

Acoustic measurement of periodic noise generation in a hydrodynamical vocal fold model

Malte Kob, Sebastian Krämer, Alice Prévot

Department of Phoniatrics, Pedaudiology and Communication Disorders,
RWTH Aachen University – University Hospital
D-52074 Aachen, Pauwelsstr. 30, Germany,
e-mail: {mkob,skraemer}@ukaachen.de, prevot@ensta.fr

Michael Triep, Christoph Brücker

Chair of Fluid Mechanics and Institute of Aerodynamics, Wüllnerstr. zw. 5 u. 7, D-52062 Aachen,
e-mail: {michaelt,bruecker}@aia.rwth-aachen.de

During phonation of voiced sounds periodic noise is generated due to the repeated evolution of shear layer vortices and turbulence downstream the glottis. In healthy persons, this noise sound integrates with the harmonic signal and forms a natural voice. In patients with voice disorders the amount of noise is often increased and contributes to the symptom of hoarseness. For experimental investigation of the noise generation process a 3:1 up-scaled model of mechanically driven vocal folds and the vocal tract has been built in a water channel. The periodic modulation of the flow is achieved by rotating cylinders of variable cross-section which allows to simulate the effect of voice disorders such as bi-phonation or an unilateral palsy. The part of the project presented herein deals with the direct acoustic measurement of the harmonic and non-harmonic components of the sound field to compare the acoustic output of this model to the voice components measured in humans during phonation. The set-up for acquisition of these acoustic signals is described and first results are presented.

1 Introduction

The contribution of aspiration noise in the voice signal is a significant indicator for the glottal conditions during human phonation. Unfortunately, a direct assessment of the noise generation process is not possible due to the effective protection of the voice organs in the larynx. In a joint project of the Department of Phoniatrics, Pedaudiology and Communication Disorders and the Chair of Fluid Mechanics and Institute of Fluid Mechanics of the RWTH Aachen University the spatio-temporal origin and distribution of vortices in different glottis configurations were investigated in a hydrodynamical in-vitro model with the method of particle-image velocimetry (PIV) and acoustical measurements.

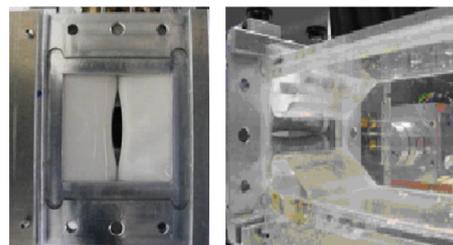


Figure 1: View on the hydrodynamic vocal fold model (left), and set-up with false vocal folds (right)

false vocal fold models is shown on the right. In figure 2 a set of four configurations of the two cams is shown for the simulation of four different phonatory conditions.

2 Vocal Fold Model

A hydro-mechanical model of the larynx has been built up in a 3:1 scale for detailed flow visualizations and measurements. The model consists of a sub- and supra glottal water channel with rectangular cross section. The vocal folds are modelled with two rotating cams of variable cross section. Over these cams a thin latex membrane is stretched such that a glottis-like opening between the cams is modulated by the cam rotation [BTK04]. In figure 1 on the left, the glottis model is shown from the supraglottal side. The white latex cover of the two cams is visible. The part of the set-up with the glottis and the

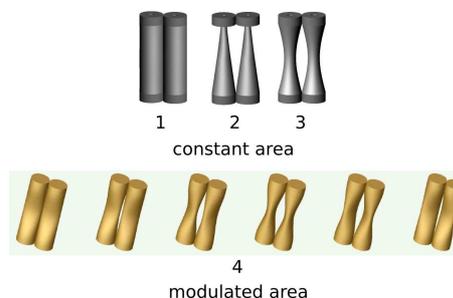


Figure 2: Different configurations of the two cams for the vocal fold simulation: closed vocal folds (1), breathing position (2), whisper position (3), and periodic oscillation (4).

Scale considerations The in-vitro model was 3-times scaled compared to the natural glottis size to reduce the characteristic flow velocities and frequencies such that time-resolved flow field measurements were possible. In addition, water was used as the carrier medium to reduce the refractive index differences between the channel walls for a better optical access into the inner flow regions in the glottis.

The dimensionless numbers to be conserved in this application are:

- the Reynold's number: $Re = \frac{uh}{\nu}$,
- the Strouhal's number: $St = \frac{fh}{u}$, and
- the Pressure coefficient: $c_p = \frac{\Delta p}{\frac{\rho}{2}u^2}$.

The conservation of the Mach-number could not be achieved simultaneous.

Taking into account these properties of air and water:

	Air	Water
Density [kg/m ³]	1.204	998
Viscosity [m ² /s]	15 · 10 ⁻⁶	10 ⁻⁶

the flow and frequency transformation can be estimated for a comparison of measurements in the hydrodynamical model (dashed variables) and actual values:

	Water
$\frac{u'}{u} = \left(\frac{h}{h'}\right) \left(\frac{\nu'}{\nu}\right)$	1/45
$\frac{f'}{f} = \left(\frac{u'}{u}\right) \left(\frac{h}{h'}\right)$	1/135

The frequency ratio (1/135) implies that in water, all frequencies are 135 smaller than in air. Thus, a sensor is needed for the detection of rather low frequency components in water. Table 1 sums up the differences between our model and a real vocal tract.

3 Aeroacoustic theory

For complex aeroacoustic problems in channel flows such as phonation a thorough theoretical description is not possible. During phonation the boundary conditions vary in space and time due to the self-excited vibrations of the folds. Therefore, the exact natural acoustic characteristics of the flow in the vocal tract could not be achieved up today. However acoustic theory allows to estimate the global properties which are applied herein to the hydrodynamic phonation model.

Since the mean flow velocity in the vocal tract is not zero, the linear approximation of the acoustic wave equation is

	Human (Air)	Model (Water)
Height of the trachea H [mm]	18 - 22	60
Length of the vocal folds L_g [mm]	14 - 20	48
Thickness of the vocal folds L [mm]	3 - 6	10
Maximal opening of the glottis h [mm]	1 - 3	8
Pressure difference upper the glottis ΔP [Pa]	500 - 2000	240 - 961
Flow speed in the glottis u [m/s]	20 - 40	0.44 - 0.88
Oscillation fundamental frequency f [Hz]	100 - 200	0.74 - 1.48

Table 1: Dimension and flow properties in the glottis

not valid. Thus, Lighthill's equation is used as a starting point. He developed a new form of the wave equation which retained the non-linear terms in the Navier-Stokes equation which reads:

$$\frac{\partial^2 \rho}{\partial t^2} - c^2 \frac{\partial^2 \rho}{\partial x_i^2} = \frac{\partial^2 T_{ij}}{\partial x_i \partial x_j} \quad (1)$$

where $T_{ij} = \rho u_i u_j + (p - c^2 \rho) \delta_{ij} - \tau_{ij}$ is the extension of the linear theory approximation in the Navier-Stokes equation. This term is called the Lighthill stress tensor. Under the assumption that the flow-induced sound emanates from a small region of fluctuating fluid flow acting on a much larger acoustic medium approximately at rest, Lighthill derived the theory that the force term on the right in (1) behaves as an external force acting on the uniform acoustic medium. Lighthill showed that the force term is equivalent to a distribution of quadrupole sources and that the acoustic power output from such a source varies with the eighth power of velocity ($u^3 M^5$, for velocity u , and Mach number, M).

Based on Lighthill's equations, Howe (1) [How75] derived an acoustic analogy which includes the effect of an irrotational mean flow (bulk flow of the acoustic medium) on sound propagation

$$\left\{ \frac{1}{c^2} \left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right)^2 - \nabla^2 \right\} B = \nabla(\omega \times v) \quad (2)$$

where B is a measure of total energy density, called stagnation enthalpy, ω is the vorticity and v is the eddy velocity. The term in brackets on the right hand side $\omega \times v$ is called Lamb vector Γ [Möh79, How75, Pow64] and represents the vortex sound source. Howe showed that this analogy which is valid for non-uniform mean flow fields at arbitrary Mach numbers can be transformed into that of Möhring. Howe used the Green's function to solve this

new wave equation and applied his solution to a contraction in a duct. He found a result in terms of the acoustic pressure which reads:

$$p = \frac{-\rho_0 A_1}{A_2(A_1 + A_2)(1 + M_1)(1 + M_2)} \int [\omega \times v \cdot \hat{U}] d^3y \quad (3)$$

where A_1 and A_2 are the duct areas upstream and downstream of the contraction, respectively, M_1 and M_2 are the Mach numbers of the stream flow at these locations, and \hat{U} is the unit vector in the direction of the mean flow velocity. This equation is only useful for potential flow. In the vocal flow, the potential flow approximation is only valid upstream of the smallest cross-section of the glottis, while further downstream, flow separation causes the creation of a free jet with shear layers at its lateral edges. The roll up of the shear layer into coherent vortex structures is one major source of noise generation when the vortices impact with solid surfaces.

Sinder [Sin99] used Howe's reformulation of Lighthill's theory in his study of vortex induced noise level in the vocal tract. Sinder applied this expression to a simple model of the shear layer vortices interacting with the false vocal folds. He considered a shedding of ring-type vortices with a shedding period proportional to the Strouhal number. He further assumed that the direction of the local vortex path velocity is always tangential to the surface along the obstacle. From these assumptions, and by using an approximation for the circulation based upon the roll-up time of the jet shear layer vortices, Sinder established an expression of the vortex-induced sound pressure level in the vocal tract downstream the vocal folds which reads:

$$p = \frac{-2\pi r_\omega \rho_0}{A} [\Gamma \times v \cdot \hat{U}] \quad (4)$$

where A is the duct area at the vortex impact location at the false vocal folds, r_ω is the radius of the vortex ring, and Γ is the circulation. This approximation is used herein to estimate the sound pressure level in the hydrodynamical model.

3.1 Application to the model

In order to estimate the noise level created by the vortices in our model, we try to approximate equation 4 to find a mean value of the pressure in the vocal tract. The unknown circulation is estimated from the shear layer roll-up of the axi-symmetric jet. A rough estimation yields $\Gamma = 0.5 \cdot U_j^2 T_{shed}$ in which U_j is the centre velocity of the jet. The shedding period T_{shed} is the inverse of the roll-up frequency which is determined from the Strouhal number of $Sr = \frac{1}{T_{shed}} \frac{D}{U_j} = 0.48$ for typical laminar, axi-symmetric jets.

The following assumptions apply: $U_j = 0.5 \text{ ms}^{-1}$, $A =$

$$1.8 \cdot 10^{-3} \text{ m}^2, \rho_0 = 1000 \text{ kg m}^{-3}, r_\omega = 0.005 \text{ m}, D = 0.01 \text{ m}.$$

Sinder assumes that the direction of the vortex velocity is tangential to the surface along the obstacle, which is the false vocal folds. We assume that the main contribution of the sound is produced in the downstream part of the false vocal folds, which have an opening angle of $\alpha = 35^\circ$, corresponding to the vortex path direction. The vortex drift velocity magnitude is roughly half the jet velocity: $v = 0.6 \cdot U_j$.

From these assumptions, for our model the following values can be calculated:

$$\begin{aligned} T_{shed} &= 4.16 \cdot 10^{-2} \text{ s} \\ \Gamma &= 5.2 \cdot 10^{-3} \text{ m}^2 \text{ s}^{-1} \\ v &= 0.3 \text{ ms}^{-1} \end{aligned}$$

With these assumptions, a rough value of the mean sound pressure created by one vortex in the shear layer interacting with the false vocal fold can be calculated:

$$p = 27 \text{ Pa} \approx 122 \text{ dB re. } 2 \cdot 10^{-5} \text{ Pa} \quad (5)$$

This rather high sound pressure would be produced by one vortex which interferes with the false vocal folds. Since several vortices are created within the life time of one vortex, a non-coherent addition of the sound from co-existing vortices must be performed. A simulation of the noise contribution has been implemented in a computational voice model [Kob02]. In normal phonatory conditions a simultaneous contribution from about 20 vortices was observed which corresponds to an increase of approximately 13 dB.

4 Methods

4.1 PIV measurements

First results from particle velocimetry measurements have been previously published [BTK04]. They indicate that significant vorticity can be observed, and differences between measurements with and without false vocal folds can be visualised. In figure 3 an analysis of a PIV measurement is shown. From these measurements we determined the spatio-temporal behaviour of these vortices and their interaction with the false vocal folds.

4.2 Sound pressure measurements

Acoustic measurements of the sound pressure at the location of the vortex production can give a quantitative information about the periodic sound and the noise content. One part of voiced sounds is due to the harmonic oscillations of the vocal folds, and another is due to noise produced by the vortices created in the vocal tract by the flow passing through the vocal folds.

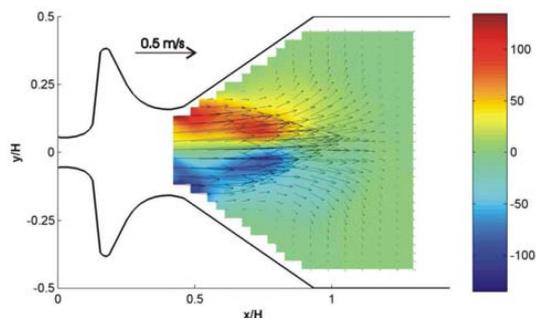


Figure 3: PIV analysis

For a comparison between the noise production in the model and the noise generation during human phonation, SPL measurements in the model and of the human voice were performed.

4.2.1 Selection of the sensors

The measurement of the sound field in the water channel requires the application of special sensors since the specifications of sensors used for sound field measurements in air usually do not comply with the conditions under water. Therefore, finding sensors that work under the given circumstances is problematic. The evolving problems are:

- the constant pressure level is high so the sensor needs to operate at a high stationary pressure
- because of the underwater situation corrosion is harmful to the sensor
- to avoid corrosion the sensor must be covered, which implies that it is surrounded by a layer which is made of a non-corrosive, non-conductive medium. This medium changes the frequency characteristics and overall sensitivity of the sensor

For the acoustical measurements different sensors were used. In first tests, four kinds of sensors have been tested: capsuled miniature air pressure microphones, commercial and self-made hydrophones and a special sound velocity probe (microflow). In the set-up only a miniature sensor could be used which is sensitive for small pressure changes and which does not disturb the flow too much. The hydrophones fulfilled these requirements. The best results were obtained with the hydrophone made by Brüel & Kjær Model 8103 with a sensitivity of $-211.6 \text{ dB re } 1 \text{ V}/\mu\text{A}$.

4.2.2 Measurement set-up

Several parallel measurements were taken to achieve a separation of the sound generated by the vortex dynamics

and turbulence from the always present background noise created from the set-up. For an estimation of the background noise level and frequencies in the set-up room, simultaneous recordings of the sound signals were taken in the water channel and out of the channel. A first set of recordings was done with the flow channel at rest, and then with only the cam motor switched on without rotation. A second set of measurements was done to measure the vortex induced sound with different forms of the vocal folds. All measurements were performed twice: once with the cams at rest and then with rotating cams as shown in fig. 2. Since the cams configurations 1-3 in figure 2 always have cams with constant radius along the circumference, there is no difference in the glottis cross-section if the cams rotate or not.

1. Closed vocal folds: The cams are full cylinder with constant cross-section in contact so that the flow does not pass through
2. Wide opened vocal folds (breathing position): The cams are conical, the flow is not modulated.
3. Partially opened vocal folds: The cams are cylindrical but with a smaller radius in the middle, the flow passes continually with higher volume velocity. No flow modulation occurs.
4. Rotating vocal folds (phonation position): The cams are cylindrical at both ends of each cam and have a semi-elliptical cross-section in the middle such that along the circumference, the radius is smaller on one side. The cams are rotating and hence create a phonation cycle, the flow is periodically modulated.

With the cam configuration (1) corresponding to the closed vocal folds we get an estimate of the background noise including the cam motor and any other noise sources not related to flow within the glottis. The noise created by the glottis during breathing is measured with the configuration (2) with conical cams creating a large orifice area and therefore yielding only relative low velocities and low turbulence levels within the vocal tract. By comparing measurements with cams of different opening areas, we expect to find differences in noise generation depending on the phonatory conditions.

4.2.3 Sensor placement

Figure 4 shows three different sensor positions. Position 1 is located as a reference position upstream in the uniform laminar inflow region. Position 2 is the main recording position, derived from the PIV measurements where the highest fluctuations of vorticity are expected. Position 3 is used as downstream reference measurement where already most of the fluctuations are died out and uniform flow is re-established.

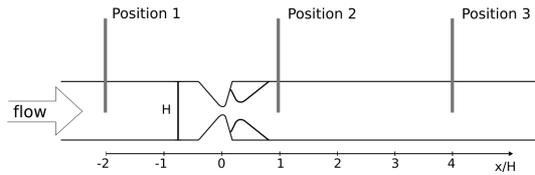


Figure 4: Positions of sensors

The results given below were taken with the B&K hydrophone in position 2 downstream of the false vocal folds.

5 Results

5.1 Voice analysis

First, sound pressure level measurements were taken for a natural human voice as a reference. The human sustained vowel /a:/ was recorded in 20 cm distance from the mouth of a healthy, male subject. By using the pitch-scaled harmonic filter (PSHF) algorithm by Jackson [JS98] the harmonic part is separated from the noise part in the voice signal. The PSHF analysis gave a harmonic/noise ratio of about 30, what means that for the signal of 55 dB SPL, an upper limit of the noise content would be approximately 25 dB SPL or less, since the sound at the glottis is significantly attenuated on its way to the microphone.

With the parameters given in table 1, the frequency distribution of the aspiration noise can be found in a broad range between 500 and 3000 Hz [Ste71]. Corresponding to our in-vitro model, this would translate to frequencies of the order of 5-22 Hz.

5.2 Background noise

A difficult task is to isolate the sound of the vortices and turbulences which are generated at the glottis from the different background noise sources (vibrations induced by the cam drive, water noise at the in- and outlets of the water channel, noise due to non-laminar flow outside the glottis section, electrical noise interfering with the measurement set-up etc.). In order to reduce the sound which is generated by the power supply, the motor and the gearbox, sound absorbing boxes were built around these sound sources.

Measuring conditions	dB(A)	dB(C)
Nothing	38	48
Power supply	39	49
Motor without axle gear	43	51
Motor with axle gear	46	53

In table 5.2 the results from SPL measurements are shown. Noise levels in air at 20 cm distance from the glottis model, in dB(A) and dB(C), with the indicated sources of background noise were measured.

5.3 Measurements with constant flow

5.3.1 Triangular cams

In figure 5 the spectra of the noise within the water channel with and without flow (flow velocity $v = 1.085 \text{ ms}^{-1}$) are shown. In the expected range between 0 and

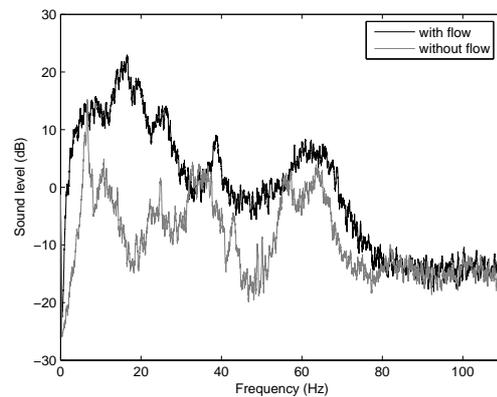


Figure 5: Comparison of the signal with the triangular cams with and without flow, without motor

80 Hz the noise increase due to the flow through the glottis can be seen, with a peak at 20 Hz. With the difference between figures 5 and 6 we can clearly see that even with maximum damping of the motor sound its noise is observable in the spectrum. In figure 5 we can see a differ-

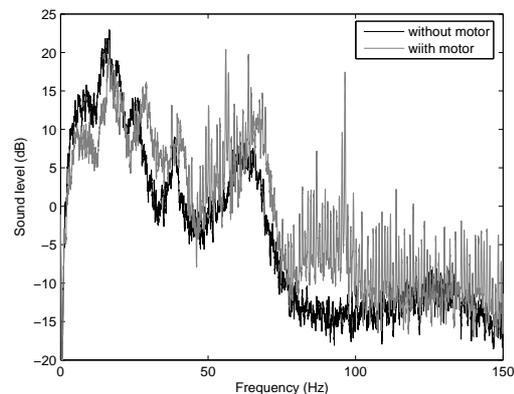


Figure 6: Comparison of the signal with the triangular cams with and without motor, with flow

ence in the noise around 20 Hz, this is where we expect

this noise. Since the aspiration noise components in voice are expected between 2000 and 5000 Hz, the corresponding frequencies in water are between 15 and 40 Hz. This noise is present in both cases with or without motor, because the opening of the glottis with these cams is the same when its rotating and when it is not.

5.3.2 Phonation cams

Some preliminary measurements have been performed with the phonation cams in static positions. Without rotation, we compared the signals when the flow goes through wide opened and half opened cams. In figure 7 we ob-

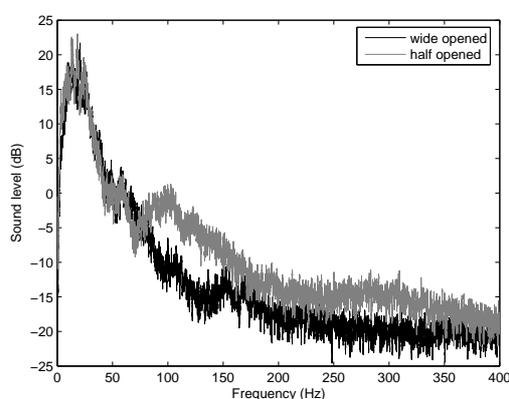


Figure 7: Comparison of the signals with the phonation cams, with stationary flow, without motor

serve that with the half opened cams the noise is significantly louder at frequencies between 100 and 150 Hz. An increase in noise is expected with the half opened cams because there might be more and stronger vortices than with cams wide opened due to the higher volume flow.

6 Discussion

The aim of our study is a comparison of the vortex induced sound in a hydrodynamical model of the human larynx and vocal tract to a theoretical approach for the noise generation in a duct. Acoustic measurements of the noise spectrum were performed in parallel to optical flow field measurements to detect and quantify the vortex dynamics in the model. Significant differences in the noise pattern could be observed through comparison of different glottal conditions. The dependence of the mean frequency range matches well the theoretical values. However, the sound pressure levels are not comparable.

Future work will concentrate on the separation of the pulsed noise from the background noise sources. This

will be achieved through simultaneous measurement in water at several positions (see figure 4) and in air.

7 Acknowledgements

This project was kindly supported by the DFG funds BR 1494/4-1 and KO 2874/1-1.

References

- [BTK04] C. Brücker, M. Triep and M. Kob, *Study of the vortex dynamics in a mechanical model of the vocal folds using Particle-Image Velocimetry*, Proceedings of the International Conference on Voice Physiology and Biomechanics in Marseille, France, 2004, pp. 11–17.
- [How75] M. S. Howe, *Contributions to the Theory of Aerodynamic Sound, with Application to Excess Jet Noise and the Theory of the Flute*, J. Fluid Mechanics **67** (1975), 597–610.
- [JS98] P. J. B. Jackson and C. H. Shadle, *Pitch-synchronous Decomposition of Mixed-source Speech Signals*, Proc. Joint Int. Cong. on Acoust. and Acoust. Soc. Am. (Seattle, WA 1), vol. 1, ICA, 1998, pp. 263–264.
- [Kob02] M. Kob, *Physical modeling of the singing voice*, Ph.D. thesis, Aachen University (RWTH), 2002.
- [Möh79] W. Möhring, *Modelling low Mach number noise*, Springer, 1979.
- [Pow64] A. Powell, *Theory of vortex sound*, J. Acoust. Soc. Am. **36** (1964), no. 1, 177–195.
- [Sin99] D. J. Sinder, *Speech Synthesis Using an Aeroacoustic Fricative Model*, Ph. D. thesis, University of New Jersey, 1999.
- [Ste71] K. N. Stevens, *Airflow and Turbulence Noise for Fricative and Stop Consonants: Static Considerations*, J. Acoust. Soc. Am. **50** (1971), 1180–1192.