A combined FE and multiple-mass model for numerical simulation of phonatory maneuvers

Andreas Gömmel, Christoph Butenweg
RWTH, Department of Structural Statics and Dynamics,
D-52074 Aachen, Mies-van-der-Rohe-Str. 1,
e-mail: **{goemmel,butenweg}@lbb.rwth-aachen.de**,

Sebastian Krämer, Malte Kob
Klinikum der RWTH, Department of Phoniatrics, Pedaudiology and Communication Disorders,
D-52074 Aachen, Pauwelsstr. 30, Germany,
e-mail: **{skraemer,mkob}@ukaachen.de**

The accuracy and realism of numerical models for voice synthesis is limited by incomplete understanding and hence implementation of physical and physiological details of phonation. One aspect is the missing knowledge of the dynamics of the intra- and extralaryngeal muscles that determine the boundary conditions for the self-sustained vocal fold oscillation. Whereas in-vivo measurements of the dynamic and elastic parameters of most laryngeal muscles is not possible, a finite element approach can be useful to determine geometrical and force values of the muscles which contribute to the movement of the vocal folds for specific phonatory tasks, such as adduction, abduction or register changes. The output of the FE-model controls the boundary conditions of a second model, which simulates the oscillation of the vocal folds, represented by a multiple mass model (VOX). The application of the combined models to the simulation of different phonatory maneuvers is demonstrated including visualization of the laryngeal cartilage movement and auralization of the corresponding sound output.

1 Introduction

For a better understanding of human phonation divers approaches are made. One way is the observation of phonatory processes with the help of imaging methods such as laryngoscopy or magnetic resonance imaging (MRI). Another way is the simulation of phonation by physical models, which represent the functional components of the voice organ.

Multiple-mass models provide good visual and acoustical results of the vocal fold oscillation. However, the control of such models is very difficult due to the huge number of control parameters. On the other hand, the boundary conditions for the positioning of the vocal folds for phonatory manœuvres such as adduction or abduction is determined by the forces of a rather small number of intra- and extralaryngeal muscles. To close this gap between vocal fold oscillation and control, a step towards a complete model of phonatory manœuvres consists of bringing together the vocal fold models and the muscle controlled movements of the laryngeal cartilages [9].

In this study, a finite element model of the intralaryngeal cartilages and muscles is combined with a multiple-mass model of the oscillating vocal folds. The links between both of the models are the position coordinates of the attachment points between vocal folds and cartilages on one hand, the other parameter which links both models is the tension of the vocal folds.

2 Methods

2.1 Finite-element model

The developed finite-element model consists of the cricoid cartilage, both of the arytenoid cartilages, the intralaryngeal muscles, and two link elements modeling the vocal folds. Since the main attention was drawn to the muscle control, the cartilages were dimensioned to fit the muscle lengths according to literature [1, 3, 4, 7, 13, 19]. Other necessary lengths and shapes were taken from general anatomical literature [16, 18]. The cartilages have been modeled as stiff beam elements (cricoid cartilage) and shell elements (arytenoid cartilages). An anatomical sketch and the model are given in figure 1.

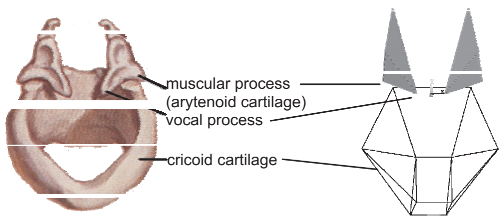

Figure 1: Left side: Anatomical sketch of the cricoid and arytenoid cartilages [15]. Right side: Finite-element model of cricoid and arytenoid cartilages.

The following muscles were included (see figure 2 for their positions):

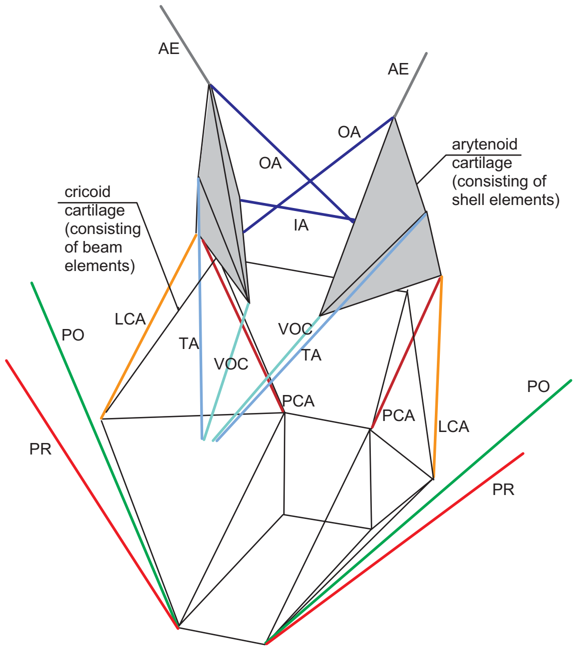

Figure 2: FE-model with all the implemented muscles (view in medial-caudal direction).

1. Cricothyroid muscle (CT), devided into pars recta (PR) and pars obliqua (PO)
2. Lateral cricoarytenoid muscle (LCA)
3. Posterior cricoarytenoid muscle (PCA)
4. Transverse arytenoid muscle (IA for intra-arytenoid, only unpaired muscle)
5. Oblique arytenoid muscle (OA)
6. Aryepiglottic muscle (AE)
7. Vocalis muscle (VOC)
8. Thyroarytenoid muscle (TA)

Congruent to the vocalis muscle, a pair of nonlinear springs has been implemented which include a strain-force relation according to Kob [11]:

$$F \;=\; -\,th\;d\;\tau\,\ln\,(1-\frac{\varepsilon}{\varepsilon_{max}}), \qquad (1)$$

with F being the force in the vocal fold, ε being the actual strain, ε_{max}=0.3 being the maximum strain, th=1.0 mm being the thickness, and d=1.0 mm being the width of the vocal folds. The value of τ was defined to 52.3 kPa.

To represent the laryngeal muscles, a two-node link element fitting to the finite-element program ANSYS [2] was programmed and implemented (see fig. 3). The formulation of the finite element consists of an inner contractile (CE) and a parallel element (PE) whereas the CE represents the active and the PE the passive behavior of the muscle. Each node of the finite element has three degrees of freedom (displacements in direction of X, Y, and Z).

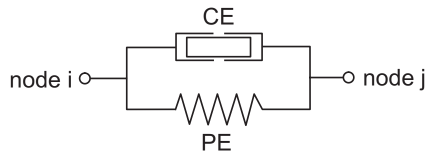

Figure 3: Schematical sketch of the programmed two-node finite element with its two components “contractile element” (CE) and “parallel element” (PE).

Most of the equations of the CE were implemented according to Günther [6]. The main parameters which have to be defined for the individual muscle are the optimum length L_{opt} and the maximum force P_0. The optimum length of a muscle is the length at which the muscle can generate its maximum active force. The active force depends on the muscle’s velocity and length. The force of the CE and its stiffness are calculated with respect to its activation level (act, see Equation 2). The main advantage of controlling the muscle activity by its stimulation rate is the possibility of deriving the stimulation rate directly from EMG (electromyography) measurements.

$$\dot{act}(t) \;=\; (stim(t)+k(t)\cdot act(t))/\tau_{act}, \qquad (2)$$

with:

$k(t)$	$= -stim(t)+stim(t)\cdot\beta-\beta$
act(t):	activation
stim(t):	stimulation
τ_{act}:	time constant of full stimulation (0.02)
τ_{deact}:	time constant of full deactivation
β:	$=\tau_{act}/\tau_{deact}$ constant (0.2)

The passive comportment of the muscle was adapted taking Alipour and Titze [1] as a source. The assumption was made that the cross-sectional area of a muscle is proportional to both, the active and the passive force. If a significant non-analogy of active and passive force is measured in future experiments, the factor P^{PE}_{max} was added to be able to adapt the element behavior. Passive force P_{PE} is calculated as follows:

$$\begin{aligned} P_{PE} \;&=\; P^{PE}_{max}\cdot P_0\cdot(0.003296\,+ \\ &\quad +\,0.5972\varepsilon-3.8444\varepsilon^2+10.87\varepsilon^3) \end{aligned} \qquad (3)$$

The muscles’ input values (P_0, L_{opt}, etc.) were calculated according to literature. In many cases—like the prestress of the PCA—assumptions had to be made.
Because of the muscles’ possibility to shorten very highly, large deformations have been taken into account. The mass of the muscle is lumped with half of the mass on each node. Non-linear stress-stiffening was respected

by a geometric stiffness matrix which is calculated by the ratio of the muscle forces and its length.
To describe phonatory manoeuvres, the cricoarytenoid joint has to be scrutinized. The joint capsule, the ligamentous attachments, and the joint facies are responsible for the complex parts of motion. These were modeled consisting of a sliding, a rocking and a rotational part (according to Kasperbauer [8]). The chosen way to limit the motions was using additional beam elements. The connection between cricoid and arytenoid cartilages consists of five beams that model the joint movements plus a structure of auxiliary beams to fix the "joint" at the cricoid cartilage. The **gliding movement** is ensured by two long beams (red in figure 4) that have high torsional and bending stiffness, but only low tension and compression stiffness. Between the node in the middle of these two beams and a node in the center on the bottom of the arytenoid cartilage three congruent beams with different cross-section values were implemented orthogonally to represent the rotational parts of movement (green in figure 4). The **rocking movement** is modelled with a beam of high stiffness in tension and compression, low torsional and bending stiffness around the rotation axis and a defined bending stiffness of 0.0096 Nm/rad around the rocking axis. Another beam is fitted with high stiffness in tension and compression, low bending stiffnesses (against rocking) and a rotational stiffness of 0.070 Nm/rad to describe the **rotational movement**. A last beam prevents the arytenoid cartilage from bending around the axis orthogonally to the rotation and the rocking axes by high bending stiffness around this axis and low stiffness around all the other axes.
The ranges were interpolated linearly to fit to the finite-element program in which equations of movement (see equation 4), which connects global degrees of freedom, can be applied.

$$\begin{aligned} 1.0\, u_{ventral} &= 2.411\, \varphi_{rotation} \\ 1.0\, \varphi_{rotation} &= 1.175\, \varphi_{rocking} \end{aligned} \tag{4}$$

with u_i being a displacement and φ_i being a rotation. This equation has been applied to the node which connects the beams which model rocking and rotation with the arytenoid cartilage. Figure 4 shows a lateral-posterior view of the model. The auxiliary beams which hold the joint beams are colored black. This solution assures an future integration into an extralaryngeal model because any restriction of movement is independent from a global coordinate system. Due to the small number of elements a fast and stable transient dynamic calculation run can be performed. Each simulation generates a file with the displacements of the attachments of the vocal folds and their tension stresses in discrete time steps as input file for the multiple-mass model.

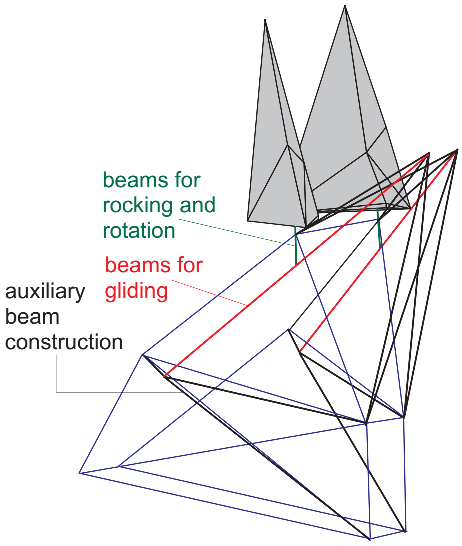

Figure 4: Finite-element model with auxiliary beams structure for the rocking, gliding, and rotating movements (view in medial-ventral direction).

2.2 Multiple-mass model

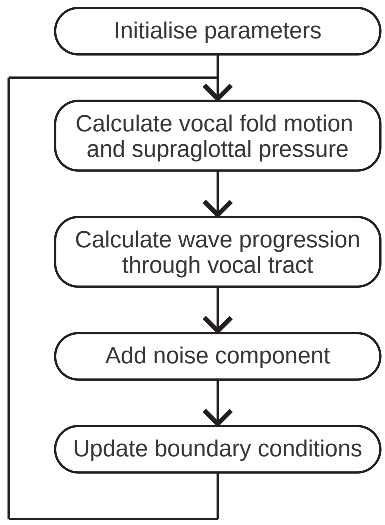

Figure 5: Signal flow of the combined model VOX [11].

The complete voice production process can be understood as the interaction of the functional components such as the respiratory apparatus, the tone generating vocal folds, and the sound encoding vocal tract. Numerical modeling of these components in the time-domain allows an interactive simulation of the voice generation process [11]. With such a model the vocal fold movement and the values of the aerodynamic and Newton forces within the laryngeal muscles can be visualized, and the related acoustic output can be made audible.

The computer program VOX includes a multiple-mass vocal fold model and modules for noise generation, wave propagation through the vocal tract, and radiation at the mouth (see figure 5). When the program was developed at first, only basic control mechanisms were implemented, with the aim to simulate main features of the singing voice [10].

For the modelling of the voice generation, the most important part of the complete model is the design of the vocal fold function. The used model is based on the 16-mass model developed by I. R. Titze [20]. However, it includes some profound modifications which are based upon recent research [14, 17]:

- Modelling of different voice registers
- Variation of individual masses or spring stiffness in an arbitrary number of sections
- Numerous parameters related to the phonatory organ can be modified during calculation
- The glottal movement and arbitrary pressure or flow values can be monitored during calculation

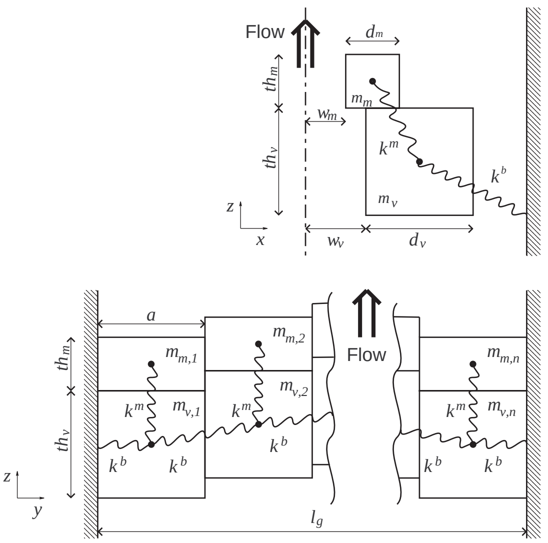

Figure 6: Sectional view (top) and side view (bottom) of the implemented vocal fold model [11].

The arrangement of n masses $m_{m,i}$, with $i = 1..n$ represents the vocalis muscle, $m_{v,i}$ represents the mucosa membrane. Each of the big masses $m_{v,i}$ is connected to the boundary by a spring with stiffness k^b and damping D^b, and to the small mass $m_{m,i}$ by a spring with stiffness k^m and damping D^m.

The most important difference between this model and classical two-mass models is the segmentation of the VF in n segments. In Titze's model, the number of 16 masses

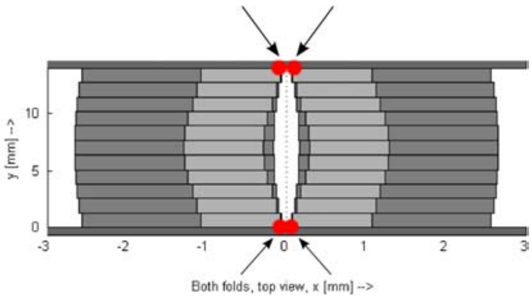

Figure 7: The marked points are where the FE-model and the multiple-mass-model connect

is fixed. The sum of the arbitrary number of n segments of width a is the longitudinal length l_g of the VF. There are no springs between the border and the small masses m_m, which is more close to the actual configuration of the vocal folds, because the mucosa membrane is not directly attached to the *cricoid cartilage*.

2.3 Interface

The FE-model produces a text file with the calculated 3D-positions of the four vocal folds edges and the stress on the folds. The sampling rate of the value sets is 10 ms. Currently, this output is loaded off-line into the vocal fold module of VOX, which requires a subsequent execution of the programs. During calculation of the vocal fold oscillation, the actual set of values is used to update the boundary conditions.

3 Results

3.1 Muscle forces

Regarding the muscle forces, the course over time is more important than the absolute values since most of them are based on assumptions and can be adapted easily when new data is got. Abduction and adduction were simulated with two different angles of the cricoid cartilage ("lp" stands for a resting angle (0^o), "hp" for an angle of 15^o). All simulations started and ended in the intermediate position (IP). The exact stimulation rate pattern was:

⟨1⟩ IP, lp (0.0-0.1s) → abduction, lp (0.1-1.1s) → IP, lp (1.1-1.2s) → adduction, lp (1.2-2.2) → IP, lp (2.2-3.0s)

⟨2⟩ IP, lp (0.0-0.1s) → IP, hp (0.1-0.3s) → abduction, hp (0.3-1.3s) → IP, hp (1.3-1.4s) → adduction, hp (1.4-2.4s) → IP, lp (2.4-3.0s)

Figure 8 shows the muscle forces of PCA, LCA, TA, CT, and OA in both of the simulations. The angle (in caudal view) between the vocal folds as indication of the resulting movement ist given, too. The results of the simulation without moving the cricoid cartilage is shown by the line with the bold line. The thin line shows the results of the simulation in which the cricoid cartilage was turned 15^o. The maximum forces were 2.8 N (PCA, TA, OA), 0.35 N (LCA), and 4.5 N (CT).

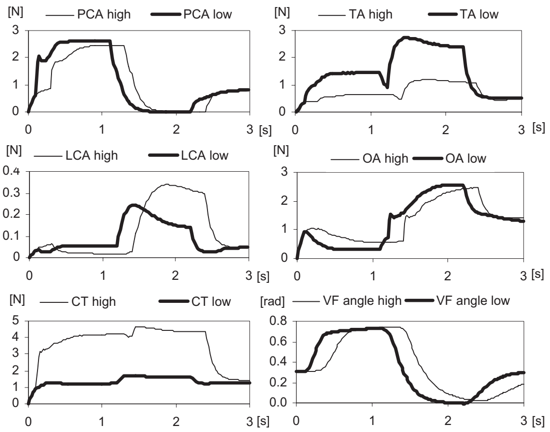

Figure 8: Forces of PCA, LCA, TA, CT, and OA and angle between the vocal folds (in caudal view) during simulations ⟨1⟩ (bold line) and ⟨2⟩ (thin line).

3.2 Vocal fold tensions

A precondition to validate the movements is a realistic tension in the vocal folds. The tension (see equation 1) was examined with the following stimulation rate pattern:

⟨3⟩ IP, lp (0.0-0.1s) → adduction, lp (0.1-1.1s) → adduction, hp (1.1-2.1s) → adduction, lp (2.1-3.1) → IP, lp (3.1-4.0s)

⟨4⟩ IP, lp (0.0-0.1s) → IP, hp (0.1-0.3s) → adduction, hp (0.3-1.3s) → adduction, lp (1.3-2.3s) → IP, lp (2.3-3.0s)

As results tensions between 40 and 160 kPa were obtained during an adducted vocal fold position. In figure 9 the stimulation rate of the CT, the angle of the cricoid cartilage, the angle between the vocal folds in caudal view, and the tension of the vocal folds can be seen.

3.3 Acoustical output

The sound output of the multiple mass model, together with other mechanical and acoustical values, are stored after calculation. The sound produced at the glottis or at the mouth is saved as a wav file.

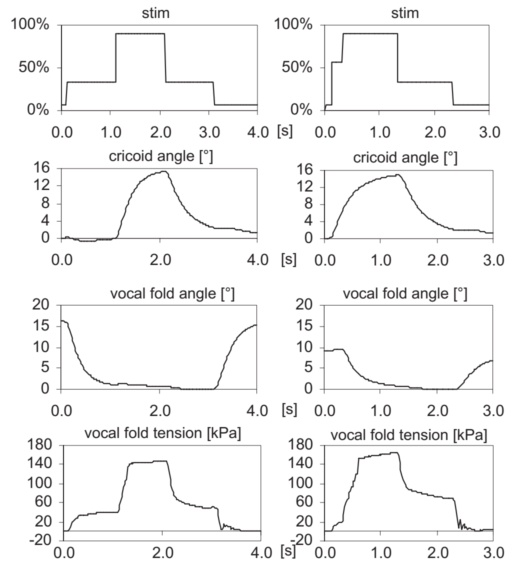

Figure 9: Stimulation rate of the CT, angle of the cricoid cartilage, angle between the vocal folds (in caudal view), and tension in the vocal folds during simulation ⟨3⟩ (left side) and ⟨4⟩ (right side).

Figure 10 shows the vocal fold tension (top) and a sonagram (bottom, produced with [12]) of the sound at the mouth for the sustained vowel /a:/, produced by the subsequent calculations with the FE-model in ANSYS and the VOX program.

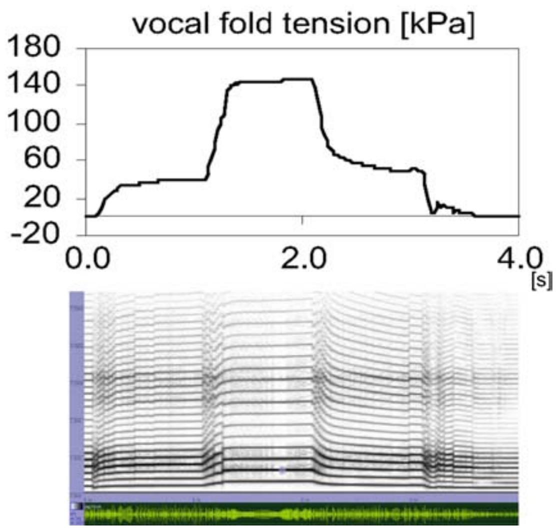

Figure 10: Sonagram and vocal fold tension of the simulated sound

The pitch change due to the change of the vocal fold tension can be read from the sonagram image.

4 Discussion

The obtained stimulation sets show that according to literature the PCA is the main abductor muscle and LCA, IA, OA, and—partially—TA are the main adductor muscles. Arranging the sets in series to get a motion sequence can only be an intermediate step an the way to a realistic simulation of muscle controlled abduction and adduction. In reality, the stimulation is effected by the nerve supply which—at the same time—controls the accurateness of the movement. Thus, in any step of the movement, the stimulation is individually adapted for each muscle.

An important next step is the improvement of the model by applying muscle data from further experiments on *all* laryngeal muscles of human specimen as input of the finite muscle elements. Canine data can be an important indication but although its comparable size values from human larynges are necessary because the completely different anatomy of the canine neck.

Our formulation of the cricoarytenoid joint is sufficient for normal phonatory manoeuvres. Other positions like a whispering position are not possible to adjust. Simulating such positions with finite-elements could be done by modeling accurately *all* the real limiting structures (particularly the cricoarytenoid ligaments) and the joint facies including a contact problem simulating the gliding motions between them.

References

[1] Alipour, F. and Titze, I. R., 1999, Active and Passive Characteristics of the Canine Cricoarytenoid Muscles. *J. Voice*, **13**, 1–10.

[2] ANSYS 9.0, 2004, *FE-Software, SAS IP Inc.*

[3] Cox, K. A., Alipour, F. and Titze, I. R., 1999, Geometric Structures of the Human and Canine Cricoarytenoid and Thyroarytenoid Muscles For Biomechanical Applications. *Ann. Otol. Rhinol. Laryngol*, **108**, 1151–1158.

[4] Farley, G. R., 1994, A biomechanical laryngeal model of voice F0 and glottal width control. *J. Acoust. Soc. Am.*, **100**, 3794–3812.

[5] Fung, Y. C.,1993, *Biomechanics.*, (New York: Springer, 2nd edt.).

[6] Günther, M., 1997, Computersimulation zur Synthetisierung des muskulär erzeugten menschlichen Gehens unter Verwendung eines biomechanischen Mehrkörpermodells. *PhD thesis, EKU Tübingen*.

[7] Hunter, E. J., Titze, I. R., and Alipour, F., 2004, A three-dimensional modal of vocal fold abduction/adduction. *J. Acoust. Soc. Am.*, **115**, 1747–1759.

[8] Kasperbauer, J. L., 1998, A biomechanical Study of the Human Cricoarytenoid Joint. *Laryngoscope*, **108**, 1704–1711.

[9] Kob, M. and Butenweg C., 2004, A finite element model of the interaction between intra- and extralaryngeal muscles. *Proceedings of the International Conference on Voice Physiology and Biomechanics in Marseille, France*, 91-92.

[10] Kob, M., 2004, Singing voice modelling - as we know it today. *Acta Acustica united with Acustica*, **90**, 649-661.

[11] Kob, M., 2002, Physical modeling of the singing voice. *PhD thesis, RWTH Aachen.*

[12] Lauer, C., 2002, Sonogram – Acoustical Frequency Analysis Toolset, URL: http://www.dfki.de/˜clauer/software.html

[13] Mineck, C. W., Tayama, N., Chan, R., and Titze, I. R., 2002, Three-Dimensional Anatomic Characterization of the Canine Laryngeal Abductor and Adductor Musculature. *Ann. Otol. Rhinol. Laryngol.*, **109**, 505–513.

[14] Pelorson, X., Hirschberg, A. , van Hassel, R. R., Wijnands, A. P. J. and Auregan, Y., 1994, Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model. *J. Acoust. Soc. Am.*, **96**, 3416-3431.

[15] Sataloff, R. T., 1997, *Professional Voice, The Science and Art of Clinical Care.* (Singular Publishing Group)

[16] Selbie, W. S., Gewalt, S. L., and Ludlow, C. L., 2002, Developing an anatomical model of the human laryngeal cartilages from magnetic resonance imaging. *J. Acoust. Soc. Am.*, **112**, 1077–1090.

[17] Sinder, D. J., 1999, Speech Synthesis Using an Aeroacoustic Fricative Model. *Ph.D. Thesis, University of New Jersey*.

[18] Sobotta, J., 1997, *Atlas der Anatomie.* (CD-ROM version: Urban & Schwarzenberg)

[19] Titze, I. R., Coupling of Neural and Mechanical Oscillators in Control of Pitch, Vibrato and Tremor. In: Davis, P. J. and Fletcher N. H. (Eds) *Vocal Fold Physiology: Controling Complexity and Chaos*, (San Diego: Singular Pub. Group), pp. 47–62.

[20] Titze, I. R., 1973, The Human Vocal Cords: A Mathematical Model, Part I. *Phonetica*, **28**, pp.- 129-170