Modeling asymmetric vocal fold vibrations

R. Schwarz[1], M. Schuster[1], T. Wurzbacher[1], U. Eysholdt[1], U. Hoppe[2], and J. Lohscheller[1]

[1]Department of Phoniatrics and Pediatric Audiology
Erlangen University Hospital, Germany
raphael.schwarz@phoni.imed.uni-erlangen.de

[2]Institute of Biomedical Engineering and Informatics
Ilmenau Technical University, Germany
ulrich.hoppe@tu-ilmenau.de

Abstract

Hoarseness in case of unilateral recurrent laryngeal nerve paralysis is caused by irregular, asymmetric vocal fold oscillations [1]. Using a high-speed camera system, these oscillations are observable in real-time. An image processing algorithm extracts the motion of the vocal folds. This work concerns the problem of reproducing the vocal fold dynamic using a biomechanical model in case of one-sided recurrent laryngeal nerve paralysis. Such a procedure, which adapts the model parameters to match the extracted dynamic of the vocal folds is the solving of an inverse problem. The values of the model parameters describe qualitatively the physiological properties of the vocal folds. The results of the inversion procedure are compared with perturbation measures defined for the extracted vocal fold oscillations.

1. Introduction

The properties of the voice signal, e.g. pitch and sound pressure level, depend on the subglottal pressure and on the vocal folds' positions and tensions. The recurrent laryngeal nerve (RLN) provides motor control signals for the intrinsic laryngeal muscles. They are important for a properly symmetric adjustment of the vocal folds and thus for normal voice production.

An injury of this nerve causes a recurrent laryngeal nerve paralysis (RLNP). Its predominant symptom is dysphonia which is called 'paralytic'. In consequence of an unilateral RLNP a symmetric adjustment of the vocal folds is inhibited. Asymmetries within the larynx morphology and physiology due to unilateral paralysis entail perturbations of the vocal fold oscillations which results into hoarse voice [2].

The detection and quantification of laryngeal asymmetries are important issues in medical voice assessment. This work addresses the approach of reproducing vocal fold oscillations, achieved from high-speed recordings, with a biomechanical model of the vocal folds in case of unilateral RLNP, i.e. solving an inverse problem. The aim of the inversion procedure is to determine parameter values which provide data about vocal fold tension and other physiological vocal fold parameters, e.g. vibrating masses or damping coefficients. Solving an inverse problem with a numerousness of parameters makes high demands on the computing power. Thus, complex and high-dimensional biomechanical models of the vocal folds are less convenient concerning the inversion procedure. Therefore, the more simple lumped-element model after [3], the two-mass model (2MM) is used. The 2MM demands only a small parameter set to describe the most dominant modes of the vocal fold oscillations, also in unilateral RLNP. Owing to its conceptual simplicity and the interpretative power of this low-dimensional model, it was successfully applied to obtain inversion results in normal voice subjects [4].

2. Methods

2.1. Two-mass model and optimization parameters

The most widely used biomechanical model of the vocal folds, the 2MM was presented by Ishizaka&Flanagan [3]. Here, the simplified version by Steineke&Herzel [5] is used. The examined motions of the vocal folds in case of unilateral RLNP show asymmetric characteristics in amplitude and phase. For modeling asymmetric vocal fold oscillations the model parameters have to be varied purposefully and asymmetry has to be introduced within the model. This demands the definition of adequate factors to change the standard parameter set of the 2MM [6]. These factors, in the following called optimization parameters, serve to adapt the oscillations of the left and right lower mass of the 2MM (theoretical curves) to the motions of the pathological vocal folds, extracted from the high-speed recordings (experimental curves). The first optimization parameter Q serves to adjust symmetrically on both sides the lower masses and the corresponding spring constants. The fundamental frequency of the theoretical curve is adapted with Q. Further on, the coupling spring constant between upper and lowers mass of the paralyzed side is scaled with the optimization parameter Q_{kc}. Thus, the necessary asymmetry for modeling one-sided vocal fold paralysis is introduced within the 2MM. The parameter Q_{kc} enables to determine the amplitude ratio between left and right oscillation. For $Q_{kc} = 1$ both sides of the 2MM oscillate with the same amplitude. The third optimization parameter Q_{k1} is used to adjust the spring constant of the lower mass belonging to the paralyzed side. With Q_{k1} the phase shift between the two sides of the 2MM is determined. The case $Q_{k1} = 1$ means a symmetric oscillation, i.e. there is no phase difference. Again, Q_{k1} affects the model parameters asymmetrically. The last optimization parameter is the subglottal pressure P_s. Due to the aerodynamic properties of the 2MM, this parameter has a symmetric influence on the

theoretical curves. P_s is used to adjust the amplitude of the model oscillations.

2.2. Inversion procedure

The aim of the inversion procedure is to reproduce the experimental curves with the theoretical curves. For this, it delivers adequate optimization parameter values. The inversion procedure is subdivided into to two steps. First, a set of initial values for the optimization parameters is determined (initial value search). The knowledge-based initial value search is based on investigated insights into the model behavior during the variation of the model parameters. In the second step, the identified initial values are used within a direct search method to determine the optimal parameter set to reproduce the experimental curves (optimization procedure). The used objective function within the optimization procedure is non-convex and contains numerous local minima within the space spanned by the parameters P_s, Q, Q_{kc} and Q_{k1} [7]. Therefore, the crucial factor to obtain reliable results with the inversion procedure is the initial value search procedure described below.

2.2.1.   Initial value search procedure

The flow diagram of the initial value search procedure is depicted in Fig. 1. The superscript i within the flow diagram indicates that *initial values* are calculated for the optimization parameters. The algorithm starts with a symmetric parameter configuration. Therefore, Q_{kc} and Q_{k1} are set to one.
For each side of the 2MM a factor Q_α (α=l for the left side and α=r for the right side) is calculated after the simple mass-spring oscillator equation to enable different fundamental frequencies. The superscript *exp* of the fundamental frequency $f_{0\alpha}$ indicates the values of the experimental curves, m_{10} and k_{10} are the standard values for the lower mass and corresponding spring constant after [6].
In the next step the subglottal pressure P_s is determined to adjust the oscillation amplitude A_h for the theoretical curve representing the healthy side. P_s is stepwise increased until the difference between the amplitude of the theoretical and the experimental curve is minimized.
In step four a value for Q_{k1} is defined to adapt the phase shift $\Delta\phi$ between the left and the right theortical curve to the value of the experimental curves.
The fifth step determines the parameter Q_{kc}. This is done by minimizing the difference in the amplitude ratio between healthy and paralyzed side of the theoretical curves R_a and the corresponding value for the experimental curves. If it is not possible to fall below a threshold value, the value for Q_{k1} is varied as depicted in Fig. 1. This step is iterated until the difference is below a desired threshold.
After the adaptation of the amplitude ratio, it is verified in step seven, if the difference in the amplitudes of the healthy side of the theoretical and experimental curve is still below a certain threshold. If this is not the case the algorithm is iterated starting from step 3.
The final result of the algorithm is the set of initial parameter values used for the optimization.

2.2.2.   Optimization procedure

The objective function Γ used within the optimization procedure is defined in the frequency domain and represents the difference in magnitude and phase between experimental and theoretical curves [3]. As Γ is non-convex within the parameter space, the Nelder-Mead algorithm, a direct search method, is initialized with the set of initial parameter values to start the optimization near the global minimum of Γ. The optimization procedure delivers the final results for the optimization parameters. The inversion procedure is applied to 15 subjects with unilateral RLNP.
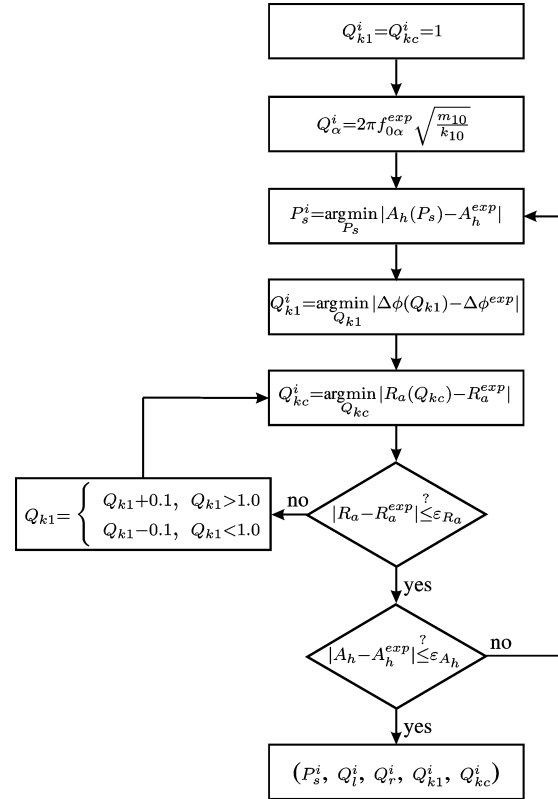

Figure 1: *Initial value search algorithm.*

2.3. Glottis signal analyses

In case of one-sided vocal fold paralysis a reduced voice quality occurs due to the asymmetric dynamic. The distance of the opposing vocal fold edges over the time modulates the air streaming into the vocal tract. In this way the voice quality is substantially affected. The time-varying distance of the edges in the medial glottal third is called glottis signal. Perturbation measures, namely jitter and harmonics-to-noise ratio (HNR), defined for the glottis signal allow to quantify the reduced voice quality due to unilateral RLNP.
Jitter describes a short-time perturbation of the fundamental frequency of the glottis signal, thus a fluctuation of the periodic time from cycle to cycle, defined after [8].
The HNR describes the ratio between the power of the harmonic portion of the glottis signal and noise signal power [9].

3. Results

From the 15 processed high-speed recordings, Fig. 2 exemplarily shows the inversion results for five subjects. The theoretical curves determined by the inversion procedure are in compliance with fundamental characteristics of the experimental curves as amplitude and phase as well as phase difference and amplitude ratio (each between healthy and paralyzed side). An insufficient glottal closure, e.g. observable for subject f3, is reproduced successfully.
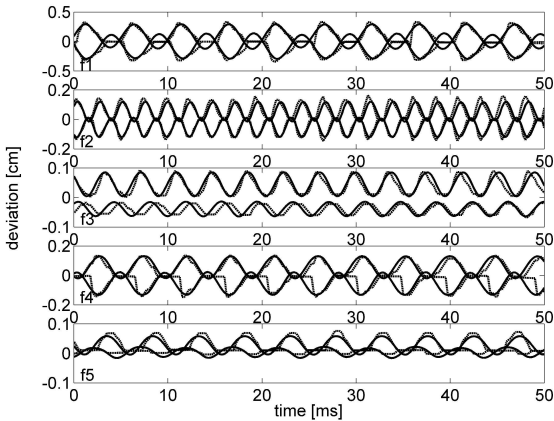

Figure 2: *Results for the subjects f1 to f5. Experimental curves (dotted) extracted from high-speed recordings and theoretical curves (solid line) generated with the two-mass model. The upper curves in each chart represent the healthy side and the lower curves the paralyzed side.*
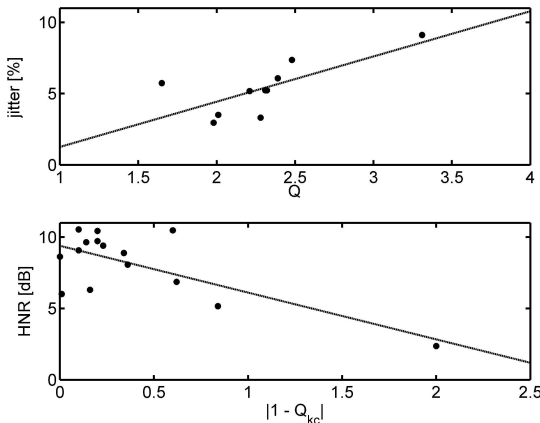

Figure 3: *Top: Optimization parameter Q plotted against the jitter of the glottis signal. The solid line depicts a linear approximation. Bottom: Deviation of the optimization parameter Q_{kc} from 1 plotted against the HNR of the glottis signal. The solid line represents a linear approximation.*

The calculated values for the jitter of the glottis signal vary between 2.96 % and 9.11 %. These values are in accordance to a study of acoustic voice quality [7]. In Fig. 3 (top) the jitter of the glottis signal is plotted against the optimization parameter Q. The solid line represents a linear approximation. For all 15 subjects the fundamental frequencies of the healthy and paralyzed side are identical. Therefore, $Q_l = Q_r = Q$ is assumed. There is a linear relationship between the parameter Q and the fundamental frequency of the theoretical curves. In accordance to [7], an increased fundamental frequency leads to increased jitter.

The HNR of the glottis signal varies between 2.37 dB and 10.57 dB. These values are in accordance to [10]. In Fig. 3 (bottom) the HNR in dB is plotted against the deviation of the optimization parameter Q_{kc} from 1. Again, the solid line depicts a linear approximation. The parameter Q_{kc} serves to adjust the amplitude ratio between healthy and paralyzed side. An increased distance from 1 entails an increased amplitude for the healthy side compared to the paralyzed side. An increased deviation of the amplitude ratio from 1 results in a reduced HNR.

4. Discussion

An inversion procedure is presented which adjusts the dynamic of a biomechanical two-mass model to the vocal fold oscillations in case of unilateral vocal fold paralysis. The motions of the vocal folds are extracted from high-speed recordings. The inversion procedure is successfully applied to 15 subjects with unilateral recurrent laryngeal nerve paralysis. The asymmetries within the vocal fold dynamics concerning amplitude ratio and phase difference are detected. The results for the optimization parameters are examined concerning an interrelation with the voice quality. As the glottis signal, extracted from the high-speed recordings, substantially affects quality of voice, the values for jitter and HNR of the glottis signal are compared with the optimization parameters. An increased value for the parameter Q, which means an increased fundamental frequency for the vocal fold oscillations, leads to an increased jitter. Consequently, this means a rise for the relative fluctuations of the fundamental frequency. This is in accordance to [7]. The more the value for Q_{kc} deviates from 1, the more distinct is the asymmetry within the theoretical curves. The amplitude difference between healthy and paralyzed side increases. Following the assumption that an increased asymmetry in the vocal fold dynamic results in decreased voice quality, the HNR is decreased respectively. The optimization parameters reveal a physiological relevance concerning the voice quality.

The clinical value of the inversion results and their applicability for therapy in unilateral recurrent laryngeal nerve paralysis is investigated in further studies. In this context, importance is ascribed to the obtained interrelation between the quality of the glottis signal and thus the voice quality and the optimization parameters.

5. Acknowledgements

6. References

[1] Eysholdt, U.; Rosanowski, F; Hoppe, U., 2003. Vocal fold vibration irregularities caused by different types of laryngeal asymmetry. *Eur. Arch. Otorhinolarygol.* 260, 412-14.

[2] Isshiki, N.; Tanabe, M.; Ishizaka, K.; Broad, D., 1977. Clinical significance of asymmetrical vocal cord tension. *Ann. Otol. Rhinol. Laryngol.* 86, 58-66.

[3] Ishizaka, K.; Flanagan, J., 1972. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Syst. Techn. J.* 51, 1233-68.

[4] Doellinger, M.; Braunschweiger, 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, 271-6.

[5] Steineke, I.; Herzel, H., 1995. Bifurcations in an asymmetric vocal fold model. *J. Acoust. Soc. Am.* 97, 1571-78.

[6] Doellinger, M.; Hoppe, U.; Hettlich, F.; Lohscheller, J.; Schubert, S.; Eysholdt, U., 2002. Vibration parameter extraction from endoscopic image series of the vocal folds. *IEEE Trans. Biomed. Eng.* 49, 773-81.

[7] Doellinger, M., 2002. Parameter estimation of vocal fold dynamics by inversion of a biomechanical model. *Kommunikationsstörungen – Berichte aus Phoniatrie und Paedaudiologie*. Aachen, Deutschland: Shaker Verlag.

[8] Ludlow, C.; Bassich, C.; Conner, N.; Coulter, D.; Lee, Y, 1991. The validity of using phoniatry jitter and shimmer to detect laryngeal pathology. In *Laryngeal function in phonation and respiration*, T. Bear, C. Sasaki, K. Harris (ed.). Singular Publishing Group INC, 492-508.

[9] Huang, D.; Kasuya, F.; Lin, S., 1995. Measure of vocal function during changes in vocal effort level. *J. Voice* 9, 429-38.

[10] Ferrand, C., 2002. Harmonics-to-noise ratio: An index of vocal aging. *J. Voice* 16(4), 480-87.