

RESEARCH ARTICLE

Biomechanical simulation of vocal fold dynamics in adults based on laryngeal high-speed videoendoscopy

Michael Döllinger^{1*}, Pablo Gómez¹, Rita R. Patel², Christoph Alexiou³, Christopher Bohr¹, Anne Schützenberger¹

1 Division of Phoniatrics and Pediatric Audiology, Department of Otorhinolaryngology, Head and Neck Surgery, Medical School, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany,

2 Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana, Indiana, United States of America, **3** Section of Experimental Oncology and Nanomedicine (SEON), Department of

Otorhinolaryngology, Head and Neck Surgery, Medical School, Else Kröner-Fresenius-Stiftung-Professorship, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany

* Michael.doellinger@uk-erlangen.de



OPEN ACCESS

Citation: Döllinger M, Gómez P, Patel RR, Alexiou C, Bohr C, Schützenberger A (2017) Biomechanical simulation of vocal fold dynamics in adults based on laryngeal high-speed videoendoscopy. PLoS ONE 12(11): e0187486. <https://doi.org/10.1371/journal.pone.0187486>

Editor: Sean D. Peterson, University of Waterloo, CANADA

Received: August 15, 2016

Accepted: October 18, 2017

Published: November 9, 2017

Copyright: © 2017 Döllinger et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available in the paper and its Supporting Information file.

Funding: This study was supported by the Deutsche Krebshilfe (<http://www.krebshilfe.de/nc/startseite.html>) under grant no. 111332. The funder had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Motivation

Human voice is generated in the larynx by the two oscillating vocal folds. Owing to the limited space and accessibility of the larynx, endoscopic investigation of the actual phonatory process in detail is challenging. Hence the biomechanics of the human phonatory process are still not yet fully understood. Therefore, we adapt a mathematical model of the vocal folds towards vocal fold oscillations to quantify gender and age related differences expressed by computed biomechanical model parameters.

Methods

The vocal fold dynamics are visualized by laryngeal high-speed videoendoscopy (4000 fps). A total of 33 healthy young subjects (16 females, 17 males) and 11 elderly subjects (5 females, 6 males) were recorded. A numerical two-mass model is adapted to the recorded vocal fold oscillations by varying model masses, stiffness and subglottal pressure. For adapting the model towards the recorded vocal fold dynamics, three different optimization algorithms (Nelder–Mead, Particle Swarm Optimization and Simulated Bee Colony) in combination with three cost functions were considered for applicability. Gender differences and age-related kinematic differences reflected by the model parameters were analyzed.

Results and conclusion

The biomechanical model in combination with numerical optimization techniques allowed phonatory behavior to be simulated and laryngeal parameters involved to be quantified. All three optimization algorithms showed promising results. However, only one cost function seems to be suitable for this optimization task. The gained model parameters reflect the phonatory biomechanics for men and women well and show quantitative age- and gender-specific differences. The model parameters for younger females and males showed lower

subglottal pressures, lower stiffness and higher masses than the corresponding elderly groups. Females exhibited higher subglottal pressures, smaller oscillation masses and larger stiffness than the corresponding similar aged male groups.

Optimizing numerical models towards vocal fold oscillations is useful to identify underlying laryngeal components controlling the phonatory process.

Introduction

The human voice represents an essential aspect of oral communication between human beings. Voice is formed by the interaction and coordination of applied air flow, vocal fold tissue and vocal fold movements. Accurate and precise physiologic interaction of several laryngeal muscle group movements is the basis for normal voice production [1]. The acoustic voice signal originates in the larynx where the two opposing vocal folds are excited by an airflow generated by the lungs (Fig 1). When starting the voice production process (i.e., phonation), the vocal folds are positioned close to each other. Airflow produced from the lungs streams upwards and increases the subglottal pressure below the vocal folds. After exceeding a certain subglottal pressure level, the vocal folds first start to produce small oscillatory motions that then result in a steady-state oscillation (i.e., periodic opening and closing of the vocal folds). A healthy voice signal is normally produced by periodic and symmetric vocal fold oscillations. Also, the closure of the glottis, where the vocal folds almost or entirely close, is considered to be an important part of the normal phonation process [2].

Depending on gender and age, vocal folds oscillate between approximately 100 and 350 times per second during normal phonation [3]: women (~ 200 Hz– 250 Hz), men (~ 100 Hz– 150 Hz), children (~ 200 Hz– 350 Hz). As these movements are so fast, vocal fold dynamics are best captured and visualized by laryngeal high-speed videoendoscopy (HSV) with recording frame rates between 4000 Hz and 20 000 Hz [4–8] (Fig 1A).

Since HSV was first introduced, image processing methods have been proposed [9–11] to allow the quantification of vibratory behavior with objectively computed perturbation measures [12–14]. The signals extracted from HSV and analyzed are either the glottal area waveform (i.e., glottis area function over time) or the vocal fold trajectories at a specific vocal fold location, preferably at mid-membranous position [15–17]. Both signals represent the oscillatory behavior of the vocal folds; i.e., the opening and closing process. Quantitative analysis based on HSV has added substantial knowledge regarding normal and pathological vocal fold dynamics [4,18,19].

Vocal fold dynamics are highly sensitive towards anatomic tissue changes [2], dysfunctions of the involved muscles [20] and subglottal air pressure [21]. Alterations of these parameters may yield disturbed dynamics, resulting in hoarseness. Typically disturbed dynamics are left-right asymmetries, aperiodicities or glottis closure insufficiency, where the vocal folds do not entirely close [2]. It is highly desirable to early diagnose and quantify pathologic dynamic laryngeal alterations to prevent severe laryngeal tissue damage [22,23].

However, because of the limited space in the larynx, it is difficult to measure and quantify the mechanical laryngeal tissue characteristics directly in-vivo [24]. Additionally, HSV evaluation alone only allows vocal fold dynamics to be described but does not give quantitative information on biomechanical parameters such as tissue elasticity and occurring subglottal air pressure. Hence indirect analysis methods based on the adaptation of numeric biomechanical larynx models towards in-vivo HSV recorded vocal fold dynamics were suggested [25,26].

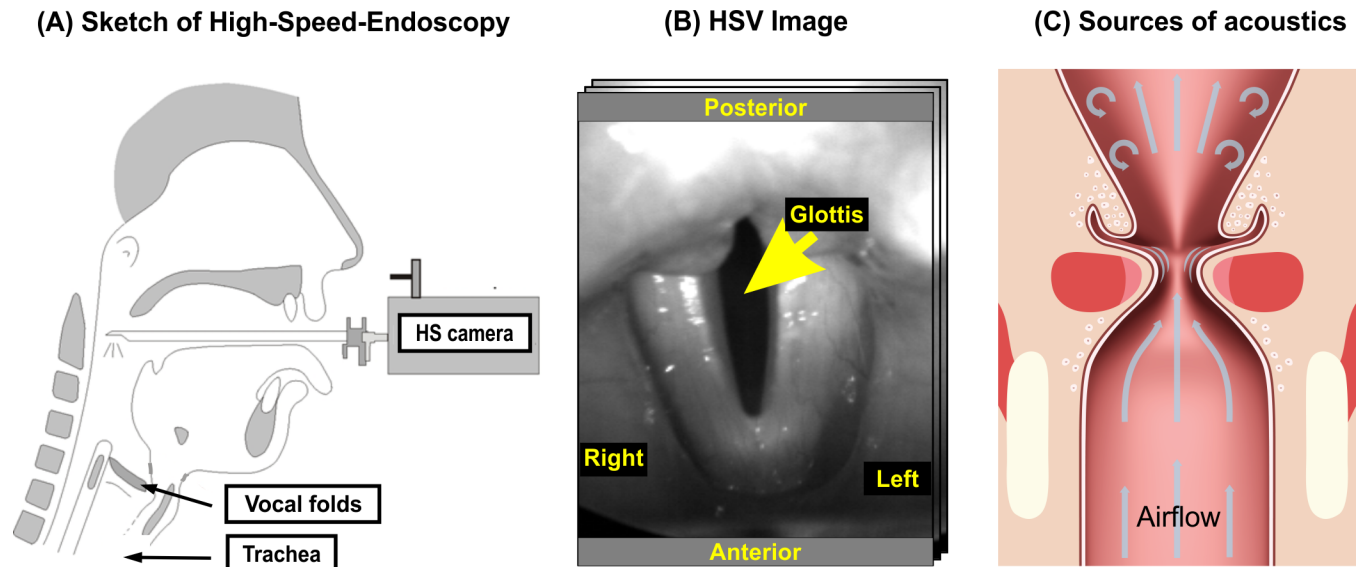


Fig 1. (A) Sketch of the sagittal section of the head and neck, indicating the vocal folds and the rigid endoscope. (B) Glottis (dark) and on the left and right sides the two vocal folds as seen through the HSV. (C) Larynx with highlighted acoustic sound sources (arrows) resulting from the vocal fold vibrations.

<https://doi.org/10.1371/journal.pone.0187486.g001>

Initially, these biomechanical models were used to simulate the underlying processes during phonation. Models were developed to allow the investigation of parameter effects such as applied subglottal air pressure, vibrating masses, tissue stiffness and elongation characteristics with respect to the dynamic vocal fold behavior [27–29]. These so-called lumped mass models (LMMs) are fairly simple but still enable many dynamic characteristics in the larynx to be reproduced [30]. In the most basic models, the vocal folds are simulated by a self-vibrating source consisting of two spring-coupled masses (2MMs) for each vocal fold (Fig 2A) [31]. Since these 2MMs allowed the simulation of only one trajectory at one vocal fold position, models with more masses (multi-mass models, MMMs) were suggested to permit the simultaneous simulation of the vocal fold dynamics at different positions (Fig 2B) [32]. However, the 2MMs and MMMs only focused on the simulation of the lateral (i.e., horizontal) vocal fold

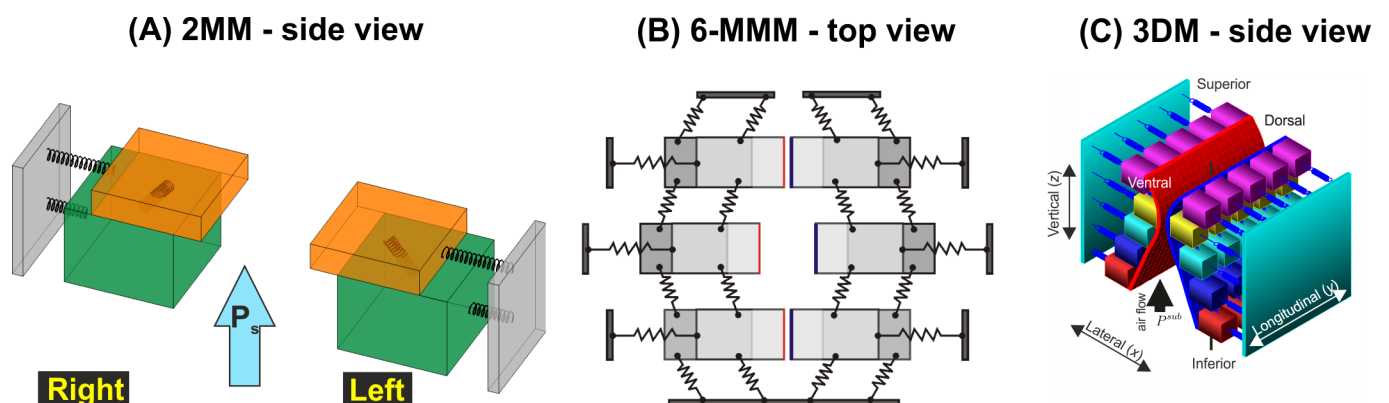


Fig 2. (A) Two-mass model as used in this work with indicated subglottal pressure P_s . (B) Six-mass model allowing the simulation of the vocal fold trajectories along three positions, i.e. at posterior, medial and anterior positions of the vocal folds. (C) Three-dimensional multi-mass model that additionally allows the simulation of vertical dynamics and the vocal fold medial surface at 25 positions along each vocal fold.

<https://doi.org/10.1371/journal.pone.0187486.g002>

displacements and on the phase differences along the inferior-superior plane. To simulate the often neglected vertical vocal fold movements [33] (i.e., vertical tissue displacement in inferior-superior direction), enhanced and more complex three-dimensional LMMs were introduced [34] (Fig 2C).

After the development of the LMMs to simulate human vocal fold dynamics by manually adapting the model parameter settings [28,35], the automatic optimization of these numerical models towards HSV recorded vocal fold oscillations was suggested in order to acquire information on parameters responsible for chaotic behavior, certain dynamic conditions and left-right asymmetric oscillations.

The first fully automatic optimization method was realized by using the Nelder–Mead algorithm to automatically optimize the parameters of a 2MM to reproduce HSV recorded human vocal fold dynamics during sustained phonation [25,26]. Three parameters (vibrating mass, stiffness and subglottal pressure) were varied in the so called cost function Γ to minimize the involved periodic oscillatory components, represented by discrete Fourier transformation (DFT) coefficients, between the model trajectories and trajectories extracted from HSV recordings. Based on this work, the 2MM and a genetic algorithm for the optimization were successfully employed to reproduce the trajectories of patients suffering from unilateral vocal fold paralysis [36] and for ex-vivo larynx experiments [37]. Combining a genetic algorithm with a quasi-Newton method a 2MM successfully reproduced vocal fold dynamics in three subjects [38]. As cost function the Euclidean distance of the glottal area waveform was chosen. Lately, statistical methods like a non-stationary Bayesian estimation approach were suggested for optimizing 2MMs but was only tested on theoretical (i.e., simulated) vocal fold oscillations [39]. Further, a time-dependent 2MM was successfully adapted for 20 healthy and pathological voices [40]. For the optimization an adaptive Simulated Annealing approach was chosen to minimize the Euclidean distance between the model and experimental vocal fold trajectories.

Schwarz et al (2008) [41] successfully optimized a MMM with six coupled masses (6-MMM) towards six disordered voices and two normal voices. They applied a Genetic algorithm and split the optimization process into several sub-steps. As cost function Γ , they used a combination of glottal area and vocal fold trajectory consistency measure. Also, MMMs were used to reproduce vocal fold dynamics within pitch rise paradigms [42]. Wavelet coefficients in the cost function were chosen to consider the time dependency of the system. Powell's Direction Set optimization algorithm was applied to the interval divided (i.e., four oscillation cycles) signal. Thereby, 30 healthy and pathological adult voices were adapted.

Simulating the phonatory process after total laryngectomy, due to cancer, was achieved through coupling eight two-mass models arranged in a circle [43]. For the optimization they selected a combination of Simulated Annealing for a preliminary global search and Powell's Direction Set method for the final local approximation. As cost function a combination of area difference, intersection and distance measures was chosen. The method was tested on 75 synthetic data sets and four human subjects.

Finally, a three-dimensional model (3DM) was developed and applied to ex-vivo human vocal fold dynamics produced in a hemi-larynx setup [44]. This model allowed the simulation of the entire vocal fold surface from inferior to superior including the vertical dynamics [44]. Owing to the increased number of masses in the 3DM (i.e., 25 on each side) and the rather unfortunate topology of the cost function (i.e., many local minima), the optimization was divided into several coarse steps, followed by fine optimization processes. It started with a global optimization combining Particle Swarm Optimization and Simulated Annealing algorithms. The local optimization was achieved by Powell's direction set method [45]. By adapting this model to ex-vivo human larynx dynamics, information was obtained on the distribution of vibrating masses and stiffness along the vocal fold surface. The optimization results

accurately matched the actual fundamental frequencies and also the experimental measured subglottal pressure values [44].

Overviews of current applied LMMs can be found in [46,47]. Owing to the increase in computational power, more accurate and therefore more computationally intensive so called Finite Element Methods (FEM) and Finite Volume Methods (FVM) simulating 2D and 3D laryngeal dynamics and airflow became popular during the last decade [31,48–50]. However, the complexity of these models and therefore the computational costs do not allow them to be used for optimization purposes yet.

As described above, numerical optimization of LMMs has so far focused only on healthy adults and compared model parameters for healthy vs. disordered phonation processes.

To the best of our knowledge no studies have applied LMMs to analyze gender specific differences. Previous ex-vivo and in-vivo studies on human larynges showed the following, partly contradictory, results: Regarding vocal fold stiffness, an ex-vivo study reported, at medial position slightly above the vocal fold edge, smaller stiffness for an elderly women compared to an elderly male [51]. In contrast, it was suggested that vocal folds in men are less stiff than in women [52]; unfortunately the authors did not report the age of the subjects. Gender related subglottal pressure differences were reported for young adults [53]; another study, for example showed no gender specific differences [54, 55]. However, male vocal folds are known to be larger (i.e., increased mass) than female vocal folds [56].

Further, to the best of our knowledge no studies have compared biomechanical LMM parameters between younger and elderly subjects. Previous ex-vivo and in-vivo studies on human larynges reported the following results: Glottal parameters extracted from HSV recordings were successfully applied to differentiate age groups; however no details on the parameters quantities were presented [57]. Age-related morphological changes of the vocal folds influencing the viscoelasticity have been described [58–61]. Previous studies suggested lower stiffness for younger males compared to elderly males [51,62]. Histologic analysis reported reduced lamina propria thickness and reduced epithelial cell density for elderly subjects [63]. However, an increase in vocal fold volume (i.e., mass) for 28.7% of analyzed elderly women was also reported [62]. Furthermore, a thickening of mucosa and vocal fold cover in elderly women was described [64]. Higher subglottal pressure in elderly males compared to younger male subjects was suggested [65]. However, other studies reported no differences [66,67]. It was also suggested that aging effects on phonatory behaviors differ in degree and kind for men and women [65]. In summary, a better understanding of the elderly voice and accompanying voice disorders is desired [66,68].

Since it is also not obvious, which optimization algorithms and cost functions are the most appropriate ones, we have decided to apply different approaches in our study. Therefore, by optimizing a LMM towards vocal fold oscillations recorded by endoscopic high-speed imaging, the aims of our study are:

1. to analyze the performance of three different optimization algorithms and three different cost functions (Γ_1 , Γ_2 , Γ_3) to automatically adapt a fairly simple LMM (i.e., in our study 2MM) to vocal fold dynamics.
2. to quantify gender related biomechanical differences in the larynx expressed by 2MM parameters. We hypothesize that healthy young males would have greater masses, lower stiffness and less subglottal pressure compared to healthy young females.
3. to quantify age related biomechanical differences in the larynx expressed by 2MM parameters. We hypothesize that healthy young adults would have greater vocal fold masses, lower vocal fold stiffness and lower subglottal pressure compared to elderly atrophic subjects.

Also, it is expected to find higher kinematic asymmetry in the elderly subjects as reported previously [69,70].

By investigating these objectives, we want to illustrate and emphasize the informative value of LMMs towards vocal fold physiology by analyzing interrelations between underlying laryngeal components such as vibrating mass, stiffness and applied subglottal pressure. Further, we want to illustrate the potential for the differentiation of vocal fold vibratory characteristics based on LMMs.

Methods

Subjects

Four groups of subjects were investigated. Two younger age groups were analyzed with 16 healthy females (18–24 years) and 17 healthy males (19–38 years). The subjects were recruited at the FAU-Erlangen-Nürnberg and University of Kentucky. The subjects had to fulfill the following criteria: negative history of vocal pathology, not a professional vocal user and having a normal voice, as confirmed by a speech–language pathologist. Adults with history of smoking were excluded.

Further six older male subjects (62–86 years) and five older female subjects (72–91 years) were analyzed. These subjects were recruited during regular office hours at the Division of Phoniatics and Pediatric Audiology of the department Otolaryngology, Head- and Neck surgery at the University Hospital Erlangen. The subjects were diagnosed with vocal fold atrophy and had no further vocal fold pathologies.

The participants gave informed, written consent prior to the participation and this consent procedure was approved by the corresponding local ethics committees (Ethik-Kommission der Medizinischen Fakultät FAU-Erlangen-Nürnberg and Office of Research Integrity Expedited Review Board at the University of Kentucky). Experiments were performed in accordance with the Declaration of Helsinki (1964).

Data acquisition

To record the vocal fold vibrations, a PENTAX Medical (Montvale, NJ, USA) Model 9710 digital gray scale high-speed (HS) camera was used at both universities. The applied temporal resolution of 4000 fps enables the vocal fold oscillations to be captured [4]. The HS camera provides a series of pictures with an image resolution of 512×256 pixels with a maximum duration of 4 seconds to visualize the laryngeal dynamics.

The recordings were performed during sustained phonation of the vowel /i:/ with a PENTAX Medical 70° endoscope containing a 300 Watt Xenon light source. For each subject, one sample of typical phonation was recorded and analyzed. For the optimization, a sequence length of 100 ms ($N = 400$ frames) of sustained phonation was chosen. This interval covered between 10 and 50 oscillation cycles depending on the fundamental frequency. The choice of this interval length is a compromise between a sufficient observation time and computational costs and lies in the range of previous reported interval lengths.

Data processing

First, image processing was performed to determine the glottis area and vocal fold edges [11], Fig 3. The glottis axis was determined using the methods being described in [71,72], Fig 3. The in-house developed software tool “Glottis Analysis Tool” (GAT) was used. The GAT tool has already proven its validity and applicability within several studies [8,73] and is also used by other voice groups [74,75].

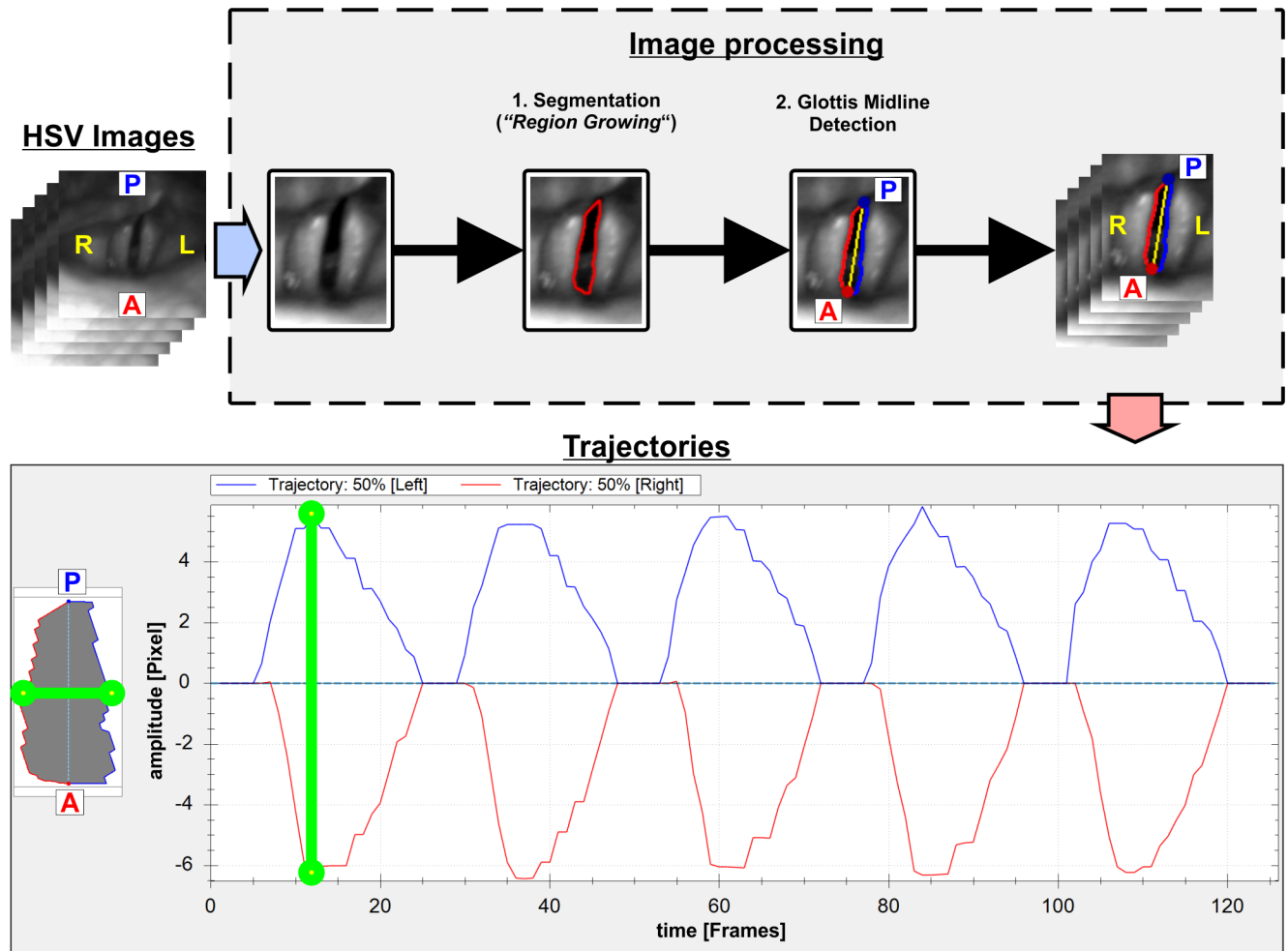


Fig 3. Performed steps for image processing yielding the experimental vocal fold trajectories.

<https://doi.org/10.1371/journal.pone.0187486.g003>

For the adaptation of the 2MM, the trajectories from the mid-membranous position, 50% position between anterior and posterior, of the vocal folds were chosen (Fig 3), since the largest vocal fold amplitudes are expected in this region [76]. The trajectories at the medial vocal fold position are also expected to be the most useful [16]. To guarantee the extraction of the vocal fold trajectories at the 50% glottal mid-line only HSV recordings were used, where no obstructions of the view by the epiglottis was present and the most anterior and the visible parts of the posterior glottis were in view.

Two-mass model (2MM)

The human voice is generated by three-dimensional vocal fold oscillations [76]. During phonation, oscillations of the vocal fold mucosa occur in the anterior–posterior, medio-lateral and vertical directions [20]. Anterior–posterior movements are fairly small and can therefore be neglected [76]. The vertical displacements (up to 2.4 mm) are approximately two-thirds of the dominant medio-lateral displacements [33,76]. However, the vertical component cannot be reconstructed based on the current HSV imaging techniques, owing to the lack of a second imaging tool (e.g., a second camera or laser projection system [77]). So far, extended studies on three-dimensional reconstruction of vocal fold surface dynamics have only been possible in

ex-vivo or in synthetic experiments [77,78]. For three-dimensional in-vivo reconstruction, only case studies [79] and proof of concepts [80] were performed lately. Therefore, we focused on the dominant medio-lateral (i.e., horizontal) displacement characteristics at one vocal fold position that can be simulated by the 2MM (Fig 4).

Important laryngeal parameters that influence the oscillations are vibrating vocal fold masses (m), tension, elasticity, damping (r), stiffness (k) and subglottal air pressure (P_s). The 2MM enables these parameters to be varied and the effects on the medio-lateral oscillatory behavior at one specific vocal fold location to be observed. Since the here applied 2MM has already been extensively described [28], we will give only information necessary to understand the functionality of the 2MM and the optimization procedure. The 2MM is based on [27,28] and assumes that each vocal fold is formed by two vertically arranged coupled masses: A larger, lower mass ($m_{1\alpha}$) and a smaller, upper mass ($m_{2\alpha}$). Each part consists of a simple mechanical oscillator with a mass and two springs at which the two masses of one vocal fold are connected by one spring. The 2MMs driving force is the subglottal pressure P_s below the masses. The 2MM is described by a system of eight ($\alpha = l,r$) ordinary differential equations [25]:

$$\frac{d}{dt} \begin{pmatrix} x_{1z} \\ v_{1z} \\ x_{2z} \\ v_{2z} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\left(\frac{k_{1z} + k_{cz}}{m_{1z}}\right) & -\frac{r_{1z}}{m_{1z}} & \frac{k_{cz}}{m_{1z}} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_{cz}}{m_{2z}} & 0 & -\left(\frac{k_{2z} + k_{cz}}{m_{2z}}\right) & -\frac{r_{2z}}{m_{2z}} \end{pmatrix} \begin{pmatrix} x_{1z} \\ v_{1z} \\ x_{2z} \\ v_{2z} \end{pmatrix} + \begin{pmatrix} 0 \\ I_{1z}(x_{1l}, x_{1r}) + \frac{1}{m_{1z}} F_1(x_{1r}, x_{2r}, x_{1l}, x_{2l}) \\ 0 \\ I_{2z}(x_{2l}, x_{2r}) \end{pmatrix} \quad (1)$$

The indices (i, α) represent lower ($i = 1$) and upper ($i = 2$) masses; $\alpha = l,r$ represents the left and right vocal fold. Ishizaka & Flanagan (1972) introduced a standard parameter value set that represented the standard vocal fold vibration pattern [81]. The parameters from Eq (1)

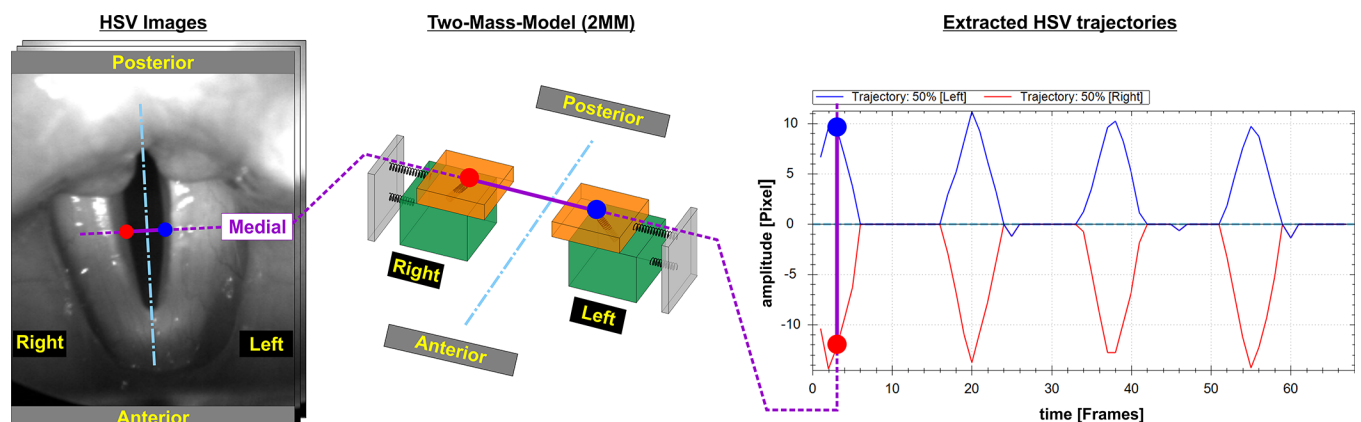


Fig 4. HSE image with indicated glottal axis (vertical blue line) and medial positions (~50% position between anterior and posterior) on left (blue dot) and right (red dot) vocal folds where the trajectories were extracted (left figure). The 2MM used (middle figure). Extracted trajectories for left (blue) and right (red) vocal folds (right figure).

<https://doi.org/10.1371/journal.pone.0187486.g004>

Table 1. Standard parameters of the 2MM. In this study, the chosen vocal fold lengths l were 10 mm for women and 16 mm for men. The rest positions x_{01} , x_{02} for the 2MM optimization were computed based on the mean amplitudes of the HSV trajectories yielding also individual rest areas a_{0i} [26]. During the 2MM optimization, the m_i , k_i and P_s values are varied.

m_1 [g]	m_2 [g]	k_1	k_2	k_c	r_1	r_2	c_1
0.125	0.025	0.08	0.008	0.025	0.02	0.02	$3k_1$
c_2	x_{01} [cm]	x_{02} [cm]	a_{01}	a_{02}	d_1 [cm]	l [cm]	P_s [cmH ₂ O]
$3k_2$	0.0179	0.0179	0.05	0.05	0.25	1.4	8.0

<https://doi.org/10.1371/journal.pone.0187486.t001>

and their standard values as originally introduced [28,81] are given in Table 1. The nonlinear components ($I_{1\alpha}$, $I_{2\alpha}$ and F_1) describe the impact forces ($I_{1\alpha}$, $I_{2\alpha}$) and the subglottal pressure function (F_1), described also in detail in earlier work [25].

For the optimization, the 2MM trajectories ($T_{M\alpha}$, $\alpha = l,r$) from each side are a combination of the displacements of the lower and upper masses. The mass ($m_{1\alpha}$, $m_{2\alpha}$) that is closer to the glottal midline (i.e., visible from above) contributes to the corresponding vocal fold trajectory $T_{M\alpha}$ [40].

The HSV recordings do not allow the extraction of the vocal fold oscillations in metric units but only in pixels. Owing to the known differences in vocal fold length, we chose mean values for the vibrating vocal fold length to be 10 mm in women and 16 mm in men. These values are the average membranous vocal fold lengths being reported in previous work [82–86]. Due to the metric mapping of the vocal fold lengths, we can calculate an approximate metric equivalent of the length of one pixel and convert the extracted trajectories to metric units.

During the optimization, the parameters vocal fold mass ($m_{1\alpha}$), stiffness ($k_{1\alpha}$), and subglottal pressure (P_s) are varied. This is based on the study by Steinecke and Herzel (1995) [28], who also introduced a scaling factor Q_α ($\alpha = l,r$) to vary the standard masses ($m_{1\alpha 0}$) and spring variables ($k_{1\alpha 0}$). Laryngeal asymmetry is expressed by the scaling factors Q_α . The scaling factors Q_l (left vocal fold) and Q_r (right vocal fold) influence the masses and spring constants in the following way [27,28]:

$$\begin{aligned}
 k_{lx} &= Q_x k_{lx0}, & k_{rx} &= Q_x k_{rx0} \\
 c_{lx} &= Q_x c_{lx0}, & m_{lx} &= m_{lx0}/Q_x
 \end{aligned}
 \tag{2}$$

This reciprocal relationship between vibrating masses $m_{1\alpha}$ and springs $k_{1\alpha}$ is based on the assumption that the larger the vibrating mass, the smaller is the stiffness of the vocal folds [28]. The 2MM oscillates symmetrically provided that Q_l and Q_r are equal or only slightly different. If the differences between the Q_i are too large, the 2MM vibrations become left-right asymmetric.

Optimization procedure

With the variation of Q_l and Q_r reflecting mass and stiffness for each vocal fold and the subglottal pressure P_s , it is possible to reproduce physiologic and pathologic vocal fold oscillations [26,36]. The goal of the optimization is to vary these parameters so that the resulting 2MM trajectories (T_M) accurately recreate the HSV recorded and extracted vocal fold trajectories (T_E). This is realized by a combination of several steps within the optimization algorithm (Fig 5). The cost function judges the quality of the 2MM optimization and compares the model trajectories T_{Ml} and T_{Mr} with the recorded vocal fold trajectories T_{El} and T_{Er} . Three different cost functions ($\Gamma_1, \Gamma_2, \Gamma_3$), as described below, are used to match the extracted vocal fold trajectories as closely as possible.

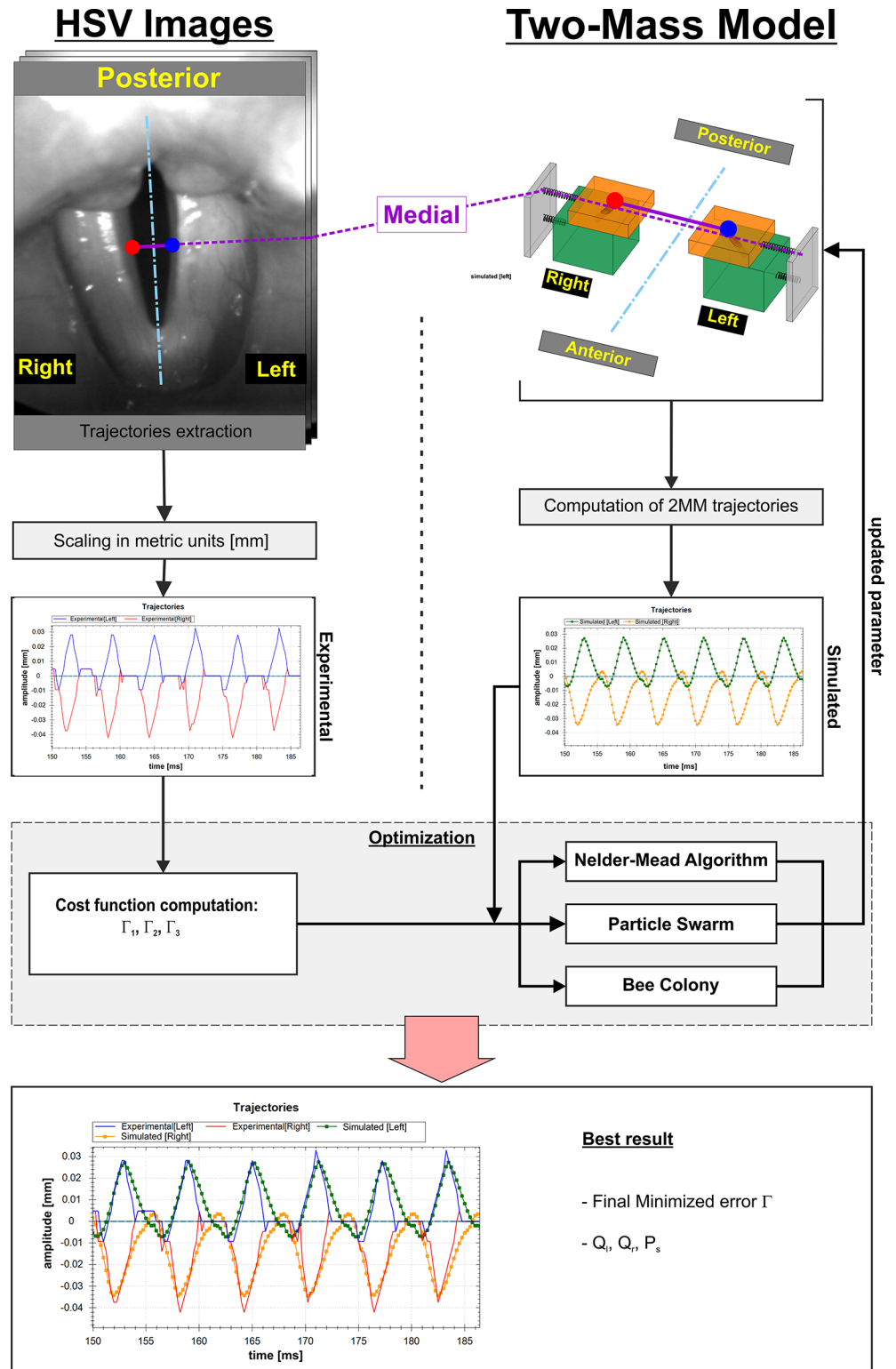


Fig 5. Flow diagram of the optimization procedure.

<https://doi.org/10.1371/journal.pone.0187486.g005>

Frequency domain (Γ_1). The absolute and phase values of the dominant harmonics between T_M and T_E are considered as suggested previously [25]. First, a preprocessing step with a Fourier transform for identification of the harmonics involved in the left experimental vocal fold trajectory (T_{El}) and right trajectory (T_{Er}) is performed. The experimental trajectory's Fourier spectrum is dominated by only a small number of harmonics, represented by their Fourier coefficients $e_{i\alpha}$ (i = number of harmonic, α = left (l) or right (r)). Only harmonics $e_{i\alpha}$ that exhibit at least 25% of the absolute value of the largest coefficient $e_{1\alpha} = \max(e_{i\alpha})$ representing the fundamental frequency are taken into account. For irregular or aperiodic vibrations the coefficient $e_{1\alpha} = \max(e_{i\alpha})$ corresponds to the dominant frequency. Additionally, the next left and right neighbors of the coefficients are selected to consider slight variations in the harmonics [25]. For the simulated 2MM trajectories T_{Ml} and T_{Mr} , the equivalent Fourier coefficients ($s_{j\alpha}$, $i = j$) are chosen and considered in the cost function Γ_1 :

$$\begin{aligned} \Gamma_1(Q_l, Q_r, P_s) := & \\ & s \cdot \| (|e_{1l}|, \dots, |e_{Ll}|) - (|s_{1l}|, \dots, |s_{Ll}|) \|_2 \\ & + \| (\arg e_{1l}, \dots, \arg e_{Ll}) - (\arg s_{1l}, \dots, \arg s_{Ll}) \|_2 \\ & + s \cdot \| (|e_{1r}|, \dots, |e_{Rr}|) - (|s_{1r}|, \dots, |s_{Rr}|) \|_2 \\ & + \| (\arg e_{1r}, \dots, \arg e_{Rr}) - (\arg s_{1r}, \dots, \arg s_{Rr}) \|_2 \end{aligned} \tag{3}$$

In Γ_1 , L corresponds to the number of coefficients for the left and R for the right vocal fold. Also, a scaling factor s is included in Γ_1 to balance the influence of phase and absolute values of the Fourier coefficients as described in [25]. Γ_1 was constructed to reduce the number of local minima and to therefore potentially yield improved optimization results [87].

Time domain (Γ_2). The Euclidean distance between the model trajectories (T_{Ml} , T_{Mr}) and the recorded trajectories (T_{El} , T_{Er}) is computed.

Normalized frequency domain (Γ_3). This cost function consists of Γ_1 , except that the absolute Discrete Fourier Transform (DFT) coefficients are normalized to 1 and that an additional regularization term is added: the Euclidean distance of the absolute values of the DFT coefficients representing the fundamental frequency.

Optimization was rated successful when the following three error criteria were achieved. (1) Frequency deviation $\leq 5\%$; (2) glottis closure as seen in the HSV recordings was achieved for the optimized 2MM trajectories; (3) amplitudes of optimized 2MM trajectories were within the amplitude variations of HSV trajectories; see e.g., Fig 4 where the right trajectory (red) varies between 12.5 and 15 pixel. Optimization would be defined successful for this trajectory when the corresponding right 2MM amplitude was also between 12.5 and 15 pixel.

For optimization, three algorithms are run separately on all three cost functions Γ_1 , Γ_2 , Γ_3 : (1) the Nelder–Mead (NM) algorithm [25], (2) the Particle Swarm Optimization (PSO) algorithm [45] and (3) the Simulated Bee Colony (SBC) optimization [88]. This yielded altogether nine optimized parameter sets (Q_l, Q_r, P_s) approximating the recorded vocal fold trajectories $T_{E\alpha}$. Since, the cost functions $\Gamma_1, \Gamma_2, \Gamma_3$ are computed in different ways and in different domains (frequency and time) their absolute values cannot be compared to judge which cost function is actually better. Hence, the final and best parameter set (Q_l, Q_r, P_s) was determined as the parameter set having the smallest normalized Euclidian distance Γ between the

experimental $T_{E\alpha}$ and simulated model trajectories $T_{M\alpha}$:

$$\Gamma = \frac{1}{2} \left(\sqrt{\frac{\sum_{i=1}^N (T_{El}[i] - T_{Ml}[i])^2}{\sum_{i=1}^N (T_{El}[i])^2}} + \sqrt{\frac{\sum_{i=1}^N (T_{Er}[i] - T_{Mr}[i])^2}{\sum_{i=1}^N (T_{Er}[i])^2}} \right) \quad (4)$$

Prior to the actual optimization, for each subject an initial value search for Q_l , Q_r , and P_s was performed for the NM and PSO algorithms to reduce the potential search space and computational time for the actual optimization process [25]. An initial search was not performed for the stochastic based SBC algorithm.

The entire optimization process was performed on an Intel® Core™ i5-4590 Processor (3.30 GHz) using an in-house developed software written in C#. The software contained a GUI for improved handling and visually reviewing the results.

Parameter analysis

For judging the left-right asymmetry in the 2MM and therefore in the vocal fold oscillations a factor Q_{lr} (≥ 1) is used, adapted from [28]. The closer the Q_{lr} quotient is to 1, the higher is the dynamic left-right symmetry:

$$Q_{lr} = \frac{\max(Q_l, Q_r)}{\min(Q_l, Q_r)} \quad (5)$$

Pairwise group differences (young females vs. young males) for the computed parameters (Q_l , Q_r), Q_{lr} , P_s and Γ were statistically investigated. (Q_l , Q_r) are merged to one data pool, since the absolute differences between both the groups are of interest. Initially, to test for normal distribution, the Shapiro–Wilk test was used. All four parameters were not normally distributed: (Q_l , Q_r) ($df = 50$, $p = 0.000$), Γ ($df = 25$, $p = 0.005$), Q_{lr} ($df = 25$, $p = 0.011$) and P_s ($df = 25$, $p = 0.001$). Hence, Mann–Whitney U-tests were applied for the four group comparisons; the significance level was set to $p = 0.05$ and no Bonferroni correction was applied.

For comparing the younger vs. elderly subjects and gender specific differences in the elderly groups, only descriptive statistics were applied, due to the small number of elderly subjects. Hence, these observations are limited and have no statistical evidence. Statistical analysis was done using IBM SPSS Statistics 21.

Results and discussion

Applicability of the 2MM optimization

Results of the optimization procedure for the 2MM were only deemed correct, when all three above introduced error criteria were met. Altogether 12 HSV recordings (27.3%) could not be correctly optimized by violating one or more of these error criteria: (1) the fundamental frequency could not be matched (two times); (2) the HSV trajectories and the optimized 2MM amplitudes did not match (nine times); (3) the glottis closure or the glottis closure insufficiency was not reproduced (five times). In Fig 6, typical examples for failed optimization results are given: (A) For the young female the Γ value was within the range of the correct rated optimizations, however glottis closure was not achieved. (B) For the young male the Γ value is in the upper range of the correctly rated optimizations, however the amplitudes did not match. (C) For the elderly female the Γ value was higher than for the correctly rated optimizations, glottis closure was not achieved and the amplitudes did not match. (D) For the elderly male the Γ value was higher than for the correctly rated optimizations, glottis closure was not achieved and the amplitudes did not match.

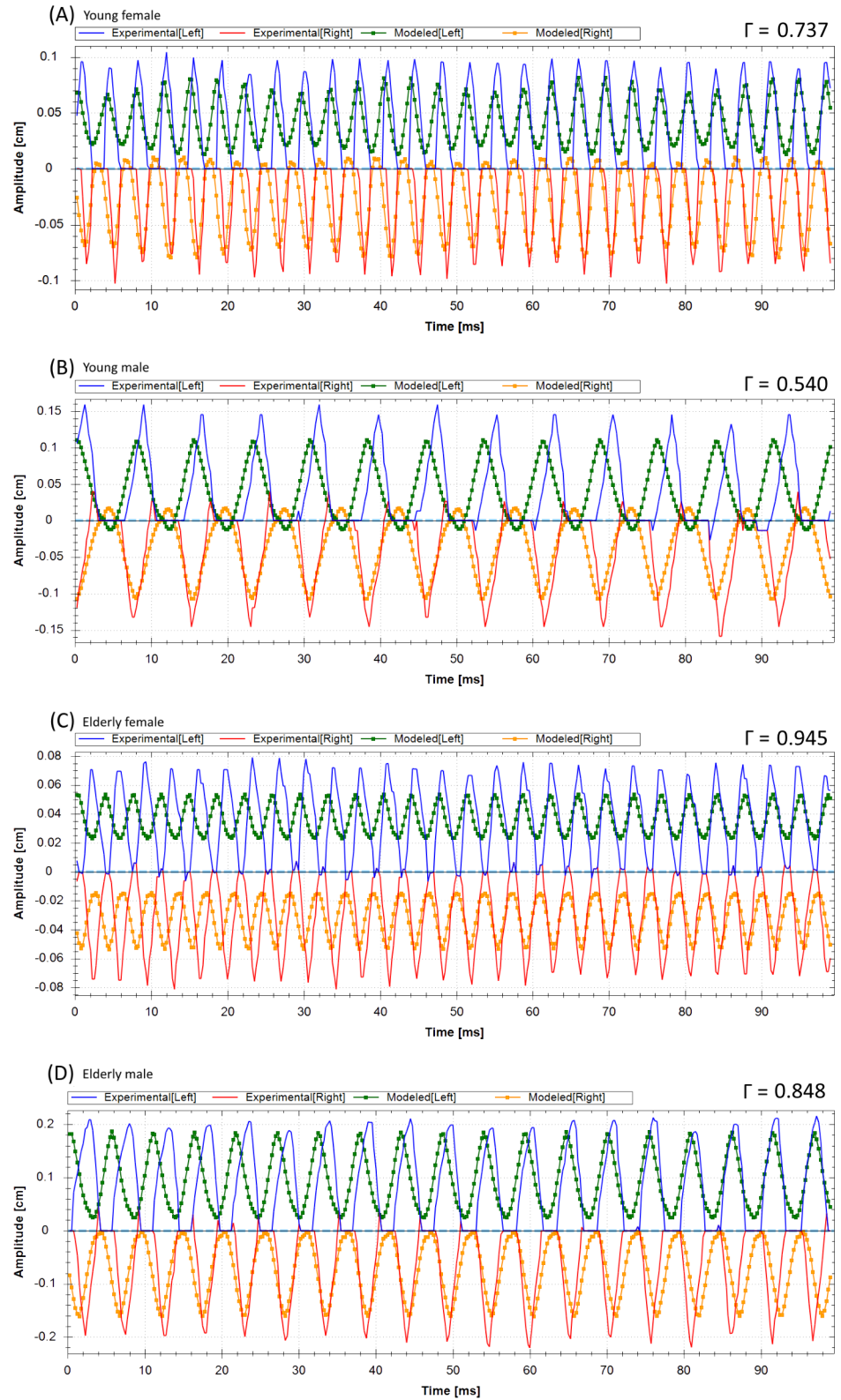


Fig 6. Examples for a young female (unmatched glottis closure and left amplitude), young male (unmatched amplitude), elderly female (unmatched glottis closure and amplitude) and elderly male (unmatched amplitudes and glottis closure) that illustrate the extracted trajectories and the incorrectly optimized trajectories of the 2MM for the left and vocal fold right side.

<https://doi.org/10.1371/journal.pone.0187486.g006>

Altogether 72.7% of the HSV trajectories were successfully optimized: 68.8% (11 out of 16) of the young females, 82.4% (14 out of 17) of the young males, 80.0% (4 out of 5) of the older females, and 50.0% (3 out of 6) of the older males. In the following, only these 32 successfully optimized HSV recordings will be considered and discussed.

As can be seen in Table 2, all three applied optimization algorithms yielded optimal results (i.e., lowest Γ value). However, the NM (38%) and SBC (43%) algorithms yielded more often the best results than the PSO (19%) algorithm. For the cost function, Γ_2 definitely showed the most promising results. Γ_2 yielded the best results most often (94%) followed by Γ_3 (6%). Γ_1 never yielded the best approximation. With regard to our first aim, the results suggest that all three optimization algorithms are suitable for the 2MM optimization but as cost function only Γ_2 seems to be promising.

The low values of the objective function Γ confirm the applicability of the 2MM for all four groups, Table 3. A value of $\Gamma = 0$ would correspond to a perfect optimization without any discrepancies between experimental and simulated curves. The highest mean Γ values are for elderly males (0.63 ± 0.09), followed by elderly women (0.59 ± 0.18). The best and lowest values are found for young men (0.45 ± 0.06) followed by the young women (0.57 ± 0.20). The difference in Γ (young women vs. young men) is not statistically different ($p = 0.434$).

The values suggest that the varied masses and stiffness parameters of the 2MM might adapt slightly better to the two younger subject groups than they do when elderly subjects are considered. However, higher values of Γ were expected for the elderly groups in comparison with young adults, as elderly subjects were reported to have lower laryngeal dynamic periodicity compared with younger adults [89,90]. This means that the glottis and therefore the extracted vocal fold trajectories oscillate not as periodically as they do in young adults. The 2MM does not allow for the simulation of slight changes in oscillation period length and slight amplitude changes between oscillation cycles (i.e., Jitter and Shimmer) hence yielding consequently higher Γ values for the elderly subjects.

The accuracy of the optimization regarding the fundamental oscillation frequencies of the vocal folds is illustrated in Fig 7, where the experimental trajectory frequencies (f_{EB} , f_{Er}) are plotted against the optimized 2MM frequencies (f_{MB} , f_{Mr}). The highest accuracy is given for young women, where the model and experimental frequencies match for all subjects. For young males, the frequencies match perfectly for eight subjects. For three subjects the frequencies deviate for one vocal fold side with $\Delta = \{6.7, 3.8, 3.8 \text{ Hz}\}$ and for three subject the frequencies deviate for both vocal folds with $\Delta = \{3.8, 3.2, 2.6 \text{ Hz}\}$. For older women, the frequencies match for three subjects. For one subject the frequencies deviate for both vocal folds with $\Delta = \{10.5 \text{ Hz}\}$. For older men, the frequencies match for two subjects. For one subject the frequencies deviate for one vocal fold with $\Delta = \{7.2 \text{ Hz}\}$.

Table 2. Overview of how often each optimization algorithm (Nelder Mead–NM, Particle Swarm Optimization–PSO, Simulated Bee Colony–SBC) and cost function Γ_1 , Γ_2 , and Γ_3 yielded the best optimization result; i.e., smallest Γ value—see Eq (4).

	Algorithms yielding best performance			Cost functions yielding best performance		
	NM	PSO	SBC	Γ_1	Γ_2	Γ_3
Young Males	6	2	6	0	13	1
Young Females	4	2	5	0	10	1
Older Males	1	0	2	0	3	0
Older Females	1	2	1	0	4	0
Sum	12	6	14	0	30	2

<https://doi.org/10.1371/journal.pone.0187486.t002>

Table 3. Mean values, standard deviations and range (minimum–maximum) of Γ , the optimized parameters P_s [cmH₂O], Q_1 , Q_r , and the symmetry quotient Q_{lr} , for the four subject groups are given.

	Young Males		Young Females		Elderly Males		Elderly Females	
	Mean \pm SD	Range	Mean \pm SD	Range	Mean \pm SD	Range	Mean \pm SD	Range
Γ	0.45 \pm 0.06	0.37–0.57	0.57 \pm 0.20	0.32–0.90	0.63 \pm 0.09	0.58–0.74	0.59 \pm 0.18	0.40–0.82
P_s [cmH ₂ O]	16.49 \pm 7.13	10.10–32.31	21.12 \pm 7.16	13.20–36.10	22.61 \pm 6.50	15.14–27.00	28.30 \pm 12.17	18.70–45.70
Q_1	1.12 \pm 0.32	0.76–1.99	2.61 \pm 0.38	2.03–3.32	1.32 \pm 0.15	1.19–1.49	2.81 \pm 0.93	1.56–3.70
Q_r	1.15 \pm 0.28	0.78–1.83	2.46 \pm 0.34	1.93–3.05	1.68 \pm 0.62	1.15–2.36	3.23 \pm 1.35	1.60–4.87
Q_{lr}	1.07 \pm 0.04	1.02–1.17	1.12 \pm 0.08	1.01–1.24	1.32 \pm 0.24	1.10–1.59	1.12 \pm 0.13	1.02–1.32

<https://doi.org/10.1371/journal.pone.0187486.t003>

In summary, 81% of the successful optimized trajectories $f_{M\alpha}$ matched the fundamental frequencies of the experimental trajectories $f_{E\alpha}$. This value is similar to that reported previously [26], where 80% of the original fundamental frequencies were correctly reproduced. Fig 8 shows examples for correctly reproduced vocal fold trajectories.

Optimized parameters

The computed values for P_s , Q_{lr} , Q_1 and Q_r and found group differences (gender and age) confirm our hypotheses as formulated in aims (2) and (3):

Table 3 gives an overview of the determined P_s , Q_1 , Q_r and Q_{lr} parameter values. The **symmetry quotient Q_{lr}** shows high symmetry for the two young healthy groups, confirming the previously performed medical diagnosis of normal voice production. In our study, young women and men showed similar, statistically not significant differences with $p = 0.202$, and highest symmetry with Q_{lr} values of 1.07 ± 0.04 for young males. Young ($Q_{lr} = 1.12 \pm 0.08$) and older females ($Q_{lr} = 1.12 \pm 0.13$) exhibit equal symmetry. Deviations of dynamic left–right symmetry of up to 20% (i.e., $Q_{lr} \approx 1.20$) were reported previously [42] and can still be considered as entirely physiologic. Further, slight physiologic and anatomic asymmetries were reported in healthy young and elderly subjects [69,70]. However, the older the subjects, the more prominent and larger the vocal fold dynamic asymmetries might become [91]. This was reflected only for the older male group ($Q_{lr} = 1.28 \pm 0.20$) that showed increased asymmetry values (i.e., higher Q_{lr} values). The older female group was much more symmetric than the elderly male group. These findings confirm our hypothesis of higher kinematic asymmetries in elder subjects for women but not for men.

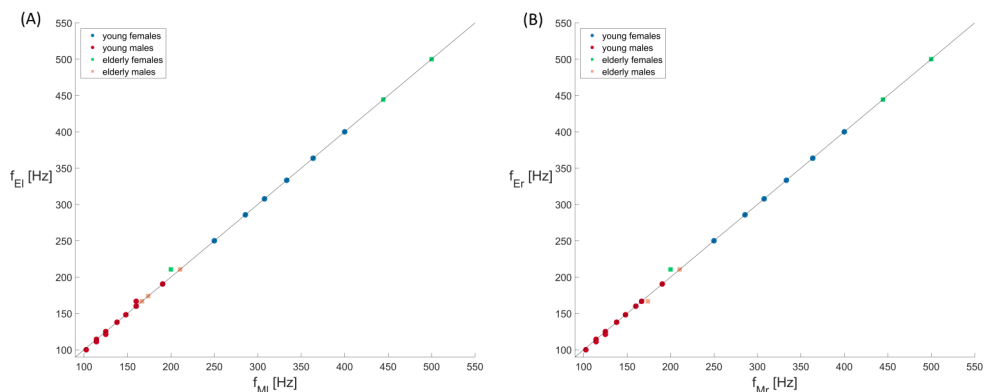


Fig 7. Fundamental frequencies (f_{El} , f_{Er}) of the experimental HSE-recorded trajectories versus the frequencies (f_{Ml} , f_{Mr}) of the optimized model trajectories shown separately for left and right vocal folds.

<https://doi.org/10.1371/journal.pone.0187486.g007>

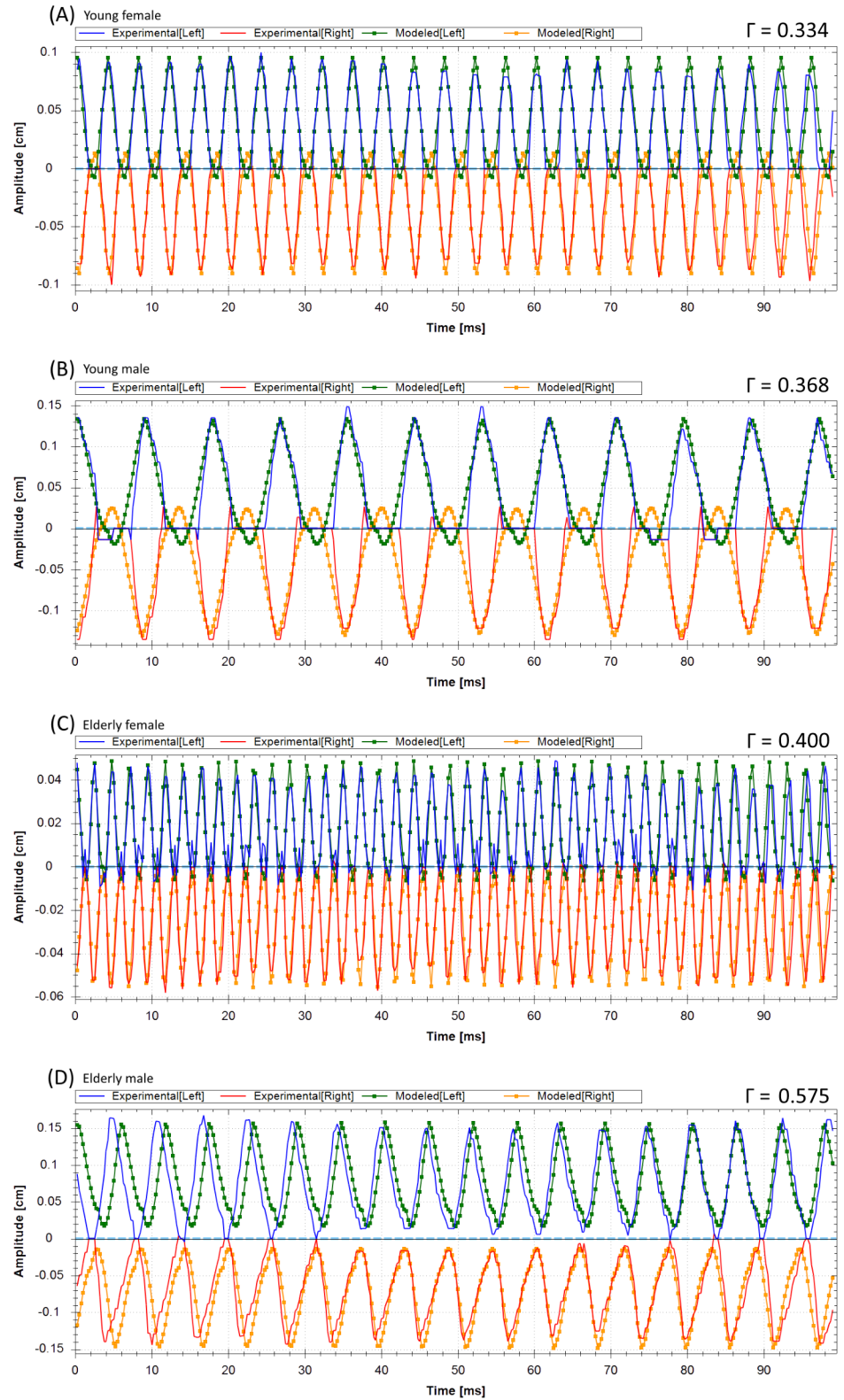


Fig 8. Examples for a young female, young male, elderly female and elderly male that illustrate the extracted trajectories and the correctly optimized trajectories of the 2MM for the left and vocal fold right side. The values for the cost function Γ are given.

<https://doi.org/10.1371/journal.pone.0187486.g008>

Table 4. Overview on subglottal pressure vales (cmH₂O) as reported for healthy subjects during normal and loud phonation in the literature.

Study	Females		Males	
	Normal	Loud	Normal	Loud
Holmberg et al [93]	3.6–8.1	6.4–13.7	4.3–9.7	6.2–16.4
Perkell et al [55]	5.5 ± 1.3	7.6 ± 1.8	5.9 ± 1.1	8.7 ± 2.4
Sulter & Wit [54]	11.2 ± 4.1	18.3 ± 6.9	11.8 ± 3.7	20.5 ± 5.9
Baken & Orlikoff [94]	3.5–12.6	6.5–27.7	4.2–12.8	5.9–24.4
Hertegard et al [95]	-	-	5–24	

<https://doi.org/10.1371/journal.pone.0187486.t004>

Young men showed the lowest **subglottal pressure** P_s with a mean value of 16.49 ± 7.13 cmH₂O, but also had the lowest fundamental frequency (147 ± 38 Hz). In contrast, older males showed clearly higher subglottal pressure at 22.61 ± 6.50 cmH₂O. Also the fundamental frequency for elderly men was increased (182 ± 22 Hz) confirming earlier studies [62]. Also, for the elderly females the fundamental frequencies (380 ± 117 Hz) were increased compared to young women (328 ± 40 Hz)—contradicting previous observations [92]. Also the subglottal pressures (28.30 ± 12.17 cmH₂O) were higher for the elderly compared to the younger female group (21.12 ± 7.16 cmH₂O).

The two elderly male and female groups showed both higher subglottal pressure and higher fundamental frequencies compared to their corresponding younger groups. The subglottal pressure for the male groups was smaller compared to the corresponding female groups. Comparing the young gender groups revealed statistically significant differences ($p = 0.021$); young men showed smaller P_s than young women. This is in contrast to previous studies where males and females showed similar values for both groups (Table 4). Overall, the computed subglottal pressures (10.10–45.70 cmH₂O) were much higher compared to previously reported in-vivo value ranges (normal phonation: 3.5–12.8 cmH₂O, loud phonation: 5.9–27.7 cmH₂O), Table 4. However, the computed P_s values are in the same range as in previous studies ($11.6 \text{ cmH}_2\text{O} \leq P_s \leq 46.3 \text{ cmH}_2\text{O}$) that optimized the 2MM towards human in-vivo [26] and a 3DM model towards human ex-vivo [44] vocal fold dynamics. In [44], the computed P_s values very well approximated the applied and measured P_s values indicating that the here presented values may not be entirely off. High P_s values were also reported for human ex-vivo larynx experiments (up to 44.0 cmH₂O in [96] and up to 35 cmH₂O in [97]). However, the computed P_s values in our study most likely overestimated the actual applied P_s values but are still in reported ranges.

Q_l and Q_r are investigated with regard to their absolute values. Clear differences between the four groups were apparent. Young men showed the lowest values for Q_l and Q_r (0.76–1.99), followed by the older males (1.15–2.36). The computed values are in the same range as reported previously [26]. Older women had the highest values (1.56–4.87) followed by young women (1.93–3.32). Transferring this to the vocal fold physiologically means that younger men and women have higher oscillating masses with smaller stiffness than their older comparison groups. Also, this means that men have larger masses and lower stiffness than the corresponding female groups. It has to be mentioned that the relation “increasing mass–decreasing stiffness” is induced by the modeling parameter Q_α as can be seen in Eq (2) [28]. However, it is generally understood that the vibrating portion of the vocal fold masses usually becomes smaller when vocal fold tension is increased [81, 98]. For young women vs. young men the difference for (Q_l , Q_r) was found statistically significant with $p = 0.000$.

For stiffness, the found gender differences confirm an earlier study where the amplitude quotient (AQ) as an indirect measure of the viscoelastic stiffness of vocal folds was used [15]. The amplitude quotient is determined by the shape and amplitude of the glottal area

waveform. A smaller absolute value of amplitude quotient in young women in comparison with young males was reported, indicative of increased stiffness for the young females. However, it should be noted that the amplitude quotient is not an explicit measure of elasticity. Further, larger absolute values of maximum area declination rate in young women compared to young men were reported [15]. This is indicative of larger absolute peak velocity during the closing phase in young women, hinting to increased stiffness in young women and being again confirmed by the computed larger Q_l , Q_r values, Fig 9A.

Relationships between parameters

It is notable that the computed lower P_s values (10–20 cmH₂O, being 28% of the entire P_s range) account for 97% of all occurring fundamental frequencies (100 Hz– 500 Hz), Fig 9B. This suggests that the subglottal pressure might play a minor role in frequency changes, as observed before [99]. Hsiao et al (2001) showed that the relationship between fundamental frequency and subglottal pressure depends on the tension of the larynx [99]. This means for our results that a lower tension or stiffness (small Q_a), as computed for young and elderly males, also means lower fundamental frequencies compared to young and elderly females, as confirmed in Fig 9C, whereas in contrast the P_s values were only slightly reduced (see means in Table 3) and almost in the same range (Fig 9B). In contrast, higher tension, as computed for both female groups, presents higher fundamental frequencies (Fig 9C) at only slightly increased P_s values. The high dependency between f_0 and stiffness is also expressed by a high Pearson correlation coefficient of 0.986 ($p = 0.000$). This relationship was also seen before

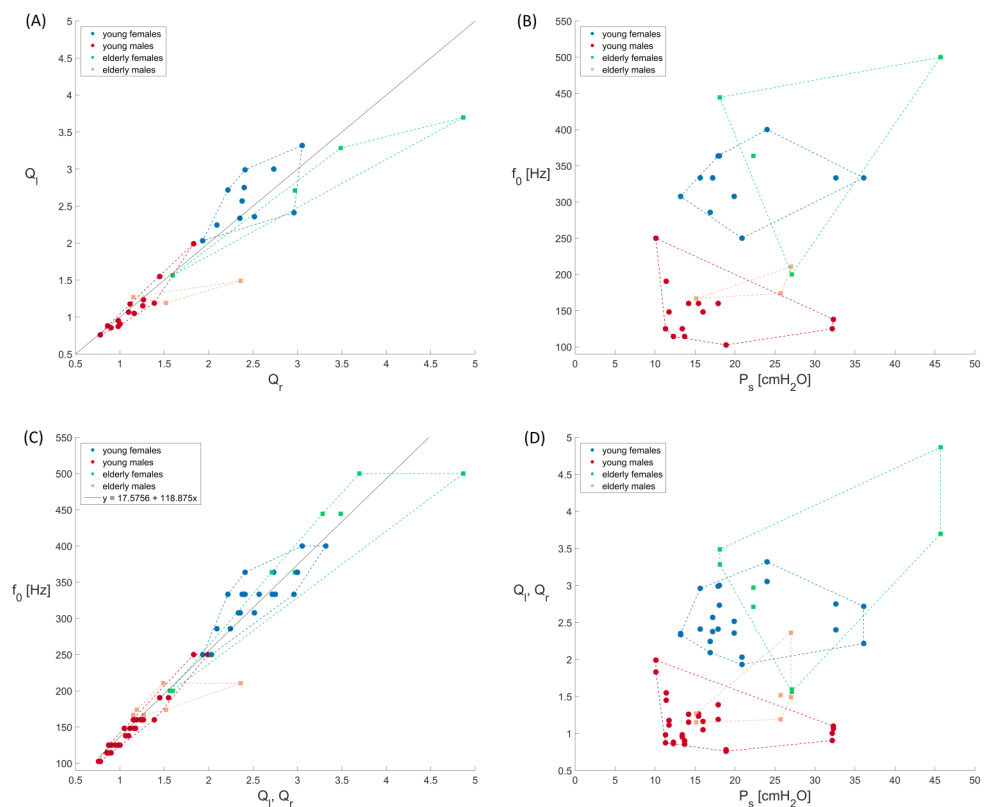


Fig 9. Scatterplots for the distribution of the four groups relating (A) Q_l vs. Q_r , (B) the fundamental frequencies f_0 vs. P_s , (C) f_0 vs. (Q_l, Q_r) and (D) (Q_l, Q_r) vs. P_s .

<https://doi.org/10.1371/journal.pone.0187486.g009>

when for a male and female group different loudness levels (soft–normal–loud: i.e., increasing stiffness) were analyzed [93]. However, this study reported slightly lower P_s values for women compared to men. In summary, the computed P_s values in our study (Table 3) and also the values presented by [93] do overlap for different analyzed subject groups and tasks showing a high inter-individual variability for P_s .

Fig 9D shows the relationship of absolute stiffness and vibrating masses (Q_l , Q_r) to the subglottal pressure P_s . Young males are clearly separated from young and old females. Older males slightly overlap with both female groups. Further, the Fig 9D shows that the values for both young groups are more centered whereas the values for both elderly groups are more spread out and seem not to be as consistent.

Study limitations and outlook

This study has clear limitations due to the sample size. When comparing the optimized 2MM parameters statistical tests were only performed when comparing young men vs. young women. When comparing age related differences and elderly men vs. elderly women no statistical tests were performed and only non-statistical tested trends were described. This lack of statistical significance is clearly a major limitation. Also only healthy young and atrophic elderly subjects were considered. However, the study yielded clear trends and initial group data for younger healthy and elderly atrophic subjects. Future studies should also investigate how the study parameters vary in elderly and young adults during modified phonation (e.g., pitch raise [40]).

Model and optimization limitations. The applied 2MM allows simulation of vocal fold oscillations only in the medio-lateral and not in the vertical direction, as reported before for a three-dimensional model [44]. Also, vibrational characteristics and changes along the vocal fold length (anterior-posterior) cannot be captured by the 2MM since only the trajectories at mid-membranous position (50% of the vocal fold length) are simulated. Hence, anterior-posterior phase differences [100] and typical posterior gaps for female phonation [69] are not captured. For analyzing these characteristics the 6-Mass-Model should be applied and optimized [41]. However, investigating such phenomena was not the focus of this study and will be taken into account in our future work. Also the considered trajectories were always extracted at the standardized 50% vocal fold length position with the assumed largest amplitudes. However, the largest amplitudes vary around this position (from posterior to anterior: females ($41.1\% \pm 10.8\%$) and males ($46.5\% \pm 18.0\%$)) as reported in [16]. Hence in further studies the influence of this assumption and the potential discrepancy towards the exact individual largest amplitude should be investigated for the optimization results.

The computed and optimized P_s values seem to be overestimated by the 2MM since the found values are much higher than assumed and reported for in-vivo measurements (see Table 4), although such high P_s values were reported for ex-vivo studies. However, this issue has to be clarified in future work. In this context, it has to be noted that the goal of investigating and optimizing LMMs towards vocal fold dynamics is not to directly transfer the quantities of computed masses, stiffness and subglottal pressure but to uncover underlying biomechanical differences between vocal fold dynamics [26,28,81,101].

No subgradient-based algorithms were applied for the optimization [102]. Applying such algorithms may enhance the number of correct optimization results.

The success of the optimization procedure was assessed by three objective criteria. However, an explicit objective measure of how the shape of the trajectories was reproduced is a question for future studies.

A dependent variation of masses and stiffness parameters as initially suggested was performed [28]. However, the independent variation and optimization of masses and stiffness

within the 2MM should be considered, since otherwise an increase of mass always goes along with a reduction of stiffness. This dependency might reduce the applicability to certain vocal fold oscillations and also might not reflect certain biomechanical constellations within the vocal folds. Time-dependent parameters should be taken into account, since the 2MM with time-independent parameters does not allow for entirely correct simulation of inter-cycle changes as seen in Fig 8D; i.e., vocal folds show closure during a few cycles and then they do not.

For future classification purposes (i.e., normal vs. pathologies), it might be interesting to vary additional biomechanical parameters like collision and contact forces [103], frequency dependent stiffness [104] and glottal flow [105].

Finally, to enable the clinical application of LMMs in the daily clinical routine in the future, the computational time has to be reduced. Currently, the optimization including the initial value search for one HSV recording takes approximately 60 minutes on a desktop computer.

HSV imaging limitations. HSV imaging projects the three-dimensional vocal fold vibrations and surfaces onto two-dimensional pictures and movies. The image processing detects the dark region between the two vocal folds as glottis. The positions of the most medial edges of the vocal fold tissue from image processing are taken as experimental trajectories ($T_{E\alpha}$). The trajectories ($T_{M\alpha}$) of the 2MM are built from the positions of the upper and lower mass (m_α) depending on which mass is more medial; i.e., visible from above. Since the exact vertical positions of the trajectories within the HSV images cannot be determined, it is unclear if the vertical position (upper or lower mass) of the model trajectories actually corresponds to the same vertical region (superior or inferior vocal fold edge) of the extracted trajectories. Applying HSV in combination with a laser projection unit that allows the reconstruction of the three-dimensional positions of the entire visible vocal fold surfaces would solve this shortcoming [80,106]. The use of a laser projection unit with HSV would further allow for the extraction of the vertical trajectory components of the vocal folds (96); however then a more complex three-dimensional LMM will have to be applied for optimization; e.g. [34].

Owing to the lack of metric specifications in the HSV recordings, average vocal fold lengths were used for males and females, whereas the individual length of the vocal folds was not taken into consideration. Because of the absence of metric units, the recorded vocal fold trajectories are initially scaled in pixels and then transferred to metric units using the averaged vocal fold length. Hence, the amplitudes may not match the actual oscillation quantities accurately. This shortcoming will also be solved in future studies when using HSV in combination with a laser projection unit [106] allowing for the extraction of metric units for vocal fold trajectories and using individual vocal fold lengths in the optimization procedure.

Conclusion

This study is the first approach to use a LMM for comparing age and gender related differences based on vocal fold dynamics recorded with endoscopic high-speed imaging. The parameter optimization objectively quantified biomechanical differences in terms of dynamic symmetry, subglottal pressure, vocal fold masses and stiffness, across gender and age. The results show promising findings for quantifying vocal fold dynamics and for differentiating normal from disordered voice as well as in differentiating between vocal fold pathologies. However, the 2MM does not have one-to-one correspondence to the actual values of the vocal fold masses, stiffness, and subglottal pressure, but allows for objectively evaluating the biomechanical inter-relationships between these variables.

Three different optimization algorithms were tested including three different cost functions. For future studies, the results do not favor a specific optimization algorithm but clearly

show that the Euclidian Distance of the trajectories (Γ_2) should be chosen as cost function to achieve best results.

Supporting information

S1 Data. Individual successful and failed optimization results.
(XLSX)

Author Contributions

Conceptualization: Michael Döllinger, Rita R. Patel, Anne Schützenberger.

Data curation: Pablo Gómez.

Formal analysis: Michael Döllinger, Pablo Gómez.

Funding acquisition: Michael Döllinger, Christoph Alexiou.

Investigation: Rita R. Patel, Christoph Alexiou, Christopher Bohr.

Methodology: Michael Döllinger, Pablo Gómez, Anne Schützenberger.

Resources: Rita R. Patel, Christoph Alexiou, Christopher Bohr.

Software: Michael Döllinger, Pablo Gómez.

Supervision: Michael Döllinger, Anne Schützenberger.

Validation: Pablo Gómez, Christopher Bohr.

Visualization: Pablo Gómez, Christopher Bohr.

Writing – original draft: Michael Döllinger, Rita R. Patel, Anne Schützenberger.

Writing – review & editing: Michael Döllinger, Pablo Gómez, Rita R. Patel, Christoph Alexiou, Christopher Bohr, Anne Schützenberger.

References

1. Chhetri DK, Neubauer J (2015) Differential roles for the thyroarytenoid and lateral cricoarytenoid muscles in phonation. *Laryngoscope* 125(12): 2772–2777. <https://doi.org/10.1002/lary.25480> PMID: 26198167
2. Inwald EC, Döllinger M, Schuster M, Eysholdt U, Bohr C (2011) Multiparametric analysis of vocal fold vibrations in healthy and disordered voices in high-speed imaging. *J Voice* 25(5): 576–590. <https://doi.org/10.1016/j.jvoice.2010.04.004> PMID: 20728308
3. Patel R, Dixon A, Richmond AM, Donohue KD (2012) Pediatric high speed digital imaging of vocal fold vibration: A normative pilot study of glottal closure characteristics. *Int J Pediatr Otorhinolaryngol* 76(7): 954–959. <https://doi.org/10.1016/j.ijporl.2012.03.004> PMID: 22445799
4. Döllinger M (2009) The next step in voice assessment: High-speed digital endoscopy and objective evaluation. *Curr Bioinform* 4(2):101–111. <https://doi.org/10.2174/157489309788184774>
5. Petermann S, Kniesburges S, Ziethe A, Schützenberger A, Döllinger M (2016) Evaluation of Analytical Modeling Functions for the Phonation Onset Process. *Comput Math Methods Med* 2016:8469139. <https://doi.org/10.1155/2016/8469139> PMID: 27066108
6. Echternach M, Döllinger M, Sundberg J, Traser L, Richter B (2013) 2013 Vocal fold vibrations at high soprano fundamental frequencies. *J Acoust Soc Am* 133(2): EL82–87. <https://doi.org/10.1121/1.4773200> PMID: 23363198
7. Warhurst S, McCabe P, Heard R, Yiu E, Wang G, Madil C (2014) Quantitative measurement of vocal fold vibration in male radio performers and healthy controls using high-speed videoendoscopy. *PLoS One* 9(6): e101128. <https://doi.org/10.1371/journal.pone.0101128> PMID: 24971625

8. Patel RR, Dubrovskiy D, Döllinger M (2014) Characterizing vibratory kinematics in children and adults with high-speed digital imaging. *J speech Lang Hear Res* 57(2): 674–686. https://doi.org/10.1044/2014_JSLHR-S-12-0278 PMID: 24686982
9. Gloger O, Lehnert B, Schrade A, Völzke H (2015) Fully automated glottis segmentation in endoscopic videos using local and shape features of glottal regions. *IEEE Trans Biomed Eng* 62(3): 795–806. <https://doi.org/10.1109/TBME.2014.2364862> PMID: 25350912
10. Andrade-Miranda G, Godino-Llorente JI, Moro-Velazquez L, Gomez-Garcia JA (2015) An automatic method to detect and track the glottal gap from high speed videoendoscopic images. *Biomed Eng Online* 14: 100. <https://doi.org/10.1186/s12938-015-0096-3> PMID: 26510707
11. Lohscheller J, Toy H, Rosanowski F, Eysholdt U, Döllinger M (2007) Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos. *Med Image Anal* 11(4): 400–413. <https://doi.org/10.1016/j.media.2007.04.005> PMID: 17544839
12. Patel R, Unnikrishnan H, Donohue KD (2016) Effects of vocal nodules on glottal cycle measurements derived from high-speed videoendoscopy in children. *PLoS One* 11(4): e0154586. <https://doi.org/10.1371/journal.pone.0154586> PMID: 27124157
13. Yokonishi H, Imagawa H, Sakakibara K, Yamauchi A, Nito T, Yamasoba T, et al (2016) Relationship of Various Open Quotients With Acoustic Property, Phonation Types, Fundamental Frequency, and Intensity. *J Voice* 30(2): 145–157. <https://doi.org/10.1016/j.jvoice.2015.01.009> PMID: 25953586
14. Unger J, Schuster M, Hecker DJ, Schick B, Lohscheller J (2016) A generalized procedure for analyzing sustained and dynamic vocal fold vibrations from laryngeal high-speed videos using phonovibrograms. *Artif Intell Med* 66: 15–28. <https://doi.org/10.1016/j.artmed.2015.10.002> PMID: 26597002
15. Patel RR, Dubrovskiy D, Döllinger M (2014) 2014 Measurement of glottal cycle characteristics between children and adults: physiological variations. *J Voice* 28(4): 476–486. <https://doi.org/10.1016/j.jvoice.2013.12.010> PMID: 24629646
16. Lohscheller J, Svec JG, Döllinger M (2013) Vocal fold vibration amplitude, open quotient, speed quotient and their variability along glottal length: kymographic data from normal subjects. *Logoped Phoniatr Vocol* 38(4): 182–192. <https://doi.org/10.3109/14015439.2012.731083> PMID: 23173880
17. Döllinger M, Dubrovskiy D, Patel RR (2012) Spatiotemporal analysis of vocal fold vibrations between children and adults. *Laryngoscope* 122(11):2511–2518. <https://doi.org/10.1002/lary.23568> PMID: 22965771
18. Ziethe A, Patel RR, Kunduk M, Eysholdt U, Graf S (2011) Clinical analysis methods of voice disorders. *Curr Bioinform* 6(3): 270–285. <https://doi.org/10.2174/157489311796904682>
19. Patel RR, Pickering J, Stemple J, Donohue KD (2012) A Case Report in Changes in Phonatory Physiology Following Voice Therapy: Application of High-Speed Imaging. *J Voice* 26(6):734–741. <https://doi.org/10.1016/j.jvoice.2012.01.001> PMID: 22717492
20. Patel RR, Liu L, Galatsanos N, Bless DM (2011) Differential vibratory characteristics of adductor spasmodic dysphonia and muscle tension dysphonia on high-speed digital imaging. *Ann Otol Rhinol Laryngol* 120(1): 21–32. <https://doi.org/10.1177/000348941112000104> PMID: 21370677
21. Deguchi S (2011) Mechanism of and threshold biomechanical conditions for falsetto voice onset. *PLoS One* 6(3): e17503. <https://doi.org/10.1371/journal.pone.0017503> PMID: 21408178
22. Horáček J, Lakkanen AM, Sidlof P, Murphy P, Svec JG (2009) Comparison of acceleration and impact stress as possible loading factors in phonation: a computer modeling study. *Folia Phoniatr Logop* 2009: 61(3):137–145. <https://doi.org/10.1159/000219949> PMID: 19571548
23. Unger J, Lohscheller J, Reiter M, Eder K, Betz CS, Schuster M (2015) A noninvasive procedure for early-stage discrimination of malignant and precancerous vocal fold lesions based on laryngeal dynamics analysis. *Cancer Res* 75(1): 31–39. <https://doi.org/10.1158/0008-5472.CAN-14-1458> PMID: 25371410
24. Miri AK (2014) Mechanical characterization of vocal fold tissue: a review study. *J Voice* 28(6): 657–667. <https://doi.org/10.1016/j.jvoice.2014.03.001> PMID: 25008382
25. Döllinger M, Hoppe U, Lohscheller J, Hettlich F, Schubert S, Eysholdt U (2002) Vibration Parameter Extraction From Endoscopic Image Series of the Vocal Folds. *IEEE Trans Biomed Eng* 49(8): 773–781. <https://doi.org/10.1109/TBME.2002.800755> PMID: 12148815
26. Döllinger M, Braunschweig T, Lohscheller J, Eysholdt U, Hoppe U (2003) Normal voice production: computation of driving parameters from endoscopic digital high speed images. *Methods Inf Med* 42(3): 271–276. <https://doi.org/10.1267/METH03030271> PMID: 12874661
27. Ishizaka K, Isshiki N (1976) Computer simulation of pathological vocal-cord vibration. *J Acoust Soc Am* 60(5): 1193–1198. PMID: 977846
28. Steinecke I, Herzog H (1995) Bifurcations in an asymmetric vocal-fold model. *J Acoust Soc Am* 97(3): 1874–1884. PMID: 7699169

29. Zhang Y, Regner MF, Jiang JJ (2011) Theoretical Modeling and Experimental High-Speed Imaging of Elongated Vocal Folds. *IEEE Trans Biomed Eng* 58(10): 2725–2731. <https://doi.org/10.1109/TBME.2010.2095012> PMID: 21118763
30. Robertson D, Zanartu M, Cook D (2016) Comprehensive, Population-Based Sensitivity Analysis of a Two-Mass Vocal Fold Model. *PLoS One* 11(2): e0148309. <https://doi.org/10.1371/journal.pone.0148309> PMID: 26845452
31. Alipour F, Brücker C, Cook DD, Gömmel A, Kaltenbacher M, Mattheus W, et al (2011) Mathematical Models and Numerical Schemes for the Simulation of Human Phonation. *Curr Bioinform* 6(3): 323–343. <https://doi.org/10.2174/157489311796904655>
32. Schwarz R (2007) Model-based quantification of pathological voice production. PhD thesis. FAU-Erlangen-Nürnberg, Technical Faculty, Shaker Verlag.
33. Döllinger M, Berry DA (2006) Computation of the three-dimensional medial surface dynamics of the vocal folds. *J Biomech* 39(2): 369–374. <https://doi.org/10.1016/j.jbiomech.2004.11.026> PMID: 16321641
34. Yang A, Lohscheller J, Berry DA, Becker S, Eysholdt U, Voigt D, et al (2010) Biomechanical modeling of the three-dimensional aspects of human vocal fold dynamics. *J Acoust Soc Am* 127(2): 1014–1031. <https://doi.org/10.1121/1.3277165> PMID: 20136223
35. Mergell P, Herzel H, Titze IR (2000) Irregular vocal-fold vibratio—high-speed observation and modeling. *J Acoust Soc Am* 108(6): 2996–3002. PMID: 11144591
36. Schwarz R, Hoppe U, Schuster M, Wurzbacher T, Eysholdt U, Lohscheller J (2006) Classification of unilateral vocal fold paralysis by endoscopic digital high-speed recordings and inversion of a biomechanical model. *IEEE Trans Biomed Eng* 53(6): 1099–1108. <https://doi.org/10.1109/TBME.2006.873396> PMID: 16761837
37. Tao C, Zhang Y, Jiang JJ (2007) Extracting physiologically relevant parameters of vocal folds from high-speed video image series. *IEEE Trans Biomed Eng* 54(5):794–801. <https://doi.org/10.1109/TBME.2006.889182> PMID: 17518275
38. Pinheiro AP, Stewart DE, Maciel CD, Pereira JC, Oliveira S (2012) Analysis of nonlinear dynamics of vocal folds using high-speed video observation and biomechanical modeling. *Digital Signal Processing* 22(2): 304–313. <https://doi.org/10.1016/j.dsp.2010.11.002>
39. Haldwin PJ, Galindo GE, Daun KJ, Zanartu M, Erath BD, Cataldo E, et al (2016) Non-stationary Bayesian estimation of parameters from a body cover model of the vocal folds. *J Acoust Soc Am* 139(5): 2683. <https://doi.org/10.1121/1.4948755> PMID: 27250162
40. Wurzbacher T, Schwarz R, Döllinger M, Hoppe U, Eysholdt U, Lohscheller J (2006) Model-based classification of nonstationary vocal fold vibrations. *J Acoust Soc Am* 120(2): 1012–1027. PMID: 16938988
41. Schwarz R, Döllinger M, Wurzbacher T, Eysholdt U, Lohscheller J (2008) Spatio-temporal quantification of vocal fold vibrations using high-speed videoendoscopy and a biomechanical model. *J Acoust Soc Am* 123(5): 2717–2732. <https://doi.org/10.1121/1.2902167> PMID: 18529190
42. Wurzbacher T, Döllinger M, Schwarz R, Hoppe U, Eysholdt U, Lohscheller J (2008) Spatio-temporal classification of vocal fold dynamics by a multi mass model comprising time-dependent parameters. *J Acoust Soc Am* 123(4): 2324–2334. <https://doi.org/10.1121/1.2835435> PMID: 18397036
43. Hüttner B, Luegmair G, Patel RR, Ziethe A, Eysholdt U, Bohr C, et al (2015) Development of a time-dependent numerical model for the assessment of non-stationary pharyngoesophageal tissue vibrations after total laryngectomy. *Biomech Model Mechanobiol* 14(1): 169–184. <https://doi.org/10.1007/s10237-014-0597-1> PMID: 24861998
44. Yang A, Berry DA, Kaltenbacher M, Döllinger M (2012) Three-dimensional biomechanical properties of human vocal folds: Parameter optimization of a numerical model to match in vitro dynamics. *J Acoust Soc Am* 131(2): 1378–1390. <https://doi.org/10.1121/1.3676622> PMID: 22352511
45. Yang A, Stingl M, Berry DA, Lohscheller J, Voigt D, Eysholdt U, et al (2011) Computation of physiological human vocal fold parameters by mathematical optimization of a biomechanical model. *J Acoust Soc Am* 130(2): 948–964. <https://doi.org/10.1121/1.3605551> PMID: 21877808
46. Erath BD, Zanartu M, Stewart KC, Plesniak MW, Sommer DE, Peterson SD (2013) A review of lumped-element models of voiced speech. *Speech Commun* 55(5): 667–690. <https://doi.org/10.1016/j.specom.2013.02.002>
47. Cveticanin L (2012) Review on mathematical and mechanical models of the vocal cord. *J Appl Math* Article ID 928591. <http://dx.doi.org/10.1155/2012/928591>
48. Döllinger M, Kniesburges S, Kaltenbacher M, Echternach M (2016) Current methods for modelling voice production. *HNO* 64(2): 82–90. <https://doi.org/10.1007/s00106-015-0110-x> PMID: 26746639

49. Döllinger M, Kaltenbacher M (2016) Preface: Recent Advances in Understanding the Human Phonatory Process. *Acta Acustica united with Acustica* 102(2): 195–208. <http://dx.doi.org/10.3813/AAA.918936>
50. Mittal R, Byron DE, Plesniack MW (2013) Fluid dynamics of human phonation and speech. *Annual Reviews of Fluid Mechanics* 45: 437–467. <https://doi.org/10.1146/annurev-fluid-011212-140636>
51. Döllinger M, Berry DA, Hüttner B, Bohr C (2011) Assessment of local vocal fold deformation characteristics in an in vitro static tensile test. *J Acoust Soc Am* 130(2): 977–985. <https://doi.org/10.1121/1.3605671> PMID: 21877810
52. Hsiao TY, Wang CL, Chen CN, Hsieh FJ, Shau YW (2002) Elasticity of human vocal folds measured in vivo using color Doppler imaging. *Ultrasound Med Biol* 28(9): 1145–1152. PMID: 12401384
53. Awan SN, Novaleski CK, Yingling JR (2013) Test-retest reliability for aerodynamic measures of voice. *J Voice* 27(6): 674–684. <https://doi.org/10.1016/j.jvoice.2013.07.002> PMID: 24119644
54. Sulter AM, Wit HP (1996) Glottal volume velocity waveform characteristics in subjects with and without vocal training, related to gender, sound intensity, fundamental frequency, and age. *J Acoust Soc Am* 100(5): 3360–3373 PMID: 8914317
55. Perkell JS, Hillman RE, Holmberg EB (1994) Group differences in measures of voice production and revised values of maximum airflow declination rate. *J Acoust Soc Am* 96(2): 695–698. PMID: 7930069
56. Titze IR. *Principles of Voice Production*. University of Michigan, Prentice Hall; 1994.
57. Forero Mendoza LA, Cataldo E, Vellasco MM, Silva MA, Apolinario JA (2014) Classification of vocal aging using parameters extracted from the glottal signal. *J Voice* 28(5): 532–537. <https://doi.org/10.1016/j.jvoice.2014.02.001> PMID: 24880675
58. Awd Allah RS, Dkhil MA, Farhoud E (2009) Fibroblasts in the human vocal fold mucosa: an ultrastructural study of different age groups. *Singapore Med J* 50(2): 201–207. PMID: 19296037
59. Ishii K, Zhai WG, Akita M, Hirose H (1996) Ultrastructure of the lamina propria of the human vocal fold. *Acta Otolaryngol* 116(5): 778–782. PMID: 8908260
60. Hirano M, Sato K, Nakashima T (2000) Fibroblasts in geriatric vocal fold mucosa. *Acta Otolaryngol* 120(2): 336–340. PMID: 11603802
61. Sato K, Hirano M, Nakashima T (2002) Age-related changes of collagenous fibers in the human vocal fold mucosa. *Ann Otol Rhinol Laryngol* 111(1): 15–20. <https://doi.org/10.1177/000348940211100103> PMID: 11800365
62. Pontes P, Brasolotto A, Behlau M (2005) Glottic characteristics and voice complaint in the elderly. *J Voice* 19(1):89–94. <https://doi.org/10.1016/j.jvoice.2004.09.002> PMID: 15766853
63. Ximenes Filho JA, Tsuji DH, do Nascimento PH, Sennes LU (2003). Histologic changes in human vocal folds correlated with aging: a histomorphometric study. *Ann Otol Rhinol Laryngol* 112(10): 894–898. <https://doi.org/10.1177/000348940311201012> PMID: 14587982
64. Hirano M, Kurita S, Sakaguchi S (1989) Ageing of the vibratory tissue of human vocal folds. *Acta Otolaryngol* 107(5–6): 428–433. PMID: 2756834
65. Higgins MB, Saxman JH (1991) A comparison of selected phonatory behaviors of healthy aged and young adults. *J Speech Hear Res* 34(5): 1000–1010. PMID: 1749230
66. Ramig LO, Gray S, Baker K, Corbin-Lewis K, Buder E, Luschei E, Coon H, et al (2001) The aging voice: a review, treatment data and familial and genetic perspectives. *Folia Phoniatr Logop* 53(5): 252–265. 52680 <https://doi.org/10.1159/000052680> PMID: 11464067
67. Goozee JV, Murdoch BE, Theodoros DG, Thompson EC (1998) The effects of age and gender on laryngeal aerodynamics. *Int J Commun Disord* 33(2): 221–238. PMID: 9709440
68. Roy N, Kim J, Courey M, Cohen SM (2016) Voice disorders in the elderly: A national database study. *Laryngoscope* 126(2): 421–428. <https://doi.org/10.1002/lary.25511> PMID: 26280350
69. Yamauchi A, Imagawa H, Yokonishi H, Nito T, Yamasoba T, Goto T, et al (2012) Evaluation of vocal fold vibration with an assessment form for high-speed digital imaging: comparative study between healthy young and elderly subjects. *J Voice* 26(6): 742–750. <https://doi.org/10.1016/j.jvoice.2011.12.010> PMID: 22521532
70. Pontes P, Yamasaki R, Behlau M (2006) Morphological and functional aspects of the senile larynx. *Folia Phoniatr Logop* 58(3): 151–158. <https://doi.org/10.1159/000091729> PMID: 16636563
71. Dubrovskiy D (2017) Bildverarbeitung bei endoskopischen Hochgeschwindigkeits-aufnahmen der Stimmlippenbewegungen. PhD thesis, Faculty of Engineering at FAU Erlangen-Nürnberg, Germany.
72. Chen J (2014) Vocal fold analysis from high speed videoendoscopic data. PhD thesis, Department of Electrical & Computer Engineering, LSU Baton Rouge, LA, USA.

73. Bohr C, Kräck A, Dubrovskiy D, Eysholdt U, Svec J, Psychogios G, et al (2014) Spatiotemporal analysis of high-speed videolaryngoscopic imaging of organic pathologies in males. *J Speech Lang Hear Res* 57(4): 1148–1161. https://doi.org/10.1044/2014_JSLHR-S-12-0076 PMID: 24686496
74. Chen G, Kreiman J, Gerratt BR, Neubauer J, Shue YL, Alwan A (2013) Development of a glottal area index that integrates glottal gap size and open quotient. *J Acoust Soc Am* 133(3): 1656–1666. <https://doi.org/10.1121/1.4789931> PMID: 23464035
75. Dippold S, Voigt D, Richter B, Echternach M (2015) High-Speed Imaging Analysis of Register Transitions in Classically and Jazz-Trained Male Voices. *Folia Phoniatri Logop* 67(1): 21–28. <https://doi.org/10.1159/000381095> PMID: 25967736
76. Bössenecker A, Berry DA, Lohscheller J, Eysholdt U, Döllinger M (2007) Mucosal wave properties of a human vocal fold. *Acta Acust United Acust* 93(5):815–823.
77. Luegmair G, Mehta DD, Kobler, Döllinger J, M. (2015) Three-dimensional optical reconstruction of vocal fold kinematics using high-speed videomicroscopy with a laser projection system. *IEEE Trans Med Imaging* 34(12): 2572–2582. <https://doi.org/10.1109/TMI.2015.2445921> PMID: 26087485
78. Luegmair G, Kniesburges S, Zimmermann M, Sutor A, Eysholdt U, Döllinger M (2010) Optical reconstruction of high-speed surface dynamics in an uncontrollable environment. *IEEE Trans Med Imaging* 29(12): 1979–1991. <https://doi.org/10.1109/TMI.2010.2055578> PMID: 21118756
79. Tokuda IT, Iwakawa M, Sakakibara KI, Imagawa H, Nito T, Yamasoba T, et al (2013) Reconstructing three-dimensional vocal fold movement via stereo matching. *Acoust Sci & Tech* 34(5): 374–377. <https://doi.org/10.1250/ast.34.374>
80. Semmler M, Kniesburges S, Birk V, Ziethe A, Patel RR, Döllinger M (2016) 3D Reconstruction of Human Laryngeal Dynamics Based on Endoscopic High-Speed Recordings. *IEEE Trans Med Imaging* 35(7): 1615–1624. <https://doi.org/10.1109/TMI.2016.2521419> PMID: 26829782
81. Ishizaka K, Flanagan JL (1972) Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Syst Tech J* 51: 1233–1268.
82. Su MC, Yeh TH, Tan CT, Lin CD, Linne OC, Lee SY (2002) Measurement of adult vocal fold length. *J Laryngol Otol* 116(6):447–449. <https://doi.org/10.1258/0022215021911257> PMID: 12385357
83. Titze IR. *The Myoelastic Aerodynamic Theory of Phonation*. National Center for Voice and Speech; 2006
84. Hertegard S, Hakansson A, Thorstensen Ö (1993) Vocal fold measurements with computer tomography. *Logoped Phoniatri Vocol* 18(2–3): 57–63. <https://doi.org/10.3109/14015439309101350>
85. Patel RR, Donohue KD, Johnson WC, Archer SM (2011) Laser projection imaging for measurement of pediatric voice. *Laryngoscope* 121(11):2411–2417. <https://doi.org/10.1002/lary.22325> PMID: 21993904
86. Patel RR, Donohue KD, Lau D, Unnikrishnan H (2013) In vivo measurement of pediatric vocal fold motion using structured light laser projection. *J Voice* 27(4):463–472. <https://doi.org/10.1016/j.jvoice.2013.03.004> PMID: 23809569
87. Döllinger M (2002) Parameter estimation of vocal fold dynamics by inversion of a biomechanical model. PhD thesis, Faculty of Engineering at FAU Erlangen-Nürnberg, Germany.
88. Pham DT, Castellani M (2009) The Bees Algorithm: Modelling foraging behavior to solve continuous optimization problems. *Proc. IMechE—J. Mechanical Engineering Science* 223 Part C: 2919–2938. <https://doi.org/10.1243/09544062JMES1494>
89. Ahmad K, Yan Y, Bless D (2012) Vocal fold vibratory characteristics of healthy geriatric females—analysis of high-speed digital images. *J Voice* 26(6): 751–759. <https://doi.org/10.1016/j.jvoice.2011.12.002> PMID: 22633334
90. Lundy DS, Silva C, Casiano RR, Lu FL, Xue JW (1998) Cause of hoarseness in elderly patients. *Otolaryngol Head Neck Surg* 118(4): 481–485. <https://doi.org/10.1177/01945989811800409> PMID: 9560099
91. Yamauchi A, Yokonishi H, Imagawa H, Sakakibara K, Nito T, Tayama N, et al (2015) Quantitative analysis of digital videokymography: a preliminary study on age- and gender-related difference of vocal fold vibration in normal speakers. *J Voice* 29(1): 109–119. <https://doi.org/10.1016/j.jvoice.2014.05.006> PMID: 25228432
92. Honjo I, Isshiki N (1980) Laryngoscopic and voice characteristics of aged persons. *Arch Otolaryngol* 106(3): 149–150 PMID: 7356434
93. Holmberg EB, Hillman RE, Perkell JS (1988) Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am* 84(2): 511–529. PMID: 3170944
94. Baken RJ, Orlikoff RF. *Clinical Measurement of Speech and Voice*. 2nd Edition, Delmar Cengage Learning; 2010

95. Hertegard S, Gauffin J, Lindestad PA (1995) A comparison of subglottal and intraoral pressure measurements during phonation. *J Voice* 9(2):149–155. PMID: [7620537](#)
96. Döllinger M, Berry DA, Kniesburges S (2016) Dynamic vocal fold parameters with changing adduction in ex-vivo hemilarynx experiments. *J Acoust Soc Am* 139(5): 2372–2385. <https://doi.org/10.1121/1.4947044> PMID: [27250133](#)
97. Döllinger M, Gröhn F, Berry DA, Eysholdt U, Luegmair G (2014) Preliminary results on the influence of engineered artificial mucus layer on phonation. *J Speech Lang Hear Res* 57(2):637–647. https://doi.org/10.1044/2014_JSLHR-S-12-0277 PMID: [24686925](#)
98. Hollien H, Colton RH (1969) Four laminagraphic studies of vocal fold thickness. *Folia Phoniatr* 21(3): 179–198. PMID: [5380734](#)
99. Hsiao TY, Liu CM, Luschei ES, Titze IR (2001) The effect of cricothyroid muscle action on the relation between subglottal pressure and fundamental frequency in an in vivo canine model. *J Voice* 15(2): 187–193. [https://doi.org/10.1016/S0892-1997\(01\)00020-0](https://doi.org/10.1016/S0892-1997(01)00020-0) PMID: [11411473](#)
100. Neubauer J, Mergell P, Eysholdt U, Herzel H (2001) Spatio-temporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes. *J Acoust Soc Am* 110(6):3179–3192. PMID: [11785819](#)
101. Mehta DD, Zanartu M, Quatieri TF, Deliyiski DD, Hillmann RE (2011) Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videodendoscopy. *J Acoust Soc Am* 130(6):3999–4009. <https://doi.org/10.1121/1.3658441> PMID: [22225054](#)
102. Alt W. *Nichtlineare Optimierung*. 1st ed. Friedr. Vieweg & Sohn Verlagsgesellschaft, Braunschweig/Wiesbaden; 2002
103. Sommer DE, Erath BD, Zanartu M, Peterson SD (2012) Corrected contact dynamics for the Steinecke and Herzel asymmetric two-mass model of the vocal folds. *J Acoust Soc Am* 132(4): EL271–276. <https://doi.org/10.1121/1.4734013> PMID: [23039564](#)
104. Rupitsch SJ, Ilg J, Sutor A, Lerch R, Döllinger M (2011) Simulation Based Estimation of Dynamic Mechanical Properties for Viscoelastic Materials Used for Vocal Fold Models. *Journal of Sound and Vibration* 330(18–19): 4447–4459. <https://doi.org/10.1016/j.jsv.2011.05.008>
105. De Vries MP, Schutte HK, Veldman AE, Verkerke GJ (2002) Glottal flow through a two-mass model: comparison of Navier-Stokes solutions with simplified models. *J Acoust Soc Am* 111(4): 1847–1853. PMID: [12002868](#)
106. Semmler M., Kniesburges S., Parchent J., Jakubaß B., Zimmermann M., Bohr C., et al (2017) Endoscopic laser-based 3D imaging for functional voice diagnostics. *Appl. Sci.* 2017, 7, 600; <https://doi.org/10.3390/app7060600>