

Experimental Investigation of the Flow-Acoustic Interaction in Human Phonation

Christoph Näger^{1,*}, Stefan Kniesburges², Stefan Schoder³, Stefan Becker¹

1: Institute of Fluid Mechanics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

2: Division of Phoniatics and Pediatric Audiology, Department of Otorhinolaryngology, Head & Neck Surgery, University Hospital Erlangen, Medical School, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

3: Aeroacoustics and Vibroacoustics Group, Institute of Fundamentals and Theory in Electrical Engineering, Graz University of Technology, Austria

*Corresponding author: christoph.naeger@fau.de

Keywords: PIV, Aeroacoustics, Fluid-Structure-Acoustic Interaction, Human Phonation.

ABSTRACT

In the human phonation process, acoustic standing waves in the vocal tract can influence the fluid flow through the glottis. To investigate the amount of two-way coupling between the flow field and the acoustics, the supraglottal flow field has been recorded via high-speed particle image velocimetry (PIV) in a synthetic larynx model for several configurations with different vocal tract lengths. Based on the obtained velocity fields, aeroacoustic source terms were computed. Furthermore, the hydrodynamic pressure in the vocal tract was recorded via wall pressure measurements. The recordings revealed that near a vocal tract resonance frequency f_R , the focal fold oscillation frequency f_o jumps onto f_R . This is accompanied by a substantial relative increase in aeroacoustic sound generation efficiency. At the same time, the glottal volume flow needed for stable vocal fold oscillation decreases strongly. The pressure measurements showed that acoustic anti-resonances of the vocal tract directly dampen the harmonic content of the hydrodynamic pressure field when matching frequencies. This result shows that back-coupling of the acoustics onto the flow field does not only occur when $f_o \approx f_R$, but also under more other conditions.

1. Introduction

The human voice is generated by a complex physiological process that is described by the fluid-structure-acoustic interaction (FSAI) between the tracheal fluid flow, structural vibration of laryngeal tissue (i.e., the vocal folds), and the sound generation and modulation in the larynx and vocal tract (Mittal et al., 2013). In this process, the two vocal folds are aerodynamically stimulated to vibrate by the airflow that arises from the lungs. In turn, this vibration leads to a modulation of the airflow, generating a pulsating jet flow in the supraglottal region which is above the vocal folds (Mittal et al., 2013).

Within this dynamic process of tissue-flow interaction, the basic sound of the human voice is generated by the highly complex 3D field of aeroacoustic sound sources which are produced by the

turbulent jet flow in the larynx (Schoder et al., 2021). The generated basic sound is further filtered by the vocal tract and radiated through the mouth, exhibiting the typical spectral characteristics of the human voice composed of tonal harmonic components of the fundamental frequency and additional tonal components called formants originating from resonance effects in the vocal tract (I. R. Titze & Story, 1997).

In early times, a linear behavior between sound source (vocal folds) and filter (vocal tract) was assumed within the linear source-filter theory which excluded the influence of the acoustic filter signal back on the source (Fant, 1971). However, this simplified representation turned out to be invalid, especially when the fundamental oscillation frequency of the vocal folds f_o is close to a resonance frequency f_R of the vocal tract (I. Titze et al., 2008). Since then, the three-way coupling of the fluid flow in the glottis, structural vibration of the vocal folds and acoustics in the vocal tract has been studied with several different methods. For example, McGowan & Howe (2012) used theoretical modelling to investigate the influences of source-tract interaction with prescribed vocal fold movement. Lucero et al. (2012) applied numerical modelling to examine the role of source-tract interaction at the vocal fold oscillation onset. There have also been multiple studies investigating source-filter interaction experimentally, in-vivo (e.g. Smith et al. (2013)) as well as with synthetic larynx models (e.g. Zhang et al. (2009); Migimatsu & Tokuda (2019)). These kinds of studies often report mainly on global metrics like f_o -variation or threshold pressure needed for stable vocal fold oscillation. Especially in the case of in-vivo studies, the restricted access to the region of interest (i.e. the larynx) often prohibits a more in-detail investigation of e.g. the flow field developing in and above the glottis. In contrast to this, numerical simulations are suited for a much more in-depth analysis of the whole FSAI process. However, numerical models mostly have an issue with accurate modeling of the contact between the two vocal folds during phonation. Here, one common method circumventing this is leaving a small gap between the vocal folds in the "closed" phase (e.g. Schoder et al. (2020)), although there are also recent approaches trying to actually model the contact, e.g. with fictitious porosity (Valášek & Sváček, 2024). Combining the advantages of good accessibility and being able to do experiments with full vocal fold contact, synthetic larynx models are suited well to study the full FSAI in human phonation.

In this context, Particle Image Velocimetry (PIV) enables to measure the unsteady flow field in the aeroacoustic source region to gain a deep insight into the FSAI process of phonation. This technique has already been applied successfully to study aerodynamic effects in synthetic as well as ex vivo larynx models that showed typical vocal fold vibrations similar to phonation (e.g. Oren et al. (2014); Lodermeier et al. (2015); de Luzan et al. (2020); Lodermeier et al. (2021)).

However, none of the studies described above have investigated the effects of supraglottal acoustics on the glottal aerodynamics and the aeroacoustic source field yet. This problem has been investigated recently by Näger, Kniesburges, Tur, et al. (2023). In a similar way, the present study provides highly resolved data of the entire process to analyse the complete FSAI between vocal

tract acoustics and laryngeal aerodynamics. Based on high-speed PIV measurements in combination with aerodynamic and acoustic pressure data, the influence of the resonance effects formed in the vocal tract on the supraglottal flow field is studied. In contrast to the study published by Näger, Kniesburges, Tur, et al. (2023), here, a special focus will also be put on the interaction of the vocal tract acoustic properties with single harmonic frequencies of the aeroacoustic source field. For this, different vocal tract models have been applied with an incremental increase in length. Changing this length is a simple way of being able to freely vary the acoustic resonance frequencies of the vocal tract. This procedure enabled us to systematically study the interaction between the laryngeal flow field and supraglottal acoustics.

We want to answer three research questions this way: 1. Can we detect a two-way flow-acoustic interaction? 2. If so, which effects does it have on the harmonic contents in the phonation process? 3. Which effects does it have on the aeroacoustic efficiency in generation of the voice?

2. Methodology

2.1. Experimental setup

The test rig used in this study is shown in fig. 1 and is the same one as in previous publications (Näger, Kniesburges, Tur, et al., 2023; Kraxberger et al., 2023). The flow direction through the setup is as indicated from left to right. Synthetic single-layer vocal folds cast from silicone were glued into their mounting and mounted into the test rig at the position indicated in the figure. Upstream from the vocal folds, a subglottal channel with a length of 210 mm was attached. This length was chosen small enough to prevent interaction of the vocal folds' oscillation with the subglottal acoustic resonances (Zhang et al., 2006). Further upstream, a silencer prevented the formation of standing acoustic waves in the air supply system to ensure well-defined subglottal acoustic boundary conditions. The vocal tract (also supraglottal channel) was made of two sections. Directly downstream of the vocal folds, a rectangular section with a glass window providing optical access followed. After that, a circular section made up of two telescopic tubes was attached. These tubes allowed a systematic variation of the vocal tract length L in the range $L \in [170 \text{ mm}, 930 \text{ mm}]$. This led to a variation in the first vocal tract resonance frequency f_{R1} in the range $f_{R1} \in [616 \text{ Hz}, 101 \text{ Hz}]$. A mass flow generator with a supercritical valve (Durst et al., 2003) produced a constant volume flow \dot{V} through the setup.

Two main measurement tasks were performed in this study. The first one was a measurement series of planar 2D-2C PIV-measurements in the coronal plane midway along the vocal fold length. The basic setup is shown in fig. 2. Here, a total of nine measurements with varying L were conducted in the supraglottal region just downstream of the vocal folds. For this, a double-pulse, frequency-doubled Nd:YLF Continuum Terra PIV high-speed laser was chosen as the light source.

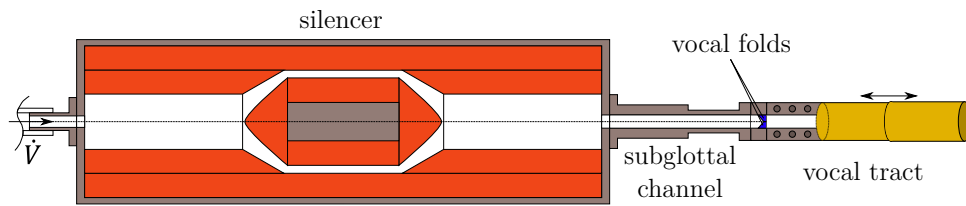


Figure 1. The experimental setup used in this study. The vocal fold position is highlighted in blue. A silencer is placed upstream to prevent standing acoustic waves in the air supply system. The vocal tract length can be varied via telescopic tubes.

One Vision Research Phantom v2511 high-speed camera in combination with a Canon Macro Lens EF 180 mm Ultrasonic lens was applied for recording purposes. The measurement frequency was set to 2×5 kHz with a pulse distance of $4 \mu\text{s}$, realized with an ILA synchronizer. Image pre-processing was done by a combination of background removal via proper orthogonal decomposition (Mendez et al., 2017) and anisotropic diffusion (Adatrao & Sciacchitano, 2019) (for details refer to Näger, Kniesburges, Tur, et al. (2023)). Velocity vectors were extracted from the image pairs via the software package PIVview2C 3.6.23.

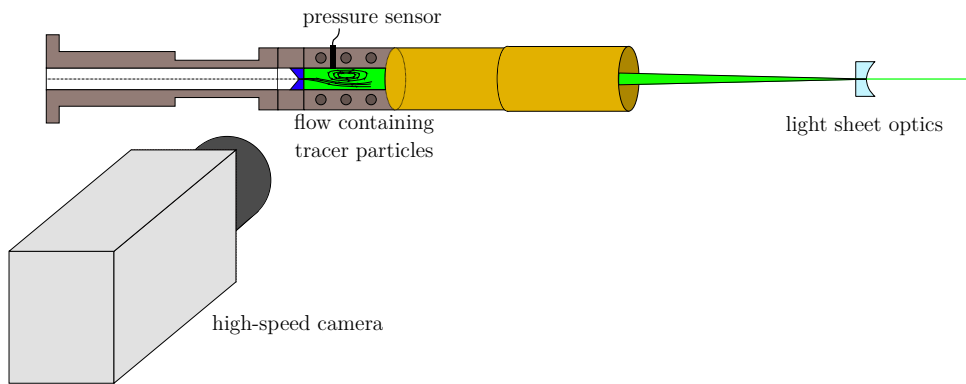


Figure 2. The measurement setup used in this study. The rectangular section of the vocal tract provided optical access for flow recording with a high-speed camera. Furthermore, a pressure sensor was placed in the supraglottal region.

For the second measurement series, a Kulite XCS-093 open gauge pressure sensor was mounted in the supraglottal channel at a position 50 mm downstream of the vocal folds. This sensor allows a time-resolved recording of the static pressure at the sensor position. The recording frequency for this series was set to 50 kHz. Here, the vocal tract length was varied in 10 mm steps over the whole range of possible length settings, resulting in a total of 77 single measurements. In all measurements, the volume flow rate \dot{V} was set to the minimal value producing a stable vocal fold oscillation with a distinct contact phase between the two vocal folds.

2.2. Data Processing

To be able to study the power in single harmonic components of the measured data, the power spectral density P_ξ of a measured quantity ξ is computed from a discrete fourier transformation. For the PIV measurements, this is then integrated over the whole region of interest (ROI), yielding the total power of the measured flow field as a function of frequency:

$$S_x(f) = \iint_{\text{ROI}} P_{u_x}(x, y, f) dx dy \quad (1)$$

and

$$S_y(f) = \iint_{\text{ROI}} P_{u_y}(x, y, f) dx dy, \quad (2)$$

where u_x and u_y indicate the measured velocity components in x - and y - direction respectively. As the volume flow rate \dot{V} needed for stable vocal fold oscillation strongly varies with changing vocal tract length L , also the average supraglottal velocity oscillation and therefore P_{u_x} and P_{u_y} show large differences between the different measurements. For comparison of the power content of single harmonic frequencies, a normalization with respect to the total power over the whole frequency range is performed:

$$S_x^*(f) = \frac{S_x(f)}{\frac{1}{f_{Ny}} \int_{f=0}^{f_{Ny}} S_x(f) df} \quad (3)$$

and

$$S_y^*(f) = \frac{S_y(f)}{\frac{1}{f_{Ny}} \int_{f=0}^{f_{Ny}} S_y(f) df} \quad (4)$$

Here, f_{Ny} denotes the Nyquist-frequency. Similarly, the supraglottal pressure measurement is normalized as well:

$$S_p^*(f) = \frac{P_p(f)}{\frac{1}{f_{Ny}} \int_{f=0}^{f_{Ny}} P_p(f) df}. \quad (5)$$

2.3. Aeroacoustic Source Computation

One important aspect of understanding the human voice production process is to evaluate the aeroacoustic sources. In contrast to the work given by Näger, Kniesburges, Tur, et al. (2023), here the aeroacoustic sources are evaluated following the theory of vortex sound (Howe, 2002), reducing the Lighthill source term to

$$T = \rho_0 \nabla \cdot (\boldsymbol{\omega} \times \mathbf{u}) \quad (6)$$

with the velocity \mathbf{u} measured via PIV, the vorticity ω and the ambient density ρ_0 . This distributed source term is aggregated in a summed source strength based on Lighthill's analogy (Lighthill, 1952) by neglecting retarded time effects in this acoustically compact 2D ROI

$$\phi(t) = \frac{1}{4\pi c_0^2(x_1 - x_0)(y_1 - y_0)} \int_{(x_0, y_0)}^{(x_1, y_1)} T(\mathbf{x}, t) \frac{\pi y dx dy}{r} = \frac{1}{4c_0^2 N_x N_y} \sum_i T(\mathbf{x}_i, t) |y_i| . \quad (7)$$

The coordinate locations x_0, x_1, y_0, y_1 are the bounding coordinates of the ROI respectively, c_0 the isentropic speed of sound, r the direction of a virtual observer point at 1 m distance. It is assumed that the jet is rotationally symmetric around the rotation axis in x -direction, pointing in the flow direction and being centered in the middle of the vocal folds. From this equation, the root-mean-squared value is computed

$$\Phi = \sqrt{\frac{1}{t_1 - t_0} \int_{t_0}^{t_1} (\phi(t))^2 dt} , \quad (8)$$

being a measure of the ability to generate aerodynamic sound. The equation is applied to the measured 2D mid-section, where the velocity's principal direction is recorded. In addition, the aerodynamic input energy is quantified by

$$p_{tot} = \Delta p + \frac{1}{2} \rho_0 U^2 , \quad (9)$$

using the subglottal pressure difference to the ambient pressure Δp . With the input energy, the efficiency of the aeroacoustic sound generation yields

$$\eta = \frac{\Phi c_0^2 (x_1 - x_0)(y_1 - y_0)}{p_{tot}} . \quad (10)$$

3. Results and Discussion

Fig. 3 shows the acoustic vocal tract input impedance as a function of the frequency f and the vocal tract length L . Z_{in} has been computed via transmission line model (Näger, Kniesburges, Tur, et al., 2023; Sondhi & Schroeter, 1987), where the vocal tract walls are assumed to be rigid. For a fixed vocal tract length L , the maxima in Z_{in} correspond to the acoustic resonance frequencies of the vocal tract (Story et al., 2000). In the colormap of fig. 3, the yellow lines therefore show the evolution of the vocal tract resonance frequencies as a function of L . As expected, the resonance frequencies decrease with increasing L . Further, there is also one anti-resonance located between each two resonances, indicated by the dark-red line, that also decreases in frequency with increasing L .

The markers and dashed line on top of the contour for Z_{in} show the vocal fold oscillation frequency f_o as determined from the PIV and pressure sensor measurements respectively via peak picking. It can be seen that the development of f_o with varying L shows a very similar behavior between

the two measurement techniques, as it is expected. There is a region of $L < 380$ mm, where f_o is almost constant at $f_o \approx 158$ Hz, with variations of approximately ± 1 Hz. These small variations can be attributed to reproducibility issues in the measurements. A second region with constant f_o can be found at $430 \text{ mm} < L < 480$ mm. In between these, there is a frequency jump to $f_o \approx 225$ Hz. For $L \geq 480$ mm there is a slight increase and then decrease in f_o . This increase and decrease is attributable to a tuning of the vocal fold oscillation frequency to the first resonance frequency f_{R1} of the vocal tract. This is also clearly visible in fig. 3, as here, an alignment of the data with the yellow line corresponding to f_{R1} is visible. This is in agreement with the findings of Näger, Kniesburgers, Tur, et al. (2023) and Kraxberger et al. (2023). In the region $430 \text{ mm} < L < 480$ mm, the frequency jump can also be explained by an alignment of f_o with f_{R1} . Here however, according to Kraxberger et al. (2023), this alignment is facilitated by a mode shift in the vocal fold oscillation. As these effects are visible in the PIV as well as the pressure measurements, it can be concluded that a two-way interaction exists between the supraglottal flow field and the standing acoustic waves in the vocal tract, at least in the regions of $430 \text{ mm} < L < 480$ mm and $L \geq 480$ mm, i.e. when $f_o \approx f_{R1}$. All of this is confirmation of the results shown in Näger, Kniesburgers, Tur, et al. (2023) and Kraxberger et al. (2023). In these works, no back-coupling could be detected from the acoustics back on the flow field and structural vibration of the vocal folds could be detected in the remaining cases, i.e. when $f_o \neq f_{R1}$.

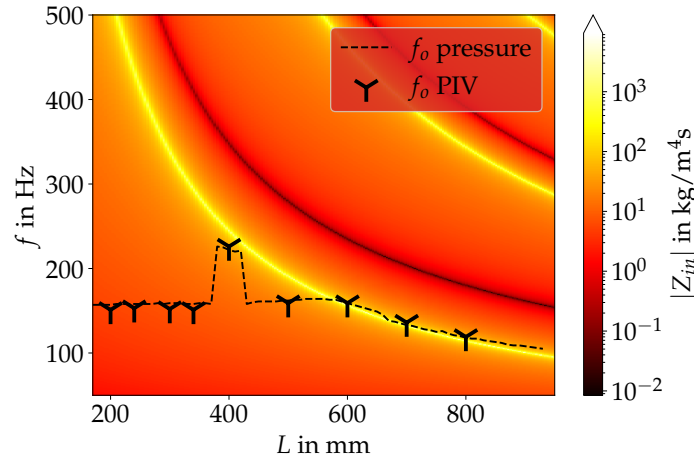


Figure 3. The vocal fold oscillation frequency f_o for different vocal tract lengths as recorded via PIV and pressure measurements. The colored background shows the acoustic vocal tract input impedance as a function of vocal tract length and frequency.

To further investigate the two-way interaction of supraglottal acoustics and flow field, a spectral analysis of the measurement data was performed. The spectra S_x and S_y as defined by equations 1 and 2 are shown in fig. 4 A and B respectively for the different vocal tract lengths. It can be seen that all the spectra are dominated by their harmonic contents. Especially, the vocal fold oscillation frequency f_o as well as its higher harmonic frequencies are prominent. In the measurements with $L < 400$ mm and $L = 500$ mm, also a more broadband peak is visible in the region around

$f \approx 45$ Hz. This one is attributable to a large-scale vortex in the supraglottal region switching its rotating direction with this frequency, which has been shown by Näger, Kniesburgs, & Becker (2023). This vortex leads to a deflection of the supraglottal jet flow. For longer vocal tracts, the vortex stabilizes and therefore does not change direction anymore. This way, the jet is permanently deflected towards one of the vocal tract walls, depending on the initial deflection at phonation onset. Generally, the spectra show vastly different levels, which is explained by the different volume flow rates \dot{V} set in the experiments. Those values are summarized in tab. 1. Normalization leads to S_x^* and S_y^* as defined by equations 3 and 4. These are shown in fig. 4 D and E. It can be seen that using S_x^* and S_y^* leads to a much more similar level for all the different L , making a comparison of harmonic contents between the different measurements easier. In the same way a normalization of the pressure measurements has been performed. Fig. 4 C and F show the power spectral density of the pressure as well as the normalized S_p^* according to equation 5 respectively. Here, for better visibility only the spectra of the measurements with the same L as the PIV experiments are shown. While fig. 4 D and E show a very similar broadband level between the single measurements, the same is not the case in fig. 4 F. This is because the harmonic components in the pressure are much more dominant compared to the general noise level than it is in the velocity. Therefore, the normalization leads instead to a similar peak level in f_o in the spectra rather than noise level.

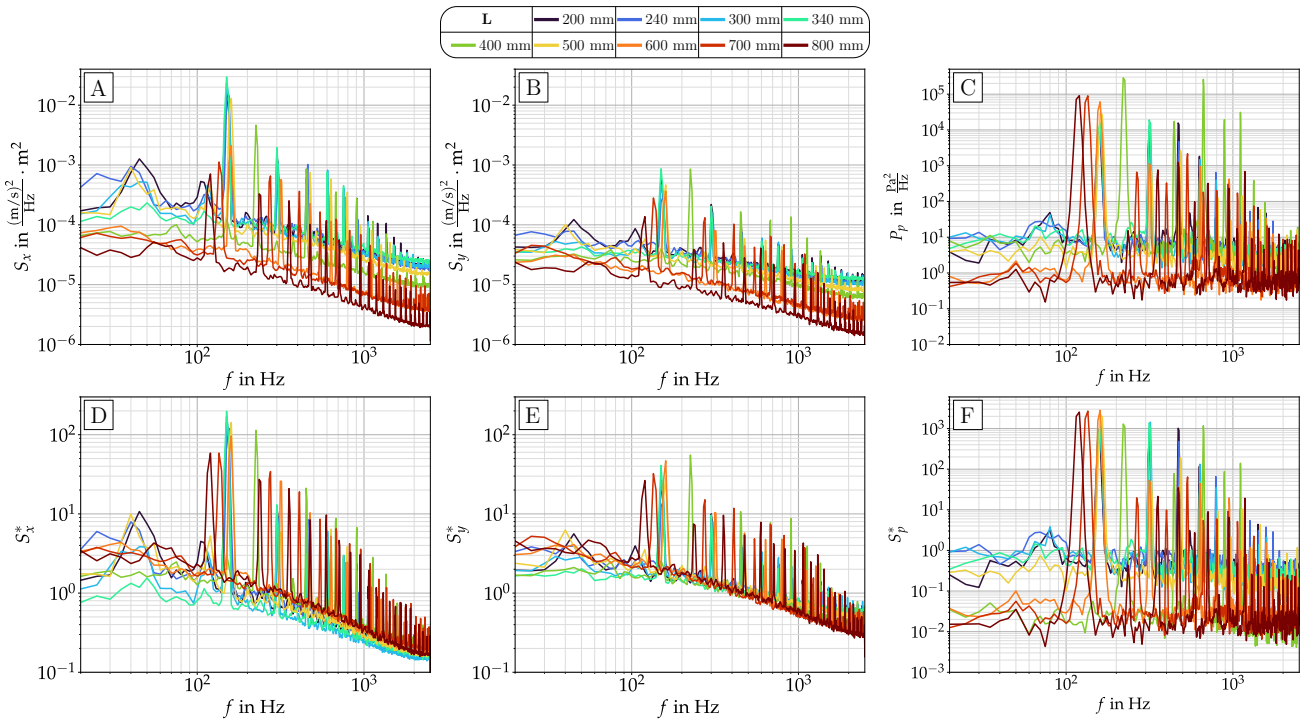


Figure 4. Velocity spectra in x - and y -direction. (A) and (B): non-normalized PSD of the PIV measurements integrated over whole spatial domain. (C): non-normalized PSD of the pressure measurements. (D) and (E): normalized PSD of the PIV measurements integrated over whole spatial domain. (F): normalized PSD of the pressure measurements.

From the normalized properties, the peak levels can be extracted as a function of L . This is shown

Table 1. The volume flow rate \dot{V} set in the PIV experiments depending on the vocal tract length L .

L in mm	200	240	300	340	400	500	600	700	800
\dot{V} in l/min	124	120	123	120	107	99	58	55	46

for the first four harmonic frequencies (f_o , $2f_o$, $3f_o$ and $4f_o$) in the bottom half of fig. 5 for the x -component of the velocity as well as the pressure measurements. The markers correspond to the PIV measurements, while the dashed and dotted lines correspond to the pressure measurements. The top half of the figure shows again a contour of Z_{in} , but this time overlaid with f_o and its higher harmonics. Looking at the harmonic content of the pressure measurements, it can be seen that the level of f_o is almost constant over the whole range of vocal tract lengths. This is a confirmation of the previously explained fact that the normalization leads to a similar peak level in the pressure measurements in f_o . In the higher harmonics, there are some very distinct minima in the levels. In $2f_o$ there are two minima at $L = 430$ mm and $L = 810$ mm. In $3f_o$ there is a minimum at $L = 340$ mm and in $4f_o$ there are again two minima at $L = 230$ mm and $L = 480$ mm. Comparing the top and the bottom half of fig. 5 reveals that these minima occur exactly at crossings of the respective harmonic frequency with an anti-resonance of the vocal tract indicated by the dark red lines in the contour. The data show that every crossing of a harmonic frequency with an anti-resonance of the vocal tract corresponds to a dampening of the respective harmonic frequency. The one exception is the crossing of $2f_o$ with the first anti-resonance, where the minimum amplitude seems to be shifted towards higher frequencies. This might have to do with the fact that the transmission-line model used to compute Z_{in} is a simplified one-dimensional model assuming rigid vocal tract walls. It can therefore be attributed to the accuracy limitations of the simplified acoustic model. However, apart from that the relationship between supraglottal pressure harmonic intensities and the acoustic anti-resonance frequencies of the vocal tract seems clear. This leads to the conclusion that a two-way interaction between the supraglottal flow field and acoustics is not only present when $f_o \approx f_{R1}$ but also at crossings of higher harmonics with the anti-resonances of the vocal tract. Interestingly, the same relationship is not found when only looking at the harmonic intensity in the flow fields measured by PIV. This gives rise to the theory that the two-way interaction caused by anti-resonances of the vocal tract and the supraglottal flow is mainly acting on the hydrodynamic pressure field and not so much on the actual velocity fields.

For the examination of the influence of the flow-acoustic interaction on aeroacoustic sound generation, aeroacoustic source terms were computed from the measured velocity fields according to equation 6. One example velocity snapshot and the corresponding aeroacoustic source field are shown in fig. 6 A and B respectively. It can be seen that according to the theory of vortex sound, the aeroacoustic sources are mainly located along the shear layer between the jet flow coming out from the vocal folds and the surrounding air in the vocal tract. These sources have been evaluated for every snapshot of every measurement and integrated according to equation 8. The resulting

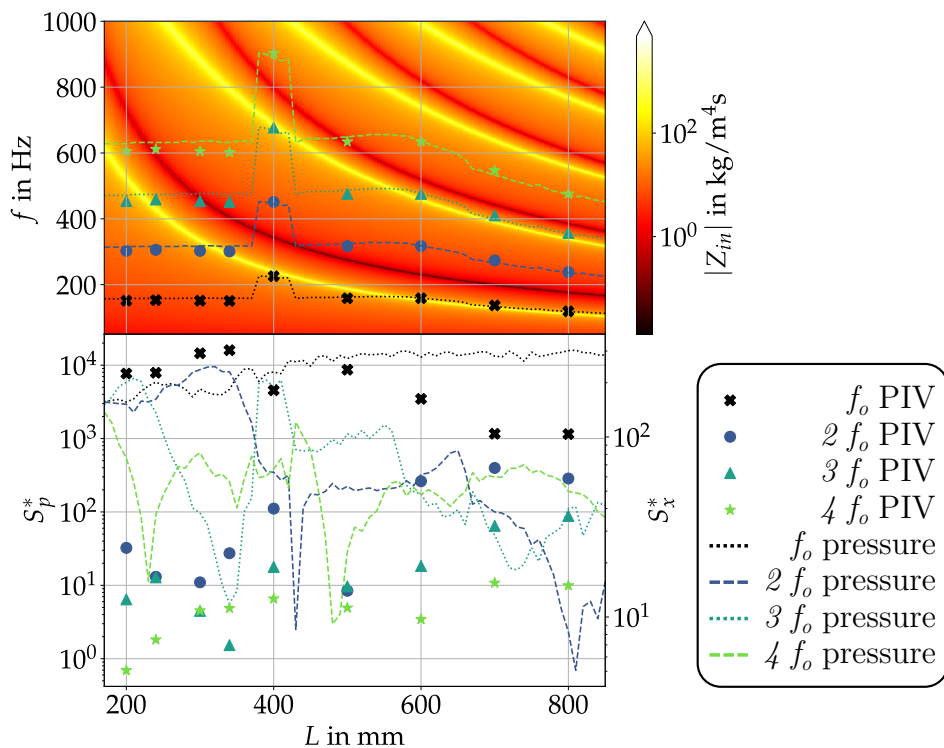


Figure 5. The peak level of S_x^* and S_p^* in the first four harmonic frequencies of the vocal fold oscillation f_o , $2f_o$, $3f_o$ and $4f_o$ (bottom). The top half of the plot shows the relationship between these harmonic frequencies and the vocal tract input impedance Z_{in} .

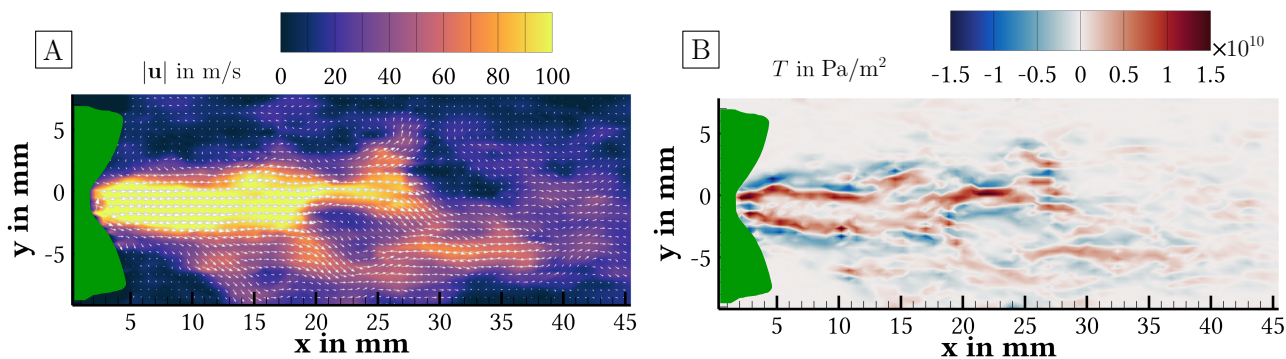


Figure 6. (A): Velocity contour of one snapshot of the flow field for $L = 200$ mm. (B) Contour of the vortex sound source field T computed from the same snapshot. The vocal fold position in this snapshot is indicated by the green surface.

values for Φ are shown in fig. 7 A. It can be seen that Φ is rather constant up to $L = 500$ mm and then sees a strong drop for larger L . This is to be expected, as for $L > 500$ mm, there is also a strong drop in the volume flow rate set in the experiment, as shown by table 1. The aeroacoustic source intensity is proportional to the flow velocities occurring, which are dependent on the volume flow rate set in the experiment. The total subglottal pressure p_{tot} displayed in fig. 7 C also shows the proportionality to \dot{V} . Furthermore, the aeroacoustic efficiency η , computed via eq. 10 and shown in fig. 7 also shows a slight decrease with increasing L , but to a lesser extent than the other quantities. In theory, Lighthill (1952) showed that the efficiency of sound generation in free turbulent flows without influence of solid walls in the flow domain scales with the fifth power of the Mach number. To compare our results to this scaling law, fig. 7 B also shows a theoretical computation of the aeroacoustic efficiency $\eta_{theor,1}$, which makes use of this fifth power law, using the bulk Mach number as the scaling variable. For the channel lengths $L \leq 340$ mm, this law shows a reasonable agreement with the measurement data. It starts deviating from the data for larger L , and shows a strong difference for $L \geq 600$ mm. From fig. 3 we know that this is also the length region where f_o is close to f_{R1} . This suggests that the acoustic resonance frequency of the vocal tract increases the aeroacoustic source intensity strongly by more than one order of magnitude. One explanation for this might be the changed vocal fold oscillation in the resonance case. Looking at high speed camera recordings of the vocal fold movement, Näger, Kniesburges, Tur, et al. (2023) found that in the resonance case, the vocal folds oscillate with a much reduced amplitude, also decreasing the glottal gap area during the phonation process. Due to continuity principles, the velocities of the jet flow therefore have to increase. This counteracts on the fact that the velocities in the supraglottal flow field overall decrease due to the decreased \dot{V} set in the experiment. Therefore, the bulk Mach number might not be the correct quantity to apply the fifth power law on. Instead, shown in fig. 7 B another theoretical computation $\eta_{theor,2}$ is computed, where the Mach number of the different experiments is computed by taking an average field velocity measured by PIV. Using this new scaling law gives a better agreement between theory and reality in most cases, while there still remain some notable differences in the cases $L \geq 600$ mm. This remaining difference opens up the question, if maybe the fifth power law is even applicable in this case of human phonation, as this is technically not a free turbulent flow but one bounded by the vocal folds as well as the vocal tract. Such walls are known to increase the efficiency of aeroacoustic sound production (Howe, 1998). Despite these uncertainties, the η -values found are still useful for a relative comparison between the configurations. The fact that the aeroacoustic efficiency still remains very large in the cases of f_o - f_{R1} -matching might also be explained by the fact that the acoustic back-coupling onto the flow increases the aeroacoustic sound generation in this case.

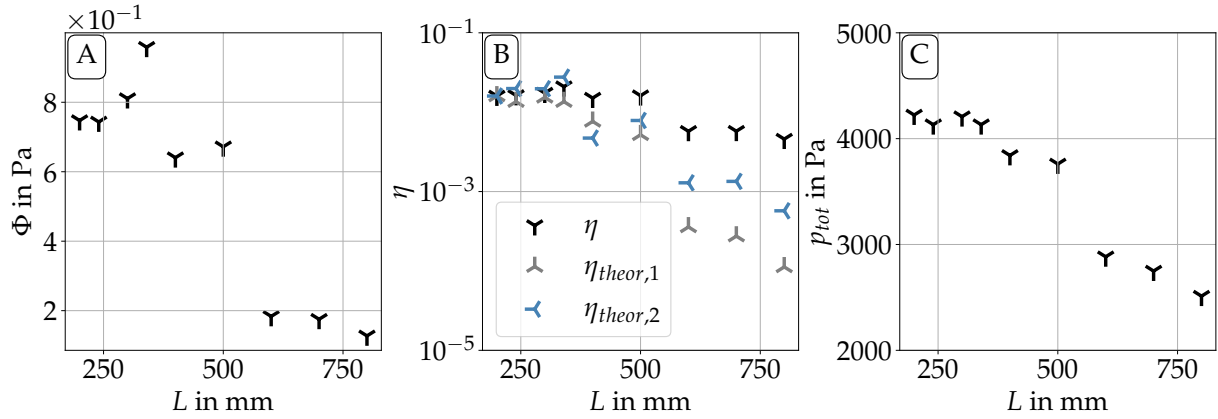


Figure 7. (A): RMS aeroacoustic source level Φ (B) Aeroacoustic efficiency η (C) total pressure p_{tot} of all vocal tract lengths L investigated.

4. Conclusion

The two-way interaction between fluid flow and acoustics in the human phonation process has been studied on a synthetic larynx model. PIV as well as wall pressure measurements have been performed. The vocal tract length was varied in the range of $L \in [170 \text{ mm}, 930 \text{ mm}]$ to study the relation between the supraglottal flow field and the acoustic resonances of the vocal tract. It could be seen that having f_o in vicinity of a resonance frequency leads to a tuning of f_o to the resonance frequency. Furthermore, a mode shift of the vocal fold oscillation induced by the acoustic standing waves in the vocal tract was found. Looking at the harmonic contents in the flow field revealed that acoustic anti-resonances led to a dampening of the higher harmonic frequencies of f_o in the wall pressure when they matched in frequency with the anti-resonances. This influence of the vocal tract acoustics could however not be found in the velocity fields measured by PIV. In the study of the aeroacoustic sound generation, it was found that especially in the cases of f_o - f_{R1} -matching, the aeroacoustic efficiency was unusually high comparing to common Mach-number scaling laws, leading to the conclusion that acoustic back-coupling onto the flow increases the efficiency in aeroacoustic sound generation. As the aeroacoustic sources according to the theory of vortex sound are solely computed from the measured velocity fields, an influence of the acoustics onto singular harmonic frequencies in the source field could not be found.

There are still some limitations to the current methodology of aeroacoustic source analysis that need consideration in the future. The vortex sound theory reduces the Lighthill source term to the principal sources of sound in the case of low Mach number flows (Kaltenbacher, 2015). An even better way of computing the aeroacoustic sources would however be making use of a classical perturbation ansatz like done in the PCWE acoustic analogy, completely separating hydrodynamics and acoustics (Kaltenbacher, 2018). For this, a time resolved pressure field would however be required, as the aeroacoustic sources in this approach are computed by the total differential of the

hydrodynamic pressure field w.r.t. time. This has been applied to 2D2C planar PIV measurements of the phonation process in combination with a Poisson-solver to obtain the pressure fields before (Lodermeyer et al., 2021). Here, some assumptions about the flow field had to be made however to obtain the pressure field. In the future, a more accurate approach could be to do three-dimensional tomographic PIV measurements to get rid of the assumptions and obtain a reliable pressure field for the computation of the pure aeroacoustic source field. Then, maybe also an influence of the acoustics back onto harmonic components of the aeroacoustic source field might be found. This would then lead to reduced sound generation in these harmonics, directly influencing the tone of the voice.

Acknowledgements

This work is funded by the German Research Foundation (DFG) through the project “Tracing the mechanisms that generate tonal content in voiced speech”. Project number: 446965891.

References

- Adatrao, S., & Sciacchitano, A. (2019, feb). Elimination of unsteady background reflections in PIV images by anisotropic diffusion. *Measurement Science and Technology*, 30(3), 035204. doi: 10.1088/1361-6501/aafca9
- de Luzan, C. F., Oren, L., Maddox, A., Gutmark, E., & Khosla, S. M. (2020, feb). Volume velocity in a canine larynx model using time-resolved tomographic particle image velocimetry. *Experiments in Fluids*, 61(2). doi: 10.1007/s00348-020-2896-x
- Durst, F., Heim, U., Ünsal, B., & Kullik, G. (2003, may). Mass flow rate control system for time-dependent laminar and turbulent flow investigations. *Measurement Science and Technology*, 14(7), 893–902. doi: 10.1088/0957-0233/14/7/301
- Fant, G. (1971). *Acoustic theory of speech production*. DE GRUYTER. doi: 10.1515/9783110873429
- Howe, M. S. (1998). *Acoustics of fluid-structure interactions*. Cambridge University Press. doi: 10.1017/cbo9780511662898
- Howe, M. S. (2002). *Theory of vortex sound*. Cambridge University Press. doi: 10.1017/cbo9780511755491
- Kaltenbacher, M. (2015). *Numerical simulation of mechatronic sensors and actuators: Finite elements for computational multiphysics*. Springer Berlin Heidelberg. doi: 10.1007/978-3-642-40170-1

- Kaltenbacher, M. (2018). *Computational acoustics*. Springer International Publishing. doi: 10.1007/978-3-319-59038-7
- Kraxberger, F., Näger, C., Laudato, M., Sundström, E., Becker, S., Mihaescu, M., ... Schoder, S. (2023, November). On the alignment of acoustic and coupled mechanic-acoustic eigenmodes in phonation by supraglottal duct variations. *Bioengineering*, 10(12), 1369. doi: 10.3390/bioengineering10121369
- Lighthill, M. J. (1952, mar). On sound generated aerodynamically i. general theory. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 211(1107), 564–587. doi: 10.1098/rspa.1952.0060
- Lodermeyer, A., Bagheri, E., Kniesburges, S., Näger, C., Probst, J., Döllinger, M., & Becker, S. (2021, nov). The mechanisms of harmonic sound generation during phonation: A multi-modal measurement-based approach. *The Journal of the Acoustical Society of America*, 150(5), 3485–3499. doi: 10.1121/10.0006974
- Lodermeyer, A., Becker, S., Döllinger, M., & Kniesburges, S. (2015, mar). Phase-locked flow field analysis in a synthetic human larynx model. *Experiments in Fluids*, 56(4). doi: 10.1007/s00348-015-1942-6
- Lucero, J. C., Lourenço, K. G., Hermant, N., Hirtum, A. V., & Pelorson, X. (2012, jul). Effect of source–tract acoustical coupling on the oscillation onset of the vocal folds. *The Journal of the Acoustical Society of America*, 132(1), 403–411. doi: 10.1121/1.4728170
- McGowan, R. S., & Howe, M. S. (2012, apr). Source-tract interaction with prescribed vocal fold motion. *The Journal of the Acoustical Society of America*, 131(4), 2999–3016. doi: 10.1121/1.3685824
- Mendez, M., Raiola, M., Masullo, A., Discetti, S., Ianiro, A., Theunissen, R., & Buchlin, J.-M. (2017, jan). POD-based background removal for particle image velocimetry. *Experimental Thermal and Fluid Science*, 80, 181–192. doi: 10.1016/j.expthermflusci.2016.08.021
- Migimatsu, K., & Tokuda, I. T. (2019, aug). Experimental study on nonlinear source–filter interaction using synthetic vocal fold models. *The Journal of the Acoustical Society of America*, 146(2), 983–997. doi: 10.1121/1.5120618
- Mittal, R., Erath, B. D., & Plesniak, M. W. (2013, jan). Fluid dynamics of human phonation and speech. *Annual Review of Fluid Mechanics*, 45(1), 437–467. doi: 10.1146/annurev-fluid-011212-140636
- Näger, C., Kniesburges, S., & Becker, S. (2023, January). Investigation of acoustic back-coupling in human phonation via particle image velocimetry. In *Proceedings of the 10th convention of the*

- european acoustics association forum acusticum 2023*. European Acoustics Association. doi: 10.61782/fa.2023.0987
- Näger, C., Kniesburges, S., Tur, B., Schoder, S., & Becker, S. (2023, November). An investigation of acoustic back-coupling in human phonation on a synthetic larynx model. *Bioengineering*, 10(12), 1343. doi: 10.3390/bioengineering10121343
- Oren, L., Khosla, S., & Gutmark, E. (2014, apr). Intraglottal pressure distribution computed from empirical velocity data in canine larynx. *Journal of Biomechanics*, 47(6), 1287–1293. doi: 10.1016/j.jbiomech.2014.02.023
- Schoder, S., Maurerlehner, P., Wurzinger, A., Hauser, A., Falk, S., Kniesburges, S., ... Kaltenbacher, M. (2021, mar). Aeroacoustic sound source characterization of the human voice production-perturbed convective wave equation. *Applied Sciences*, 11(6), 2614. doi: 10.3390/app11062614
- Schoder, S., Weitz, M., Maurerlehner, P., Hauser, A., Falk, S., Kniesburges, S., ... Kaltenbacher, M. (2020, feb). Hybrid aeroacoustic approach for the efficient numerical simulation of human phonation. *The Journal of the Acoustical Society of America*, 147(2), 1179–1194. doi: 10.1121/10.0000785
- Smith, B. L., Nemcek, S. P., Swinarski, K. A., & Jiang, J. J. (2013, may). Nonlinear source-filter coupling due to the addition of a simplified vocal tract model for excised larynx experiments. *Journal of Voice*, 27(3), 261–266. doi: 10.1016/j.jvoice.2012.12.012
- Sondhi, M., & Schroeter, J. (1987, jul). A hybrid time-frequency domain articulatory speech synthesizer. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(7), 955–967. doi: 10.1109/tassp.1987.1165240
- Story, B. H., Laukkanen, A.-M., & Titze, I. R. (2000, dec). Acoustic impedance of an artificially lengthened and constricted vocal tract. *Journal of Voice*, 14(4), 455–469. doi: 10.1016/s0892-1997(00)80003-x
- Titze, I., Riede, T., & Popolo, P. (2008, apr). Nonlinear source–filter coupling in phonation: Vocal exercises. *The Journal of the Acoustical Society of America*, 123(4), 1902–1915. doi: 10.1121/1.2832339
- Titze, I. R., & Story, B. H. (1997, apr). Acoustic interactions of the voice source with the lower vocal tract. *The Journal of the Acoustical Society of America*, 101(4), 2234–2243. doi: 10.1121/1.418246
- Valášek, J., & Sváček, P. (2024, August). Aeroacoustic simulation of human phonation based on the flow-induced vocal fold vibrations including their contact. *Advances in Engineering Software*, 194, 103652. doi: 10.1016/j.advengsoft.2024.103652

Zhang, Z., Neubauer, J., & Berry, D. A. (2006, sep). The influence of subglottal acoustics on laboratory models of phonation. *The Journal of the Acoustical Society of America*, 120(3), 1558–1569. doi: 10.1121/1.2225682

Zhang, Z., Neubauer, J., & Berry, D. A. (2009, apr). Influence of vocal fold stiffness and acoustic loading on flow-induced vibration of a single-layer vocal fold model. *Journal of Sound and Vibration*, 322(1-2), 299–313. doi: 10.1016/j.jsv.2008.11.009