cases, WP4 will be started, which is basically an extension of case study 2 where in addition to the aspects of WP3 (symmetry of the VF motion, glottal closure and periodicity of the oscillation) the impact of the subglottal pressure is assesed. Furthermore, all these configurations are additionally performed with a reduced size of the vocal tract yielding altogether 96 configurations.

In order to reach the overall goal of the project, which is the clinical applicability of the numerical model of human phonation, further fundamental advances in computational efficiency are essential. With the current implementations, the CFD simulation takes about 20 hours (15 oscillation cycles) despite using a high-performance compute cluster, whereas the remaining aeroacoustic workflow (conservative interpolation, low-pass filtering, source term computation, and actual acoustic simulation) can be performed in about two hours on a conventional PC. Moreover, a massive amount of data is generated due to the required fine spatial resolution of the flow field. Consequently, the crucial factor of the clinical applicability lies in the CFD simulation. A potential way to circumvent the costly flow simulation would be to utilize neural networks that compute the acoustic source terms based on characteristic parameters of the VF motion and the laryngeal geometry. Thereby, the required training data could be obtained by simulations as well as by measurements. Of course, this approach is quite revolutionary and would require substantial investigations and the acquisition of know-how as aeroacoustic source term computation based on artificial intelligence (AI) can not be found in literature yet. Nevertheless, the potential of AI is huge and it has already been applied successfully in the field of health care [66–68], aerodynamics [69], aeroacoustics [70–72]) and phonation (e.g. speaker adaption for automatic speech recognition [73] or assessment of voice quality [74]). A less substantial adaption of the applied methodology to reduce the computational effort would be the extrapolation of the acoustic results (e.g. acoustic spectrum at a microphone point) from a reduced number (ideally one) to an arbitrary number of oscillation cycles. Therewith, the computational cost could be reduced by up to 93 % based on the current implementation that considers about 15 cycles. Again, neural networks could be employed and training data could be provided by measurements and simulations.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Bibliography

[1] Elaine Smith, Katherine Verdolini, Steven Gray, Sara Nichols, Jon Lemke, Julie Barkmeier, Heather Dove, and Henry Hoffman. Effect of voice disorders on quality of life. *Journal of Medical Speech-Language Pathology*, 4(4):223–244, 12 1996.

[2] Seth M Cohen. Self-reported impact of dysphonia in a primary care population: An epidemiological study. *The Laryngoscope*, 120(10):2022–2032, 2010.

[3] Ray M Merrill, Allison E Anderson, and Arielle Sloan. Quality of life indicators according to voice disorders and voice-related conditions. *The Laryngoscope*, 121(9):2004–2010, 2011.

[4] Robert J Ruben. Redefining the survival of the fittest: communication disorders in the 21st century. *The Laryngoscope*, 110(2):241–241, 2000.

[5] M Döllinger and M Kaltenbacher. Preface: Recent advances in understanding the human phonatory process. *Acta Acustica united with Acustica*, 102(2):195–208, 2016.

[6] Alistair G. Brown, Yubing Shi, Alberto Marzo, Cristina Staicu, Isra Valverde, Philipp Beerbaum, Patricia V. Lawford, and D. Rodney Hose. Accuracy vs. computational time: Translating aortic simulations to the clinic. *Journal of Biomechanics*, 45(3):516 – 523, 2012.

[7] Paul D Morris, Andrew Narracott, Hendrik von Tengg-Kobligk, Daniel Alejandro Silva Soto, Sarah Hsiao, Angela Lungu, Paul Evans, Neil W Bressloff, Patricia V Lawford, D Rodney Hose, et al. Computational fluid dynamics modelling in cardiovascular medicine. *Heart*, 102(1):18–28, 2016.

[8] Takuya Ueda, Hiroshi Suito, Hideki Ota, and Kei Takase. Computational fluid dynamics modeling in aortic diseases. *Cardiovascular Imaging Asia*, 2(2):58–64, 2018.

[9] J Flanagan and Lois Landgraf. Self-oscillating source for vocal-tract synthesizers. *IEEE Transactions on Audio and Electroacoustics*, 16(1):57–64, 1968.

[10] Kenzo Ishizaka and James L Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell system technical journal*, 51(6):1233–1268, 1972.

[11] Anxiong Yang, David A Berry, Manfred Kaltenbacher, and Michael Döllinger. Three-dimensional biomechanical properties of human vocal folds: Parameter optimization of a numerical model to match in vitro dynamics. *The Journal of the Acoustical Society of America*, 131(2):1378–1390, 2012.

[12] Anxiong Yang, Jörg Lohscheller, David A Berry, Stefan Becker, Ulrich Eysholdt, Daniel Voigt, and Michael Döllinger. Biomechanical modeling of the three-dimensional aspects of human vocal fold dynamics. *The Journal of the Acoustical Society of America*, 127(2):1014–1031, 2010.

[13] M Döllinger, S Kniesburges, M Kaltenbacher, and M Echternach. Aktuelle methoden zur modellierung des stimmgebungsprozesses. *HNO*, 64(2):82–90, 2016.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[14] Fariborz Alipour, Christoph Brucker, Douglas D Cook, Andreas Gommel, Manfred Kaltenbacher, Willy Mattheus, Luc Mongeau, Eric Nauman, Rudiger Schwarze, Isao Tokuda, et al. Mathematical models and numerical schemes for the simulation of human phonation. *Current Bioinformatics*, 6(3):323–343, 2011.

[15] Manfred Kaltenbacher, Stefan Zörner, and Andreas Hüppe. On the importance of strong fluid-solid coupling with application to human phonation. *Progress in Computational Fluid Dynamics, an International Journal*, 14(1):2–13, 2014.

[16] Fariborz Alipour, David A. Berry, and Ingo R. Titze. A finite-element model of vocal-fold vibration. *The Journal of the Acoustical Society of America*, 108(6):3003–3012, 2000.

[17] Marcelo de Oliveira Rosa, José Carlos Pereira, Marcos Grellet, and Abeer Alwan. A contribution to simulating a three-dimensional larynx model using the finite element method. *The Journal of the Acoustical Society of America*, 114(5):2893–2905, 2003.

[18] Comer Duncan, Guangnian Zhai, and Ronald Scherer. Modeling coupled aerodynamics and vocal fold dynamics using immersed boundary methods. *The Journal of the Acoustical Society of America*, 120(5):2859–2871, 2006.

[19] Xudong Zheng, Steve Bielamowicz, Haoxiang Luo, and Rajat Mittal. A computational study of the effect of false vocal folds on glottal flow and vocal fold vibration during phonation. *Annals of biomedical engineering*, 37(3):625–642, 2009.

[20] Willy Mattheus and Christoph Brücker. Asymmetric glottal jet deflection: Differences of two- and three-dimensional models. *The Journal of the Acoustical Society of America*, 130(6):EL373–EL379, 2011.

[21] Chao Tao, Yu Zhang, Daniel G Hottinger, and Jack J Jiang. Asymmetric airflow and vibration induced by the coanda effect in a symmetric model of the vocal folds. *The Journal of the Acoustical Society of America*, 122(4):2270–2278, 2007.

[22] Rajat Mittal, Haibo Dong, Meliha Bozkurttas, FM Najjar, Abel Vargas, and Alfred Von Loebbecke. A versatile sharp interface immersed boundary method for incompressible flows with complex boundaries. *Journal of computational physics*, 227(10):4825–4852, 2008.

[23] Rajat Mittal and Gianluca Iaccarino. Immersed boundary methods. *Annu. Rev. Fluid Mech.*, 37:239–261, 2005.

[24] 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. *The Journal of the Acoustical Society of America*, 130(1):404–415, 2011.

[25] X Zheng, R Mittal, and S Bielamowicz. A computational study of asymmetric glottal jet deflection during phonation. *The Journal of the Acoustical Society of America*, 129(4):2133–2143, 2011.

[26] Weili Jiang, Xudong Zheng, and Qian Xue. computational modeling of fluid–structure–acoustics interaction during voice production. *Frontiers in bioengineering and biotechnology*, 5:7, 2017.

[27] Brad H Story, Ingo R Titze, and Eric A Hoffman. Vocal tract area functions from magnetic resonance imaging. *The Journal of the Acoustical Society of America*, 100(1):537–554, 1996.

[28] Daniel Aalto, Olli Aaltonen, Risto-Pekka Happonen, Päivi Jääsaari, Atle Kivelä, Juha Kuortti, Jean-Marc Luukinen, Jarmo Malinen, Tiina Murtola, Riitta Parkkola, et al. Large scale data acquisition of simultaneous mri and speech. *Applied Acoustics*, 83:64–75, 2014.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[29] Marc Arnela, Saeed Dabbaghchian, Oriol Guasch, and Olov Engwall. MRI-based vocal tract representations for the three-dimensional finite element synthesis of diphthongs. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(12):2173–2182, 2019.

[30] Lukas Schickhofer, Jarmo Malinen, and Mihai Mihaescu. Compressible flow simulations of voiced speech using rigid vocal tract geometries acquired by mri. *The Journal of the Acoustical Society of America*, 145(4):2049–2061, 2019.

[31] Institute of Mechanics and Mechatronics: Research Unit of Technical Acoustics. Homepage of TU Wien, 2020.

[32] Manfred Kaltenbacher. *Computational Acoustics*. Springer, 2018.

[33] Gerhard Link, M Kaltenbacher, Michael Breuer, and M Döllinger. A 2d finite-element scheme for fluid–solid–acoustic interactions and its application to human phonation. *Computer Methods in Applied Mechanics and Engineering*, 198(41-44):3321–3334, 2009.

[34] P Šidlof and S Zörner. Computational aeroacoustics of human phonation. In *EPJ Web of Conferences*, volume 45, page 01085. EDP Sciences, 2013.

[35] S Zörner, M Kaltenbacher, and M Döllinger. Investigation of prescribed movement in fluid–structure interaction simulation for the human phonation process. *Computers & fluids*, 86:133–140, 2013.

[36] Rémi Blandin, Marc Arnela, Rafael Laboissière, Xavier Pelorson, Oriol Guasch, Annemie Van Hirtum, and Xavier Laval. Effects of higher order propagation modes in vocal tract like geometries. *The Journal of the Acoustical Society of America*, 137(2):832–843, 2015.

[37] Marc Freixes, Marc Arnela, Joan Claudi Socoró, Francesc Alías Pujol, and Oriol Guasch. Influence of tense, modal and lax phonation on the three-dimensional finite element synthesis of vowel [a]. In *IberSPEECH*, pages 132–136, 2018.

[38] Marc Arnela, Saeed Dabbaghchian, Rémi Blandin, Oriol Guasch, Olov Engwall, Annemie Van Hirtum, and Xavier Pelorson. Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds. *The Journal of the Acoustical Society of America*, 140(3):1707–1718, 2016.

[39] Brian B Monson, Eric J Hunter, Andrew J Lotto, and Brad H Story. The perceptual significance of high-frequency energy in the human voice. *Frontiers in psychology*, 5:587, 2014.

[40] Judith Probst, Alexander Lodermeyer, Sahar Fattoum, Stefan Becker, Matthias Echternach, Bernhard Richter, Michael Döllinger, and Stefan Kniesburges. Acoustic and aerodynamic coupling during phonation in mri-based vocal tract replicas. *Applied Sciences*, 9(17):3562, 2019.

[41] Roland Ewert and Wolfgang Schröder. Acoustic perturbation equations based on flow decomposition via source filtering. *Journal of Computational Physics*, 188(2):365–398, 2003.

[42] Manfred Kaltenbacher. *Numerical simulation of mechatronic sensors and actuators: Finite elements for computational multiphysics*. Springer, 3 rd edition, 2013.

[43] JC Hardin and DS Pope. An acoustic/viscous splitting technique for computational aeroacoustics. *Theoretical and Computational Fluid Dynamics*, 6(5-6):323–340, 1994.

[44] W. Z. Shen and J. N. Sørensen. Aeroacoustic modelling of low-speed flows. *Theoretical and Computational Fluid Dynamics*, 13(4):271–289, 1999.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[45] J. Seo and Y. J. Moon. Perturbed compressible equations for aeroacoustic noise prediction at low mach numbers. *AIAA journal*, 43(8):1716–1724, 2005.

[46] M. Munz, C.-D.and Dumbser and S. Roller. Linearized acoustic perturbation equations for low Mach number flow with variable density and temperature. *Journal of Computational Physics*, 224(1):352–364, 2007.

[47] M. Kaltenbacher, M. Escobar, S. Becker, and I. Ali. Numerical simulation of flow-induced noise using les/sas and lighthill's acoustic analogy. *International journal for numerical methods in fluids*, 63(9):1103–1122, 2010.

[48] Stefan Schoder. *Aeroacoustic analogies based on compressible flow data*. PhD thesis, TU Wien, 2019.

[49] Stefan Kniesburges. *Fluid-structure-acoustic interaction during phonation in a Ssynthetic larynx model*. PhD thesis, FAU Erlangen-Nürnberg, 2014.

[50] Pavel Holoborodko. Smooth noise robust differentiators. http://www.holoborodko.com/pavel/numerical-methods/numerical-derivative/smooth-low-noise-differentiators/, 2008.

[51] Hossein Sadeghi, Stefan Kniesburges, Sebastian Falk, Manfred Kaltenbacher, Anne Schützenberger, and Michael Döllinger. Towards a clinically applicable computational larynx model. *Applied Sciences*, 9(11):2288, 2019.

[52] Hossein Sadeghi, Stefan Kniesburges, Manfred Kaltenbacher, Anne Schützenberger, and Michael Döllinger. Computational models of laryngeal aerodynamics: Potentials and numerical costs. *Journal of Voice*, 33(4):385–400, 2019.

[53] Barbara Kaltenbacher, Manfred Kaltenbacher, and Imbo 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*, 235:407–422, 2013.

[54] Alexander Hauser. *Efficient FEM model for human phonation*. Bachelor thesis, TU Wien, 2019.

[55] Stefan Schoder, Michael Weitz, Paul Maurerlehner, Alexander Hauser, Sebastian Falk, Stefan Kniesburges, Michael Döllinger, and Manfred Kaltenbacher. Hybrid aeroacoustic approach for the efficient numerical simulation of human phonation. *The Journal of the Acoustical Society of America*, 147(2):1179–1194, 2020.

[56] S Zörner, P Šidlof, A Hüppe, and M Kaltenbacher. Flow and acoustic effects in the larynx for varying geometries. *Acta Acustica united with Acustica*, 102(2):257–267, 2016.

[57] Q Xue, R Mittal, X Zheng, and S Bielamowicz. A computational study of the effect of vocal-fold asymmetry on phonation. *The Journal of the Acoustical Society of America*, 128(2):818–827, 2010.

[58] Monika Tigges, Patrick Mergell, Hanspeter Herzel, Thomas Wittenberg, and Ulrich Eysholdt. Observation and modelling of glottal biphonation. *Acta Acustica united with Acustica*, 83(4):707–714, 1997.

[59] Ulrich Eysholdt, F Rosanowski, and U Hoppe. Vocal fold vibration irregularities caused by different types of laryngeal asymmetry. *European Archives of Oto-rhino-laryngology*, 260(8):412–417, 2003.

[60] Elisabeth C Inwald, Michael Döllinger, Maria Schuster, Ulrich Eysholdt, and Christopher Bohr. Multiparametric analysis of vocal fold vibrations in healthy and disordered voices in high-speed imaging. *Journal of Voice*, 25(5):576–590, 2011.

[61] Christopher Bohr, Angelika Kraeck, Ulrich Eysholdt, Anke Ziethe, and Michael Döllinger. Quantitative analysis of organic vocal fold pathologies in females by high-speed endoscopy. *The Laryngoscope*, 123(7):1686–1693, 2013.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[62] Christopher Bohr, Angelika Kräck, Denis Dubrovskiy, Ulrich Eysholdt, Jan Svec, Georgios Psychogios, Anke Ziethe, and Michael Döllinger. Spatiotemporal analysis of high-speed videolaryngoscopic imaging of organic pathologies in males. *Journal of Speech, Language, and Hearing Research*, 57(4):1148–1161, 2014.

[63] Anne Schützenberger, Melda Kunduk, Michael Döllinger, Christoph Alexiou, Denis Dubrovskiy, Marion Semmler, Anja Seger, and Christopher Bohr. Laryngeal high-speed videoendoscopy: sensitivity of objective parameters towards recording frame rate. *BioMed research international*, 2016, 2016.

[64] V Birk, A Sutor, M Döllinger, C Bohr, and S Kniesburges. Acoustic impact of ventricular folds on phonation studied in ex vivo human larynx models. *Acta Acustica united with Acustica*, 102(2):244–256, 2016.

[65] Georg Luegmair, Daryush D Mehta, James B Kobler, and Michael Döllinger. Three-dimensional optical reconstruction of vocal fold kinematics using high-speed video with a laser projection system. *IEEE transactions on medical imaging*, 34(12):2572–2582, 2015.

[66] Geoffrey Hinton. Deep learning—a technology with the potential to transform health care. *Jama*, 320(11):1101–1102, 2018.

[67] Kun-Hsing Yu, Andrew L Beam, and Isaac S Kohane. Artificial intelligence in healthcare. *Nature biomedical engineering*, 2(10):719–731, 2018.

[68] Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, and Yongjun Wang. Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4):230–243, 2017.

[69] Jeffrey P Thomas, Earl H Dowell, and Kenneth C Hall. Nonlinear inviscid aerodynamic effects on transonic divergence, flutter, and limit-cycle oscillations. *AIAA journal*, 40(4):638–646, 2002.

[70] Spyros Kamnis, Konstantina Malamousi, Alex Marrs, Bryan Allcock, and Konstantinos Delibasis. Aeroacoustics and artificial neural network modeling of airborne acoustic emissions during high kinetic energy thermal spraying. *Journal of Thermal Spray Technology*, 28(5):946–962, 2019.

[71] Sajjad Beigmoradi, Hassan Hajabdollahi, and Asghar Ramezani. Multi-objective aero acoustic optimization of rear end in a simplified car model by using hybrid robust parameter design, artificial neural networks and genetic algorithm methods. *Computers & Fluids*, 90:123–132, 2014.

[72] K Förner and W Polifke. Nonlinear aeroacoustic characterization of Helmholtz resonators with a local-linear neuro-fuzzy network model. *Journal of Sound and Vibration*, 407:170–190, 2017.

[73] George Saon, Hagen Soltau, David Nahamoo, and Michael Picheny. Speaker adaptation of neural network acoustic models using i-vectors. In *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, pages 55–59. IEEE, 2013.

[74] RT Ritchings, M McGillion, and CJ Moore. Pathological voice quality assessment using artificial neural networks. *Medical engineering & physics*, 24(7-8):561–564, 2002.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN