Oral Presentations

From *in silico* to *in vivo*: Applying Simulation to Laryngeal Surgery Planning

Ted Mau, MD PhD [1], Anil Palaparthi, MS[2], Tobias Riede, PhD[2,3], Ingo R. Titze, PhD [2,4]

[1]University of Texas Southwestern Medical Center, [2]National Center for Voice and Speech,University of Utah, [3]Midwestern University,[4]The University of Iowa

ted.mau@utsouthwestern.edu

Introduction

The vocal consequence of laryngeal surgery can be unpredictable. This reflects the fact that surgery often has a singular aim, but sound production by flow-induced vocal fold oscillation is a complex process. A single acoustic property (e.g. pitch) has a dynamic dependence on multiple vocal fold morphologic parameters. Conversely, surgically changing one morphologic parameter may lead to simultaneous changes in multiple acoustic parameters (e.g. pitch, pitch range, SPL). To tackle the problem of hard-to-predict vocal outcomes, surgeons may benefit from voice simulation, which by its nature integrates the complex mapping between vocal fold morphology, muscle activation, and acoustic output. The specific aim of this study is to illustrate the application of voice simulation to two surgical problems. First, the relative acoustic merits of two types of cordectomies used to resect small vocal fold cancers are compared. Second, a pitch-lowering laryngeal framework surgery (type IIIB thyroplasty) is simulated to investigate its effect on pitch range.

Methods

The National Center for Voice and Speech simulator using a self-oscillating vocal fold model [1,2] was used to evaluate the acoustic output of alternative vocal fold morphologies over physiologic ranges of subglottal pressure and tissue properties. In Experiment 1, two morphologies corresponding to different depths of cancer resection were simulated: resection of the superficial lamina propria (SLP) alone (subepithelial cordectomy), vs. resection of the SLP plus the vocal ligament (subligamental cordectomy). In Experiment 2, a type IIIB thyroplasty designed to decrease pitch was simulated by reducing the vocal fold length by 25% and reducing the longitudinal shear moduli of all three layers of the vocal fold (muscle, ligament, SLP). The primary outcome measures were the mean F0, F0 range, and SPL. The area in the voice range profile (F0 vs. SPL plot) produced by viable simulations was also compared.
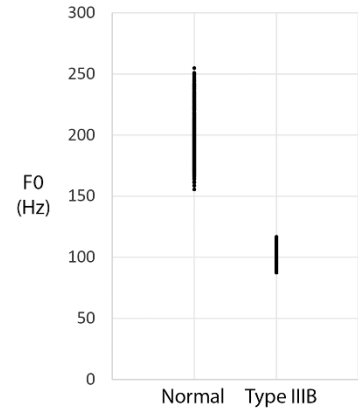
Results

In Experiment 1, subligamental cordectomy, which removes more non-cancerous tissue than subepithelial cordectomy, produced a larger F0-SPL range than subepithelial cordectomy.
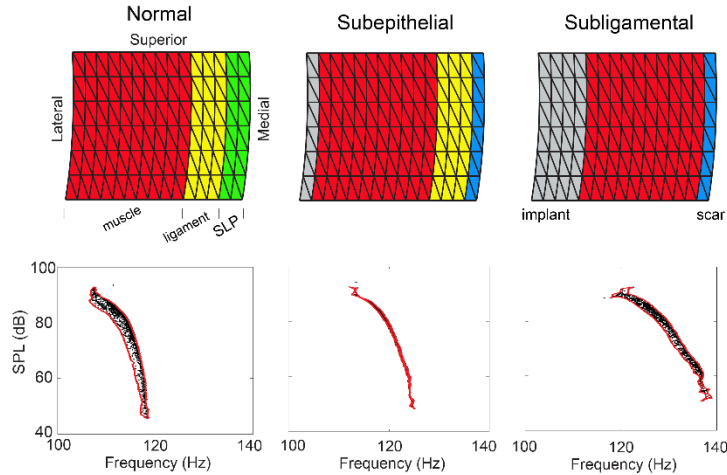
In Experiment 2, the type IIIB thyroplasty lowered the mean F0 from 199.2 Hz to 102.5 Hz as the surgery was intended to accomplish. However, the F0 range decreased from 0.71 octave to 0.42 octave.

Conclusions

These results suggest that in the resection of small vocal fold cancers, the conventional wisdom of maximally preserving normal tissue can paradoxically lead to a worse voice. A better voice outcome may be achieved by removing the vocal ligament rather than preserving it, because removing the vocal ligament leads to a greater F0-SPL range. This finding is consistent with clinical reports [3]. In type IIIB thyroplasty to lower pitch, the desired decrease in pitch is achieved at the expense of a reduced pitch range. The patient should be counseled about this expected outcome. These two examples illustrate the use of voice simulation to forecast voice outcome that may not have been predicted based on a surgeon's intuition alone.

References

1. Titze, I. R., Palaparthi A, & Smith, S. L. (2014). Benchmarks for time-domain simulation of sound propagation in soft-walled airways: Steady configurations. J. Acoust. Soc. Am. 128(2), 828–838.

2. Alipour, F., Berry, D. A., & Titze, I. R. (2000). A finite-element model of vocal-fold vibration. J. Acoust. Soc. Am. 108(6), 3003-3012.

3. Hillel, A. T., Johns, M. M., Hapner, E. R., Shah, M., Wise, J. C., & Klein, A. M. (2013) Voice outcomes from subligamentous cordectomy for early glottic cancer. Ann. Otol. Rhinol. Laryngol. 122(3), 190-196.

Surgical Modeling Using Dynamic Voice CT to Improve Voice Post Airway Reconstruction

Alessandro de Alarcon[1,2], Liran Oren[2], Robert Fleck[3], Goutham Mylavarapu[4], Ephraim Gutmark[4], Dimitar Deliyski, [5] and Sid Khosla[2]

[1]Pediatric Otolaryngology-Head and Neck Surgery, Cincinnati Children's Hospital Medical Center, [2] Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati, [3]Department of Radiology, Cincinnati Children's Hospital Medical Center [4]Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, [5]Department of Communication Sciences and Disorders, Michigan State University

alessndro.dealarcon@cchmc.org

Introduction

Up to 50% of children who undergo complex airway reconstruction develop dysphonia. Evaluation and treatment of these patients is difficult as many have severe voice disorders characterized by short phonation times, aperiodic signals, and supraglottic vibration. Combing Dynamic Voice CT and computational modeling is a new modality that provides information that may be useful in surgical planning to improve these patients' voices.

The feasibility of this modality to predict surgical success is shown using two case studies; aphonic patient who underwent an endoscopic posterior glottic reduction and supraglottic patient who underwent open laryngoplasty with anterior commisureplasty.

Methods

The pre-operative geometry of the models was set up using dynamic CT scans. The scans were taken while the patients where attempting to phonate. Compressible large eddy simulation (LES) was employed to numerically investigate airflow characteristics in the model.

Post-operative simulations were perform using virtual surgery on the reconstructed pre-surgery models. The virtual surgery process generated another set of airway models with the airway anatomy changed to account for the particular surgery that was performed. LES simulations were performed on the new models and the results were compared with the post-operative data taken using high-speed video endoscopy.

Results

Pre-operative model of the supraglottic patient predicted formation of localized vortical structure above the false vocal folds and showed higher wall shear stress (compared with the aphonic patient). Nominal changes in the flow field were predicted post-operative. Post-operative model of the aphonic patient predicted higher flow unsteadiness, formation of vortical structure in the flow and increase wall shear stress. These flow characteristics are favorable for glottic vibrations, which were confirmed using high-speed video data.

The potential of using dynamic CT scans for flow structure interaction (FSI) in virtual surgery simulations in order to predict surgical success of airway reconstruction will be further discuss.

Conclusions

Flow features predicted in post-operative model using on virtual surgery technique are similar to flow features that observed in normal folds vibrations.

Comparing arytenoid adduction and infraglottal medialization to glottal medialization alone in an excised canine larynx model.

Sid Khosla[1], Liran Oren[1], Ephraim Gutmark[2]

[1]Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati

[2]Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati

Khoslam@ucmail.uc.edu

Introduction

In unilateral vocal fold paralysis (UVFP), there is usually structural vocal process asymmetry in the medial-lateral direction. There can also be vocal process asymmetries in the superior-inferior and anterior-posterior directions in addition to vocal fold stiffness asymmetries. The most common procedures for UVFP are thyroplasty type I (TT1) and injection laryngoplasty. Two anecdotal observations from many laryngologists are as follows. First, "better" voice is produced when there is greater medialization of the infraglottal and inferior glottis than there is of the superior aspect (which will be defined as increased inferior stiffness gradient (IISG)). Second, if there are moderate to severe asymmetries, an arytenoid adduction produces a "better voice". From our previous work we have seen that intraglottal vortices form in a divergent duct and produce negative pressures during closing in the superior half of the glottis. The strength of the vortices (SV) is highly correlated with the maximum divergence angle(MDA), maximum flow declination rate (MFDR), acoustic intensity (SPL), and acoustic energy in the higher harmonics (HH).

Our hypotheses are that MDA, SPL, and HH will be increased in an excised canine larynx UVFP model if:

1. IISG in a thyroplasty type I produces greater MDA, SPL, and HH relative to equal medialization of inferior and superior aspects of the fold.

2. Arytenoid adduction (AA) is added to TT1 relative to TT1 alone.

Methods.

High-speed 2D particle imaging velocimetry (PIV) is completed during phonation of the excised larynx. 3 excised canine larynges are used. No vocal tract is used, thus minimizing the effects of inertance. During each phonation trial, intraglottal velocity fields and intraglottal geometry are obtained for the mid-membranous plane. Pressures during closing are then computed from intraglottal velocity and geometry for multiple phases during

measure. EGG, high speed videography of the vocal fold vibrations and acoustics are also measured. For each larynx, baseline (both folds adducted to midline) tests are run. Then one fold is adducted to midline, and the other fold is in the non-adducted lateralized position – this simulates minimal reinnervation in UVFP. The following two conditions are tested: TT1 with IISG, TT1 without IISG or any other stiffness gradient. The optimal TT1 is then tested relative to TT1 + AA.

Results.

Up to a certain point, increasing IISG in a TT1 produces stronger vortices and greater MDA, SPL, HH, MFDR relative to a TT1 with no stiffness gradient. TT1 + AA also increases MDA of at least one fold relative to TT1 alone.

Conclusions

In cases of moderate-severe vocal process asymmetries in UVFP, increasing the inferior stiffness gradient and adding an AA to a TT1 results in improving parameters that are likely to lead to less vocal fatigue and higher intelligibility in noise. However, our current model can only determine correlations. Computational models are needed to determine causation and optimal conditions.

Using the *in vivo* canine model to assess and treat laryngeal disorders

Dinesh Chhetri [1]

[1] University of California, Los Angeles

dchhetri@mednet.ucla.edu

Introduction

Voice disorders impair quality of life and are a significant disease burden in our society. The intrinsic laryngeal muscles (ILMs) of the larynx are one of the important input variables controlling voice production. In treating disorders of voice production, laryngeal surgeons often perform procedures that mimic ILM activation (e.g. thyroplasty type 1 and arytenoid adduction). Thus, investigations of the effects of ILM activation on voice production are relevant for understanding voice production as well as treatment of dysphonia.

Methods

An *in vivo* canine model was used and a methodology for studying symmetric and asymmetric activation states of laryngeal muscles was developed. Surgical dissection of the larynx was performed to identify distal nerve branches for activation of individual laryngeal muscles. Up to three nerves on each side of the larynx could be concurrently and independently stimulated. Symmetric and asymmetric activation conditions were tested for effects on acoustics, aerodynamics, and vibration.

Results

There were complex interactions between the cricothyroid (CT) and thyroarytenoid (TA) muscles in fundamental frequency (F0) and vocal intensity (SPL) control, and a surprisingly special role for the lateral cricoarytenoid (LCA) muscles in facilitating the CT/TA interactions. Laryngeal asymmetry led to a predicable pattern of vibratory phase asymmetry. LCA/IA activation facilitated the full range of TA and CT interactions.

Conclusions

Graded activation of individual laryngeal muscles via distal neuromuscular activation allows for comprehensive study of the effects of symmetric and asymmetric activation of the larynx in voice production. The ILMs can interact in complex ways to reach the same F0, SPL, and vocal efficiency. The implications for laryngeal surgery will be discussed.

Modeling voice production with time-delay systems: the larynx tube

F. Silva [1] G. Artana [2], D. Sciamarella [3]

[1] LMA-CNRS, [2], LFD-FIUBA, [3]LIMSI-CNRS

denisse.sciamarella@limsi.fr

Introduction

Time-delay systems (shortly, TDS) are also called systems with aftereffect or dead-time, hereditary systems, equations with deviating argument or differential-difference equations [1]. In voice production, TDS play a role in voice-production modeling when source and tract are coupled allowing for delayed feedback on the vocal fold dynamics [2,3]. This work undertakes the incorporation of the larynx tube to this modeling scheme, following an approach inspired in the assimilation of the larynx tube to a Helmholtz resonator, introduced four decades ago to study the singing formant [4].

Methods

The glottal source is modeled using a TDS approach that takes up the approach developed in Ref. [3]. The vocal tract is modeled in two components: the larynx tube is described using lumped elements accounting for the acoustic compliance and mass, while a classic lossy delay line accounts for the remaining downstream airways. This enables the separate control of the dynamic characteristics of the acoustic resonator. Extending the acoustical study in Ref. [4], this resonator is coupled with the mucosal wave model of the vocal folds vibration and time domain simulations are performed.

Results

The TDS including the larynx tube differs from previous approaches in the appearance of combined delay terms and derivative terms in the acoustical equation. This system is used as a voice simulator for different control parameter values, in particular the characteristics of the acoustic resonator. Numerical examples illustrate the richness of the solutions introduced by the new delay terms in the equations.

Conclusion

sModeling voice production with time delay systems can be improved if the larynx tube is included in the scheme. Time delay effects are enhanced through the incorporation of a resonator connecting the vocal source with the vocal tract. The model has the feature of directly mapping the variation of the volume of the Morgagni sinus in a low-order model, allowing to test the articulatory interpretation of the signing formant within a theoretical framework accounting for the glottal source.

References

1 Jean-Pierre Richard, Time-delay systems:an overview of some recent advances and open problems. Automatica 39 (10) pp 1667–1694 (2003)

2 R. Laje, T. Gardner, G. Mindlin. Continuous model for vocal fold oscillations to study the effect of feedback. Phys. Rev. E 64, 056201 (2001)

3 D. Sciamarella, G. Artana. Relaxation to one-dimensional postglottal flow in a vocal fold model. Speech Communication 66, 176-181 (2015)

4 J. Sundberg. Articulatory interpretation of the "singing formant". Journal of Acoustical Society of America 55 (4), 838-844 (1974

Modeling the pathophysiology of vocal hyperfunction

Matias Zañartu[1], Gabriel E. Galindo[1], Sean D. Peterson[2],
Byron D. Erath[3], and Robert E. Hillman[4].

[1] Universidad Técnica Federíco Santa María, [2] University of Waterloo,
[3] Clarkson University, [4] Massachusetts General Hospital.

matias.zanartu@usm.cl

Introduction

Many of the most common voice disorders are chronic or recurring conditions that are believed to result from repeated detrimental patterns of vocal behavior, referred to as vocal hyperfunction (VH). This can sometimes lead to the formation of benign vocal fold lesions (e.g., nodules), which is referred to as phonotraumatic VH. Subjects with these disorders often attempt to compensate for their vocal deficits, for example by increasing subglottal pressure in response to reduced loudness due to excessive breathiness, which can trigger a vicious cycle of worsening pathology.

In spite of the significant prevalence of these voice disorders, very little is known about the underlying biomechanical and aero-acoustic mechanisms of VH. The aim to this study is to describe the underlying biomechanical deficiencies and faulty compensations related to VH, using a multi-physics numerical modeling approach. In particular, the goal is to assess if VH leads to higher risk factors for the development of benign vocal fold lesions.

Methods

The numerical model includes kinematic, aerodynamic, and acoustical modules that are interconnected via VF posturing, fluid-structure interactions, and nonlinear source-filter interactions. An extended asymmetric implementation of a six-mass body-cover model that incorporates a triangular glottis and a posterior glottal opening (PGO) was used. These additions allow for a more realistic glottal configuration that is critical for studying VH. Thus, incomplete glottal closure in the model can occur due to a membranous glottal opening (MGO) and a PGO.

In this study, VH is conceptualized as "incorrect" glottal configurations that are corrected to restore a target output through "faulty" compensations. The selected glottal configuration is incomplete closure with both MGO and PGO linked together through a model of vocal posturing. Faulty compensatory mechanisms include vocal tract constriction, increased VF tensioning via cricothyroid muscle activation, and increased lung pressure. The objective target is radiated sound pressure level (SPL), wherein the Nelder Mead minimization technique was used to find the minimum RMS error between the model output and the target. The effects of VH are illustrated in the resulting maximum declination flow (MFDR), unsteady flow (AC flow), and peak collision forces.

Results

The degree of incomplete glottal closure in both the membranous portion of the folds (MGO and PGO) consistently leads to a reduction in SPL. All three compensatory mechanisms can counterbalance this effect, but they also lead to overly increased AC Flow, MFDR, and collision forces, which are thought to be key feature of patients with phonotraumatic VH. Individual compensatory mechanisms indicate that lung pressure exhibits the largest contribution, followed by vocal tract constriction, and muscle tensioning. Furthermore, the range over which the compensatory mechanisms can operate is more limited for vocal tract constriction and muscle tensioning. These effects are illustrated in Fig 1 for AC flow.
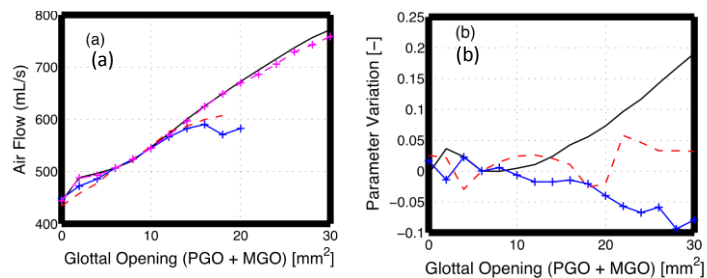

Fig. 1. Compensatory mechanisms to regain SPL on AC Flow for increasing degrees of posterior (PGO) and membranous (MGO) glottal opening. Compensatory mechanisms: lung pressure (black), vocal tract constriction (blue), CT muscle activation (red), and combined compensation (magenta).

Conclusions

Current findings support that the pathophysiology of VH is associated with biomechanical deficiencies that are exacerbated by faulty compensations. We will continue exploring onset conditions (e.g., heightened/uncoordinated activity of perilaryngeal musculature), and other optimization schemes for the compensatory mechanisms.

Broadband synchronization of asymmetric vocal fold oscillators

Jorge C. Lucero[1], Jean Schoentgen[2]

[1] Dept. Computer Science, University of Brasilia, Brazil, [2]Laboratories of Image, [2]Signal Processing and Acoustics, Université Libre de Bruxelles, Belgium

lucero@unb.br

Introduction

The right and left vocal folds constitute a pair of coupled oscillators that, in normal conditions, oscillate with perfect in-phase synchrony. Right-left asymmetries caused by pathological conditions handicap the synchronization and may induce complex entrainment regimes, quasi-periodic oscillations and other nonlinear phenomena (e.g., Mergell et al, 2000). A recent theoretical study showed that a natural frequency (or stiffness) asymmetry increases the threshold subglottal pressure for 1:1 phase synchronization (Lucero et al., 2015). The study also showed that the introduction of a damping asymmetry causes an extension of the region of 1:1 synchronization (broadband synchronization; Kuznetsov et al., 2007) at a low subglottal pressure. In the extended region, the fold with the lowest damping is dominant and the other fold follows with a large phase difference and small amplitude. This phenomenon matches experimental observations of vocal fold oscillations (Zhang and Luu, 2012) and requires further investigation.

Methods

The vocal folds are represented as a coupled system of two one-degree-of-freedom oscillators of the form

$$\ddot{x}_r + \beta(1 - \gamma/2)(1 + x_r^2)\dot{x}_r + (1 - \Delta/2)x_r = \alpha(\dot{x}_r + \dot{x}_l)$$
$$\ddot{x}_l + \beta(1 + \gamma/2)(1 + x_l^2)\dot{x}_l + (1 + \Delta/2)x_l = \alpha(\dot{x}_r + \dot{x}_l)$$

where x is the normalized tissue displacement and the subindices designate the right and left folds. Coefficient α models the aerodynamic coupling, β is the damping, and Δ and γ introduce a natural frequency and a damping asymmetry, respectively. Regions of phase synchronization are characterized in terms of bifurcations diagrams and kymograph plots of the solutions.

Results

Fig. 1 shows an example of regions of n:m phase entrainment, represented in different tones of gray (the larger the product nm, the darker the tone). The 1:1 region occupies the upper left half of the figure and to the right a strip, between lines (a) and (b), involving small values of the coupling α (low subglottal pressure). The lower boundary of that strip extends accross

the full range of the natural frequency asymmetry Δ. Lines (a) and (b) are the oscillation thresholds of the right and left oscillators, respectively, when they are uncoupled, and roughly define the lower and upper boundaries of the extended strip, when the intrusive 1:3 zone is ignored. The lower boundary of the 1:1 region has a minimum at Δ = 0 and is almost horizontal at large values of Δ, in agreement with measures on a vocal fold replica (Haas et al., in press).
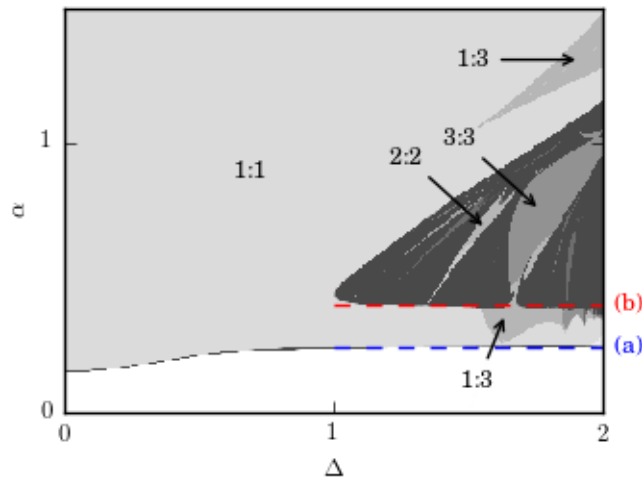

Figure 1. Entrainment regions for $\beta = 0.32$, $\gamma = 0.5$. Lines (a) and (b) are given by $\alpha = \beta(1 \pm \gamma/2)$.

Conclusions

1:1 synchronization of the right and left vocal folds may occur at a low subglottal pressure even when the left and right natural frequencies are very different. This appears to be a new finding in vocal fold oscillation dynamics, and its implications for abnormal phonation and the production of disordered voices should be the subject of further studies.

References

Haas, J., Luizard, P., Pelorson, X., & Lucero, J. C. (in press). Study of the effect of a moderate asymmetry on a replica of the vocal folds. Acta Acustica united with Acustica.

Kuznetsov, A. P., Paksyutov, V. I. & Roman, R. Y. (2007). Features of the synchronization of coupled van der Pol oscillators with nonidentical control parameters. Tech. Phys. Lett. (33), 636–638.

Lucero, J. C., Schoentgen, J., Haas, J., Luizard, P., & Pelorson, X. (2015). Self-entrainment of the right and left vocal fold oscillators. J. Acoust. Soc. Am. (137), 2036–2046.

Mergell, P., Herzel, H., & Titze, I. R. (2000). Irregular vocal-fold vibration – High-speed observation and modeling. J. Acoust. Soc. Am. (108), 2996–3002.

Zhang, Z. & Luu, T. H. (2012). Asymmetric vibration in a two-layer vocal fold model with left-right stiffness asymmetry: Experiment and simulation. J. Acoust. Soc. Am. (132), 1626–16

Predicting Achievable Fundamental Frequency ranges in Vocalization Across Species

lngo Titze[1], Tobias Riede[2], and Ted Mau[3]

[1] National Center for Voice and Speech, University ofUtah,

[2] Department of Physiology, Midwestern University,

[3] Department of Otolaryngology-Head and Neck Surgery, University of Texas Southwestern

ingo.titze@utah.edu

Introduction

We propose that vocal output can be conceptualized as the dependent variable in a multi-dimensional morphospace. We further hypothesize that surfaces exist in the morphospace to define functionally equivalent morphologic variations. Such alternative morphologies could be clinically important as viable solutions for vocal fold repair.

Methods

Vocal folds are used as sound sources in various species, but it is unknown how vocal fold morphologies are optimized for different acoustic objectives. Here we identify two main variables affecting vocal fold vibration frequencies and develop a simple theory to predict their ranges across different vocal fold sizes.

Results

Average fundamental frequencies are predominantly determined by vocal fold length (larynx size), but range of oscillation is facilitated by laryngeal muscles that control length change and tissue fiber tension. One adaptation that would increase fundamental frequency range is greater freedom in rotation or gliding of two cartilages (thyroid and cricoid), so that vocal fold length change is maximized. Alternatively, tissue layers can develop to bear a disproportionate fiber tension (i.e., a ligament with high density collagen fibers), increasing the fundamental frequency range and thereby vocal versatility.

Conclusions

The range of fundamental frequency across species is thus not simply a function of vocal fold length as previously thought.

Generation of diphthongs using finite elements in three-dimensional simplified vocal tracts

Marc Arnela[1], Saeed Dabbaghchian[2], Oriol Guasch[1]and Olov Engwall[2],

[1] GTM – Grup de recerca en Tecnologies Mèdia, La Salle, Universitat Ramon Llull

[2] Department of Speech, Music, and Hearing, School of Computer Science & Communication, KTH Royal Institute of Technology

marnela@salle.url.edu

Introduction

Finite Element Methods (FEM) have been extensively used to generate vowel sounds (see e.g., [1]), yet little attention has been paid to the simulation of diphthongs (see e.g., [2, 3] for some recent exceptions). Generating diphthongs requires dealing with acoustic waves propagating through dynamic three-dimensional (3D) vocal tracts that transition, for instance, from vowel /a/ to vowel /i/ to produce the diphthong /ai/. Indeed, one of the main problems is that of obtaining the corresponding time evolving 3D finite element meshes where to solve the acoustic wave equation. This is especially challenging when using very detailed 3D vocal tract geometries generated e.g. from Magnetic Resonance Imaging (MRI). Since the acquisition times for sufficiently geometrically detailed MRI does not allow for real-time imaging, an interpolation between the meshed geometries of the starting and ending vowels is needed to obtain the articulation of a diphthong, but this is far from easy.

As an alternative, in this work it is proposed to generate good quality diphthong sounds from simplified versions of MRI-based vocal tracts. The dynamics of these simplified geometries has the advantage to be driven only by a few parameters, which facilitates the interpolation between the initial and the final vowel geometries.

Methods

The MRI-based vocal tract geometries for different vowel sounds in [4] are used as a starting point. These are first discretized in a finite set of cross-sections. For each cross-section we extract the area and shape, together with the location and orientation of the vocal tract midline through the cross-section. Simplified 3D vocal tract geometries with different degrees of realism can then be reconstructed combining different cross-sectional shapes (realistic, elliptical and circular) and vocal tract midlines (bent or straight) [5], see Fig. 1. Finally, to generate a dynamic vocal tract, the area, shape and location of each cross-section are linearly interpolated from those of the starting vowel to the target one (see Fig. 1).

Acoustic wave propagation in the constructed dynamic 3D vocal tracts is then simulated. To do so, two problems are numerically solved using FEM (see e.g., [2, 3]). The first one deals with the propagation of acoustic waves in a moving computational domain. In particular, the mixed wave equation for the acoustic pressure and particle velocity is expressed in an

Arbitrary Eulerian-Lagrangian (ALE) frame of reference. The second problem consists in determining the evolution of the inner nodes of the vocal tract volume mesh according to the vocal tract wall movement. This is done by solving the Laplacian equation for the node displacements, which smoothly translates the prescribed motion at the boundary nodes (vocal tract walls) to the inner ones.
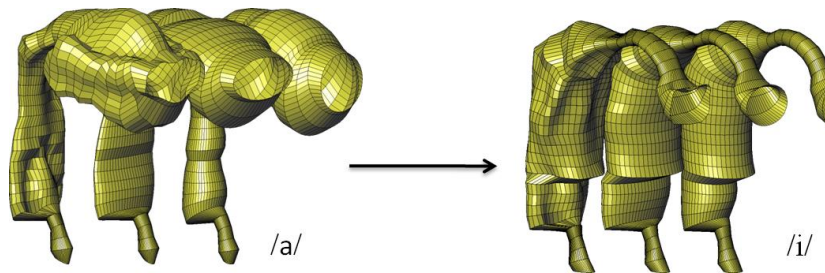

Figure 1 – Simplified vocal tracts with realistic, elliptical and circular cross-sections used to simulate diphthong /ai/.
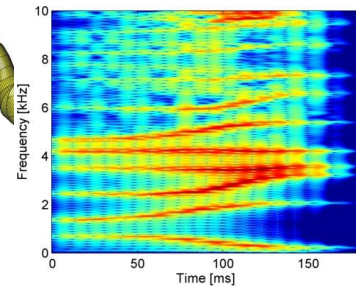
Figure 2 – Spectrogram of diphthong /ai/ computed using FEM.

Results

In Fig. 2 we show the spectrogram of diphthong /ai/ obtained from simplified vocal tracts with circular cross-sections. It can be observed how the formants smoothly transition from those of vowel /a/ to those of vowel /i/.

Conclusions

In this work, diphthong sounds have been numerically generated using FEM. Three-dimensional simplified vocal tracts have been used, which allow one to easily interpolate between static geometries.

Acknowledgments

References

1 Arnela, M., Guasch, O., Alías, F. (2013). Effects of head geometry simplifications on acoustic radiation of vowel sounds based on time-domain finite-element simulations. The Journal of the Acoustical Society of America 134(4), 2946-54.

2 Arnela, M., Guasch, O., Codina, R., Espinoza, H. (2014), Finite element computation of diphthong sounds using tuned two-dimensional vocal tracts, 7th Forum Acusticum, September 7-12, Krakow, Poland.

3 Guasch, O., Arnela, M., Codina, R. Espinoza, H. (2016); A stabilized finite element method for the mixed wave equation in an ALE framework with application to diphthong production. Acta Acustica united with Acustica, 102 (1), 94-106.

4 Aalto, D., Aaltonen, O., Happonen, R.-P., Jääsaari, P.,Kivelä, A., Kuortti, J., Luukinen, J.-M., Malinen, J., Murtola, T., Parkkola, R., Saunavaara, J., T. Soukka, T., and Vainio, M. (2014). Large scale data acquisition of simultaneous MRI and speech, Appl. Acoust. 83, 64–75.

5 Dabbaghchian, S., Arnela, M., Engwall, O. (2015), Simplification of vocal tract shapes with different levels of detail, 18th International Congress of Phonetic Sciences (ICPhS), August 10-14, Glasgow, Scotland, UK

Sensitivity Analysis of Agent-Based Model of Vocal Fold Inflammation and Repair

Trickey, K. [(1)], Shung, C. [(2)], Yuen, S. 1, Seekhao, N. [(3)], JaJa, J. [(3)], Mongeau, L. [(2)], Li, N. Y. K.[(1)]

[(1)] School of Communication Sciences and Disorders, McGill University, [(2)] Department of Mechanical Engineering, McGill University, [(3)] Department of Electrical & Computer Engineering, University of Maryland-College Park

nicole.li@mcgill.ca

Introduction

An agent-based model (ABM) has been developed to simulate cellular-level inflammation processes in vocal folds following acute phonotrauma [1,2]. Current ABMs were partially calibrated and validated against a limited set of empirical vocal fold data. The purpose of this study was to perform parameter screening and sensitivity analysis to (1) identify the most and least significant parameters on output variables and (2) better estimate parameter values. The information will be important in exact calibration of key variables, preserving key cellular and molecular mechanisms while saving computational resources.

Methods

The iterative Morris screening method (SIMLAB 2.2, European Community, Brussels) was used to identify parameters with negligible effects on the output. This method yields the overall importance of each parameter in affecting the model output, μ, and the degree of non-linear effects of each parameter and interactions with other parameters, σ. Parameters which have a low μ and σ are considered uninfluential [3].

Three acute phonotrauma and treatment ABMs (voice rest, resonant voice, spontaneous speech respectively) were used as a case study for this sensitivity analysis. Each model had 213 parameters associated with cytokine/growth factor synthesis by different cell types as well as parameters related to ECM synthesis, cell proliferation, cell death, cell activation/deactivation, cell recruitment (sprouting), impact/vibratory stress, thresholds for cytokine damage, and cytokine/growth factor half life. Each ABM model was evolved in 1180 – 1290 runs with 20 random seeds and 4 levels. Each run simulated an area 17.4 mm x 24.9 mm over 48 hours.The output of each run was the total levels for each one of 8 inflammatory mediators (TNF, TGF, FGF, MMP8, IL-1β, IL-6, IL-8, IL-10).

Results

The Morris screening analysis yielded 8 values of μ and 8 values of σ for each parameter. We computed the sum of the normalized values of μ (μ_{ns}) and sum of the normalized values of σ (σ_{ns}) for each parameter and ranked them based on these values. The most influential parameters across all treatments were involved in cytokine/growth factor synthesis (Figure 1).

The first 20 parameters with the highest values of normalized μ sum and the first 20 parameters with the highest values of normalized σ sum for each treatment were ranked and considered as the most influential parameters (Figure 2). 21 parameters were identified for voice rest, 23 for resonant voice, and 20 for spontaneous speech case as the most influential and will be used for further sensitivity analysis.
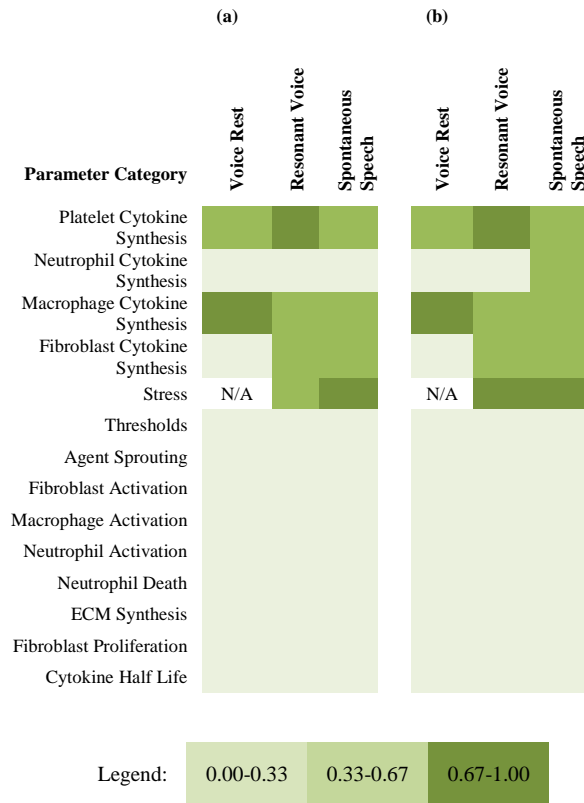

Figure 1: (a) Average μ_{ns} and (b) average σ_{ns} as measurements of relative influence of parameters across treatment options.
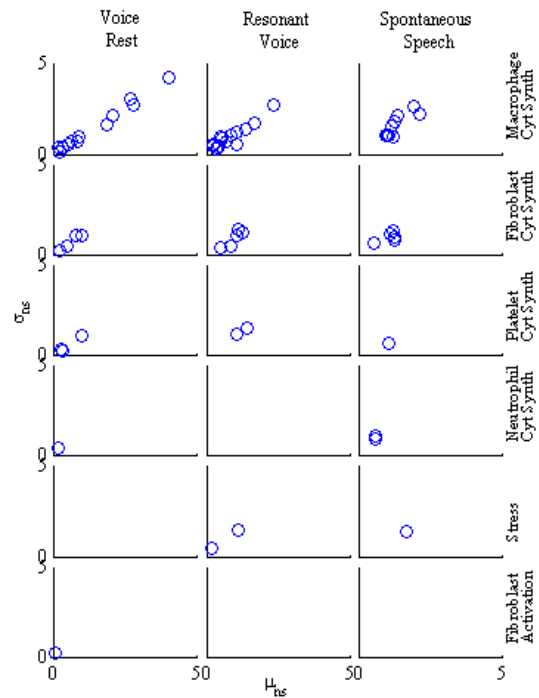

Figure 2: Sensitivity indices (μ_{ns} and σ_{ns}) of most influential parameters for each treatment group from Morris Screening

Conclusions

Parameters of current acute phonotrauma ABMs that have the most influence on inflammatory mediator outputs were identified. These parameters are largely involved in cytokine and growth factor synthesis. These results were as expected because most of the agent-rules revolve around maintaining proper cytokine levels. Sobol's method for sensitivity analysis will be performed on the parameters identified by Morris screening to better rank and further reduce the number of key parameters used for the ultimate model calibration and validation. Exact calibration of key influential parameters is an important step in improving the ABM's prediction accuracy as a clinical tool to prescribe and optimize patient-specific phonotrauma treatment

References

1. Li NYK, Vodovotz Y, Kim KH, Mi Q, Hebda PA, Verdolini Abbott K. Biosimulation of acute phonotrauma: An extended model. *The Laryngoscope.* 2011;121(11):2418-2428.
2. Li NYK, Verdolini K, Clermont G, et al. A patient-specific in silico model of inflammation and healing tested in acute vocal fold injury. *PLoS One.* 2008;3(7):e2789.
3. Morris MD. Factorial sampling plans for preliminary computational experiments. *Technometrics.* 1991(33):161-174.

Effect of *β-Catenin* signaling on cell proliferation in developing vocal folds and its potential for vocal fold regeneration and repair in adulthood

Vlasta Lungova [(1)], Xin Sun [(2)], Susan Thibeault [(1)]

[(1)] Department of Surgery, University of Wisconsin-Madisonl, [2] Laboratory of Genetics, University of Wisconsin-Madison

lungova@surgery.wisc.edu

Introduction

Understanding vocal fold regeneration and tissue repair is one of the key issues in the voice biology. Since the vocal folds (VFs) are located within the larynx, at the entrance to the lower airway structures, they are exposed to numerous daily irritants e.g. pollutants, smoke and/or acid reflux. They also vibrate during phonation and undergo stresses and collisions. Knowledge of the proximate mechanisms underlying their development could be, therefore, very useful for identifying genes and signaling pathways that maintain healthy VFs in adulthood and/or stimulate their regeneration. Recently, we provided the first systematic determination of the cellular and molecular progression of VF epithelium development in mice. We focused on early stages of the larynx and VF development, when the lateral walls of the primitive laryngopharynx (LPh) grow into the center of the lumen and temporary fuse to create the epithelial lamina (EL) at embryonic day (E) 11.5. The EL is consequently recanalized allowing VFs to separate (E18.5). The function of the EL is not known. It certainly ensures that VFs develop opposite to each other to close the glottis. In addition, we hypothesize, that epithelial cells located in the EL are reorganized after they fuse and instruct underlying mesenchyme to proliferate and/or differentiate into fibroblasts of the lamina propria, muscle cells or supporting cartilages.

Methods

Generation of *β-Catenin* mutants: mice carrying a conditional loss-of-function allele of *β-Catenin (Ctnnb1^{tm2.1Kem})* were mated to mice carrying the *Shh^{cre}* allele to generate *Shh^{cre/+}; Ctnnb1^{tm2.1Kem/tm2.1Kem} (β-Cat^{cko})* mutants. Wild-type embryos were used as controls for *β-Cat^{cko}* experiments. Embryos were dissected from time-mated mice, at the embryonic day E11.5. They were fixed, dehydrated, embedded in paraffin and cut into serial sections. Hematoxylin-eosin was used to assess morphology, TUNEL assay to detect cell death and EDU assay to detect cell proliferation. Immunohistochemistry was used for analysis of the spatial distribution of following genes: p63, cytokeratin 8, cyclin D1 and cyclin D kinase inhibitor p27^{Kip1}.

Results

Wnt/β-Catenin signaling pathway plays a critical role in the development and tissue maintenance. It controls proliferation of cells and/or their differentiation into particular cellular types (1). Here we report that ablation of *β-Catenin* in VF epithelial progenitors in *β-Catenin* conditional mutants *(β-Cat^{cko})* at E11.5 resulted in the failure of the EL formation. As a consequence, the septum separating the larynx from the esophagus is not formed. Mesenchymal cells do not condensate to induce formation of the cricoid cartilage and/or differentiate into muscle cells to give rise to dorsal intrinsic laryngeal muscles. To determine the mechanism underlying the failure in EL formation, we examined possible altered gene expression in VF epithelial progenitors, cell death and rate of cell proliferation in VF epithelium and adjacent mesenchyme. Our results indicate that the inability of epithelial cells to obliterate the ventral lumen is due to the cell cycle arrest.

D-Cyclins along with their catalytic partners, cyclin-dependent kinases (CDKs), are major regulators of the cell cycle progression in developing tissues as well as during tissue maintenance in the adulthood (2). D-cyclins also play the kinase independent role and interact with cell cycle inhibitors p21^{Cip1} or p27^{Kip1} (3), which inhibit activity of cyclin-CDK complexes. They direct cell cycle exit in developing tissues or in response to DNA damage (2). Here we report that an analysis of cyclin D1 and p27 expression showed reduced levels of cyclin D1 and increased activity of p27 in VF epithelial progenitors and underlying mesenchyme in *β-Cat^{cko}* mutants during the EL formation. Possible deletion of p27^{Kip1} gene in double-mutant *β-cat-/-; p27-/-* mice could rescue the normal phenotype in *β-Cat^{cko}* mutants and stimulate proliferation of VF epithelial progenitors.

Conclusion

Ours finding reveal a possible novel treatment technique for the wound healing in VFs, as both inhibitors p21 and p27 act also during wound healing (4). Moreover, p21^{Cip1} deficient mouse strains show regenerative capabilities and heal without scarring (5). Further investigation is needed to test whether the temporal downregulation of p21^{Cip1} or p27^{Kip1} could serve as a potential target for stimulation of reepithelization during the VF wound healing and tissue repair.

References:

1. Harris-Johnson, K.S., et al., b-Catenin promotes respiratory progenitor identity in mouse foregut. PNAS 106, 38: 16287–16292.

2. Lee Y.S., et al. 2006. A morphogenetic wave of p27Kip1transcription directs cell cycle exit during organ of Corti development. Development 133: 2817-2826.

3. Geng, Y., et al., 2001. Depletion of the p27Kip1gene restores normal development in cyclin D1-deficient mice. PNAS 98: 194-199.

4. Zhu, X., et al., 2001. Expression of cyclin-dependent kinase inhibitors, p21Cip1 and p27Kip1, during wound healing in rats. Wound Repair and Regeneration 9: 205-212.

5. Bedelbaeva, K., 2010. Lack of p21 expression links cell cycle control and appendage regeneration in mice. PNAS 107, (13): 5845–5850.

Using RNA-Seq Characterization of Differently Expressed Genes from Larynx Responding to Smoking and Reflux

Xia Chen [1], Shuyun Ye [1], Christina Kendziorski [1], Susan Thibeault [1]

[1] Division of Otolaryngology-Head and Neck Surgery, Department of Surgery,

University of Wisconsin-Madison

chenx@surgery.wisc.edu

Introduction

Chronic laryngitis is a very common disorder that represents a significant clinical problem, which can be caused by a variety of factors, such as cigarette smoking and reflux [1]. Present treatments are almost entirely empirical with uncertain effectiveness. Here, using high-throughput illumina RNA sequencing (RNA-Seq, a advance sequencing technique for quantitative entire transcriptome analysis) technology, EBseq and functional enrichment analysis, we generated a VF-specific transcriptome blueprint for persons with smoking and reflux compared to normal to understand function and chronic laryngitis in this specific tissue. This work will serve as an important reference database significant for laryngeal researchers to explore the basis of health and inflammatory disease in larynx.

Methods

Healthy, untreated volunteers (n=130) at the University of Wisconsin-Madison were enrolled this study. Clinical information (age, sex, smoking status and reflux [24hr pH probe with manometry]) was collected from all participants according to the Institutional Review Board protocol of University of Wisconsin-Madison. Total RNA was extracted from the very small biopsied tissues from false vocal folds. Among of 130 participates, RNA samples from 92 donors were qualified for RNA-Seq analysis. All RNA sequence data were analyzed using EBSeq, an empirical Bayes analysis approach [2] to create lists of differently expressed (DE) genes and further functional enrichment analysis (KEGG pathways and GO terms) [3] to describe the sets of DE genes in a biological meaningfully way.

Results

Overall, we found most DE genes related to cilium organization, cell projection assembly and microtubule-based process were down-regulated in male smokers with reflux. Under reflux condition, the DE genes related to regulation of dipeptidase activity, cAMP-mediated signaling, calcium ion and oxygen transport were down-regulated, and the DE genes related to lymphocyte chemotaxis & migration, epithelium development, and protein glycosylation

were up-regulated. In male smokers with reflux, most DE genes were related to muscle system process, tissue morphogenesis, regulation of ion transport, and tissue homeostasis down-regulated compared to females. However, in male smokers without reflux, most DE genes involved in keratinization, regulation of defense response, steroid metabolic process, demethylation, and protein localization to membrane were up-regulated compared to females.

Conclusions

Specific gene signatures of human larynx tissue that have been exposed to cigarette smoke and laryngopharyngeal reflux were identified. Among of 18969 detected genes, we obtained the DE genes responding to the status of smoke, reflux and gender, which were mainly related to defense response, oxygen transport, protein metabolism, tissue homeostasis and epithelium development. These VF-specific RNA-Seq datasets will help people to differentiate health and chronic laryngitis, guide clinical treatment, be a resource for both clinical and basic scientific research, and have potential to stimulate new hypotheses for further investigation. Further biological functional analysis of these DE genes and their downstream pathway will provide more information about chronic laryngitis with different pathogenesis.

References

1. Cohen, S. M., Kim, J., Roy, N., Asche. C., and Courey, M. (2012). Prevalence and causes of dysphonia in a large treatment-seeking population. Laryngoscope, 122,343-348.

2. Jiang, H., and Wing, W. H. (2009). Statistical inferences for isoform expression in RNA-Seq. Bioinformatics, 25(8), 1026-1032.

3. Chen, L., Chu, C., Lu, J., Kong, X., Huang, T., and Cai, Y.D. (2015). Gene Ontology and KEGG Pathway Enrichment Analysis of a Drug Target-Based Classification System. PLoS One, 10, e0126492.

Aortic arch compliance and idiopathic left-sided vocal fold paralysis

Julie Barkmeier-Kraemer[1], Reza Bekham, [2] Andrew Bierhals [3], Julia Edgar [4], Eileen Jacobs [3], Randal Paniello[3], Gayle Woodson [5], Jonathan Vande Geest [2]

[1]University of Utah, [2]University of Arizona, [3]Washington University, St Louis, [4] Truman State University, [5] Southern Illinois Medical Center

JulieB.Kraemer@hsc.utah.edu

Introduction

Between 12-34% of individuals diagnosed with unilateral vocal fold paralysis (UVP) are classified as "idiopathic"[1,2]. Of these, 2/3 exhibit left-sided UVP[1] with an average age of onset of 54 years ($\pm$ 12 years)[3]. The left recurrent laryngeal nerve (RLN) supplying innervation to the larynx is significantly longer than the right RLN and branches from the vagus nerve within the thoracic cavity to travel around and adjacent to the aortic arch before ascending to the larynx. Prior research demonstrated that the diameter and compliance of the aortic arch increases until age 10 and then decreases with aging[4,5]. Case-based evidence also shows that changes to cardiac or aortic arch vessels are associated with onset of left-sided UVP[6] suggesting that changes in the diameter and compliance of the aorta may impact left RLN function. We hypothesized that age-related changes in diameter and compliance of the aortic arch may be greater in those with left-sided UVP compared to normal controls. To investigate this question, we compared aortic arch diameter and compliance between individuals diagnosed with left-sided idiopathic UVP compared to age and gender-matched controls.

Methods

Nine individuals meeting inclusion criteria for the diagnosis of left-sided idiopathic UVP and 9 age- and gender-matched controls ages 26-82 years of age participated in the study (average age = 53 years). Ungated and gated structural 1.5T MRI scans of the thoracic aortic arch were obtained at the segment of the aortic arch where the RLN typically resides. Blood pressure measures were also acquired to calculate aortic compliance. Electrocardiogram (ECG)-gating with data acquisition during diastole and systole were used to control for cardiac motion artifacts during image acquisition.

Measures of aortic arch diameter were made at each of several MRI scanning time points so that the aortic arch could be reconstructed in each participant. Compliance was determined for the aorta from measures of relative diameter change (in the plane orthogonal to the aortic centerline) over a cardiac cycle that was normalized to the pulse pressure. Mean peak aortic arch diameter and compliance were compared between experimental groups and by age.

Results

Changes in diameter prevoiusly shown to be associated with aging were replicated[4,5] for both groups without significant differences between participant groups. However aortic arch compliance was found to be significantly greater in those with left-sided idiopathic UVP compared to normal controls (p=0.0022) If this study was replicated with an alpha level of 0.05 and power of 99%, 16 total observations would be necessary to show a significant difference among groups.

Figure 1. Aortic arch compliance across age for each group.
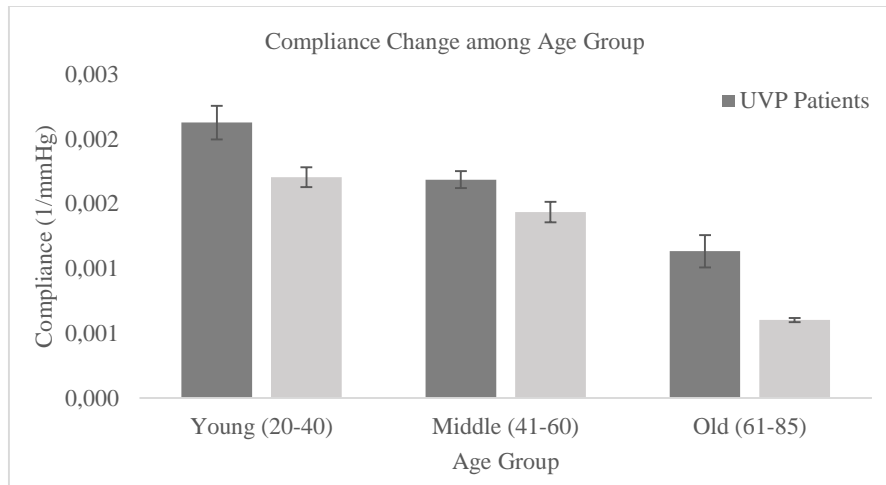
Conclusions

The findings from this study demonstrated that individuals with idiopathic left-sided UVP exhibited significantly greater aortic arch compliance than their age- and gender-matched controls. These findings suggest that hypercompliance of the aortic arch may impact RLN function resulting in left-sided UVP. Future research needs to address whether a specific aortic arch compliance level, or the interval of time during which compliance levels change induces impaired RLN function.

References

1. Spataro E.A., Grindler D.J., Paniello R.C. (2014). Etiology and time to presentation of unilateral vocal fold paralysis. Otolaryngology – Head and Neck Surgery, 151(2), 286-293.

2. Sulica L. (2008). The natural history of idiopathic unilateral vocal fold paralysis: evidence and problems. Laryngoscope, 118, 1303-1307.

3. Chen, D.W., Young, A., Donovan, D.T., Ongkasuwan, J. (2015). Routine computed tomography in the evaluation of vocal fold movement impairment without apparent cause. Otolaryngology – Head and Neck Surgery, 152(2), 308-313.

4. Lang, R.M., Cholley, B.P., Korcarz, C., Marcus, R.H., Shroff, S.G. (1994). Measurement of regional elastic properties of the human aorta. A new application of transesophageal echocardiography with automated border detection and calibrated subclavian pulse tracings. Circulation, 90(4), 1875-1882.

5. Länne, T., Hansen, F., Mangell, P., Sonesson, B. (1994). Differences in mechanical properties of the common carotid artery and abdominal aorta in healthy males. Journal of Vascular Surgery, 20(2), 218-225.

6. Subramaniam, V., Herle, A., Mohammed, N., Thahir, M. (2011). Ortner's syndrome: Case series and literature review. Braz J Otorhinolaryngol, 77, 559-562.

Electrophysiological activity of the pharyngoesophageal segment and tracheoesophageal voice and speech proficiency in total laryngectomees

Hilton Ricz[1], Telma Kioko Takeshita-Monaretti[2], Patrícia Bastos[2], Flávia Pinto Neves[1], Wilson Marques Júnior[2], Lílian Aguiar-Ricz[1]

[1] Department of Ophthalmology, Otorhinolaryngology and Head and Neck Surgery, School of Medicine of Ribeirão Preto, University of São Paulo, [2] Department of Neuroscience and Behavioral Sciences, School of Medicine of Ribeirão Preto, University of São Paulo

hricz@fmrp.usp.br

Introduction
The vocal function of total laryngectomees is dependent on the interaction between airflow, pressure and mucosal characteristics, remaining muscles of the pharyngoesophageal transition and adjacent structures, however, the conditions capable of justifying the voice and tracheoesophageal speech results have not yet been established. Although it is evident the anatomical changes after total laryngectomy, especially the sacrifice of the laryngeal nerves, few studies allow the electrophysiological understanding of the muscles that make up the new sound source. The objective of this study was to investigate the association between the electrophysiological activity of the pharyngoesophageal segment muscles and proficiency of the tracheoesophageal speech in total laryngectomees.

Methods
Attended by 34 individuals, 26 males and eight females, with average age of 62.5 years old, who underwent total laryngectomy and had primary (n = 5) or secondary insertion (n = 29) of the voice prosthesis. All participants underwent the test of voice and speech, recorded by camera camcorder. Vocal intensity, dynamic range and maximum phonation time were measured. Three judges evaluated the proficiency of the tracheoesophageal speech, using adapted protocol, classifying them into good, moderate or bad speakers. Then the patients underwent video fluoroscopy for location of pharyngoesophageal segment and identification of the region to perform directed electromyography. Electromyography was performed by percutaneous needle electrode by locating the muscles of pharyngoesophageal segment bilaterally. The electromyographic analysis was characterized as: normal, presence of neurogenic injury (moderate to severe, severe injury, severe injury to total), myopathic or inconclusive injury. The characterization of tracheoesophageal speech proficiency was determined after evaluation of agreement between the judges, using the Kappa coefficient. To investigate the association between two qualitative variables (final judgment of

tracheoesophageal speech proficiency and electromyographic analysis) the Fisher's exact test was used.

Results
In the final judgment of tracheoesophageal speech proficiency, the majority of laryngectomees was categorized as moderate speakers (n = 24) and the minority as good ones (n = 3). For all total laryngectomees, neurogenic lesions became evident in the electromyographic analysis, more frequently of severe type (n=20) and severe to total type (n=20), followed by moderate to severe one (n=4). There was no significant association between the electromyographic analysis of neurogenic injury and final judgment of tracheoesophageal speech proficiency.

Conclusions
The preservation of the motor unit in the pharyngoesophageal segment muscles during voice emission does not prevent the acquisition of alaryngeal voice, but it can make the difference to the quality of tracheoesophageal speech. The tracheoesophageal voice is often characterized as a deviation when compared to ideal vocal quality. However, it is necessary to recognize the adjustments made by total laryngectomees in the attempt to overcome the limitations of alaryngeal voice, while further studies are still needful.

Toward a Psychoacoustic Model of Spectral Noise in the Voice Source

Rosario Signorello[1], Nari Rhee, Bruce R. Gerratt[1], and Jody Kreiman[1]

[1]Bureau of Glottal Affairs, Department of Head and Neck Surgery,
University of California,

jkreiman@ucla.edu

Introduction

Previous studies (Kreiman *et al.*, 2015) have proposed a psychoacoustic model of voice quality that includes a four-parameter spectral-domain model of the harmonic voice source (H1-H2, H2-H4, H4-the harmonic nearest 2 kHz, and that harmonic to the harmonic nearest 5 kHz). In this paper we report a parallel psychoacoustic study of the spectral shape of the inharmonic source. Our goal is two-fold: to provide a framework for nuanced measurement of vocal noise, and to improve understanding of the interaction between harmonic and inharmonic aspects of the voice source in determining voice quality.

Methods: Acoustic Analysis

120 steady-state /a/ vowels (60 female speakers; 15 each normal, mildly, moderately, and severely pathologic) were copy-synthesized using analysis-by-synthesis. The noise spectrum was extracted from each synthesized vowel and downsampled to 111 points. These data were analyzed using principal component analysis (PCA) with varimax rotation to derive a set of factors describing noise spectral shapes across voices.

Results

For both female and male voices, PCA revealed four factors that accounted for most of the variance in the underlying data (females: 87%; males: 91%). For female speakers, these factors divided the spectrum into components ranging from 0-1400 Hz, 1400-2800 Hz, 2800-4200 Hz, and 4200-5000 Hz. For male speakers, the spectrum was divided into components ranging from 0-600 Hz, 600-2200 Hz, 2200-3800 Hz, and 3800-5000 Hz.

Perceptual Evaluation

Two synthetic versions of each voice were created: one with the original noise spectrum, and one with the noise spectrum smoothed with a 4-piece model corresponding to the ranges described above. To determine if the four-piece model adequately captures the perceptually-important features of the inharmonic source, listeners compared these two synthetic tokens to each other and to the original natural voice sample. If the spectral model is adequate, then

the two synthetic tokens should be indistinguishable from one another and from the original voice sample. Comparison to the original voice sample establishes the adequacy of the synthesis on which the analysis was based; comparison of the two synthetic tokens establishes the adequacy of the 4-parameter model. This experiment is currently underway.

Conclusions

Preliminary results suggest that the psychoacoustic model is sufficient to characterize perceptually-important aspects of the spectral shape of the inharmonic voice source for normal and pathologic steady-state phonation.

Acknowledgements

References

Kreiman, J., Gerratt, B. R., Signorello, R., and Rastifar, S. (2015). "Adequacy of a four-parameter spectral model of the voice source". In *Journal of the Acoustical Society of America*, 137(4): 2266.

Brain Mapping of Laryngeal Sensorimotor Control in Normal Phonation

Maryna Kryshtopava [1], Kristiane Van Lierde [2], Iris Meerschman2, Evelien D'Haeseleer [2]

Guy Vingerhoets [3], Sofie Claeys [1]

[1] The Department of Otorhinolaryngology, University Hospital Ghent, [2] The Department of Speech, Language and Hearing Sciences, University Hospital Ghent, [3] Faculty of Psychology and Educational Sciences, Ghent University

Maryna.Kryshtopava@ugent.be

Introduction

This study investigated the brain activitiy associated with pitch adaptation during phonation by healthy women without voice disorders using fMRI. Human phonation is a laryngeal motor behavior that extends from reflexive laryngeal actions to highly skilled laryngeal sensorimotor control to support speech or singing Phonation requires integration of the sensory input and laryngeal motor output for adaptation of pitch and intensity during vocalization. Recent fMRI and PET studies have shown that in order to understand the neural control of phonation, laryngeal control must be investigated distinct from respiratory and articulatory control. However, the neural control of phonation involving pitch adaptation, isolated from articulatory and respiratory control, remains poorly investigated. In the present study, we aimed to investigate the neural control of voice pitch variation (comfortable and high) in healthy women, and to examine usability of a blocked design fMRI method in defining the neural control of phonation isolated from respiratory and articulatory control. To prevent the involvement of pharyngeal and articulatory muscles, we excluded tasks connected with laryngeal functions like coughing, swallowing and speech. Consequently, we implemented an experimental paradigm by contrasting prolonged phonation of unarticulated vowels with prolonged exhalation in order to differentiate phonation from exhalation. We focused on the superior temporal gyrus (STG) because it has been detected as an integration area of sensory input and motor output during phonation

Methods

Sixteen healthy young women (mean age: 24.3 years) underwent fMRI using a blocked design involving two phonation tasks (phonation /i/ at a high or comfortable voice pitch) and exhalation (prolonged exhalation) tasks. Brain Voyager QX software was used for fMRI data analysis (Goebel et. al., 2006). We reported the results of analyses performed on a cohort of 15 subjects (group analysis). We excluded Subject 4 (data from functional scan was missing). Using GLM analysis (q(FDR)<0.05), we tested task-related effects to contrast brain activity associated with voice pitch adaptation with brain activity during exhalation.

Results

Table 1 and Figure 1 summarize the main results of the study. Analyses showed a significant main effect of phonation with pitch adaptation compared to rest period in the bilateral precentral, superior frontal, posterior cingulate, superior and middle temporal gyrus, insula

and cerebellum, left middle and inferior frontal gyrus, right lingual gyrus, cingulate gyrus, and thalamus. Statistical results also identified a significant main effect of exhalation compared to rest period in the bilateral precentral gyrus, cerebellum, right lingual gyrus, thalamus, and left supramarginal gyrus. In addition, a significant main effect of phonation compared to exhalation was found in the bilateral STG and right insula as well as in the left brainstem for high pitch phonation only.

Conclusions

We demonstrated that a blocked design fMRI is sensitive enough to define a neural network associated with phonation involving pitch variation. Our results are corroborated by recent fMRI studies on phonation involving simple voice production tasks and provides a greater insight into the process of integration of sensory input in laryngeal motor output during voice pitch variation.

FIGURE 1. Brain activation during phonation for the contrasts of COMFORTABLE PHONATION>REST (A), HIGH PHONATION>REST (B), PROLONGED EXHALATION>REST (C), COMFORTABLE PHONATION>PROLONGED EXHALATION (D) and HIGH PHONATION>PROLONGED EXHALATION (E). The arrows indicate clusters of significant activation (q(FDR)<0,05). (Abbreviations: STG: superior temporal gyrus; PreCG: precentral gyrus; LG: lingual gyrus; PoCG: posterior cingulate gyrus; Th: thalamus; CE: cerebellum; Br: brainstem).
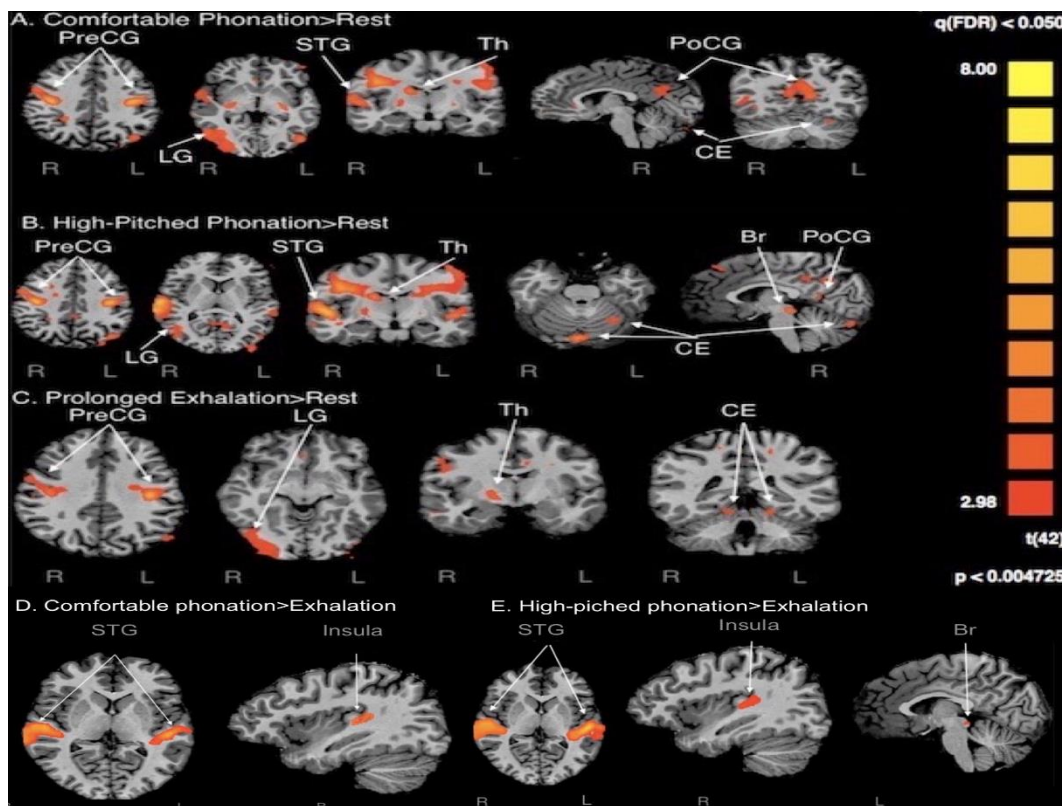
References

Ludlow, C. L. (2015). Laryngeal Reflexes: Physiology, Technique, and Clinical Use. *Journal of Clinical Neurophysiology*, 32(4), 284-293.

Ludlow, C. L. (2015). Laryngeal Reflexes: Physiology, Technique, and Clinical Use. *Journal of Clinical Neurophysiology*, 32(4), 284-293.

Parkinson, A. L., Flagmeier, S. G., Manes, J. L., Larson, C. R., Rogers, B., & Robin, D. A. (2012). Understanding the neural mechanisms involved in sensory control of voice production. *Neuroimage*, *61*(1), 314-322.

Universal mechanisms of sound production and control in birds and mammals

C.P.H. Elemans [(1-3)], J.H. Rasmussen[(1)], C.T. Herbst[(4)], D.N. Düring[(1)], S.A. Zollinger[(5)], H. Brumm[(4)], K. Srivastava[(6,7)], N. Svane[(1)], M. Ding[(8)], O.N. Larsen[(1)], S.J. Sober[(6)], J.G. Švec[(4)]

[(1)] Department of Biology, University of Southern Denmark [(2)] QuanTM program, Emory University [(3)] Université de Saint-Etienne, [(4)] Voice Research Lab, Department of Biophysics, Faculty of Science, Palacky University, [(5)] Max Planck Institute for Ornithology, [(6)] Emory University, [(7)] Georgia Institute of Technology, [(8)] Department of Orthopaedic Surgery & Traumatology, Odense University Hospital, University of Southern Denmark

herbst@ccrma.stanford.edu

Introduction

As animals vocalize, their vocal organ transforms motor commands into vocalisations for social communication. In the case of humans (van den Berg, 1958) and nonhuman mammals (Herbst *et al.*, 2012) it is known that sound is produced according to the myoelastic-aerodynamic (MEAD) theory of sound production (Titze, 1980) via passive, flow induced oscillations of laryngeal tissue. In birds, on the other hand, the physical mechanisms by which vocalisations are produced and controlled remain unresolved so far, owing to the extreme experimental difficulty in obtaining in vivo measurements.

Methods

Here, an ex vivo preparation of the avian vocal organ is introduced, which allowed simultaneous high-speed video (HSV) imaging, muscle stimulation, and kinematic and acoustic analyses, in order to reveal the mechanisms of vocal production in birds across a wide range of taxa. The investigated specimens came from seven bird species, ranging in size (~15 g – 200 kg) and vocal complexity, each with highly divergent syrinx morphologies, containing one or two paired oscillators.

Results

With the aid of tracheal endoscopy and syringeal trans-illumination, two vibratory components (VC) were documented: One vibratory component (VC1) was constituted by a medio-lateral vibratory component gating the airflow, resulting in complete syringeal closure within an oscillatory cycle in each species except for the ostrich, where full closure was never observed, more closely resembling human breathy phonation. Furthermore, in each species, regardless of syringeal morphology, the presence of a caudo-cranial component (VC2) of a travelling tissue wave was confirmed using spatiotemporal analysis of syringeal inner wall displacement, syringeal opening and/or micro-electroglottographic (μEGG) assessment. The presence of VC2 has been shown to be one essential option for self-sustaining oscillation (the

other being an inertive supraglottal vocal tract), as it helps to transfer aerodynamic energy into the tissue vibration via an asymmetric forcing function (Ishizaka and Flanagan, 1972; Titze, 1988). Data anlaysis further revealed that in each investigated species, multiple different combinations of bronchial and interclavicular air sac pressures could achieve the same target frequency.

Conclusions

The presence of the two vibratory components VC1 and VC2 leads to the conclusion that all bird species tested employ the MEAD mechanism, the same mechanism used to produce human speech and song. Furthermore, the substantial redundancies in the control of key vocal parameters *ex vivo* suggest that also *in vivo* vocalisations may not be specified by unique motor commands. We propose that such motor redundancy can accelerate vocal learning and is common to MEAD sound production across birds and mammals, including humans.

References

Herbst, C. T., Stoeger, A. S., Frey, R., Lohscheller, J., Titze, I. R., Gumpenberger, M., and Fitch, W. T. (2012). "How Low Can You Go? Physical Production Mechanism of Elephant Infrasonic Vocalizations," Science 337, 595-599

Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of voiced sounds from a two-mass model of the vocal cords," The Bell System Technical Journal 51, 1233-1268.

Titze, I. R. (1980). "Comments on the myoelastic - aerodynamic theory of phonation," J Speech Hear Res 23, 495-510.

Titze, I. R. (1988). "The physics of small-amplitude oscillation of the vocal folds," J Acoust Soc Am 83, 1536-1552.

van den Berg, J. (1958). "Myoelastic-aerodynamic theory of voice production," Journal of Speech and Hearing Research 3, 227-244.

The Effect of Social Isolation on Vocalizations and Neuromuscular Junctions of Aged Rats

Aaron M Johnson [1]

[1]University of Illinois at Urbana-Champaign

amj3@illinois.edu

Introduction

The thyroarytenoid muscle plays a primary role in vocal production and undergoes atrophic changes with advancing age. Studies of the limb musculature have shown atrophic changes are preceded by alterations in the neuromuscular junction (NMJ) (Deschenes, Roby, Eason, & Harris, 2010). Previous work from our lab and others has shown that vocal changes and atrophy in the laryngeal muscles are related to morphologic alterations at the NMJ (Johnson, Ciucci, & Connor, 2013). However, the role of reduced voice use and its interaction with advanced age is unknown. To successfully diagnose and treat age-related voice disorders, it is critical to understand the differential effects of aging versus decreased voice use. We hypothesized that reduced voice use would exacerbate the effects of advanced age on both functional vocal decline and atrophic changes in laryngeal muscles. We tested this hypothesis with a rodent vocalization model by examining vocal function (intensity and frequency of vocalizations) and morphologic signs of atrophy in NMJs from the thyroarytenoid muscles of old rats that were either socially isolated or communally housed.

Methods

A total of 20 F344/BN male rats were housed over an 8-week period, evenly divided into a social isolation group or a communally housed group. All rats were 31 months old at the conclusion of the study, representing advanced age. Rats use ultrasonic vocalizations (USVs) to communicate with each other; therefore, we hypothesized the socially isolated rats would vocalize less than the communally housed rats.

Baseline and post-intervention USVs were elicitated and recorded using an existing mating paradigm (Johnson et al., 2011). USVs were also recorded and counted from each cage for 24-hour periods throughout the 8 weeks to monitor spontaneous USV production and confirm the socially isolated rats were vocalizing less than the communally housed rats.

Larynges were excised post-mortem for NMJ imaging using immunohistochemistry and confocal microscopy. Individual NMJs were analyzed for synapse size, pre/post-synaptic overlap, and motor endplate dispersion (Johnson et al., 2013).

Results

During the 8 weeks, the socially isolated rats vocalized an average of 11 times per day, significantly less than the communally-housed rats' average of 61 vocalizations per day. This validated our model as a means of studying reduced vocal use. Acoustic comparisons of pre/post USVs showed a decrease in mean USV amplitude in socially isolated rats with the opposite effect in the communally-housed group (an increase in mean amplitude). No differences in peak frequency were found. Likewise, no differences were found in NMJ morphology, with all rats exhibiting age-related changes consistent with our previous findings.

Conclusions

This study demonstrates social isolation can cause functional declines in the voice in an aging rat model. However, the neuromuscular mechanisms underlying these functional declines are still unclear. While our previous work showed increasing voice use in old rats decreased motor endplate dispersion in the NMJ, this current study does not show decreased voice use increases motor endplate dispersion. Therefore, motor endplate dispersion is likely due to aging, not reduced vocal use.

References

Deschenes, M. R., Roby, M. A., Eason, M. K., & Harris, M. B. (2010). Remodeling of the neuromuscular junction precedes sarcopenia related alterations in myofibers. *Experimental gerontology, 45*(5), 389-393. doi:10.1016/j.exger.2010.03.007

Johnson, A. M., Ciucci, M. R., & Connor, N. P. (2013). Vocal training mitigates age-related changes within the vocal mechanism in old rats. *J Gerontol A Biol Sci Med Sci, 68*(12), 1458-1468. doi:10.1093/gerona/glt044

Johnson, A. M., Doll, E. J., Grant, L. M., Ringel, L., Shier, J. N., & Ciucci, M. R. (2011). Targeted training of ultrasonic vocalizations in aged and Parkinsonian rats. *J Vis Exp*(54). doi:10.3791/2835

Highly resolved temporal analysis of the flow field in a synthetic human larynx model

Alexander Lodermeyer[(1)], Stefan Becker[(1)], Veronika Birk[(2)], Michael Döllinger[(2)] and Stefan Kniesburges[(2)]

[(1)]Dep. of Process Machinery and Systems Engineering, Friedrich-Alexander University Erlangen-Nürnberg, [(2)]Dep. of Phoniatrics and Pediatric Audiology at the Dep. of Otorhinolaryngology, Head and Neck Surgery, University Hospital Erlangen, Medical School at Friedrich-Alexander Unversity Erlangen-Nürnberg

lo@ipat.uni-erlangen.de

Introduction

Phonation is the result of a fluid-structure-acoustic interaction. The vocal fold oscillation has been extensively examined within the last decades, as it is the source of the phonation process. The influence of the geometric boundary conditions especially in the supraglottal region of the larynx has not been fully understood yet. As current studies showed, these boundary conditions play a significant role in the phonation process. In a simplified supraglottal geometry, the formation of a large recirculation area was verified (Kniesburges et al., 2013). The existence of a supraglottal geometry also strongly influences the glottal jet flow (Lodermeyer et al., 2015) as well as the resulting acoustics (Blandin et al., 2015). It is therefore obvious that a more profound analysis of the flow in the supraglottal region would lead to a better understanding of the fluid-structure-acoustic interaction during phonation.

Methods

The synthetic vocal fold model applied in the experimental setup is based on the M5 geometry (Scherer et al., 2001) and shows flow-induced oscillations. In the experimental setup, a subglottal and a supraglottal channel are incorporated in order to replicate human scales. For an analysis of the supraglottal flow field, a standard Particle Image Velocimetry (PIV) system was extended to execute phase-locked measurements. In a post-processing procedure, phase-averaging lead to a phase-resolved flow field of a whole oscillation cycle. As this method only reveals the contents of the basic oscillation frequency, higher-frequency components within an oscillation cycle are neglected. This limitation is handled by another measurement series using high-speed PIV. Therewith, the intra-cycle devolution of the flow field can be investigated with adequate time-resolution.

Results and Conclusions

The resulting flow field of both the phase-averaging PIV and the high-speed PIV measurements reveals an asymmetric supraglottal jet flow. The jet is deflected in the medial-lateral direction to different sides of the supraglottal channel wall. The direction of the jet

deflection changes between different ocillation cycles. This bimodal behaviour vanishes if synthetic ventricular folds are added into the supraglottal channel or glottis closure is avoided during vocal fold oscillation.

An analysis of the aeroacoustical sources in the flow field according to Lighthill's acoustic analogy revealed the most intense source right at the glottis. This is similar to numerical investigations (Kaltenbacher et al. 2014). Additionally, a Pricipal Component Analysis (PCA) of both the high-speed PIV and the phase-locked PIV measurements show coherent structures in the flow field.
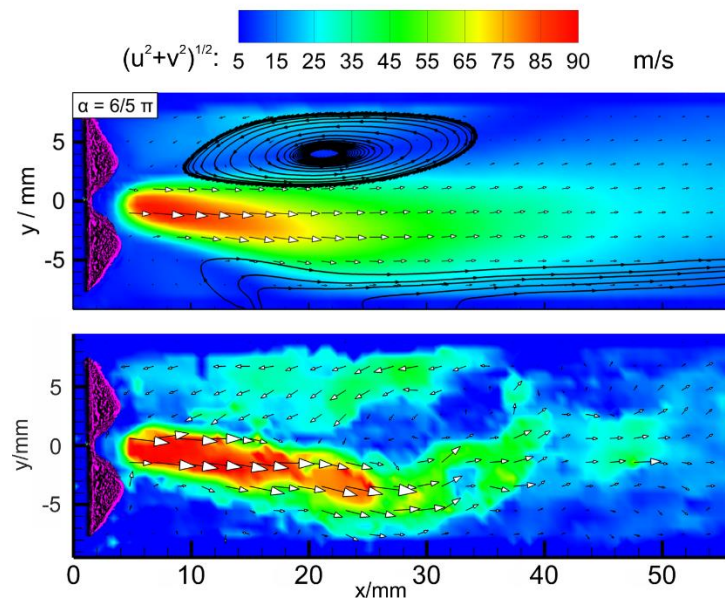

Figure 1: Exemplary supraglottal flow field as the result of the phase-averaging method (top) and a single PIV measurement (bottom). A large recirculation area leads to a deflection of the glottal jet.

As the supraglottal flow field and therewith the phonation process are influenced by the supraglottal channel, we suggest the inclusion of a supraglottal model for future studies applying synthetic larynx models.

References

Blandin R, Arnela M, Laboissière R, Pelorson X, Guasch O, Van Hirtum A, Laval X: Effects of higher order propagation modes in vocal tract like geometries (2015). J Acoust. Soc. Am. 137(2):832-843

Kaltenbacher M, Zörner S, Hüppe A, Sidlof P: 3D numerical simulations of human phonation (2014). Proceedings of the 11th world congress on computational mechanics - WCCM XI

Kniesburges S, Hesselmann C, Becker S, Schlücker E, Döllinger M: Influence of vortical flow structures on the glottal jet location in the supraglottal region (2013). J Voice 27(5):531-544

Lodermeyer A, Becker S, Döllinger M, Kniesburges S (2015): Phase-locked flow field analysis in a synthetic human larynx model (2015). Exp Fluids 56(4)

Scherer R, Witt K, Zhang C, Kucinschi B, Afjeh A: Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees (2001). J Acoust Soc Am 109(4):1616–1630

Subharmonic tone generation in an artificial vocal fold model

Stefan Kniesburges[1], Alexander Lodermeyer[2], Stefan Becker[2], Michael Döllinger[1]

[1]Dep. of Phoniatrics and Pediatric Audiology at the Dep. of Otorhinolaryngology, Head and Neck Surgery, University Hospital Erlangen, Medical School at Friedrich-Alexander Unversity Erlangen-Nürnberg, [2]Dep. of Process Machinery and Systems Engineering, Friedrich-Alexander University Erlangen-Nürnberg

Stefan.kniesburges@uk-erlangen.de

Introduction

Within a synthetic larynx model including elastic vocal folds which showed flow-induced oscillation in the physiological frequency range, Kniesburges et al. (2013) found additional subharmonic tones in the sound spectra which are similar to those found in diplophonia (Cavalli and Hirson (1999)). These subharmonics were located in between two consecutive harmonics of the oscillation frequency of the vocal folds each at exactly the medial frequency position. Although subharmonic tones are commonly associated with voice disorders originating from asymmetrical vocal fold oscillations (Cavalli and Hirson(1999), Kramer et al. (2013)), the synthetic vocal folds oscillated left-right symmetrically and periodically with glottis closure. However, the amplitudes of the subharmonic tones depended strongly on the transversal diameter of the supraglottal channel. On the basis of hydrodynamic pressure data obtained along the supraglottal channel, the spectral analysis of the hydrodynamic pressure revealed a small peak at the half fundamental frequency of the vocal fold oscillation $f_0/2$ which displays the first subharmonic peak. Furthermore, the generation of this peak could be located in the region of the pulsatile glottal jet flow. The frequency at $f_0/2$ represents a cycle-to-cycle variation of the static pressure that indicates a flow phenomenon as origin.
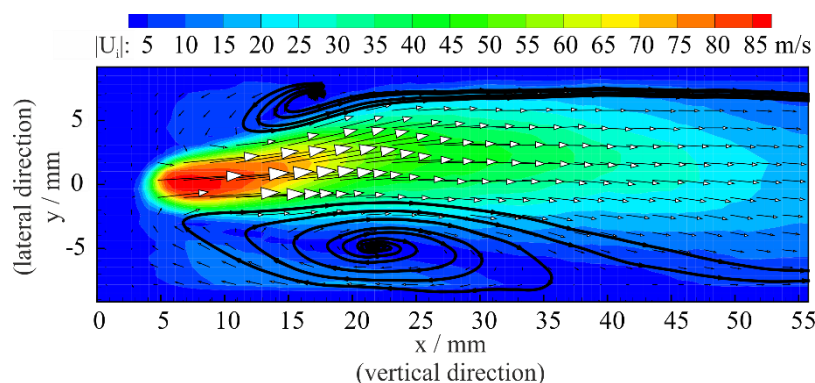

Figure 1: Phase-average flow field in the supraglottal region exhibiting the deflected glottal jet (Lodermeyer et al. 2015). For averaging, flow field data of 900 oscillation cycles at a constant phase angle were used.

Methods

A synthetic larynx model was used which included silicone vocal folds based on the M5 model. To get insight in this flow phenomenon, the two-dimensional supraglottal flow field that was measured at an instance with fully developed glottal jet was decomposed by a PCA into spatial and temporal eigenvectors and the according eigenvalues. The flow field data were obtained by a phase-locked particle image velocimetry technique. The decomposition was achieved on the basis of more than 900 instantaneous single flow velocity fields representing more than 900 oscillation cycles of the vocal folds. All dimensions of the synthetic larynx as well as details of the measuring setup are provided in Lodermeyer et al. 2015.

Results

The flow field was dominated by the first two spatial eigenvectors: the first eigenvector referred to the average flow field with the deflected glottal jet. The second spatial eigenvector showed a large recirculation area, which switched its rotation direction in different cycles. Thereby, the ratio between cycles with clockwise and counterclockwise rotating recirculation area amounted 50:50 indicating a cycle-to-cycle change of the rotation direction. Concerning the deflected glottal jet, the second spatial eigenvector superposed a velocity component orthogonal to the jet axis which resulted in a small lateral jet deflection from its average location. Under the assumption of a cycle-to-cycle change of the rotation direction in the second spatial eigenvector, the glottal jet location fluctuated cycle-wise with a frequency of $f_0/2$. This would result in cycle-wise fluctuation of the mean pressure in the immediate supraglottal flow region which potentially generates the first subharmonic peak at $f_0/2$ in the hydrodynamic pressure spectra.
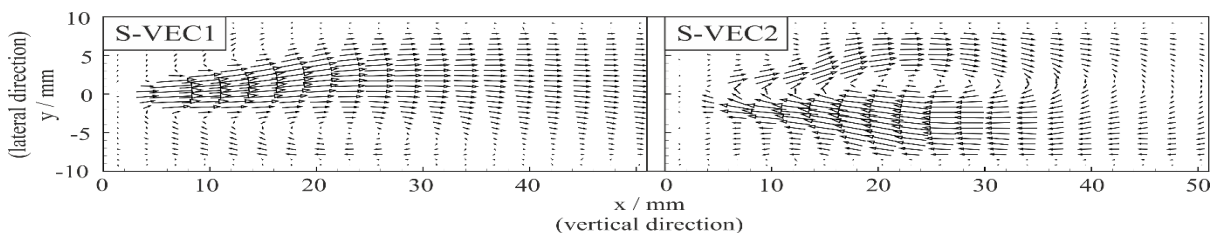

Figure 2: First two spatial eigenvectors associated to the two largest eigenvalues. The corresponding phase-averaged flow field is depicted in fig. 1.

Conclusions

We conclude that subharmonic tones in human voice might also be generated by lateral variations of the glottal jet location even in the absence of asymmetric vocal fold oscillation. In these cases, the glottal jet variation might be the result of flow instabilities produced by specific boundary conditions in the immediate supraglottal tract, e.g. a small lateral diameter.

References

Kniesburges, S., Hesselmann, C., Becker, S., Schlücker, E., & Döllinger, M. (2013). Influence of vortical flow structures on the glottal jet location in the supraglottal region. J. Voice, 27(5), 531-544.

Cavalli, L., & Hirson, A. (1999). Diplohonia reappraised. J. Voice, 13(4), 542-556.

Kramer, E., Linder, R., & Schönweiler, R. (2013). A Study of subharmonics in connected speech material. J. Voice, 27(1), 29-38.

Intraglottal vortices during phonation

Sid Khosla[1], Liran Oren[1], Ephraim Gutmark[2]

[1]Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati

[2]Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati

Khoslam@ucmail.uc.edu

Introduction.

During phonation, the vocal folds can (but not always) produce a divergent shape during closing. This produces an adverse pressure gradient that acts on the boundary layer, resulting in flow separation and intraglottal vortices. Although there is minimal controversy about the existence of intraglottal formation, there is mild debate about the cause of the vortices – for example, some noted voice scientists have speculated that intraglottal vortices form because of shear layer interactions rather than the mechanism listed above. There is significant controversy about whether the vortices affect vibration and/or sound production. Our experimental methods will be presented, multiple correlations will be shown, and findings inconsistent with current theories will be discussed. We will also describe how experiments in excised canine larynges can show correlations and provide data that can be used for validation of computationsal models. However, computational models are necessary to prove causation.

Methods

Using high speed 2D particle imaging velocimetry (PIV) and other techniques, intraglottal velocity fields and intraglottal geometry of the medial aspect of both folds can be measured during the closing phase of vibration. There is no vocal tract. The laser illuminates a 1mm coronal plane at the mid-membranous location (halfway between anterior commissure and vocal process). From the two-dimensional intraglottal geometry, divergence angle, inferior width and superior width of the glottis is measured at multiple times in the closing phase. From the superior width measures, a maximum width declination rate (MWDR) is determined. Maximum area declination rate (MADR) is determined from high speed videography. Intraglottal velocity fields can be used to characterize location, phase, and strength of the vortices. Flow rate at the glottal exit in the 1 mm plane is also determined. From the measurements intraglottal pressures during closing are determined. From the acoustic signal, SPL, higher harmonics (HH) and other acoustic measures can be determined. Stress strain curves for the inferior and superior halves of the fold are determined by the indentation method.

Results.

The following correlations are seen. Increasing the MDA increases the circulation of the vortices (VC). Increased VC is correlated with increase maximum vocal fold closing speed, increased SPL and HH. Increasing the subglottal pressure is correlated with increase of MDA. Increase VC is also correlated with increased skewing to the right of the 2D flow rate and width curves.

Conclusion.

The findings show a correlation between the strength of the intaglottal vortex and several acoustic and flow parameters. There is also a correlation between area and flow rate skewing to the right and VC. These findings will be contrasted with current findings in computational and mechanical models. Although our findings support that the vortices have a role in phonation and show multiple correlations, computational models are needed to prove causation.

Simulations of Vocal Fold Replicas Containing Liquid-Filled Cavities

Scott L. Thomson[1]

[1]Brigham Young University-Idaho, Rexburg

thomson@byu.edu

Introduction

Self-oscillating vocal fold replicas are used to study human phonation. Two types of replicas have included liquid-filled cavities; first, water-filled latex tubes that simulate the vocal folds (e.g., Ruty et al., 2007), and second, silicone models with liquid- or gel-filled interiors (e.g., Latifi et al., 2014). In the latter, live cells are inserted into the cavities and the replicas are phonated in a bioreactor. In this research the flow-induced vibratory responses of these models are studied using computational models.

Methods

Two-dimensional finite element models, each with fully-coupled fluid and solid domains, of the two abovementioned types of replicas have been developed (see Figures 1 and 2). Each simulation includes one region for the primary airflow, a second region for the solid domain (e.g., either a latex tube or a silicone model), and a third region for the liquid-filled interior. In the presentation, details regarding material properties, boundary conditions, solver setup, and moving mesh control will be discussed.

Results

Both models exhibit self-sustained oscillations when defined using parameters representative of those used in the experiments. The water-filled tube steady-state vibration pattern is distinctly different than that of the silicone model with the liquid-filled cavity. The silicone model exhibits distinct wave-like motion, whereas the tube model does not. In the presentation, motion characteristics of the models will be presented, along with results for sensitivity of each model's response to liquid-filled cavity properties (e.g., viscosity, shape, size, and location).

Conclusions

Simulations of two synthetic vocal fold replicas that feature liquid-filled cavities have been performed. The models and resulting simulations allow for exploration of the flow dynamics within the fluid-filled cavity as well as of the sensitivities of the models to parameter values.
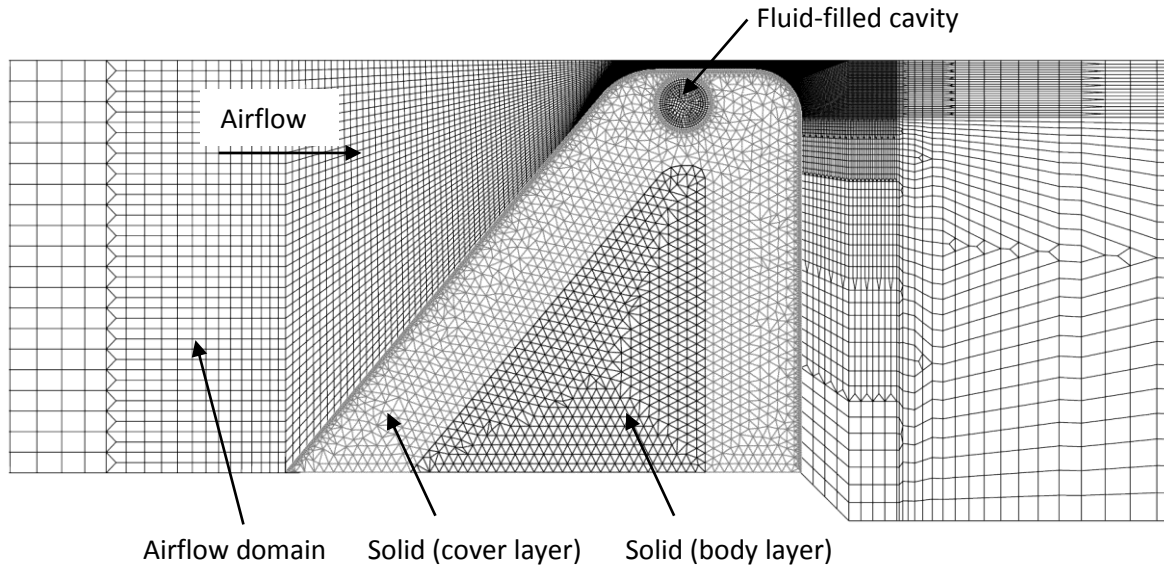
Figure 1. Fluid and solid meshes of computational model of two-layer (body-cover) silicone vocal fold replica with fluid-filled cavity.
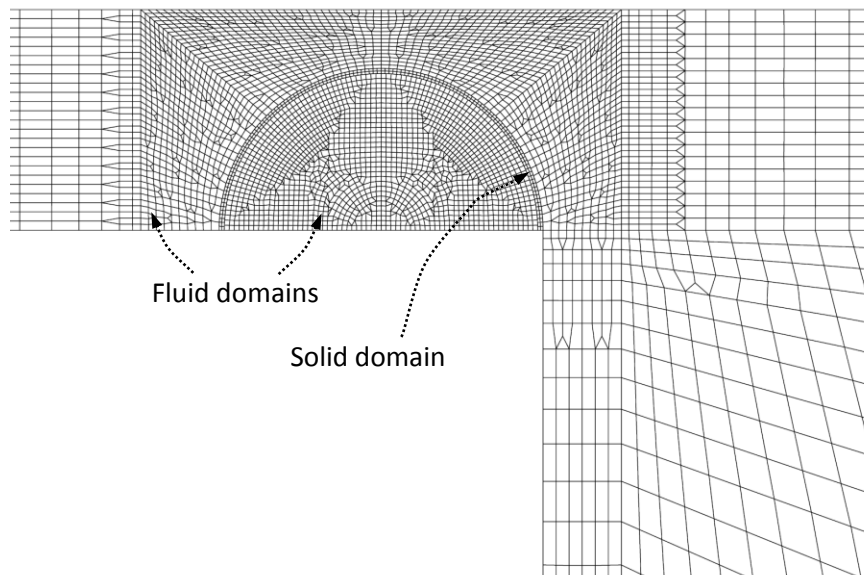
Figure 2. Computational fluid domains of liquid-filled latex tube model.

References

Latifi, N., Khadivi Heris, H., Kazemirad, S., & Mongeau, L. (2014). "Development of a Self-Oscillating Mechanical Model to Investigate the Biological Response of Human Vocal Fold Fibroblasts to Phono-Mimetic Stimulation," Proceedings of the ASME 2014 International Mechanical Engineering Congress & Exposition IMECE2014, Montreal, Quebec, Canada, November 14-20.

Ruty, N., Pelorson, X., Van Hirtum, A., Lopez-Arteaga, I., & Hirschberg, A. (2007). "An in vitro setup to test the relevance and the accuracy of low-order vocal folds models." J. Acoust. Soc. Am., 121(1):479-490.

Computational Modeling of Flow-Structure-Acoustic Interaction inside a Simplified Airway during Voice Production

Weili Jiang[1], Xudong Zheng[1], and Qian Xue [1]

[1]University of Maine

xudong.zheng@maine.edu

Introduction

Human voice quality is directly determined by the dynamic behavior of the glottal flow, vocal fold (VF) vibration and acoustics in the upper airway. Despite many years' research effort, direct relationship among the three coupled systems still remains elusive. This study aims to develop a fully coupled, first-principle based, three dimensional numerical model of flow-structure-acoustic interaction of voice production, providing a more accurate and reliable method to learn the relationship inherent the coupled dynamic systems. Numerical simulation is carried out in a simplified airway shape. Detailed descriptions of the flow structure, VF vibration and acoustic characteristics are obtained. Relationship of the three coupled systems is studied.

Methods

A sharp interface immersed boundary method based incompressible flow solver is utilized to model the glottal flow. A finite element method based solid mechanics solver is utilized to model the VF vibration. A 6th-order immersed boundary method based acoustics solver is utilized to directly compute sound. These three solvers are explicitly coupled to mimic the flow-structure-acoustic interaction. The geometry of the airway is reconstructed based on the in-vivo MRI measurement reported by Story et al.(1995). A three-layer VF model is taken from Titze and Talkin(1979). All the three layers are modeled as linear elastic, transversely isotropic and nearly incompressible material. A pressure drop is applied at the inlet and exit of the airway. Fig.1 shows the basic simulation setup. The FSI interface is the place that airway and VFs contact.
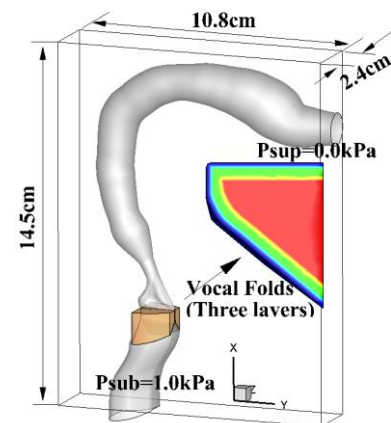
Fig.1 Simulation setup: computational domain, airway and VFs model

Results

The fundamental frequency of the flowrate waveform is 219 Hz. Mean flow rate, maximum flow rate, open quotient and skewness quotient are 129.9mL/s, 231.1mL/s, 0.73 and 1.47, respectively. Fig.2 shows the total vorticity and nondimensionalized acoustic pressure in the airway during one vibration cycle. At the beginning of the cycle (Fig.2(a)-(b)), the glottis

starts to open from the beneath and forms a convergent shape. The glottal jet is relatively uniform along the anterior-posterior direction. At the maximum flow rate (Fig.2(c)), the jet is found to hit on the the lumen wall and complex vortex structure start to appear. At the closing phase (Fig.2(d)-(e)), The glottal jet quickly dissipated. At the end of the cycle, only small vortex structure remains in airway dissipating gradually. No flow deflection is observed in this simulation. Acoustic pressure during the glottis opening and closing phases is shown in Fig.2(f) and (e). detailed analysis of the relationships between flow, VF vibration and acoustics will be presented in the conference.
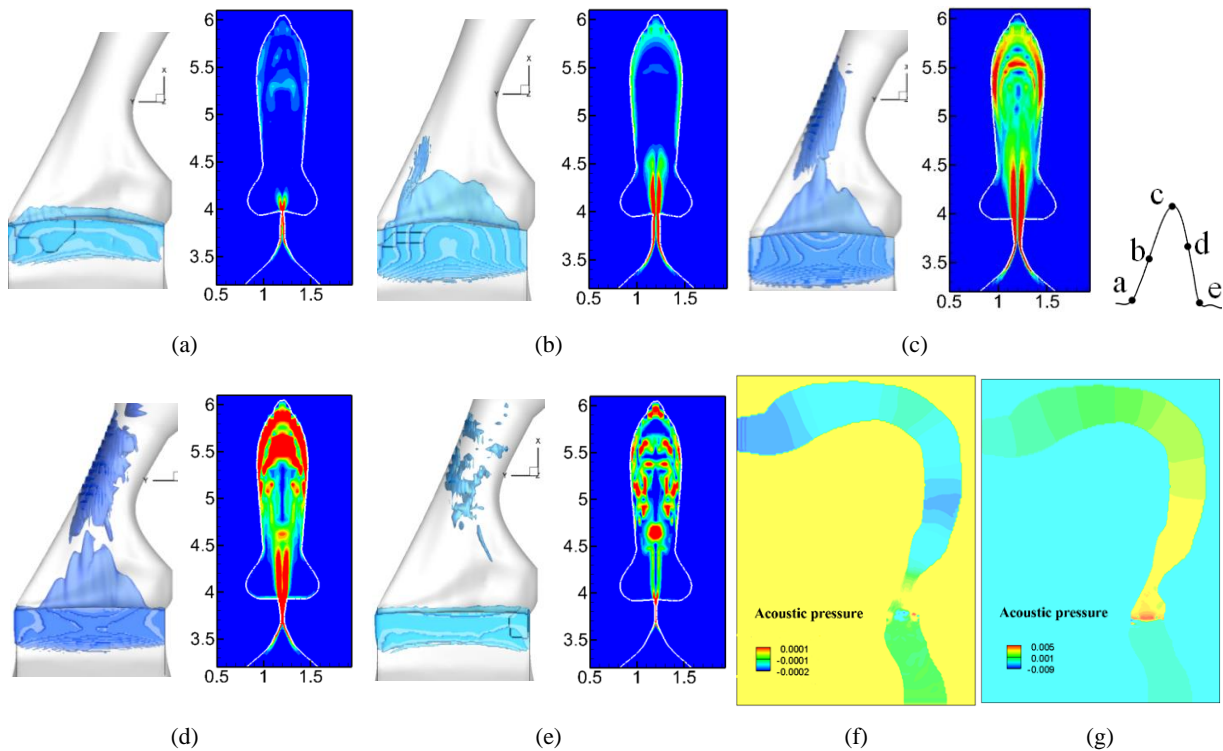
Fig.2 Total vorticity and nondimensionalized acoustic pressure in one vibration cycle: (a)-(e) corresponding to black dots a to e on the wave form curve respectively. In each fiture, the left side is isosurface of the total vorticity, the right side is the contour of the total vorticity at Y=3.5cm plane. (f) and (g) are the contours of the acoustic pressure at the Z=1.2cm plane, corresponding to instant b and d respectively.

Conclusions

A direct numerical simulation of flow-structure-acoustic interaction is carried out in the entire airway. The resulting glottal flow wave form and some important voice quality related quantities are well within the physiological range. This simulation captures the important convergent-divergent vibration pattern of the VFs. Aerodynamics pressure distribution, vorticity structures and acoustic pressure propagtation in the glottis are carefully analyzed. This simulation lays the foundation work for future study of building direct relationship between biomechanics and sound.

References

Story, B. H., and Titze, I. R. (1995). "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. 97(2): 1249–1260.

Titze, I. R. and Talkin, D. T.(1979). " A theoretical study of effects of various laryngeal configuration on the acoustics of phonation," J. Acoust. Soc. Am. 66:60-74.

Dynamic and energetic relevance of glottal jet asymmetry

Jubiao Yang[1], Michael H. Krane[2,,3,4] , and Lucy T. Zhang[(1)]

[(1)]Department of Mechanical, Aerospace and Nuclear engineering, Rensselaer Polytechnic Institute,[(2)]Applied Research Laboratory, Pennsylvania State University,

[(3)] Department of Biomedical Engineering, Pennsylvania State University[(4)] Department of Acoustics, Pennsylvania State University

zhanglucy@rpi.edu

Introduction

The asymmetry of glottal jet flow and its effects on the volume flow, sound production and vibration frequency and patterns have been questioned and debated for years (Tao et al, 2007). It has been hypotheiszed that glottal jet asymmetries affect the glottal resistance and voiced sound production. In this work, we compare measures of glottal flow and sound production for simulated phonatory flow, for a case where glottal jet symmetry is imposed, and an otherwise identical case without the said enforced flow symmetry.

Methods

Numerical simulation of the phonation process is performed using the Immersed Finite Element Method (Wang and Zhang, 2013). In order to produce a simulation with enforced symmetry, a half-space domain was used. For the case where symmetry was not imposed, a full-space simulation was performed. In all other respects, the two domains are identical.

The Immersed Finite Element Method efficiently solves fully-coupled fluid-structure interaction (FSI) problems, obviating time-consuming mesh-updating and re-meshing processes by combining fluid mesh throughout the domain in which the effect of the solid is given by appropriate interaction forces. A perfectly matching layer (PML) is also included at the inflow and outflow boundaries to control the propagation of sound waves at these boundaries.

Results

We observe strong asymmetry of the glottal jet in the full-space domain simulation (see Figure 1), and this is associated with asymmetric vocal fold motion. We estimated the terms of the momentum and energy equations in a control volume, and use these to assess the dynamic and energetic relevance of the glottal jet asymmetry. Figure 2 shows glottal volume flow rates for the two cases.

Summary

Using high-fidelity, fully-coupled fluid-structure interaction simulations of phonation, we have compared the the differences in phonation between a symmetric and an asymmetric glottal jet.

(a) no imposed symmetry

(b) imposed symmetry

Figure 1 Glottal jet flow velocity fields in cases that have (a) no imposed symmetry and (b) imposed symmetry.
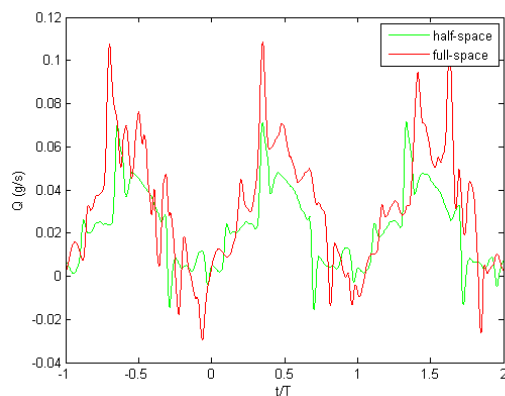
Figure 2 Comparisons of supraglottal mass flow rates over vibration cycles between half-space (symmetric) and full-space (asymmetric) models

References

Tao, C., Zhang, Y., Hottinger, D.G., and Jiang, J.J. (2007) Asymmetric airflow and vibration induced by the Coanda effect in a symmetric model of the vocal folds. *J Acoust Soc Am.* **122**(4):2270-8.

Wang, X. & Zhang L. (2013) Modified immersed finite element for fully-coupled fluid-structure interactions, *Comp. Meth. Appl. Mech. Eng.* **267** 150-169.

Rethinking Vocal Fold Contact: The Role of Viscous Dissipation

Byron D. Erath[1], Matias Zañartu [2], Sean D. Peterson[3],

[1] Clarkson University, [2] Universidad Técnica Federíco Santa María, [3] University of Waterloo,

Introduction

Vocal hydration is known to play a critical role in vocal fold health by mitigating excessive contact forces. Most interestingly, it has been reported that chronic dehydration also results in a decrease in vocal fold thickness. This finding appears to confirm previously proposed theories that PTP is inversely proportional to vocal fold thickness. We propose that modeling viscous dissipation during vocal fold collision will provide novel insights into the relationships between vocal fold tissue hydration (i.e., thickness and viscosity) and contact forces.

Methods

The body-cover orientation of the vocal folds is modeled as a solid body of mass M, with the cover being represented by a thin layer of viscous fluid supported by an elastic component that is coupled to the body. The medial surface is assumed to have a parabolic geometry, expressed in terms of a parameter ε, which specifies the height of the vocal fold surface above a nominally flat body. In this manner, the influence of viscous dissipation on contact forces is considered for both normal phonation ($\varepsilon = 0$) and the progression of bilateral nodules ($\varepsilon > 0$). Viscous dissipation during vocal fold collision is modeled as a Newtonian fluid. As such, the Navier-Stokes equations govern the dynamics of the vocal fold cover layer, with the stated assumptions producing the standard thin film lubrication equation. The governing equation of vocal fold contact is solved by considering the 1-D dynamics of the body mass and elastic coupling of the cover with the resultant viscous contact pressure. Included is elasticity in the cover during contact.

Results

Nondimensionalization of the governing equations introduces two nondimensional parameters;

$$\beta = \frac{\mu b L^2}{MVH},$$

and

$$\kappa = \frac{2EbH^2}{mV^2}.$$

In the above equations μ is the cover viscosity, b is the vocal fold width, L is the length in the inferior-posterior direction, M is the vocal fold mass, V is initial velocity, H is the initial thickness of the cover layer, and V is the initial velocity of the cover. The nondimensional

parameter β provides a ratio of viscous tissue dissipation to the rate of change of kinetic energy in the vocal fold mass, and κ prescribes a ratio of the elastic energy stored in the cover over the kinetic energy.

Figure 1 shows how the peak contact pressure varies in time (Figure 1a) as well as its dependency on κ (Figure 1a) and β (Figure 1b). As κ decreases (Figure 1a) the pressure loading becomes an impulse where the elasticity of the vocal fold cover becomes negligible. This same behavior has been observed in experimental and computational investigations of vocal fold contact (Jiang and Titze, 1994; Tao et al., 2006). Figure 1b shows how the maximum contact pressure increases with decreasing values of β. As $\beta \rightarrow 0$ it is representative of a pressure impulse arising from contact between two rigid bodies, and $\beta \rightarrow \infty$ representing instantaneous viscous energy dissipation.

(a)   (b)

Figure 2: (a) Variation of peak contact pressure versus nondimensional time for values of $\kappa = 0, 0.10, 0.25$, and 0.50.
(b) Variation of peak pressure versus κ for values of $\beta = 0.001, 0.01, 0.1, 1$, and 10.

Figure 2 shows how the contact pressure is influenced by both the presence of a protuberance on the vocal fold surface and the nondimensional parameters κ and β. As expected, the maximum pressure increases with decreasing values of κ and β. For a constant value of κ or β a slight nonlinear trend in pressure increase is observed with increasing protuberance height, ε.

(a)   (b)

Figure 3: Nondimensional pressure variation as a function of protuberance height, ε as a function of (a) κ (with values of $\kappa = 0, 0.10, 0.25$, and 0.50), and (b) β (with values of $\beta = 0.001, 0.01, 0.1, 1$, and 10).

Conclusions

A physical, causal link between contact pressure and tissue thickness, viscosity, and elasticity is established by modeling viscous dissipation of vocal fold contact. The findings have direct implications for vocal fold hydration, indicating that decreased cover thickness and increased viscosity associated with dehydration causes a significant increase in contact pressure.

References

Jiang, J. J. & Titze, I. R. (1994). Measurement of vocal fold intraglottal pressure and impact stress. Journal of Voice, 8:132-144.

Tao, C., Jiang, J. J., & Zhang, Y. (2006). Simulation of vocal fold impact pressures with a self-oscillating finite-element model. Journal of the Acoustical Society of America, 119: 3987

Assessing the influence of intraglottal vortices on vocal fold dynamics

Brian Q. Kettlewell[1], Sean D. Peterson [1]

University of Waterloo

peterson@mme.uwaterloo.ca

Introduction

The rate of closure of the vocal folds in the closing phase of the phonatory cycle is generally understood to play a role in the intelligibility of speech. In particular, rapid glottal closure contributes to the high frequency content of the resulting acoustic signal, which is perceived as a richer sounding voice. Recent studies have noted the formation and shedding of vortices within the glottis in the closing phase of the phonatory cycle (Erath and Plesniak, 2010), which has led to the hypothesis that the pressure gradients due to these vortices may promote rapid closing of the glottis (Khosla *et al*., 2009). Farahani and Zhang (2014) performed simulations of a coupled fluid-structure interaction with a one-dimensional Bernoulli flow model of the glottal flow with a superposed *ad hoc* negative pressure correction at the medial surfaces of the vocal folds to account for the influence of intraglottal vortices. In this paper, we directly study the influence of intraglottal vortices on vocal fold dynamics using a potential flow model of the glottal airflow with advecting vortices coupled to a reduced order structural model of the vocal fold tissues. Rather than imposing an *ad hoc* pressure correction on the vocal folds to account for the presence of a vortex, herein we directly model the vortices and their kinematics in a fully coupled fluid-structure interaction simulation.

Methods

The vocal folds are represented using a two-mass bar-plate model with standard properties (Titze and Story, 2002). The fluid is modeled as inviscid, incompressible, and irrotational, allowing for a two-dimensional potential flow representation. The glottal geometry, prescribed by the cover mass position of the bar-plate model, is transformed into a two-dimensional strip via a Schwarz-Christoffel transformation. The fluid flow is solved in this simpler geometry then mapped back to the physical domain. The fluid pressure loading on the medial surfaces of the vocal folds is obtained from the potential flow solution and fed back to the structural dynamics through the forcing terms in the governing equations. Since potential flow is governed solely by conservation of mass, the flow solution is linear; consequently, vortices can be added to the flow field directly through superposition. This enables a direct evaluation of the impact of fluid vortices on the vocal fold dynamics, by first simulating a case without vortices, then again with their addition. For the cases with vortices, a pair of counter-rotating vortices are added far upstream of the glottis such that they pass through the glottal entrance when it is in a diverging configuration. The vortices advect due to both a constant free stream and mutual induction

Results

By employing the aforementioned two-dimensional fluid flow solution (in the absence of acoustics), the glottal area waveform is inherently asymmetric, with closure occurring more rapidly than opening. This is in agreement with *in vivo* experimental measurements. This does not appear with one-dimensional flow models unless acoustic coupling is included. Inclusion of vortices into the fluid flow results in a pressure perturbation that moves along the length of the medial vocal fold surface, as shown in Figure 1a. The glottal area waveforms for cases both with and without vortices are shown in Figure 1b. There are some slight deviations, but overall the rates of closure for the two cases are very similar. This is quantitatively shown with the glottal area waveform derivatives for both cases, which largely overlap.
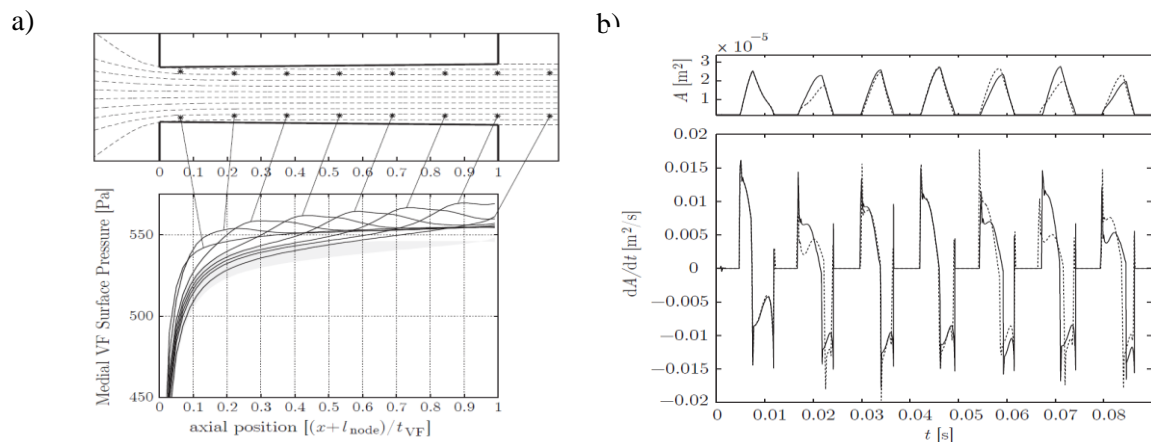

Figure 1: (a) Streamlines overlaid with vortex positions within the glottis, as well as the associated pressure distribution on the medial surface as functions of time. (b) Glottal area waveform and waveform derivative for cases with (solid lines) and without (dashed lines) vortices.

Conclusions

While intraglottal vortices do influence the pressure distribution on the medial vocal fold surfaces, their influence on the overall vocal fold dynamics appears to be modest. This may be in part due to the relatively short duration that the vortices influence the pressure loading on the vocal folds in comparison with loading from the free stream flow over the full glottal cycle. We note that the model neglects viscosity, and as such vortex roll-up is not captured, which likely plays a role in the overall vocal fold dynamics. However, in light of the work of Farahani and Zhang, in conjunction with the present work, it seems that intraglottal vortices do significantly impact glottal closure.

References

Erath, B. D. and Plesniak, M. W., "Viscous flow features in scaled-up physical models of normal and pathological phonation," *International Journal of Heat and Fluid Flow* **31**, pp. 468-481 (2010).

Farahani, M. H. and Zhang, Z., "A computational study of the effect of intraglottal vortex-induced negative pressure on vocal fold vibrations," *Journal of the Acoustical Society of America* **136**, pp. EL369-EL375 (2014).

Khosla, S. M., Murugappan, S., Paniello, R., Ying, J. and Gutmark, E. J., "Role of vortices in voice production: Normal versus asymmetric tension," *Laryngoscope* **119**, pp. 216-221 (2009).

Titze, I. R. and Story, B. H., "Rules for controlling low-dimensional vocal fold models with muscle activation," *Journal of the Acoustical Society of America* **112**(3), pp. 1064-1076 (2002).

The effect of vocal fold superior-inferior stiffness variation on sound production

Biao Geng[1], Qian Xue[1], Xudong Zheng[1]

[1] Department of Mechanical Engineering, University of Maine

qian.xue@maine.edu

Introduction

It was observed in both canine (Oren et al., 2014) and human (Chhetri et al., 2011) vocal folds (VFs) that the inferior aspect of the VF is much stiffer than the superior aspect. Such a feature, which might be due to the proximaty of the inferior aspect to the conus elaticus, is supposed to promote the divergent shape of the glottis during VF vibration and consequently faciliate the production of sound. This study aims to utilize a computational method to verify if this supposition is correct and improve the understanding of the effect of superior-inferior stiffness variation on sound production.

Methods

The computational model couples a finite element model of VF vibration (Xue et al. 2011) and and the 1D Bernoulli equation for the glottal flow. VF tissues are modeled as viscoelastic transversal isotropic incompressible material. The VF consists of two layers, as shown in Fig.1.a. Vertical stiffness gradient, defined as the ratio of the inferior-superior difference to the overall mean stiffness, is introduced in the cover layer by linearly increasing the stiffness from the superior to the inferior aspect. The overall stiffness is kept invariant while changing the stiffness gradient. A separation point at the minimum area is assumed, downstream which, the pressure is atmospheric, i.e. zero; upstream which, the pressure is solved with the Bernoulli equation. A parameteric study was conducted with subglottal presure ranging from 0.5kPa to 1.2kPa and stiffness gradient from 0 to 90%.

Fig.1: a. Finite element model of vocal fold, b. Vertical stiffness gradient.

Results

Fig.2.a shows the effect of stiffness gradient on the peak flowrate and mean flow rate when the subglottal pressure is kept at 1.0kPa. It is observed that as the stiffness gradient increases, the peak flowreate increases significantly; however, the mean flowrate remains almost the same. The trend is also observed under other suglottal pressure conditions. Fig.2.b comapres the phase-averaged waveform of the flowrate of the two cases with 90% stiffness gradient and 0% stiffness gradient (no stiffness variation), respectively, at 1.0 kPa subglottal pressure. It is observed that, for the case with stiffness gradient, the flowrate increases slower at the beginning, but it catches up quickly and
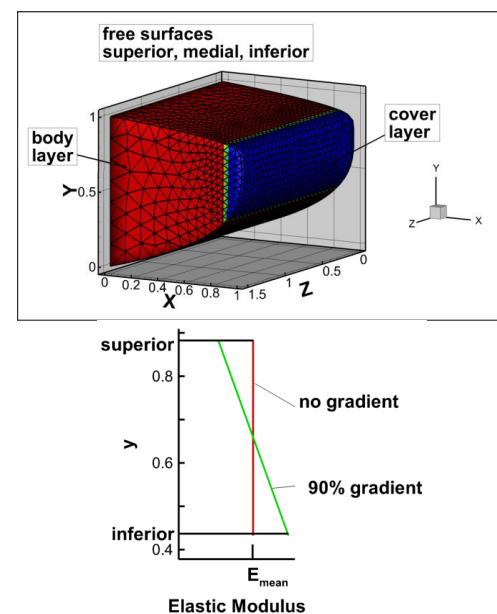
eventually reaches a much higher peak value. The deceleration of the flowrate is faster for the case with stiffness gradient. The big difference in flowrate suggests a significant effect of stiffness gradient on VF dynamics. Fig.2.c shows the VF profile at the mid-coronal plane at three different instants during a cycle of the two cases. It is found that, for the case with stiffness gradient, the glottis opens slower at the beginning. It is because the VF opens from the inferior aspect which has beomce stiffer and harder to be pushed apart. It explains why the flowrate incrases slower at the beginning. When the inferior aspect is gradully pushed apart, the superior aspect, which is softer, will be easier to be pushed apart. Therefore, a larger displacement and higher peak is seen in this case. During the closing phase, the stiffer inferior part has a larger restoring force. Therefore, it closes more quickly. At this phase, the larger displacement of the superior part and the quick closing of the inferior part result in a larger phase difference between the superior and inferior aspects. A more distinct divergent shape is observed in this case.
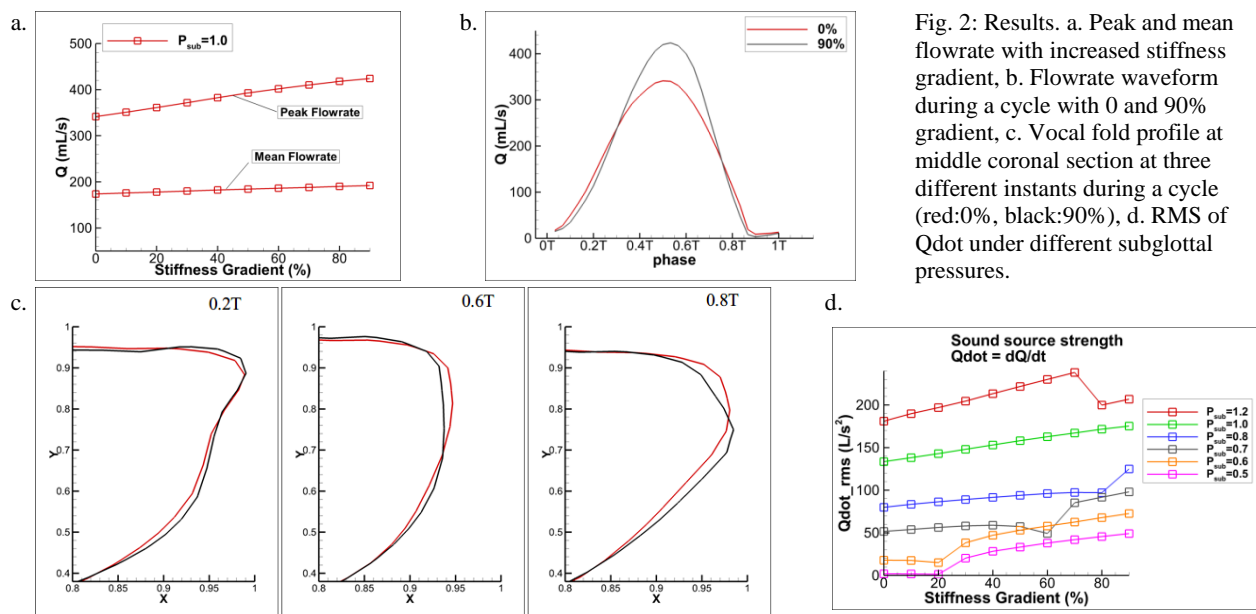

Fig. 2: Results. a. Peak and mean flowrate with increased stiffness gradient, b. Flowrate waveform during a cycle with 0 and 90% gradient, c. Vocal fold profile at middle coronal section at three different instants during a cycle (red:0%, black:90%), d. RMS of Qdot under different subglottal pressures.

Figure 2.d shows the variation of the root-mean-square value of Qdot with the stiffness gradient. Qdot is the changing rate of fowrate. It can be used as a measurement of sound source strength of the glottal flow. With increased gradient, Qdot increases significantly, indiating a stronger sound source. Futhermore, increasing stiffness gradient is found to reduce the threshold pressure. For example, for the cases with subglottal pressure at 0.5kPa, no sustained vibration is obtained when the stiffness gradient is below 20. As the gradient further increases, sustained vibration becomes possible.

Conclusions

The superior-inferior stiffness gradient is found to promote the divergent angle, increase peack flowrate, reduce threshold pressure, and produce a stronger sound source.

References

Chhetri, Dinesh K., Zhaoyan Zhang, and Juergen Neubauer. "Measurement of Young's modulus of vocal folds by indentation." Journal of Voice 25.1 (2011): 1-7.

Oren, Liran, et al. "Characterization of the vocal fold vertical stiffness in a canine model." Journal of Voice 28.3 (2014): 297-304.

Xue, Q., et al. "Sensitivity of vocal fold vibratory modes to their three-layer structure: Implications for computational modeling of phonation." The Journal of the Acoustical Society of America 130.2 (2011): 965-976.

Experimental study of the influence of a growth on a replica of the vocal folds

Xavier Pelorson [1], Paul Luizard [1]

[1] Gipsa-lab, UMR CNRS 5216, Grenoble University

xavier.pelorson@gipsa-lab.grenoble-inp.fr

Introduction

The presence of a growth at the surface of the vocal folds, such as a cyst or a nodule, is a common pathology especially amongst professionals of the voice.
The goal of this work is to try to understand how these growths affect the production and the acoustical quality of voiced sounds. For this purpose, an experimental set-up is designed in order to simulate these pathologies on a mechanical replica of the vocal folds. Different spheres can be added to the replica to simulate a growth with varying mass and diameter. Systematic threshold and sustained measurements are performed in order to quantify the perturbation induced by the presence of the growth.

Methods

The experimental set-up consists of a large pressure reservoir on which is attached a self-oscillating mechanical replica of the vocal folds. This replica is made of two layers of latex tubes between which small spheres of known mass and diameter can be inserted, in order to simulate a cyst or a nodule. The elasticity of the vocal folds replica could be controlled by inserting water under pressure inside the internal latex tube. A Laser Doppler Vibrometer (Optomet Vector Series) coupled to a scanning system (HGL Dynamics, VL Scanner) allows for complex frequency response functions estimation and thus to the determination of the first mechanical resonance frequencies and bandwidths of the vocal folds replica.

The pressure upstream of the vocal folds replica is recorded using a Kulite pressure sensor (XCS093) while the radiated acoustical pressure is recorded using a Brüel & Kjaer microphone. From these pressure measurements, physical parameters such as the fundamental frequency of oscillation, the threshold pressure (i.e. the minimum pressure needed to obtain self sustained oscillations) or the sound intensity are derived.

Three configurations are considered:

- a reference in the absence of any growth,

- a slight growth having a diameter of 2.75 mm and a mass of 0.10 g

- a large growth with a diameter of 3.75 mm and a mass of 0.34 g.

Results

An example of results is presented in figure 1.
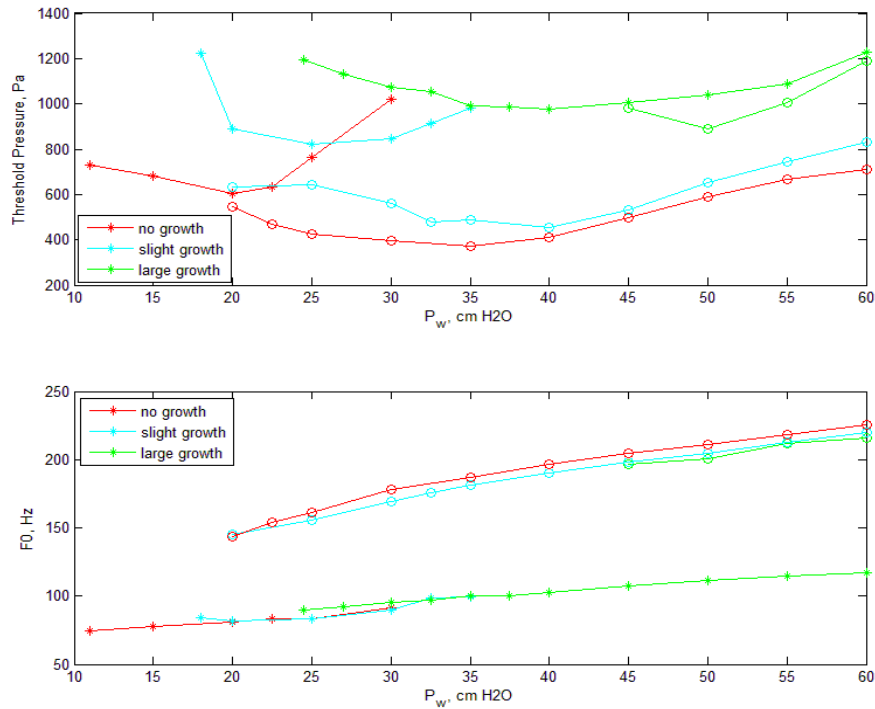
Figure 1 : Measured Threshold pressure and Fundamental frequency of vibration as a function of the elasticity of the replica controled by the pressure of water, P_w, inside the replica.

The reference case, in the absence of any growth, shows that, depending on its elasticity, the replica presents two oscillation regimes which correspond to two mechanical resonance frequencies of the vocal folds replica. As one could have expected, the presence of a growth on the vocal folds replica significantly alters the threshold pressure which increases with the mass and the diameter of the growth. The fundamental frequency of oscillation appears less affected by these parameters. However, it is observed that, as the size and mass of the growth increases, the higher frequency regime becomes more and more difficult to reach.

Conclusions

In this paper we presented an original set-up that allows to simulate and to measure quantitatively the effect of the presence of a growth at the surface of the vocal folds during phonation. The first results compare well with the literature on voice pathology. The presence of growths induces indeed an increase of the pressure thresholds, a reduction of the sound intensity and of the frequency range.

In the future, the data collected will be compared with the theoretical models of increasing complexity.

Hysteresis and Relaxation of Vocal Fold Tissue and the Difference between Phonation Onset and Offset

Lewis Fulcher[(1),] Ronald C. Scherer[(1)]

[(1)] Bowling Green State University

Fulcher@bgsu.edu

Introduction

Alipour and Vigmostad1 recently carried out an extensive set of ergometer measurements with human vocal fold tissue. In particular, they found the force necessary to elongate a tissue sample was larger than the force the tissue exerts when it returns to its original length, that is, a hysteresis effect, and they obtained numerical results for the area of the hysteresis loop. Continuing this cycle of stretching and releasing, they found that the magnitude of the forces generated during the cycle decreased and that the areas of the hysteresis loops became smaller during most of the first 10 cycles, a behavior expected for tissue relaxation or stress-softening. After 10 cycles the hysteresis cycle became reproducible, and the tissue was then described as preconditioned. A simulation of the results of Alipour and Vigmostad is shown in Figure 1.

Titze's formulation2 of the Surface Wave Model provides a framework for incorporating these effects.

Methods

The hysteresis effect is included by choosing larger stiffness parameters when the vocal fold is moving away from its equilibrium position than when it is returning to its equilibrium position. Choosing the stiffness parameters during the return to equilibrium about half those for the displacement from equilibrium gives an area for the hysteresis loop consistent with that measured by Alipour and Vigmostad during their first series of stretching and releasing measurements.

Results

Preliminary calculations for the phonation onset-offset experiments done by Chan and Titze in 2006 are shown below in Figure 2. These researchers made an extensive set of measurements with a physical model of the vocal fold mucosa, which consisted of biomaterials, such as hyaluronic acid (HA) and fibronectin, implanted under a thin silicone membrane. Fits to the onset data require a small damping parameter B in addition to the hysteresis losses described above, as well as careful attention to the behavior of the entrance loss coefficient4. In order to fit the offset pressure data, it is assumed that the number of oscillation cycles has been sufficient so that the model has reached its preconditioned stage, where the hysteresis cycle has become reproducible. In accord with the Alipour-Vigmostad

measurements, it is assumed that the area of the hysteresis loop during this preconditioned stage is smaller than that during onset, and thus less energy is lost to hysteresis.

Such an assumption provides a natural explanation to lower offset pressures than onset pressures with the change of a single parameter to describe the different areas of the hysteresis loops. Results are presented for all of the data sets collected by Chan and Titze as well as some earlier measurements.
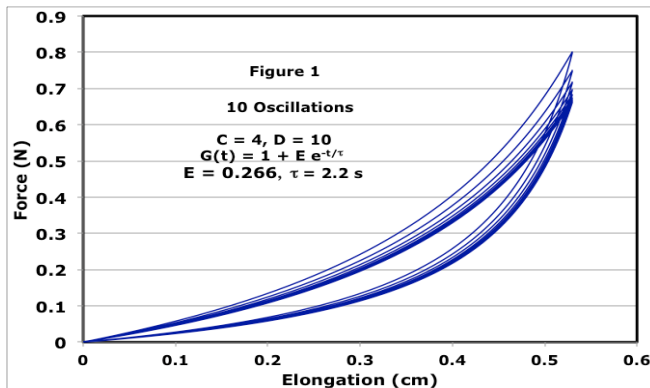
Figure 1: Simulation of the results of Alipour and Vigmostad
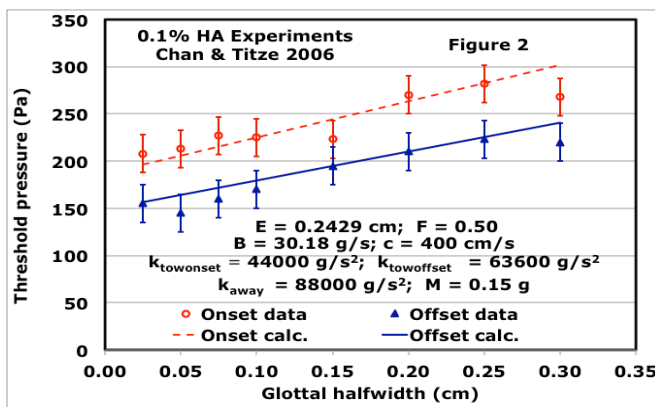
Figure 2: Experiments done by Chan and Titze

Reference

[1]F. Alipour and S. Vigmostad, "Measurement of Vocal Folds Elastic Properties in Continuum Modeling," Journal of Voice **26**, 816.e21-e29 (2012).

[2]I. Titze, "The physics of small amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**, 1536-1552 (1988).

[3]R. Chan and I. Titze, "Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics," J. Acoust. Soc. Am. **119**, 2351-2362 (2006).

[4]L. Fulcher and R. Scherer, "Phonation threshold pressure: Comparison of calculations and measurements taken with physical models of the vocal fold mucosa," J. Acoust. Soc. Am. **130**, 1597-1605 (2011).

An Empirical Equation for Posterior Glottal Flow

Ronald Scherer[1], Fariborz Alipour[2]

[1]Bowling Green State University, [2]University of Iowa

ronalds@bgsu.edu

Introduction

Glottal flow can have two parts, a cyclic modulation of flow due to vocal fold oscillation that varies the glottal area, and a more steady flow due to glottal area that changes slowly. A primary location for the latter is the posterior glottis. This project examined the posterior glottal flow using two excised canine larynges.

Methods

The anterior glottis was sutured closed at the vocal processes and a small weight was placed over the anterior glottis so that the vocal folds were prevented from vibrating. The area of the posterior glottis was varied from approximately 10 to 27 mm^2 using arytenoid stabilizing threads. Video recordings (from above) and custom software were used to measure the areas. After setting the area, subglottal pressure was increased in 2 cm H2O steps (read from a wall manometer, with a range of approximately 2 to 30 cm H2O) and the dependent flow was obtained (measured using a vertical rotameter). Pressure-flow results for each nominal glottal area were fit with empirical equations and combined across the areas to obtain general pressure-flow-area (posterior glottal area) equations.

Results

An empirical quadratic equation for Larynx 1 (6 coefficients) fit the flow data (as a function of posterior glottal area and transglottal flow) with an absolute average discrepancy of 7.6% (SD = 6.4%), and an empirical linear equation for Larynx 2 (5 coefficients) fit the flow data with an absolute average discrepancy of 7.2% (SD = 8.9%). The two equations made relatively similar predictions for flow given posterior glottal area and transglottal pressure (0.4% non-absolute average difference, SD = 18.9%, and 14.9% absolute average difference, SD = 11.5%). Given the relative similarity between the two larynges, Larynx 2 data were used to create a General Posterior Glottal Flow Empirical Equation to give glottal flow as a function of posterior glottal area and transglottal pressure by using the quadratic equation (solving for flow) and only 6 coefficients. The absolute average difference between prediction and real flow was 9.0%, SD = 11.7%. The equation also includes projected pressure-flow-area estimates for posterior glottal areas less than those studied with Larynx 1 and 2. Thus, the General Posterior Glottal Flow Empirical Equation matches the empirical data well and also provides posterior glottal flow that goes to zero as transglottal pressure or posterior glottal area goes to zero.

The general equation is:

Flow = (-b + [b^2 + 4*a*P]^0.5/(2*a),

where

a = c1*A + c2

b = c3*A + c4 for A >= 10 mm^2

b = d1*A^d2 for A <= 10 mm^2

c1 = 0.000000202

c2 = 0.000005884

c3 = -0.001064

c4 = 0.02945

d1 = 0.1037

d2 = -0.824

where P is in cm H2O, A in mm^2, and Flow in cm^3/s

Conclusions

The general equation can be used in simulations requiring estimates of the portion of glottal flow that passes through the posterior glottis (or through slowly changing glottal area, including a portion of the anterior glottis for breathy phonation where there is lack of full glottal closure).

Figures showing the pressure-flow relationships generated from the general equation:
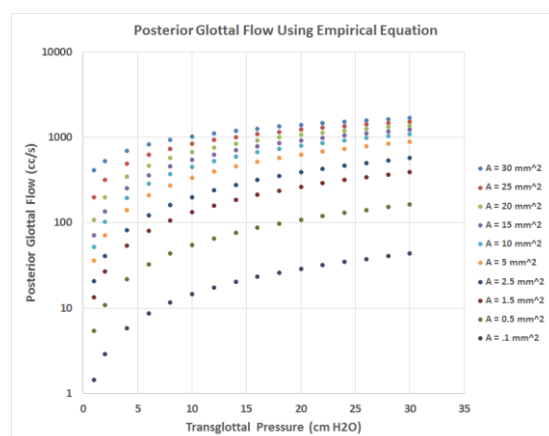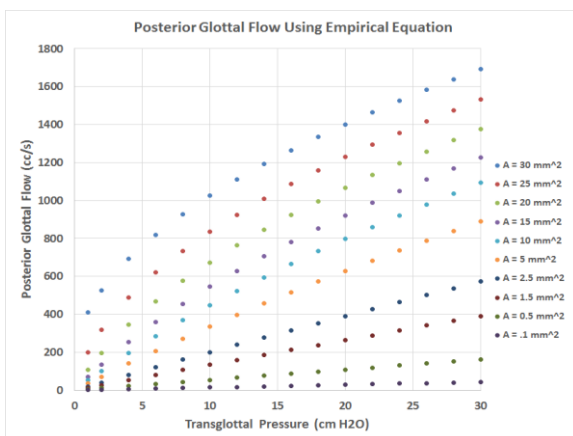
Comparison of glottal flow rate predicted by inverse filtering and direct measurements

Liran Oren[1], Sid Khosla[1], Alexandra Maddox[2], Ephraim Gutmark[2]

[1]Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati

[2]Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati

liran.oren@uc.edu

Introduction

The volumetric flow rate produced at the glottal exit during the phonation cycle is described in terms of a volume of air traveling through the glottal exit per a given unit of time, such as cm^3/sec. The rapid rate of flow reduction, which usually occurs in normal phonation during the closing of the vocal folds, is measured by quantity known as the maximum flow declination rate (MFDR). Clinically, MFDR is derived from the flow rate, or volume velocity waveform, which is most commonly measured by a technique known as inverse filtering[1]. In this technique the volume velocity is first directly measured at the mouth, and inverse filtering is then used to derive the flow rate at the glottal exit. Inverse filtering assumes linear source filter interactions, an assumption that has been questioned[2]. Nonetheless, this technique has been used in multiple clinical studies showing general agreement with theoretical models. Because there are no presently published techniques to directly and accurately measure flow at the glottal exit, inverse filtering has not yet been validated in human or canine larynges, and in particular for either low or high vocal tract inertance.

The purpose of this study is to assess the effectiveness of inverse filtering and of Rothenberg mask by comparing the predicted volume velocity waveform with direct measurements of the flow rate at the glottal exit.

Methods

Volumetric flow rate is measured directly at the glottal exit using volumetric particle imaging velocimetry (tomo-PIV), and indirectly at the mouth using inverse filtering, in an excised canine larynx model. The larynx is attached to a mechanical model of the vocal tract with false vocal folds (FVF) and a Rothenberg mask oval at the exit (Figure 1). The vocal tract inertance is varied by changing the gap between the FVF.

Results

The results show that as the vocal tract inertance increase, the differences in the flow parameters between the direct and indirect measurement of flow rate also increase.

As previously discussed, inverse filtering using a Rothenberg mask assumes linear source-tract interactions. The results are compared with other theoretical and human studies that highly suggest that linear source-tract interactions are not correct at higher vocal tract inertance.
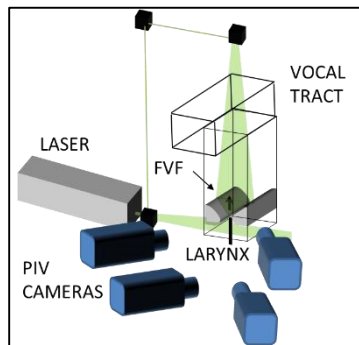
Figure 1 – Experimental setup

Conclusion

The accuracy of the Rothenberg parameters decreases as inertance increases since the linear source filter assumption is less valid at higher inertance.

References

[1] Rothenberg M. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. J Acoust Soc Am, 1973. 53: 1632-1645

[2] Titze IR. Nonlinear source-filter coupling in phonation: theory. J Acoust Soc Am, 2008. 123:2733-2749

The influence of sound emission on the *lamina propria* of the ventricular fold

André Armani[1], Maria Célia Jamur[2], Rui Celso Martins Mamede[1]

[1]Department of Otorhinolaryngology and Head and Neck Surgery, Ribeirão Preto Medical School, University of São Paulo. [2]Department of Cell and Molecular Biology and Pathogenic Bioagents, Ribeirão Preto Medical School, University of São Paulo

andrearmani@yahoo.com

Introduction

The vocal folds (VF) are unique structures, highly specialized in vibrating for sound production. This specialization is mainly due to a layered structure of the *lamina propria* (LP). This layered structure is not present at birth, and develops only after a several years of phonation. The LP of the mature vocal fold consists of three layers. The ventricular folds (VTF) are not originally vibrating structures for sound production, and its LP the layers are poorly organized. It is not known what happens to the constituents of the LP in the VTF in subjects that use VTF vibration as a source of voice production.

Methods

In the present study, the distribution and quantity of collagen and elastic fibers of the *lamina propria* from VTF of patients that use it as the main source of vibration for voice production were compared with the VTF from control subjects. Six individuals which used at least one of the VTF as source of vibration for sound production for minimum of six years, were selected. A small fragment of VTF (0.5cm^2) used as vibration source of sound production was collected from each subject. The samples were processed for histological analysis. Collagen fibers were stained with Picrosirus Red and elastic fibers were stained with *Weigert's* Resorcin-Fuchsin. A total of 54 images were obtained from the superficial layer of the LP from each VTF for each stain. After image acquisition, collagen type I, III, total collagen and elastic fibers were quantified and compared with the VTF from the control group. Quantification was done using ImagePro Plus software. Statistics were performed using an unpaired T test.

Results

The amount of total collagen in the most superficial layer of LP when the VTF was used as the source of vibration for the production of sound was significantly higher when compared to controls. The same result was seen for the amount of type I collagen in both groups. There was no difference in the quantity of type III collagen and elastic fibers between the two groups.

Conclusions

Vibration of the VTF as a source of sound, for at least six years, leads to an increase in the amount of total collagen fibers and an increase in type I collagen, but does not increase the amount of type III collagen and elastic fibers in the most superficial layer of LP. These results may help elucidate the unique development of the *lamina propria* of the vocal fold.
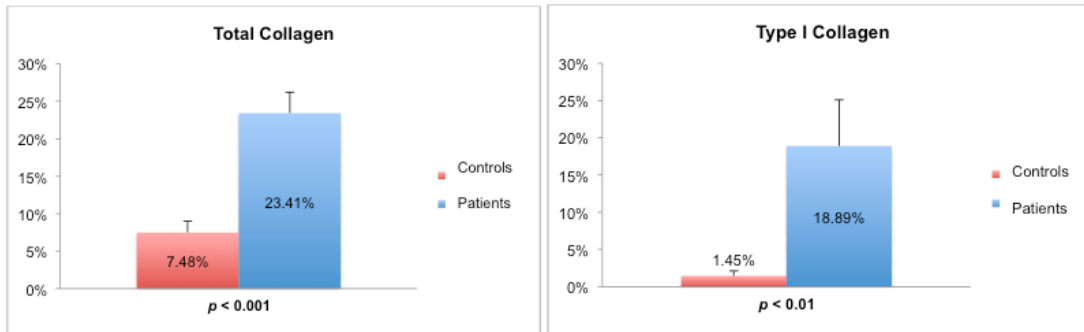

Figure 1 - Mean ± SD of total collagen and type I collagen. A 300μm^2 total area from control and patients groups was analyzed

Computer controlled set-up for automated phonation excitation in excised larynx experiments

V. Birk[1], S. Kniesburges[1], C. Bohr[1], M. Döllinger[1]

[1]University Hospital Erlangen, Department of Otorhinolaryngology, Head & Neck Surgery, Department of Phoniatrics and Pediatric Audiology

Veronika.birk@uk-erlangen.de

Introduction

Excised larynx experiments are limited in time due to tissue degeneration. Furthermore, the variation of parameters like adduction level or pretension of the vocal folds is not standardized and requires experience by the experimenter. Therefore, an optimization of the experimental procedure is desirable, first to increase the data amount per larynx and second to electromechanically control the experimental parameters.

In order to achieve the posture of the larynx for sustained phonation, manual devices are usually used in experiments, like syringes or screws for a static and manual control of the cartilage position [1]. Parameters like vocal fold pretension and adduction were selected by putting weights on the syringes which were fixed at the force application point. The procedure for experiment preparation and execution is time-consuming and requires long-time experience by the experimenter, whereas there is still a high risk of variability and errors.

Methods

We introduce a customized set-up for a combined computer controlled regulation of the arytenoid and thyroid cartilages [2, 3, 4, 5]. Figure 1 shows the developed mechanical set-ups including the sensors for force measurement and the motors for force application. Thereby it is possible to adjust the pre-tension, the adduction and the asymmetry of the vocal folds. Using electro-mechanical devices, controlled via LabView, a defined induction of force and torque on the cartilages is enabled. This allows for fast and quantified parameter control within the experiment.

Furthermore, the tracheal airflow is controlled on basis of the subglottal pressure as feedback parameter. This enables for driving the tracheal airflow by specific temporal functions.

The data analysis includes the time-resolved subglottal and acoustic signals which are captured synchronously by a National Instruments PXI-1073 system as well as a high-speed video recording of the vocal folds.

Results

The set-up was validated using excised porcine larynges. Different laryngeal asymmetry as well as vocal fold adduction and pre-tension adjustments were executed. The functionality of the set-up as well as preliminary measurement results will be presented and discussed with regard to the expenditure of time of the experiments.
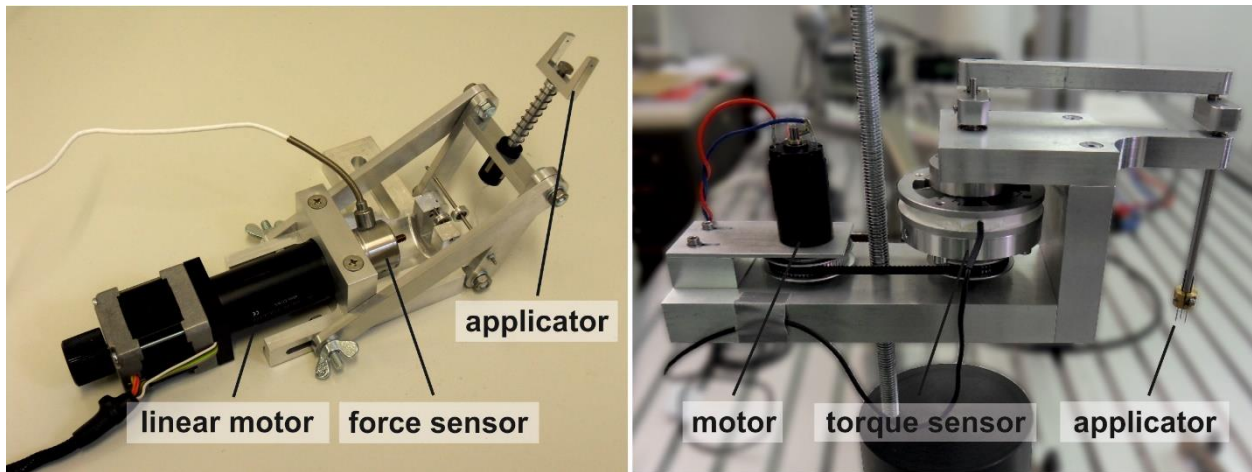
Figure 4: Developed mechanical tools for thyroid (left) and arytenoid (right) cartilage posturing.

Conclusions

The presented set-up allows for a reduction of the experimental performing duration. Due to the automated parameter control a fast experimental execution is achieved and time consuming larynx preparation like suture positioning is no longer required.

Moreover the possibility of a quantitative parameter variation during the experiment enables a standardized experimental procedure within a wide range of larynx configurations.

References

[1] Döllinger, M., Kobler, J., Berry D. A., Mehta, D. D., Luegmair, G., & Bohr, C. (2011). Experiments on Analysing Voice Production: Excised (Human, Animal) and In Vivo (Animal) Approaches. Current Bioinformatics, 6, 286-304.

[2] Luegmair, G. (2014). 3D Reconstruction of Vocal Fold Surface Dynamics in Functional Dysphonia. PhD thesis.

[3] Gedeon, D. T. (2011). Optimierung eines Messplatzes zur Untersuchung der Stimmlippendynamik während der Phonation. Diploma thesis.

[4] Schild, S. (2013). Stimmlippenelongation durch Artikulation des Schildknorpels (Konzept). Bachelor thesis.

[5] Häßler, S. (2014). Stimmlippenelongation durch Artikulation des Schildknorpels (Technische Umsetzung). Bachelor thesis.

Dynamic vocal fold imaging with combined optical coherence tomography/high-speed video endoscopy

Nicusor Iftimia[1], Gopi Maguluri[1], Ernest Chang[1], Jesung Ppark[1], James Kobler[2] and Daryush Mehta[2]

[1]Physical Sciences Inc., [2] Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital

iftimia@psicorp.com

Introduction

Voice disorders due to trauma (e.g., intubation injuries, vocal abuse) and disease (e.g., dysplasia, cancer, recurrent respiratory papillomatosis, nodules, polyps, and scar) are currently evaluated in otolaryngology clinics using endoscopic imaging techniques such as videostroboscopy [1] or high-speed videoendoscopy (HSV) [2]. Clinicians couple these visual observations of vocal fold tissue motion with auditory-perceptual judgments of voice quality as part of a comprehensive assessment of the health and function of the larynx during phonation [3]. Endoscopic imaging, however, only provides two-dimensional (2D) spatial information and thus only quantifies the lateral tissue motion, critically lacking vertical axis information. Various high-speed imaging techniques have attempted to capture vocal fold surface motion during phonation in three spatial dimensions [4–6]; however, they either lack adequate spatial resolution or have not been validated in vivo. In this paper, we present a dual modality imaging approach, where optical coherence tomography (OCT) imaging augments HSV by providing the missing depth information. OCT quantitatively measures the vertical motion of the vocal fold surface during phonation with micron scale resolution. Furthermore, it also provides the subsurface structural morphology to depth to at least 1.5 mm. Therefore, the combination of these two modalities within the same instrument seems to be a suitable approach for examining the pathology of the vocal folds.

Methods

A common optical imaging path swept source OCT/high speed video (SSOCT/HSV) endoscopy instrument was developed and used on an ex vivo study on excised animal tissue specimens. A simplified schematic of the instrument is shown in Figure 1.
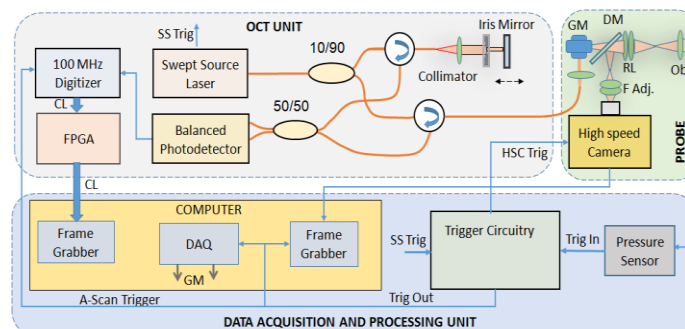

Figure 1. Simplified schematic of the SSOCT/HSV instrument

The OCT component of the system uses a swept source (SS) approach. The light source (Santec) operates at a scan rate of 20 kHz, while providing a broad spectrum light with a 3dB bandwidth of 100 nm at a center wavelength of 1310 nm. This enables subsurface tissue imaging with an axial resolution better than 10 □m. A fiber optic interferometer and a balance detector are used to generate interference fringes, by combining the retro-reflected light from the imaged sample with that from a mirror, placed in the reference arm of the interferometer. A constructive interference occurs when the path-length difference between the two arms of the interferometer is within the coherence range of the light source. Each wavelength sweep of the light source is thus used to generate what is called an OCT A-scan. An individual A-scan is thus used to generate a sample reflectivity profile. The A-scan signals from the balanced detector are fed to a custom built FPGA module that acquires data when receives A-scan triggers from a custom-designed trigger circuit. These triggers are sent only when a signal from the vocal fold pressure sensor are detected, such that each full movement cycle of the vocal folds, also called phonation cycle, is digitized on a reasonable number of points (N>10).

The same triggers are fed to the HSV camera, enabling a perfect temporal correlation of the OCT and HSV data. The digitized signals are sent through a camera link interface to a frame grabber (NI 1430), and then the acquired data are transferred to a graphic processing unit, which enables real time processing and display the OCT images. The HSV component of the system constitutes a Phantom v7.3 color camera (Vision Research Inc., Wayne, NJ), typically set to record at 5000 frames per second at a spatial resolution of 128 x 128 pixels. A photograph of the experimental setup for imaging airflow-driven phonation of an ex vivo calf larynx is shown in Figure 2. Special attention has been paid to the temporal synchronization of OCT and HSV data during a phonation cycle. Therefore, we used a pressure sensor which detects the motion of the vocal folds, such the OCT A-scans and HSV images are synchronously acquired over the entire period of a single vibratory cycle. For a typical phonating frequency of 100 to 200 Hz, at least 10 OCT scans are acquired per each HSV frame. The 3D OCT data set provides the necessary lateral and depth information needed to correlate the HSV frames and augument them with the depth info, such that a 3D video image is reconstructed.

Results

Measurements on vibrating calf vocal folds were performed combining OCT and HSV data. Both static anatomical images and dynamic images (10 phases during a phonation cycles) were recorded. Figure 3 shows the 3D reconstruction of the vocal fold surface at rest with an axial area of 10 mm x 10 mm and imaging depth of 3.5 mm, while Figure 4 shows the vertical position of the vocal fold surface during a phonation cycle. The depth information provided by OCT was used to reconstruct the HSV data in a 3D format. A representative 3D HSV frame is shown in Fig. 5. The 3D coordinates are used to provide a 3D surface contour for the HSV frames to enable 4D visualization of vocal fold kinematics.
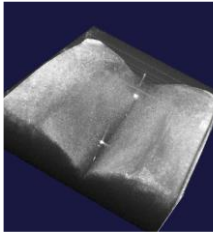
Fig. 3. 3D reconstruction of vocal fold integrating OCT and HSV images while the larynx is at rest.

Open, 0 ms   1.8ms   Closed, 3.6 ms

HSV   OCT
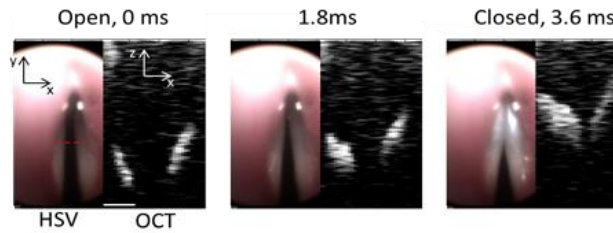
Fig. 4 Preliminary data displaying co-registered images acquired using swept-source OCT and HSV imaging modalities. The OCT beam scan path is indicated by a red dashed line. Scale bar: 1 mm.
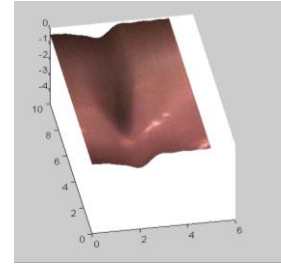
Fig.5. 3D HSV frame reconstructed based on OCT depth information

Conclusions

In conclusion, we have preliminarily tested a novel imaging modality combining optical coherence tomography (OCT) and high-speed videoendoscopy (HSV) to image laryngeal motion at high speeds and retrieve the 4D data showing the 3D movement of the vocal folds in time. This technology has the potential to enable a more in depth analysis of irregularities in vocal fold vibration. If successful, this technology will provide the otolaryngologists in the future with a novel tool for more reliable assessment of the vocal folds pathology and function.

References

[1] D. D. Mehta and R.E. Hillman. "Current role of stroboscopy in laryngeal imaging," Curr Opin Otolaryngol Head Neck Surg, **20**, 429–436 (2012).

[2] D. D. Deliyski, P. P. Petrushev, H. S. Bonilha, T. T. Gerlach, B. Martin-Harris, and R. E. Hillman, "Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and evolution," Folia Phoniatr Logop, **60**, 33–44 (2008).

[3] S. M. Zeitels, A. Blitzer, R. E. Hillman, and R. R. Anderson, "Foresight in laryngology and laryngeal surgery: A 2020 vision," Ann Otol Rhinol Laryngol, **116 Suppl 198**, 1–16 (2007).

[4] C.G. Tsai, J.H. Chen, Y.W. Shau, T.Y. Hsiao, "Dynamic B-mode ultrasound imaging of vocal fold vibration during phonation," Ultrasound Med Biol, **35**, 1812–1818 (2009).

[5] G. Luegmair, S. Kniesburges, M. Zimmermann, A. Sutor, U. Eysholdt, M. Doellinger, "Optical reconstruction of high-speed surface dynamics in an uncontrollable environment," IEEE T Med Imaging, **29**, 1979–1991 (2010).

[6] D. E. Sommer, I.T. Tokuda, S.D. Peterson, K. Sakakibara, H. Imagawa, A. Yamauchi, T. Nito, T. Yamasoba, and N. Tayama, "Estimation of inferior-superior vocal fold kinematics from high-speed stereo endoscopic data in vivo," J Acoust Soc Am, **136**, 3290–3300 (2014).

Potential of endoscopic high speed imaging – current projects

Michael Döllinger[1], Stefan Kniesburges[1], Christopher Bohr[1]

[1] Department of Phoniatrics and Pediatric Audiology at the Department of Otorhinolaryngology Head & Neck Surgery, University Hospital Erlangen

Michael.doellinger@uk-erlangen.de

Introduction

Our department is involved in the development in high-speed endoscopy (HSE) and its analysis since 1995. Since then, the hardware, imaging and data analysis techniques have made huge progress. However, the clinical acceptance of HSE is still rather little. We will present current HSE projects of our department and give insight in the efforts behind.

Methods

For HSE we use different kind of cameras: clinical high-speed cameras (KayPENTAX, Wolf) and industrial cameras (Photron SA1.1, Phantom v2511) in combination with rigid and flexible endoscopes. For larger studies we use clinical cameras that are more managable with frequency rates of 4000 fps (256 x 512 pixel). In these cases, we mainly investigate and objectively quantify the phonatory process: Phonation onset, frequency changes, sustained phonation, comparison between controls and pathologies, pre- and post treatment, potential of analysing 3D vocal fold dynamics.

We apply industrial cameras (up to 1280 x 800 pixel at 25.6 kHz) to investigate more basic aspects as: behaviour of objective HSE parameters due to frequency changes and pitch shift reflex in combination with EEG analysis.

For analyzing the large amount of produced data by HSE, software tools for image processing (Figure 1) and signal analysis are developed. These software tools always have the purpose of potential future clinical application.

Results

Objective parameters were located for differentiating between pathological voices and controls. HSE parameters were correlated to acoustics and EGG. The mucosal wave was traced and fully automatic glottis segmentation was achieved. However, developing image processing and analysis software tools is highly time-consuming.

Conclusions

HSE is useful and enables a detailed analysis of the phonatory process. During the last years, HSE added much knowledge on the phonatory process. Plenty of work has been published during the last decade (Deliyski & Hillman, 2010; Döllinger, 2009; Ziethe et al., 2011). To date many research groups but rather little clinicians apply HSE. Standards on parameter evaluation are missing. Additionally the high costs, complex analysis and the high technical effort (developing imaging and analysis tools) to be able to gain more benefit from HSE than from stroboscopy might be reasons, why HSE is still not accepted and commonly applied in clinics (Woo, 2014). To widely introduce HSE in daily clinical routine, international combined effort should be undertaken to bring HSE from the status of a "nice tool to have" to the level of "must have".
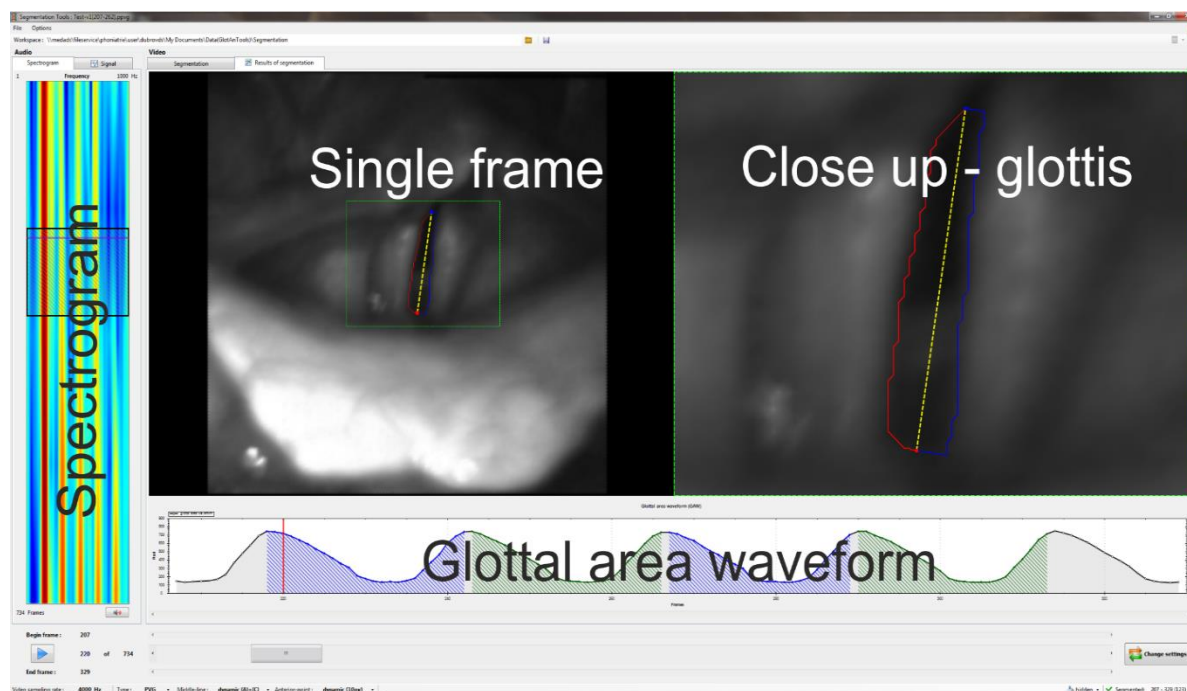
Figure 1: Screenshot of the developed image and signal processing software "Glottis Analysis Tool" (GAT). We started in 2009 and it is contiously enhanced and extended.

References

Deliyski, D.D., & Hillman, R.E. (2010). State of the art laryngeal imaging: Research and clinical implications. Current Opinion in Otolaryngology and Head and Neck Surgery, 18(3), 147-152.

Döllinger, M. (2009). The next Step in voice assessment: High-Speed digital endoscopy and objective evaluation. Curr Bioinform, 4(2), 101-111.

Woo, P. (2014) Objective measures of laryngeal imaging: What have we Learned since Dr. paul moore. J Voice, 28(1), 69-81.

Ziethe, A., Patel, R., Kunduk, M., Eysholdt, U., & Graf, S. (2011). Clinical Analysis Methods of Voice Disorders. Curr Bioinform, 6(3), 270-285.

Analysis of spatial characteristics of the larynx using high-speed digital imaging

Ken-Ichi Sakakaibara [1], Hiroshi Imagawa[2,3], Isao T. Tokuda[4],

AkihitoYamauchi[2,3], Hisayuki Yokonishi[3,5], and Niro Tayama[2]

[1]Health Sciences University of Hokkaido, [2] National Center for Global Health and Welfare,

[3]The University of Tokyo, [4]Ritsumeikan University, [5]Tokyo Metropolitan Hospital

kis@hoku-iryo-u.ac.jp

Introduction

In this presentation, the authors review studies on spatial characteristics of vocal fold vibration extracted from high-speed digital imaging and introduced our recent results using laryngotopography and stereographic high-speed imaging.

Methods and Results

High-speed digital imaging technique applied to observe the larynx was first proposed and implemented in RILP, Univ. Tokyo (Honda et al, 1985a). The first attempt using high-speed digital imaging is now referred to as digital kymography which is a time-varying function of brightness on one-dimensional line. The method of digital high-speed imaging was soon developed to on two-dimensional image (Honda et al, 1985b), and since then, many various studies on measurement time-varied characteristics, such as glottal width and glottal area, have been reported. However, for long years, it is difficult to say that rich information which is inherent in brightness function on time-varying two-dimensional images, i.e. on four-dimensional images has been sufficiently used. One breakthrough on how to use spatially rich information of high-speed images was a laryngotopographic technique (Granqvist et al., 2001). The authors have developed this new view technique and applied to modal analysis of the vocal fold vibration as mucosal propagation (Sakakibara at al., 2010; Kimura et al. , 2010; Yamauchi at al., 2013).

Another direction to analyze spatial characteristics of vocal fold vibrations has been developed as stereographic imaging in RILP, University of Tokyo. Stereo fiber scope was first proposed and next stereo endoscope was proposed for analysis of still image or normal video images (Sawashima et al., 1974; Honda et al. 1980). Due to improvement of spatial resolution of CCD image sensor, stereo graphic endoscopy could be also applied to high-speed images (Imagawa et al., 2010). Using stereo graphic high-speed imaging, vertical motion in vocal fold vibrations can be observed *in vivo*, and measured in real valued (Sommer et al., 2014).

Conclusions

Not only time-varied scolar functions, such as glottal width and glottal area, also spatial distribution of vibratory characteristcs, such as frequencies, phases, and amplitude-like properties of different vibratory modes are now extractable from high-speed digital images which essentially carry rich spatial information using relatively new method laryngotopography. Not ony two-dimensional characteristics of vocal fold vibrations projected on two-dimensional images, also three-dimensional images realized by using stereo-graphic endoscopy can make it possible to analyze virtical movement of the vocal folds. In new future, improvement of instruments and analysis method for graphical methods of analysis of vocal fold vibrations will increase the accuracy of observation and make a breakthrough contributed better understanding of human vocalization.

References

K. Honda, S. Kiritani, H. Imagawa, & H. Hirose. (1985a). High-speed digital image recording system for studying vocal fold vibration, *Proc. Annual meeting of IEICE*, p.127, in Japanese.

K. Honda, S. Kiritani, & H. Imagawa. (1985a). High-speed digital recording of vocal fold vibration using a solid-state image sensor, *Ann. Bull. RILP*, 19:47—53. http://www.umin.ac.jp/memorial/rilp-tokyo/

S. Granqvist & P.- Å. Lindestad. (2001). Fourier analysis to high-speed laryngoscopy, *J. Acoust. Soc. Am.*, 110(6):3193—3197.

K.-I. Sakakibara, H. Imagawa, M. Kimura, H. Yokonishi, & N. Tayama. (2010). Modal analysis of vocal fold vibrations using laryngotopography, *Proc. Interspeec*h, pp.917—920.

M. Kimura, T. T.Nito, H. Imagwa, K.-I. Sakakibara, R. Chan, & N. Tayama. (2010). Correction of vocal fold asymmetry by arytenoid adduction: high-speed imaging with videokymographic and laryngotopographic analysis, *Ann. Otol., Rhinol. Laryngol.* , 119(7):439—446.

A. Yamauchi, H. Imagawa, K.-I. Sakakibara, H. Yokonishi, T. Nito, T. Yamasoba, & N. Tayama. (2013). Phase difference of vocally healthy subjects in high-speed digital imaging analyzed by laryngotopography, J. Voice, 27(1):39—45.

M. Sawashima & S. Miyazaki. (1974). Stereo-fiberscopic measurement of the larynx: a preliminary experiment by use of ordinary laryngeal fiber-scopes, Ann. Bull. RILP, 8:7—11. http://www.umin.ac.jp/memorial/rilp-tokyo/

K. Honda, S. R. Hibi, S.Kiritani, S.Niimi & H. Hirose. (1980). Measurement of the laryngeal structures during phonation by use of a stereoendoscope, 14:73—78. http://www.umin.ac.jp/memorial/rilp-tokyo/

H. Imagawa, K.-I. Sakakibara, I. T. Tokuda, M. Otsuka, & N. Tayama. (2010). Estimation of glottal area function using stereo-endoscopic high-speed digital imaging. *Proc. Interspeec*h, pp.1005—1008.

D.E.Sommer, I.T.Tokuda, S.D.Peterson, K.-I. Sakakibara, H. Imagawa, A.Yamauchi, T.Nito, T.Yamazoba, & N.Tayama. (2014). Estimation of inferior-superior vocal fold kinematics from high-speed stereo endoscopic data in vivo. *J. Acoust. Soc. Am.*, 136(6):3290—3300.

Cross Sectional Imaging of Phonating Human Vocal Fold in Vivo Using VCSEL Optical Coherence Tomography

Brian J.F. Wong, MD, PhD[1,2,3], Carolyn Coughlan, MD[1,2], Li-dek Chou, MS[2,3],

Joe Jing, MS[2,3], Jason J. Chen, BS[2], Swathi Rangarajan, MS[4], Theodore H. Chang, BS[4], Christian Barnes, MD[1], Julie Goddard, MD[1], Zhongping Chen, PhD[2,3]

[1]Department of Otolaryngology - Head and Neck Surgery, University of California, Irvine, School of Medicine,[2]Beckman Laser Institute & Medical Clinic, University of California, Irvine,[3]Department of Biomedical Engineering, University of California, Irvine,[4]OCT Medical Imaging Inc., Newport Beach, CA

bjwong@uci.edu

Introduction

Vocal fold imaging has evolved over decades in step with the development of endoscopic and digital imaging technology. However, the management of suspicious vocal fold lesions has remained the same as contemporary practice still requires precise microsurgical biopsy performed in the operating room under general anesthesia. Office-based imaging and examination informs the decision to perform these operations, though current in-office technology is limited to only visualizing the surface of the vocal folds (via fiber-optic or rigid endoscopy) and inferring information about the submucosal processes (via stroboscopic or high-speed video). Over the past fifteen years, cross-sectional imaging of vocal fold *in vivo* using optical coherence tomography (OCT) has developed[1-5]. While OCT has demonstrated high-resolution capability during operative endoscopy, translation to an office-based platform has being challenging due to short imaging range, cord and examiner motion, all leading to poor image quality despite 10 um resolution. These are due to the intrinsic limitations of both conventional time and spectral domain OCT systems. The development of vertical-cavity surface-emitting laser (VCSEL) sources for OCT provides the solution to the small working distances and imaging ranges of conventional OCT systems. Here we present our experience with vocal fold imaging.

Methods

Full-field (panoramic), high-speed, and high-resolution OCT images were obtained in subjects in the office using only topical anesthetic. The OCT system uses a 1310 nm center wavelength 200 kHz VCSEL swept source, and the output light was split by a 99:1 ratio fiber-optic coupler into the sample and reference arms, respectively. A 635 nm aiming laser beam is combined with the 1310 nm sample beam to allow visualization of the targeting area. The axial resolution is 9.3 μm in tissue, and the lateral resolution is 100 μm at 10 cm. The images have a fixed lateral frame size of 1000 A-lines and are captured at 200 Hz. The working distance is 60 mm and the imaging range is approximately 1.4 mm. The probe of the OCT is paired with a 90-degree laryngoscope (Fig. 1). OCT imaging was performed in

tandem using a DV laryngeal endoscopy system. Cross-sectional OCT images were obtained, and off-line Doppler analysis of vocal fold motion was performed.

Results

High-resolution full-field images of the vocal folds during phonation were recorded using the described device. Wave propagations in both true and false vocal folds were visualized (Fig. 2).

Conclusions

VCSEL OCT offers depth-resolved, full-field cross-sectional images of both true and false vocal folds, and provides high-resolution and functional data in both static and dynamic imaging. Combined with Doppler, VCSEL OCT has the potential to characterized lesions of the vocal folds, which will revolutionize in-office imaging of the larynx.
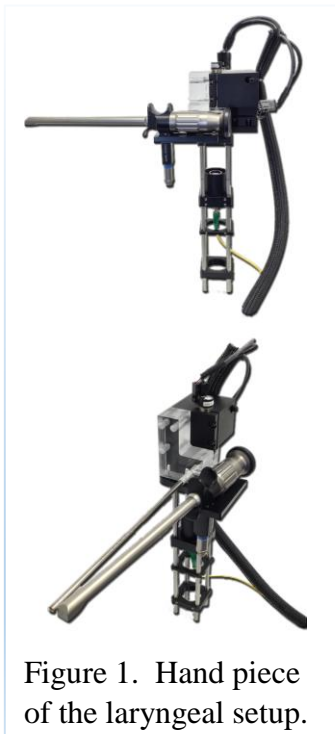
Figure 1. Hand piece of the laryngeal setup.
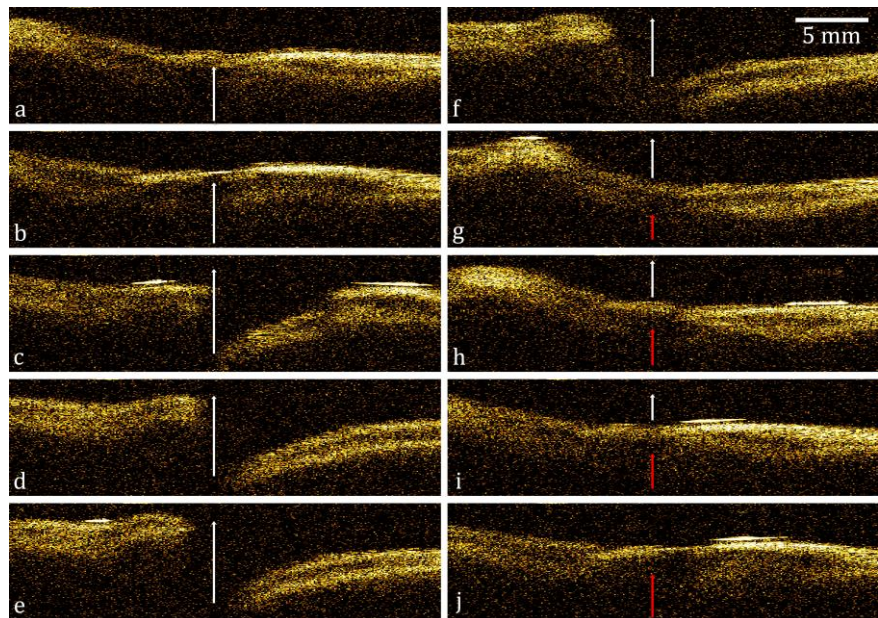
Figure 2. OCT Image Montage of the true vocal colds in a healthy adult volunteer who was phonating at approximately 250 Hz.

References

1. Wong, B. J. et al. In vivo optical coherence tomography of the human larynx: normative and benign pathology in 82 patients. Laryngoscope 115, 1904-1911, doi:10.1097/01.MLG.0000181465.17744.BE (2005).
2. Guo, S. et al. Office-based optical coherence tomographic imaging of human vocal cords. J Biomed Opt 11, 30501, doi:10.1117/1.2200371 (2006).
3. Guo, S. et al. Gradient-index lens rod based probe for office-based optical coherence tomography of the human larynx. J Biomed Opt 14, 014017, doi:10.1117/1.3076198 (2009).
4. Yu, L. et al. Office-based dynamic imaging of vocal cords in awake patients with swept-source optical coherence tomography. J Biomed Opt 14, 064020, doi:10.1117/1.3268442 (2009).
5. Jing, J., Zhang, J., Loy, A. C., Wong, B. J. & Chen, Z. High-speed upper-airway imaging using full-range optical coherence tomography. J Biomed Opt 17, 110507 (2012).

Analysis of connected speech using high-speed videoendoscopy

Dimitar D. Deliyski [1], Maryam Naghibolhosseini [1], Stephanie RC Zacharias [2,3],

Alessandro de Alarcon [2,3], Robert F. Orlikoff [4]

[1] Department of Communicative Sciences and Disorders, Michigan State University,
[2] Division of Pediatric Otolaryngology, Cincinnati Children's Hospital Medical Center,
[3] Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati,
[4] Department of Communication Sciences and Disorders, West Virginia University,

ddd@msu.edu

Introduction

Voice disorders present significant challenges to human communication. Even though people reveal their voice disorders primarily when they communicate using connected speech, due to methodological limitations, the clinical voice assessment techniques are primarily designed for evaluating sustained-vowel voice production. We have completed our first approach in utilizing high-speed videoendoscopy (HSV) for the analysis of vocal-fold vibratory behavior during connected speech production. The ability to measure and document phonatory oscillations, voice onset and offset, and adduction/abduction behavior in running speech represents an important advance in vocal and laryngeal assessment—one that more closely reflects the presenting disorder, disability, and handicap. This lecture reviews our initial data in recording, segmenting and measuring vocal-fold vibratory behavior during connected speech.

Methods and Results

Using a newly-developed color flexible fiberoptic HSV system, connected speech samples were recorded at the rate of 4,000 frames per second. Participants were asked to read the "Rainbow Passage" and other standard connected speech stimuli, as well as series of sustained vowels at varying pitch and loudness including their vocal onset and offset and pre/post-phonatory laryngeal adjustments. The maximum length of a recording was 88 seconds. Using a custom-built Matlab software; automatic temporal segmentation, glottal area waveform estimation and pitch extraction were performed (Figure 1). Simulated stroboscopy video sequences were produced. Vocal attack time, vibratory offset time, left-right vibratory phase symmetry and periodicity were measured, and mucosal wave and vocal-fold adduction and abduction were visually rated.

Results

The study demonstrated the ability to record, analyze and extract clinically-essential features of connected speech. From all connected speech reading stimuli, the "Rainbow Passage" allowed for most instances in which the vocal folds were visualized through the onset, phonation and offset of the vowels. However, due to the articulatory context and related mechanisms, in some vowels the epiglottis or other supraglottal anatomic structures intermittently obstructed the view to the vibrating vocal folds, partially or completely.
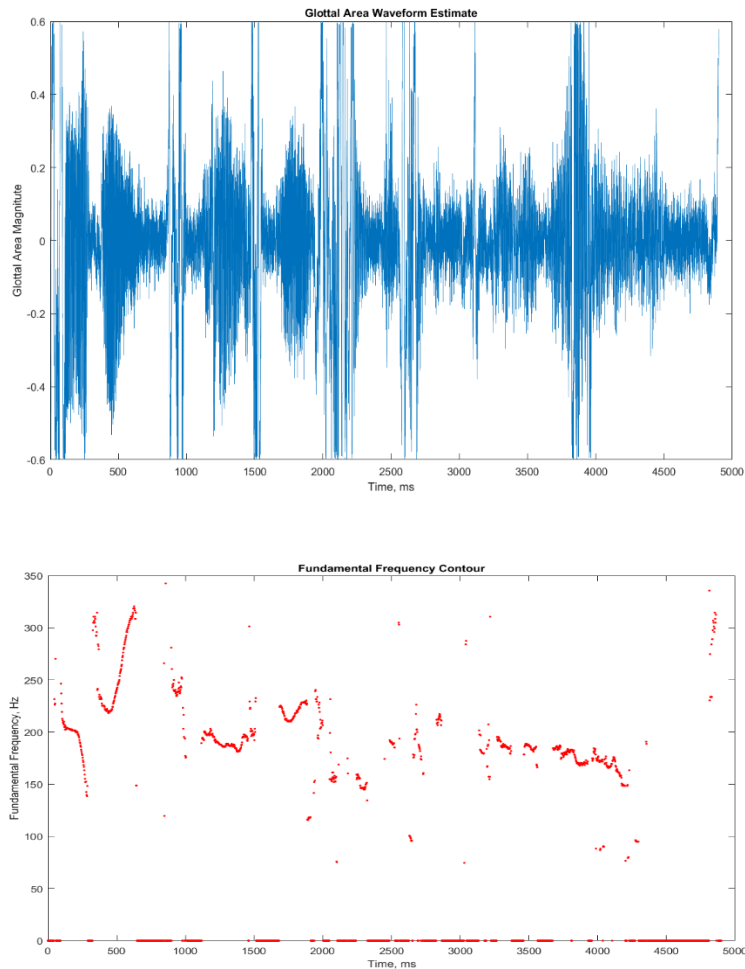
Figure 1: Glottal Area Waveform and Fundamental Frequency extracted using automated temporal segmentation of a high-speed videoendoscopic recording of the "Rainbow Passage". The figure shows only the first 5 seconds reading: "*When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow*."

Conclusions

The use of HSV for recording, analyzing and extracting clinically-relevant vibratory vocal-fold characteristics in connected speech is feasible and demonstrates enormous potential for advancing and refining instrumental voice assessment protocols. Further studies are warranted to creating connected speech stimuli most appropriate for laryngeal imaging and analysis via HSV, and to developing clinical

Vocal fold oscillation patterns in soprano singers` high fundamental frequency phonation

Matthias Echternach[1], Fabian Burk[1], Michael Döllinger[2], Michael Burdumy3,Bernhard Richter[1]

[1]Institut of Musicians' Medicine, Freiburg University Medical Center, [2] Department of Phoniatrics and Pedaudiology, Erlangen University Medical Center, [3] Dept. of Medical Physics, Freiburg University Medical Center,

Introduction

Human voice production at very high fundamental frequencies is not yet understood in detail. Before, it was hypothesized that these frequencies are produced by turbulences, vocal tract/vocal fold interactions or vocal fold oscillations without closure. In a single professional soprano subject singing a diatonic scale it was found, however, that vocal fold oscillation was detectable up to fundamental frequencies of 1568Hz with total closure of the membraneous part of the vocal folds. This study aims to characterize voice production at high fundamental frequencies in a greater amount of professional soprano subjects and to analyze oscillatory patterns in detail. Furthermore, the stability of vocal fold oscillations was also analyzed during glissando tasks.

Methods

 6 professional, vocally healthy (verified by medical examination, voice handicap index and establishment of the Dysphonia Severity Index) opera soprano singers in the high frequency range were analyzed using transnasal flexible high speed digital imaging (HSDI) technology with a frame rate of 20000fps. Simultaneously, audio and electroglottographic (EGG) signals were recorded with the same rate as the HSDI. At firstm the subjects were asked to sing a scale from C6 (1047Hz) to G6 (1568Hz) as they would do on stage. In a second step, the subjects were asked to perform a glissando from G5 (784Hz) to G6 (1568Hz) avoiding major irregularities. At last, the subjects were asked to phonate as high as possible. Since vowel conditions are limited in this high frequency range, the task was performed on the vowel /a/. All HSDI videos were segmented using the software Glottal Analysis Tools (Erlangen University) and phonovibrograms were established. Out of the different signals (Audio, EGG and Glottal Area Waveform), characteristics such as open quotient, speed quotient, perturbation values were calculated using Glottal Analysis Tools.

Results

In all of the subjects vocal fold oscillations were visible on the total length of the membraneous part of the vocal folds up to 1568Hz. Furthermore, there was total closure of the vocal folds during the glottal cycle for all pitches and no major changes in oscillation patterns. However, for both the glottal area waveform and the EGG, the open quotient was gradually decreased for the top pitches. Reaching higher pitches was in most of the subjects associated with a small narrowing of the ventricular folds, as consequence of increased vocal fold adduction. In a single subject phonation was possible up to C#7 (2217Hz). Also in this exceptional case vocal fold oscillations were visible with total vocal fold closure throughout.

Conclusion

The results confirm for 6 professional soprano subjects that voice is produced by fragmentation of the airflow by oscillating vocal folds up to fundamental frequencies of 1568Hz. In a single subject this could be seen up to 2217Hz. Furthermore, there was no evidence that the laryngeal mechanism is changing in glissando tasks from G5 to G6 in professional soprano singers` subjects, since no major changes in oscillatory patterns could be observed.

A new method to present high-speed data for laryngeal assessment based on Optical Flow computation.

Gustavo Andrade-Miranda[1], Nathalie Henrich Bernardoni[2,3], Juan I. Godino-Llorente[1]

[1]Centro de Tecnología Biomédica, Univ. Politécnica de Madrid, [2]Univ. Grenoble Alpes, GIPSA-Lab, F-38000 Grenoble, [3]CNRS, GIPSA-Lab, F-38000 Grenoble,

gxandrade@ics.upm.es

Introduction

High-speed videos (HSV) provide a huge amount of images, whose analysis requires a great deal of human intervention and observation [1]. However, with appropriate image-processing techniques, the time-varying data can be condensed into a few static images reducing the spatial-temporal dimensionality. Most of current methods rely on glottal-area segmentation, which is not a trivial task. Motion analysis is then focused only on vocal-folds edges. In this paper, we propose a new approach for visualizing the glottal dynamics in a compact manner based on Optical flow (OF) computation of the HSV sequences.

Methods

OF computation allows the possibility to track unidentified objects solely based on its motion, with no need of image segmentation. Three new playbacks are proposed to visualize the computed optical flow: OFKG, OFGVG and GOFW playbacks. The Optical-Flow Kymogram playback (OFKG) uses the same principle than Digital Kymogram (DKG) to compact high-speed information. The information used to condense the data is taken from the displacements originated from the x axis. It is coded with different tonalities of red and blue depending on the displacement direction. The Optical-Flow Glottovibrogram (OFGVG) represents the velocity of glottal movement per cycle plotted along the vocal-folds length. It is computed by averaging each row of the x component of the flow and representing it as a column vector. The Glottal Optical-Flow Waveform (GOFW) is a 1D representation of the velocity. GOFW is based on the same principle of the Glottal Area Waveform (GAW). Graphically the GOFW represents the change of velocity as a function of time

Results

To evaluate the accuracy of this image-processing technique, comparison is made with the information obtained via glottal segmentation proposed by Karakozoglou et al [2]. First, the trajectories of both approaches are carried out and their similarities are investigated (see Fig.1). The trajectories are computed considering only the displacements on the x-axis of the OF for a line that intersect perpendicularly the glottal main axis at 50% of its total length. The OF of the glottal-segmentation will be obtained via the displacements of consecutive glottal edges. The second trial consists on study the correlation between the usual playbacks and the proposed ones (illustrated in Fig.2). According to the results, a great similitude in the shape and in the glottal phases (Opening and closing) is noticed when both approaches are compared.
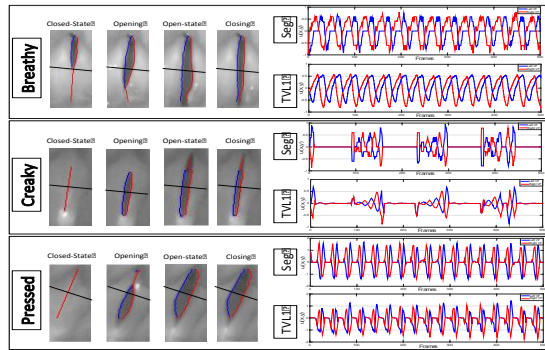
Fig.1: Comparison of vibratory trajectories derived from glottal segmentation (Seg) and OF (TVL1) for three vocal qualities (breathy, creaky, and pressed). The trajectories are computed along the line that intersect perpendicularly the glottal main axis at 50% of its total length (black line).
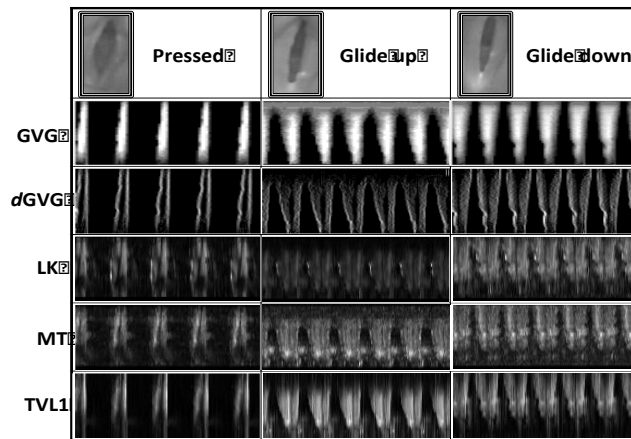
Fig.2: Illustration of several playbacks in three phonatory tasks (pressed voice, glide up and glide down). GVG : Glottovibrogram; dGVG: Derivative of GVG; LK: OF computed by Lukas Kanade Method; MT: OF computed by motion tensor; TVL1: OF computed by Total variation approach.

Conclusion

A new approach has been explored so as to condense dynamical information from HSV recordings in a compact way which would not depend on prerequisite glottal segmentation. The glottis is treated as an unidentified object, and attention is focused on the motion field produced by vocal-folds vibration. Three new playbacks are proposed to visualize the computed optical flow: OFKG, OFGVG and GOFW playbacks. Similarities are found in the information extracted either from glottal segmentation or from OF, since both methods quantify glottal motion. For the purpose of clinical diagnosis, it seems a promising approach which could complement, and eventually to replace, segmentation-based techniques.

References

. [1] Kendall, K., Leonard, R.. Laryngeal Evaluation: Indirect Laryngoscopy to High-Speed Digital Imaging. New York - Stuttgart: Thieme; 1st ed.; 2011. ISBN 9781604062731.

. [2] Karakozoglou, S.Z., Henrich, N., D'Alessandro, C., Yannis, S.. Automatic glottal segmentation using local-based active contours and application to glottovibrography. Speech Communication 2012; 54(5): 641–654.

The Impact of Glottal Closure on Speech Breathing

Elizabeth S. Heller Murray[1], Carolyn M. Michener[1], Laura Enflo[1], Meredith J. Cler[3], Cara E. Stepp[1,2,3,4]

[1] Department of Speech, Language, and Hearing Sciences, Boston University, [2]Department of Biomedical Engineering, Boston University, [3]Graduate Program for Neuroscience - Computational, Boston University, [4]Department of Otolaryngology – Head and Neck Surgery, Boston University School of Medicine, ehmurray@bu.edu

Introduction

The presence of vocal fold nodules can cause incomplete glottic closure during adduction, often resulting in the perception of a breathy vocal quality (Stemple, Glaze, & Gerdeman, 2000). Individuals with vocal nodules have been shown to speak with increased lung volume initiation and decreased lung volume termination (Sapienza & Stathopoulos, 1994; Sapienza, Stathopoulos, & Brown, 1997); this is hypothesized to be a compensatory mechanism for incomplete glottic closure or a potential mechanistic feature behind the development of hyperfunction voice disorders. The current work examines changes in lung volume initiation (*LV initiation*) and lung volume termination (*LV termination*) in healthy speakers emulating a breathy vocal quality.

Methods

Twelve healthy participants (7 female) read the "Rainbow Passage" (Fairbanks, 1960) while wearing the Inductotrace® inductive plethysmograph (Ambulatory Monitoring, INC., Ardsley, NY), which measures the excursion of the rib cage and abdomen. At a controlled reading rate, participants read the "Rainbow Passage" 5 times in their typical speaking voice (baseline phase), 10 times in an emulated breathy speaking voice (breathy phase), and 5 times again in their typical speaking voice (return phase). During each reading of the "Rainbow Passage," the kinematic data of *LV initiation* and *LV termination* were collected and subsequently averaged within each reading.

Results

All values were normalized to the baseline phase and tidal volume. A 3×2 (phase: baseline, breathy, return $\times$ kinematic measure: *LV initiation*, *LV termination*) repeated measures ANOVA indicated that there were significant main effects of phase and kinematic measure, as well as a significant interaction of phase $\times$ kinematic measure (all $p < 0.01$). Planned comparison *t*-tests revealed that *LV termination* was significantly lower in the breathy phase than in the baseline phase ($p < 0.01$). There were no significant differences between the baseline and breathy phase in *LV initiation*, as well as no significant difference between the baseline and the return phases in either *LV initiation* or *LV termination*.

Conclusions

Healthy participants spoke with decreased *LV termination* when emulating a breathy vocal quality. This provides further support that these changes in lung volume may be an attempt to compensate for an inefficient glottal configuration, either due to organic changes or volitional vocal quality modifications. Future work will examine whether these compensations are similar in individuals with a history of hyperfunctional voice disorders.
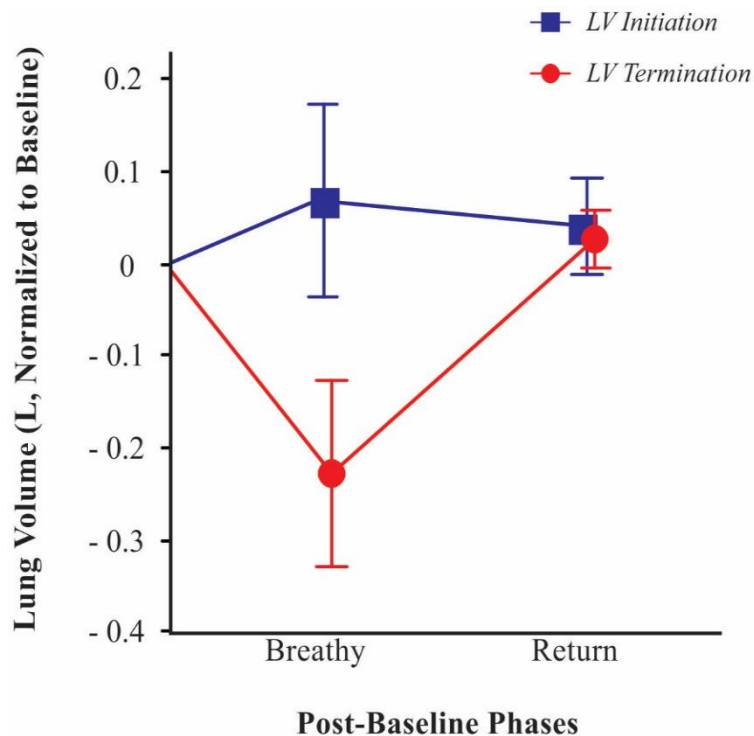
Figure 5: Means and 95% confidence intervals for the change in lung volume (liters, normalized to baseline) of LV initiation (blue squares) and LV termination (red circles). LV termination was significantly lower in the breathy phase than in the baseline phase. There were no other significant differences.

References

Fairbanks, G. (1960). *Voice and articulation drillbook (Vol. 2)*. New York, NY: Harper and Brothers.

Sapienza, C. M., & Stathopoulos, E. T. (1994). Respiratory and laryngeal measures of children and women with bilateral vocal fold nodules. *Journal of Speech, Language, and Hearing Research, 37*(6), 1229-1243.

Sapienza, C. M., Stathopoulos, E. T., & Brown, W. S. (1997). Speech breathing during reading in women with vocal nodules. *Journal of Voice, 11*(2), 195-201.

Stemple, J. C., Glaze, L. E., & Gerdeman, B. K. (2000). *Clinical voice pathology: theory and management (fourth edition)*.San Diego, CA: Plural Publishing.

Vocal fold opening and closing phase differences in children with and without bilateral lesions

Stephanie RC Zacharias [1,2,3], Alessandro de Alarcon[1,2] Dimitar Deliyski[4]

[1]Division of Pediatric Otolaryngology, Cincinnati Children's Hospital Medical Center
[2]Department of Otolaryngology- Head and Neck Surgery, University of Cincinnati,[3]Division of Speech-Language Pathology, Cincinnati Children's Hospital Medical Center [4]Department of Communicative Sciences and Disorders, Michigan State University

Stephanie.Zacharias@cchmc.org

Introduction

Vibratory characteristics of the vocal folds play essential role in voice quality. The difference between the opening and closing phases of the vibratory cycle may have important implications in understanding and characterizing the pathophysiology of bilateral lesions. This knowledge could be especially valuable in pediatric populations due to the lack of research on the topic and the high prevalence of nodules and cysts.

Objectives

Speed quotient (SQ) and speed index (SI) are quantitative measures with substantial clinical potential. The purpose of this study was to provide initial quantitative data on the measurement of vocal fold opening and closing phase differences in children with and without benign bilateral vocal fold lesions.

Methods

In a prospective study, 10 participants (6 healthy; 4 with lesions) between the ages of 5 and 10 were examined using rigid monochrome high-speed videoendoscopy at the rate of 8,000 frames per second. Participants were asked to sustain phonation at normal pitch and loudness. Using custom-built software for dynamically linking HSV to digital kymography, the following objective measures were obtained: left and right SQ, left and right SI, and open quotient. Measures were obtained at three positions along the vocal folds; for patients: at the lesions site and anteriorly and posteriorly of the lesions.

Results

The study tested the hypothesis that the SQ/SI would vary along the anterior-posterior vocal fold axis. This variation is particularly underscored in the presence of bilateral lesions. Additionally, SQ/SI differ in patients compared to the corresponding anatomic areas in children with no lesions.

Conclusions

Results of this study quantitatively demonstrate how the opening and closing phase difference varies along the anterior-posterior axis and across children with and without vocal fold lesions, which can help refine the theoretical framework in pediatric laryngeal pathology.

Recommendations

Further studies are warranted to address potential age and sex related differences and differences between bilateral vocal fold nodules and cysts with contralateral lesions.

Center of Vocal Fold Vibration during Initiation and Termination Phases

Melda Kunduk[1,2,3], Takeshi Ikuma[1], Daniel Fink[1,3],and Andrew J. McWhorter[1,3],

[1]Department of Otolaryngology–Head and Neck Surgery,Louisiana State University Health Sciences Center, [2]Department of Communication Sciences and Disorders, Louisiana State University, [3]Our Lady of the Lake Regional Medical Center- Voice Center

mkunduk@lsu.edu

Introduction

During the transient phases of phonatory vibration of vocal folds, the location and length of the glottis may be dynamically changing. Moore[1] reported that the initial oscillation of vocal folds may occur only over a small section where the vocal folds are least resistant to displacement, and the oscillation progressively extends over the rest of vocal folds. Werner-Kukuk and von Leden[2] also observed that the oscillation with a hard vocal attack took a few cycles to reach the anterior end and even more cycles for the oscillation to finally reach the posterior extremity. Such shift in the location of the vibration can be parameterized by analyzing the center of vibration (CoV), which is analogous to the center of mass of a physical object. This work introduces a CoV feature and its affine parametric model during the transients based on high-speed videoendoscope (HSV) data. The efficacy of the parameters was tested for the effects of phonation types (normal, pressed, and breathy).

Methods

Proposed CoV feature is a ternary HSV feature derived from the glottal width root-mean-square (RMS) waveforms[3], which sample represents the vibrational strength at a kymographic position on an HSV frame. The CoV is defined as the sagittal position of the centroid of the width RMS samples across all kymographic lines on a given frame. To assess the CoV behaviors during the transient phases, the dynamic CoV behavior was evaluated over an entire transient period with an affine approximation. The CoV location at vibration onset or offset ($s_{CoV,0}$) and the CoV rate of change over the transient period ($\dot{s}_{CoV}$) emerge as two transient parameters. Fig. 1 illustrates the relationships of the width RMS, CoV, and its affine approximations. These CoV parameters were evaluated for both the initiation and termination in 18 HSV recordings (6 each of normal, breathy, and pressed voice) of a female subject. Mixed model analyses were conducted using the mixed procedure of SAS (SAS Institute Inc. 2008) to determine the effect of phonation type on the CoV parameters.

Results

Table 1 summarizes the group means of the CoV parameters. The phonation type significantly affected both $s_{CoV,0}$ ($p < 0.001$) and $\dot{s}_{CoV}$ ($p < 0.05$) during the initiation of oscillation phase. No significant effect was found in the termination phase. The onset of the oscillation for the pressed phonation occurs noticeably more towards the anterior than normal or breathy phonation. The CoV of the pressed phonation then migrates towards the mid-fold

over the initiation phase. On the other hand, the vibration for both the normal and breathy voice generally onsets at mid-fold and remain at the mid-fold in the steady state.

Conclusions

Distinctive initiation CoV characteristics have been observed for imitated pressed voice in the presented preliminary evaluation. The findings warrant a future study to confirm if the same trend can be observed for general normal population. Assuming the findings are generalizable, these CoV parameters may be useful in quantifying the pre-phonatory laryngeal tension and force of the vocal fold adductory behavior associated with certain voice disorders, e.g., spasmodic dysphonia.
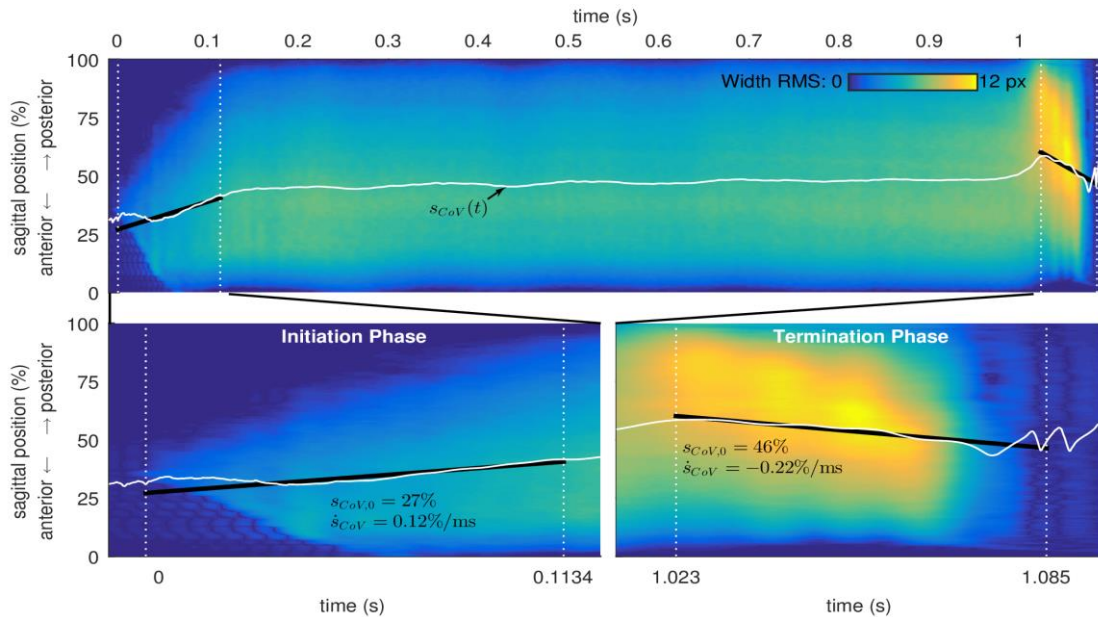
Fig. 1. Illustration (pressed phonation) of glottal width RMS (as color intensity map), CoV profile ($s_{CoV}(t)$, white line), and the affine CoV approximations in the transient phases (straight black lines).

Segment	Variable	Units	Voice Type			Table 1. Conditional Means of CoV Parameter Measurements
			Normal	Breathy	Pressed	
Initiation	$s_{CoV,0}$	%	50.7	48.1	35.3	
	$\dot{s}_{CoV}$	%/ms	0.00	0.05	0.17	
Termination	$s_{CoV,0}$	%	43.7	46.1	49.8	
	$\dot{s}_{CoV}$	%/ms	-0.06	-0.04	-0.16	

References

1. Moore P. Otolaryngology and speech pathology. *The Laryngoscope.* 1968;78(9):1500-1509.
2. Werner-Kukuk E, von Leden H. Vocal initiation: High speed cinematographic studies on normal subjects. *Folia Phoniatrica.* 1970;22(2):107-116.
3. Ikuma T, Melda K, Fink D, McWhorter AJ. A Spatiotemporal approach to the objective analysis of initiation and termination of vocal-fold oscillation with high-speed videoendoscopy. *Journal of Voice.* In Press.

Computerized tomography measures during and after artificial lengthening of the vocal tract in subjects with voice disorders

Marco Guzman[(1)], Gonzalo Miranda[(1)], Christian Olavarria[(2)], Sofia Madrid[(2)], Daniel Muñoz[(3)], Miguel Leiva [(2)]

[(1)]University of Chile, Department of communication Sciences and Disorders. Santiago, Chile.
[(2)]University of Chile Hospital, Santiago, Chile.
[(3)]University of Chile, Department of Otolaryngology, Santiago, Chile
guzmann.marcoa@gmail.com

Introduction

Several effects have been attributed to semi-occluded vocal tract postures. One of the effects that have been explored is the modifications produced in vocal tract configuration. [1, 2] The present research aimed to observe the effect on vocal tract bidimentional and tridimentional images of two type of tubes commonly used in voice therapy and training.

Methods

Ten participants were included in this study. All participants were asked to undergo flexible laryngoscopy to confirm the presence of functional dysphonia. Computerized tomography (CT) was carried out in all subjects. The subject were asked to produce the following phonatory tasks: (1) to sustain vowel [a:] (baseline, condition pre), (2) to phonate a sustained vowel-like sound into a drinking straw (tube 1) (5 mm of inner diameter and 25.8 cm in length) for fifteen minutes, immediately after that, (3) to produce another sustained vowel [a:] (condition post). After twenty minutes of complete silence (vocal rest), phonation into a plastic stirring straw (tube 2) (2.7 mm of inner diameter and 10.7 cm in length) was performed for fifteen minutes. Immediately after that, the participant was asked to produce another sustained vowel [a:] (post condition).

Five CT midsagittal images (five phonatory tasks x one repetition) were chosen to perform a series of distance measurements (mm). Moreover, three cross-sectional areas (mm^2) were measured from the same ten midsagittal CT images. Additionally, from the transversal CT images, two areas (mm^2) were measured. Furthermore, volumetric measures of total vocal tract were also performed.

Descriptive statistics were calculated for the variables, including median and interquartilic range. A generalized estimating equation (GEE) model to assess differences between pre, during and post tube use and also difference between tubes, was fitted.

Results

Result from distance measures showed that during both tube 1 and tube 2 there was an increased vertical length, tongue dorsum height increased, there was a wider oropharynx, higher velum position, and a wider hypopharynx compared to vowel phonation pre a post tube. Result from area measures showed that during both tubes, the outlet of the epilaryngeal tube, inlet to the lower pharynx, pharyngeal region, epilaryngeal region were found greater during phonation into both tubes. Results from volumetric measures showed that during both tubes there was an increased total volume.

Conclusion

Vocal tract configuration during tube phonation seems to lead to a more economic voice production. More prominent changes are reached during stirring straw than drinking straw.

References

1. Laukkanen, A-M., Horacek, J., Krupa, P., Svec, JG. (2010). The effect of phonation into a straw on the vocal tract adjustments and formant frequencies. A preliminary MRI study on a single subject completed with acoustic results. Biomed Signal Process Control, 7, 50-57.

2. Guzman, M., Laukkanen, A-M., Krupa, P., Horáček, J., Švec, J., & Geneid, A. (2013). Vocal Tract and Glottal Function During and After Vocal Exercising with Resonance Tube and Straw. J Voice, 27, 305-311.

Comparison of supraglottic activity and spectral slope between theater actors and vocally untrained subjects

Marco Guzman[1], Andres Ortega[2], Christian Olavarria[3],,Daniel Muñoz[4], Pedro Cortes[3], María Josefina Azocar[1], David Cayuleo[1], Felipe Quintana[1], Catalina Silva [1]

[1]University of Chile, Department of communication Sciences and Disorders. Santiago, Chile.
[2]Clínica las Condes, Deparment of Otolaryngology. Santiago, Chile
[3]University of Chile Hospital, Santiago, Chile.
[4]University of Chile, Department of Otolaryngology, Santiago, Chile

quintana.b.felipe@gmail.com

Introduction

Medial and anterior-posterior (A-P) laryngeal compression have commonly been described as endoscopic signs of vocal hyperfunction. [1] Previous studies, however, have shown that a perceived increase in supraglottic activity may also occur in subjects diagnosed with normal voice. [2, 3] The purpose of the present study was to assess supraglottic activity in professional theater actors and to observe whether they present differences compared to subjects with no voice training. Acoustic and perceptual analysis was also performed. We hypothesize that supraglottic activity may not necessarily be greater in untrained participants, and it could be even more prominent in vocally trained individuals.

Methods

Twenty participants were included. They were divided into two groups: An experimental group of vocally trained theater actors (n=10), and a comparative group of subject with no voice training (n=10). Absence of laryngeal pathology was confirmed by rigid videostroboscopy in all participants. During the trans-nasal endoscopic examination, each participant was instructed to produce two different phonatory tasks: 1) to read a phonetically balanced text at three loudness levels (medium, high and low), and 2) to produce sustained vowels [a:], [i:], and [u:] during five seconds at comfortable pitch and at three loudness levels (medium, high and low) each production. Voice recording was also carried out. Acoustical analysis with Long-Term Average Spectrum (LTAS) was performed. The acoustical variables in this study were 1) the energy level difference between the F1 and F0 regions, and) the alpha ratio, which is the energy level difference between 50-1000 Hz and 1000-5000 Hz. Four blinded judges were asked to assess laryngoscopic and auditory perceptual variables using a visual analog scale.

Results

Multivariate linear regression showed that trained participants had a higher degree of both laryngeal and pharyngeal activity compared to untrained participants. Moreover, phonatory tasks at high intensity showed higher activity than medium and low intensity for most phonatory tasks and laryngoscopic parameters. Vocally trained participants evidenced higher values for all spectral variables compared to untrained participants.

Conclusion

Actors have a greater degree of both laryngeal and pharyngeal activity than vocally untrained subjects. Apparently, this higher activity is associated to speaking voice training and not to a hyperfuctional vocal behavior. A-P laryngeal compression is greater than medial compression. Intensity and phonatory task have an effect on all laryngoscopic variables. Supraglottic activity during professional speaking voice may be not necessarily a hyperfunctional behavior, but a strategy to avoid vocal fold damage while producing the desired voice quality.

References

1. Stager, S., Bielamowicz, S., Regnell. J., Gupta, A., Brakmeier, J. Supraglottic activity: Evidence of hyperfunction or laryngeal articulation? J Speech Hear Res. 2000, 43, 229-238.
Behrman, A., Dahl, L., Abramson, A., Schutte, H. Anterior-Posterior and Medial Compression of the Supraglottis: Signs of Nonornic Dysphonia or Normal Postures? J Voice. 2003, 17, 403-410.

2. Sama, A., Carding, PN., Price, S., Kelly, P., Wilson, JA. The clinical features of functional dysphonia. Laryngoscope. 2001, 111, 458-463.

Simulations of child-like speech as test material for speech analysis algorithms

Brad H. Story[1],, Kate Bunton[1],

[1] Dept. of Speech, Language, and Hearing Sciences, University of Arizona

bstory@email.arizona.edu

Introduction

Children's speech is often characterized by a high fundamental frequency that generates widely spaced harmonics. The effect is an undersampling of the vocal tract frequency response, making estimation of the formant frequencies quite challenging. The purpose of this study was to develop a database of simulated child-like speech representative of ages ranging from infancy to 6 years old that can be used to test formant analysis algorithms.

Methods

The samples for the database were simulated with an airway modulation model of speech production (Story, 2013). The model includes a voice source, vocal tract, and trachea and generates sound output analogous to a microphone signal.

Three different voice source models were used to provide different levels of source-tract interaction: 1) a glottal flow pulse generator (linear), 2) a kinematic model of the vocal folds surfaces (nonlinear interaction of glottal flow), and 3) a lumped-element self-oscillating model (nonlinear interaction of glottal flow and tissue). For the latter two models, the dimensions of the vocal folds were scaled to be representative a child-like system.

The simulations include combinations of three time-varying vocal tract shapes (/iaiai/, /uaeuaeu/, /iai/) and two fundamental frequency contours. These were simulated with each of the three voice source models such that 18 audio samples were generated. Treating each simulation as if it were an actual recording of a child, the formant frequencies were analyzed with a standard linear prediction algorithm (LPC) and spectral filtering algorithm (Story & Bunton, 2015). The results were compared to the time-varying calculated resonance frequencies.

Results

Figure 1 demonstrates some results of a simulation where the vocal tract shape represented a vowel transition of /iai/ and the voice source was the lumped-element self-oscillating model. In Fig. 1a is narrow band spectrogram of output sound pressure, and Fig. 1b shows the fundamental frequency and harmonics as gray lines and the calculated (true) resonance frequencies as thick black lines. Fig. 1c is the same as 1b except that the results of the LPC and spectral filtering formant algorithms are superimposed.
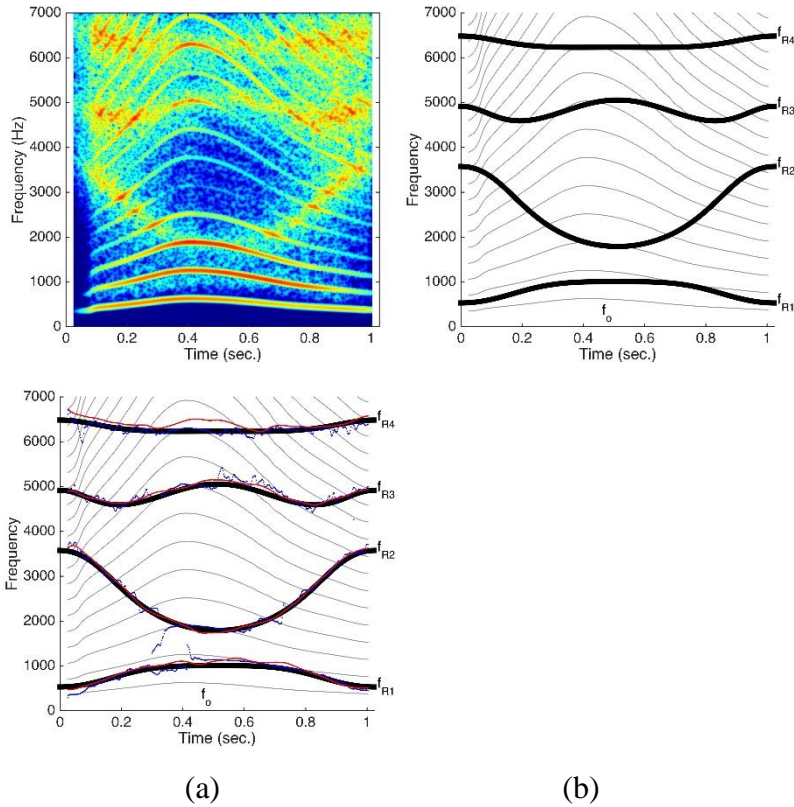
Figure 1: (a) Narrowband spectrogram of simulated utterance, (b) f_o contour and harmonics (gray) and calculated resonances f_{Rn}, and (c) same as (b) except results of formant analyses are superimposed (blue = LPC, red = spectral filtering).

Conclusions

At this point all simulations have been generated and formant analyses with different methods completed. Additional simulations will be added in the future that address new combination of vocal tract state and vocal fold configuration.

References

Story, B.H., (2013). Phrase-level speech simulation with an airway modulation model of speech production, Computer Speech and Language. 27(4), 989-1010.

Story, B. H., and Bunton, K. (2015). A spectral filtering method for tracking formants in children's speech, Proceedings of Meetings on Acoustics, 23, Technical area: Paper 2pSC23.

Effectiveness of recurrence quantification measures in discriminating patients with and without voice disorders

Leonardo Wanderley Lopes[1], Vinícius Jefferson Dias Vieira[2], Silvana Luciene do Nascimento Cunha Costa[3], Suzete Élida Nóbrega Correia[4], Mara Behlau[5]

[1] Universidade Federal da Paraíba (UFPB)/Universidade Federal de São Paulo (UNIFESP), [2] Universidade Federal de Campina Grande (UFCG), [3] Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB), [4] Centro de Estudos da Voz (CEV)/ Universidade Federal de São Paulo (UNIFESP)

lwlopes@hotmail.com

Introduction

The signal processing techniques based on nonlinear models and the use of binders with isolated and combined acoustic measurements allow new possibilities of vocal signals analysis. These techniques include recurrence plots (RPs) and their quantification measures (RQMs), which are based on the Poincare's Recurrence Theorem1. Considering that there is no consensus on the set of acoustic measurements with greater accuracy for discriminating the presence / absence of vocal disorder2,3 and there is no scientific evidence regarding a single reference standard (perceptual-auditory assessment, laryngeal visual examination or self-assessment) to be used for identification of a vocal problem in the field of voice clinical evaluation4, the objective of this research was to analyze the accuracy of recurrence quantification measures in the discrimination of patients with and without voice problem.

Methods

The total of 489 Patients, with vocal complaints and prior laryngological evaluation has participated, BOTH sexes, with 114 men and 375 women, with a mean age of 40.05 (SD = 14.63). In voice recording, we request the emission of sustained vowel /ɛ/. From the perceptual-auditory analysis of the vocal deviation intensity (GG), done by three speech therapists with extensive experience in voice assessment, using a visual analogical scale (VAS) of 100 mm, the patients were divided into two groups: 52 patients without voice problem (GG≤35.5mm) and 437 patients with voice problems (GG≥35.6mm)5. Although all patients had vocal complaints, the perceptual analysis was used in this study as a criterion for deciding the outcome studied (presence or abscense of voice disorder). For the acoustic analysis, tracks of signs with 2 s were selected, excluding the start and end of the sample. Twelve RQM's were extracted using 1% recurrence rate: determinism, medium and maximum lengths of diagonal lines, Shannon entropy, laminarity, length of stay, maximum length of vertical lines, recurrence time of type 1 and type 2, entropy of type 1 recurrence time, transitivity and divergency. Parameters related to the recurrent topology of the vocal production system (embbeding dimension, reconstruction step and neighborhood radius) were also analyzed totaling a set of 15 measurements. The average values of measurements were used in the rating process by means of quadratic discriminant analysis. To balance the

amount of healthy and abnormal signals, the latter were subdivided into eight subsets (three with 54 signals and five 55 with signals) randomly, confronting each of these subsets with 52 healthy signals, using cross-validation. The accuracy measurements, sensitivity and specificity were used to evaluate the classifier performance. Considering that, in this research, we analyze the effectiveness of acoustic measurements to differentiate patients with and without voice problem in the initial situation of vocal assessment (screening), we focus on the higher sensitivity in the data analysis.

Results

Individually, measures based on the formation of diagonal structures (Shannon entropy, averaged diagonal lines and determinism) had a higher rating potential to discriminate patients from vocally healthy subjects (Table 1). The low specificity of TRANS and DET measures and may indicate that such measures are not effective in correctly identifying the absence of deviation in voice quality. There was a gain in the classification rate when the RQM's were combined, besides better balance between sensitivity and specificity values (Table 2). There was much accuracy with the combination of eight measures, including parameters related to the formation of diagonal lines (L_{med} and ENTR), vertical lines (TT and V_{max}), the system dynamics reconstruction (τ and m), and related to the formation and evolution/transition of recurring points (T^1 and TRANS, respectively).

Conclusions

Recurrence measures, isolated or combined, perform well in the discrimination of individuals with and without voice problem. The combination of measures improves the rating effectiveness, with better balance between sensitivity and specificity values, making such useful measures for procedures of assessment and vocal screening.

Table 1 - Classification statistics for presence versus absence of voice disorder by individual recurrence measures

Medida	Accuracy (%)	Sensitivity (%)	Specificity (%)
ENTR	74.00±4.06	89.00±4.92	60.33±5.48
L_{med}	72.73±5.41	61.67±7.69	82.67±5.28
TRANS	72.73±4.98	87.00±7.31	59.00±5.80
DET	66.36±2.39	84.33±3.98	48.00±5.67

Subtitles: **ENTR** = Shannon entropy; **L_{med}** = averaged diagonal lines; **TRANS** = transivity; **DET** = determinism

Table 2 - Classification statistics for presence versus absence of voice disorder by combined recurrence measures with best accuracy and combination

Measure	Combination	Accuracy (%)	Sensitivity (%)	Specificity (%)
τ, TRANS	2x2	80.36±3.22	84.67±4.67	75.33 ±5.95
τ, RPDE, DIV	3x3	82.18±4.01	86,67±4.19	77.67±5.24
τ, m, TT, TRANS	4x4	82.18±3.25	87.00±3.80	78.33±5.84
τ, m, L_{med}, ENTR, T1	5x5	82.27±2.78	87.00±3.80	77.33±3.84
τ, m, L_{med}, ENTR, LAM, T1	6x6	83.18±3.27	84.33±5.80	81.67±4.64
τ, L_{med}, L_{max}, V_{max}, T^1, T^2, TRANS	7x7	81.00±3.14	88.00±5.33	74.00±4.24
τ, m, L_{med}, ENTR, TT, V_{max}, T^1, TRANS	8x8	83.27±3.76	86.67±5.14	80.33±4.88
τ, m, DET, L_{med}, ENTR, TT, V_{max}, RPDE, DIV	9x9	82.18±3.19	88.67±4.31	75.33±4.19
τ, m, RADIUS, DET, L_{med}, L_{max}, ENTR, LAM, V_{max}, T1, T2	10x10	83.09±3.37	84.67±5.54	80.67±7.13
τ, m, DET, L_{med}, L_{max}, LAM, TT, V_{max}, RPDE, TRANS, DIV	11x11	79,55±2.92	84.33±3.98	75.33±5.19
τ, m, RADIUS, DET, L_{med}, L_{max}, LAM, TT, V_{max}, RPDE, TRANS, DIV	12x12	79.27±4.23	84.67±4.67	74.00±7.55
τ, m, DET, L_{med}, L_{max}, ENTR, LAM, V_{max}, T^1, T^2, RPDE, TRANS, DIV	13x13	77.45±4.85	81.00±4.92	74.67±7.26
τ, m, RADIUS, DET, L_{med}, L_{max}, ENTR, LAM, TT, V_{max}, T^1, T^2, RPDE, DIV	14x14	76.27±4.68	78.33±6.31	74.00±4.91
All measures	15x15	75.09±5.68	84.67±6.96	66.00±7.68

Subtitles: **τ** = delay; **m** = embbeding dimension; **RADIUS**= neighborhood radius; **DET**= determinism; **L_{med}**= averaged diagonal length; **L_{max}**= Length of the longest diagonal line; **ENTR**= Shannon's entropy; **LAM**= laminarity; **TT**= trapping time; **V_{max}**= Length of the longest vertical line; **T^1**= recurrence time of first type; **T^2**= recurrence time of second type; **RPDE**= recurrence period density entropy; **TRANS**= transitivity

Discriminating patients with vocal fold nodules from matched controls using acoustic and aerodynamic features from ambulatory voice monitoring data

Juan P. Cortés[1], Víctor M. Espinoza[1] [2], Matías Zañartu[1], Marzyeh Ghassemi[3], John V. Guttag[3], Daryush D. Mehta[4], Jarrad H. Van Stan[4], and Robert E. Hillman[4].

[1] Universidad Técnica Federico Santa María, [2] Universidad de Chile
[3] Massachusetts Institute of Technology, [4] Massachusetts General Hospital

juan.cortes.5@sansano.usm.cl

Introduction

Clinicians need tools that can monitor daily vocal function to help determine what constitutes safe versus harmful patterns of voice use. This work investigates the best features to identify people with and without vocal fold nodules using accelerometer-based ambulatory voice monitoring. Since higher vocal fold collision forces could lead to phonotrauma, (Titze, 1994), (Hillman et al, 1989), we incorporate aerodynamic features that are associated with these forces (Hillman et al, 1989) into the classification task. We believed these features have more of a theoretical basis underlying phonotraumatic vocal behavior than standard acoustic features such as sound pressure level (SPL) and fundamental frequency (F0).

Methods

Neck-surface acceleration (ACC) was recorded in 34 adult female subjects (17 patients with nodules and 17 controls matched for age and occupation) over seven days. The accelerometer was sampled at 11025 Hz using a smartphone platform with enough memory to record ACC data for over 18 hours per day. We computed standard measures of ambulatory vocal function, such as SPL and F0 (Cheyne et al, 2003). In addition, aerodynamic features were obtained using an estimated glottal volume velocity signal by means of an impedance-based inverse filtering technique (IBIF) (Zañartu et al, 2013). Thus, we were able to compute time-domain measures such as peak-to-peak flow (AC flow) and maximum flow declination rate (MFDR) on running speech, in contrast to previous work (Zañartu et al, 2013) that used IBIF on vowels only. Following the work of (Ghassemi et al, 2014), we calculated statistical features (mean, median, 5th and 95th percentile, skewness, and kurtosis) over 5-minute non-overlapping windows that contained at least 1% phonation time. These features were fed into a logistic regression model with L2 regularization where the training phase was done by leaving one patient-control pair out while the rest of data was used for training. We selected the best model of a 5-fold cross-validation on the training data and tested it on the remaining pair.

Results

The following preliminary results on 34 subjects showed an advantage when adding AC flow and MFDR as features in the logistic regression classifier, as shown in Figure 1. From 6 subjects with nodules misclassified as normal (controls) in Figure 1(a), only 5 are misclassified in Figure 1(b). Likewise, from 3 subjects misclassified as phonotraumatic (positives) in Figure 1(a), only 1 is misclassified in figure 1(b). Overall, the improvement goes from the SPL-F0 baseline (25/34, 73.5%) to the SPL-F0 baseline + AC flow- MFDR (28/34, 82.4%). These preliminary results indicate a correlation between higher AC Flow/ MFDR values and subjects with vocal fold nodules, which has been noted for sustained vowels in previous studies. It is important to note that the margin between classes widens when we added the aerodynamic features, as more subjects are clustered in the extremes of the classification graph.

(a)   (b)
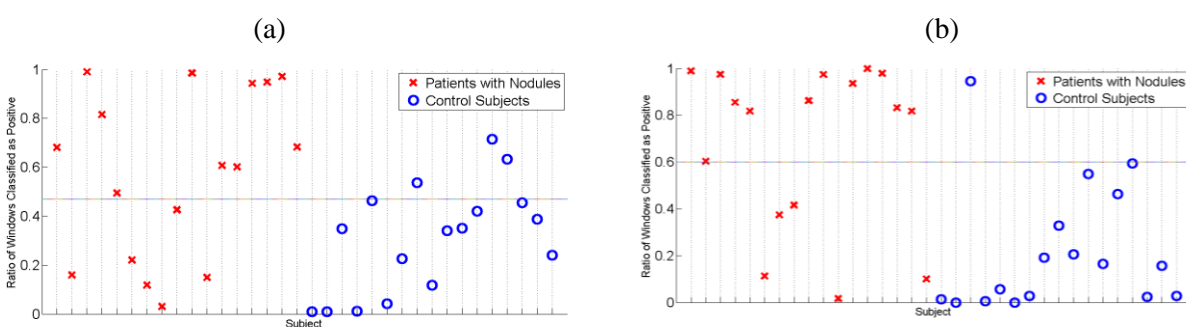
Figure 1: Classification results on 34 subjects using: (a) acoustic features (SPL and F0). (b) acoustic (SPL and F0) and aerodynamic features (AC flow and MFDR). Values closer to 1 indicate a high number of 5-min. windows classified as phonotraumatic, while values close to 0 indicate a high number of windows classified as normal.

Conclusions

Future research needs to address the robustness of the aerodynamic IBIF estimation under different conditions (soft vs. loud speech), as well as to identify different modes of voice use during a day with unsupervised learning techniques.

References

Cheyne, H. A.; Hanson, H. M.; Genereux, R. P.; Stevens, K. N.; Hillman, R. E. "Development and testing of a portable vocal accumulator," J. *Speech. Lang. Hear. Res.,* vol. 46, no. 6, pp. 1457–1467, 2003.

Ghassemi, M.; Van Stan, J.H.; Mehta, D.D.; Zanartu, M.; Cheyne, H.A.; Hillman, R.E.; Guttag, J.V. "Learning to Detect Vocal Hyperfunction From Ambulatory Neck-Surface Acceleration Features: Initial Results for Vocal Fold Nodules", *Biomedical Engineering, IEEE Transactions on,* On page(s): 1668 - 1675 Volume: 61, Issue: 6, June 2014

Hillman, R. E., Holmberg, E. B., Perkell , J. S., Walsh , M., and Vaughan, C. "Objective assessment of vocal hyperfunction:An experimental framework and initial results", *J.Speech Hear. Res.*, vol. 32, no. 2, pp.373 -392 1989

Titze, I. R., "Mechanical stress in phonation," J. Voice 8, 99–105, 1994

Zanartu, M.; Ho, J.C.; Mehta, D.D.; Hillman, R.E.; Wodicka, G.R., "Subglottal Impedance-Based Inverse Filtering of Voiced Sounds Using Neck Surface Acceleration," in *Audio, Speech, and Language Processing, IEEE Transactions on* , vol.21, no.9, pp.1929-1939, Sept. 2013

Constructing a subject-specific lumped-mass model from clinical data using Bayesian estimation

Gabriel E. Galindo[1], Matias Zañartu[1], Sean D. Peterson[2], Byron D. Erath[3]

[1] Universidad Técnica Federíco Santa María, [2] University of Waterloo, [3] Clarkson University

gabriel.galindo.12@sansano.usm.cl

Introduction

The pursuit of quantitative measures of vocal function is of high interest in speech research for enhancing the clinical assessment of voice and speech. Mapping numerical models of voice production to clinical data allows for the extraction of measures of vocal function that are difficult to obtain through direct in-vivo observation in patients (e.g., collision forces, muscle activation, etc.). However, mapping models from real data should be considered a stochastic process that involves uncertainties from the experimental information, and confidence in the estimated measure.

Stochastic identification techniques such as Bayesian estimation, could provide a link between direct observation and numerical models, obtaining subject-specific representations with a specified confidence. Such subject-specific models could produce measures of vocal function and insights on the underlying pathophysiology. Thus, the objective of this study is to estimate parameters of interest of a lumped-mass model of voice production using Bayesian estimation from clinical data.

Methods

A Bayesian framework (importance sampling and particle filter) was used to estimate parameters in a modified implementation of the body cover model (Zañartu et al. 2014) with muscle activation (Titze & Story 2002). The first 75% of a synchronized high speed video and aerodynamic information recordings was used as the input data for estimating parameters of the numerical model, while the remaining data was used to assess the model predictions. Due to the unknown nature of some of the model parameters (e.g., posterior glottal opening, subglottal pressure, etc.), probability density functions (PDFs) were generated using a Monte Carlo approach for the muscle activation parameters and size of the posterior glottal opening.

Results

The ability of the proposed framework to predict model parameters from clinical data is demonstrated in Figure 1, wherein Figure 1a shows the 95% credibility intervals of the DC flow predicted by the model. The predicted value (best estimation) is shown as a black line and the actual value that was clinically measured (experimental value), is represented by a blue line that is virtually overlaid, showing excellent agreement. Figure 1b plots the predicted time history of the radiated pressure and compares it with the clinically-acquired radiated pressure that was directly measured. It is emphasized that the predicted radiated pressure was a model output generated solely from using the oral volume velocity and glottal area in the Bayesian approach. Both the clinical and predicted pressures were low-pass filtered at 1 kHz to discard the acoustic effects of the circumferentially vented mask. Finally, Figure 1c shows the predicted thyroarytenoid activation, with 95% credibility intervals, demonstrating the capability of the approach.

Figure 1
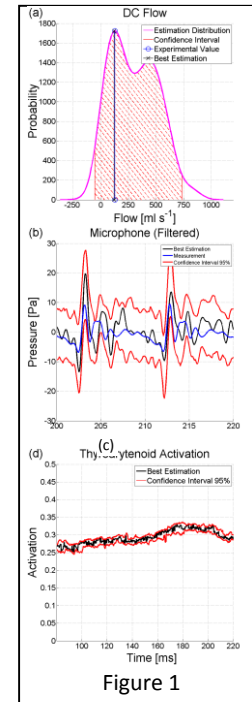
Conclusions

Bayesian methods have, so far, shown good performance correlating the vocal folds parameters over healthy vocal folds, therefore supporting the theory that a subject specific model of the vocal folds is possible. In these early stages, the Bayesian identification method requires a comprehensive analysis over its capabilities and limitations. However, as more clinical synchronized information becomes available, the capabilities of the Bayesian framework will improve, allowing more accurate description of the underlying mechanisms of the vocal folds.

References

Zañartu, M., Galindo, G. E., Erath, B. D., Peterson, S. D., Wodicka, G. R., & Hillman, R. E. (2014). Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunctiona). The Journal of the Acoustical Society of America, 136(6), 3262-3271.

Titze, I. R., & Story, B. H. (2002). Rules for controlling low-dimensional vocal fold models with muscle activation. The Journal of the Acoustical Society of America, 112(3), 1064-1076.

Steinecke, I., & Herzel, H. (1995). Bifurcations in an asymmetric vocal-fold model. The Journal of the Acoustical Society of America, 97(3), 1874-1884.

Cataldo, E., Soize, C., & Sampaio, R. (2013). Uncertainty quantification of voice signal production mechanical model and experimental updating. Mechanical Systems and Signal Processing, 40(2), 718-726.

Phonation Type as a Function of the Activation of the Intrinsic Laryngeal Muscles

D.A. Berry[1], D.K. Chhetri[1], M. Döllinger[2]

[1] Department of Head & Neck Surgery, UCLA, Los Angeles, California, USA

[2] Universitätsklinikum, Abteilung für Phoniatrie und Pädaudiologie, Erlangen, Germany

daberry@ucla.edu

Introduction

Neuromuscular control of the larynx plays a crucial role in phonation. Previously, using the in vivo canine laryngeal model, the pre-phonatory impact of the intrinsic laryngeal muscles was studied, both individually and in combination (Choi et al., 1993a, 1993b, 1995). Despite the importance of such studies in assessing the contributions of individual intrinsic laryngeal muscles, crude control over the stimulation levels prevented comprehensive studies of neuromuscular control of phonation from being conducted. In particular, these studies implemented only coarse "on-off" or "absent-low-high" settings to stimulate the individual laryngeal muscles. To obtain fine control over laryngeal posturing and phonatory output, graded stimulation of the laryngeal nerves and muscles is needed. Recently, Chhetri et al. (2010, 2012) presented a method to achieve graded stimulation in the in vivo canine larynx. In this study, we utilize this method to study phonation type as a function of the activation of the laryngeal muscles of the larynx, including the cricothyroid muscle (CT), the thyroarytenoid muscle (TA), the interarytenoid muscle (IA), the lateral cricoarytenoid muscle (LCA), and the posterior cricoarytenoid muscle (PCA).

Methods

Using methods of graded stimulation to the intrinsic laryngeal muscles described previously (Chhetri et al., 2010, 2012), we study several phonation types (chest-like, falsetto-like, etc.) as a function of the activation of the intrinsic laryngeal muscles. Specifically, we quantified the three-dimensional vibrations of the vocal folds, the spatial and temporal empirical orthogonal eigenfunctions (Berry et al., 1994), and the spectrum of the resultant acoustic output as a function of these muscular activations using three-dimensional, highspeed, stereoscopic imaging in conjuction with a hemilarynx methodology described previously (Berry et al., 2001; Döllinger & Berry, 2006).

Results

Our results quantify several phonation types in terms of the three-dimensional vibrations of the vocal folds, empirical orthogonal eigenfunctions (Berry et al., 1994), and the spectrum of the resultant acoustic output as a function of the activation of the intrinsic laryngeal muscles.

Conclusions

Despite the longstanding interest of the voice community in understanding neuromuscular control of phonation type from an *in vivo* larynx, few, if any, systematic studies of the phonation type have been performed previously as a function of the activation of the intrinsic laryngeal muscles. Thus, these data yield new insights into how activation of the intrinsic laryngeal muscles influence phonation type.

References

Berry DA, Herzel H, Titze IR, Krischer K (1994). Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions, J Acoust Soc Am 95, 3595-3604.

Berry DA, Montequin DW, Tayama N (2001). High-speed, digital imaging of the medial surface of the vocal folds, J Acoust Soc Am 110, 2539-2547.

Choi HS, Berke GS, Ye M, Kreiman J (1993a). Function of the thyroarytenoid muscle in a canine laryngeal model, Ann Otol Rhinol Laryngol 102, 769-76.

Choi HS, Berke GS, Ye M, Kreiman J (1993b). Function of the posterior cricoarytenoid muscle in phonation: in vivo laryngeal model, Otolaryngol Head Neck Surg 109, 1043-51.

Choi HS, Ye M, Berke GS (1995). Function of the interarytenoid (IA) muscle in phonation: in vivo laryngeal model, Yonsei Med J 36, 58-67.

Chhetri DK, Neubauer J, Berry DA (2010). Graded activation of the intrinsic laryngeal muscles for vocal fold posturing, J Acoust Soc Am 127, EL127-EL133.

Chhetri DK, Neubauer J, Berry DA (2012). Neuromuscular control of fundamental frequency and glottal posture at phonation onset, J Acoust Soc Am 131, 1401-1412.

Döllinger M, Berry DA (2006). Computation of the three-dimensional medial surface dynamics of the vocal folds, J Biomechanics 39, 369-374.

Relative Fundamental Frequency Distinguishes Between Phonotraumatic and Non-Phonotraumatic Vocal Hyperfunction

Elizabeth S. Heller Murray[1], Yu-An S. Lien[2] Jarrad H. Van Stan[3,4], Daryush D. Mehta[3-5], Robert E. Hillman[3-5], J. Pieter Noordzij[1,6], Cara E. Stepp[1,2,6]

[1] Department of Speech, Language, and Hearing Sciences, Boston University,[2] Department of Biomedical Engineering, Boston University, [3] Massachusetts General Hospital Institute of Health Professions, [4] Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital, [5] Department of Surgery, Harvard Medical School,[6] Department of Otolaryngology – Head and Neck Surgery, Boston University School of Medicine

cstepp@bu.edu

Introduction

The voice disorder vocal hyperfunction (VH) can further be divided into two distinct sub-types: *phonotraumatic VH* and *non-phonotraumatic VH* (Hillman, Holmberg, Perkell, Walsh, & Vaughan, 1989). Individuals with *phonotraumatic VH* present with organic changes to their vocal folds (e.g., vocal nodules), whereas individuals with *non-phonotraumatic VH* do not show signs of vocal fold damage (e.g., primary muscle tension dysphonia). Previous research has hypothesized that *phonotraumatic* and *non-phonotraumatic VH* develop due to different mechanisms of hyperfunction (Hillman et al., 1989). The current work examines the ability of an acoustic measure, relative fundamental frequency (RFF), to distinguish between these two different sub-types of VH.

Methods

Speech samples, consisting of repetitions of a vowel – voiceless consonant – vowel instance (e.g., afa), were recorded from three groups: control participants with typical voices (N = 49), individuals with *phonotraumatic VH* (N = 54), and individuals with *non-phonotraumatic VH* (N = 46). RFF was calculated using an automated algorithm in which the instantaneous fundamental frequency of the 10 vocal cycles of the vowel preceding the voiceless consonant (offset) and the 10 vocal cycles of the vowel following the voiceless consonant (onset) was obtained (Lien, 2015; Lien et al. 2015; Stepp, Hillman, & Heaton, 2010).

Results

A mixed design ANOVA was conducted between the three participant groups and each vocal cycle in the RFF offset. Results indicated a significant main effect of group ($p < 0.01$), cycle ($p < 0.01$) and a significant interaction of vocal cycle $\times$ group ($p < 0.01$). A follow-up *Tukey*

post hoc test to determine group differences, revealed that all three groups were significantly different from each other (all $p < 0.01$), with RFF offset values progressively decreasing from controls, to *non-phonotraumatic VH*, to *phonotraumatic VH* (see Figure 1). A mixed-design ANOVA on onset cycles revealed a significant main effect of cycle ($p < 0.01$) and a significant interaction of vocal cycle × group ($p < 0.01$).

Conclusions

Offset RFF was significantly different in individuals with typical voices, *non-phonotraumatic VH*, and *phonotraumatic VH*. Based on previous work (Stepp et al., 2010), we hypothesize that this difference is not just due to differences in vocal fold structure (e.g., the presence or absence of pathology), but rather to the functional differences between the two sub-types of VH.
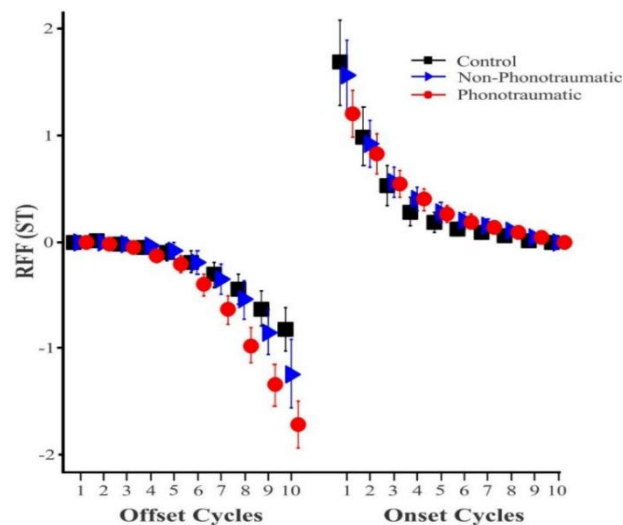
Figure 6: Relative fundamental frequency (RFF) means and 95% confidence intervals for control (black squares), non-phonotraumatic vocal hyperfunction (blue triangles) and phonotraumatic vocal hyperfunction (red circles) participants. All three groups are significantly different for offset RFF.

References

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: an experimental framework and initial results. *Journal of Speech and Hearing Research*, 32(2), 373-392.

Lien, Y. A. S. (2015). *Optimization and automation of relative fundamental frequency for objective assessment of vocal hyperfunction*. (Ph.D.), Boston University, Boston, MA.

Lien, Y. A. S., Calabrese, C. R., Michener, C. M., Heller Murray, E. S., Van Stan, J. H., Mehta, D. D., .Hillman, R. E., Noordzij, J.P., Stepp, C. E. (2015). *Automated algorithms for estimation of relative fundamental frequency in individuals with and without voice disorders.* Paper presented at the 11th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research (AQL), London.

Stepp, C. E., Hillman, R. E., & Heaton, J. T. (2010). The impact of vocal hyperfunction on relative fundamental frequency during voicing offset and onset. *Journal of Speech and Hearing Research*, 53(5), 1220-1226.

The Role of the Epilarynx in Clear Speech Production: An Acoustic Analysis

Elizabeth Godoy[(1)], Nicolas Malyska[(1)], Thomas Quatieri[(1)]

[(1)] MIT Lincoln Laboratory

elizabeth.godoy@ll.mit.edu

Introduction

The epilarynx is the subject of much study in the voice community due to its key role in the production of a speaker's/singer's formant and, more broadly, resonant voice. In production, epilarynx narrowing provides an efficient mechanism to increase vocal loudness (Titze & Story, 1996), specifically by amplifying the cavity's resonance located in a frequency region where human loudness perception is most sensitive (Honda et. al. 2010). The question investigated in this work is whether acoustics of this epilarynx narrowing are observed when humans modify their speech production to increase intelligibility, specifically in a *clear* speaking style (adopted when a listener faces a communication barrier but the speaker does not). Compared to *casual* communication, speakers adopting a clear style typically slow down and hyper-articulate (Hazan & Baker, 2010), with maximal spectral power gains previously observed of up to 5dB (Godoy et. al. 2014). The present study aims more explicitly to pinpoint the role that epilarynx acoustics play in shaping the clear speech spectrum. For this purpose, recently-proposed multi-resolution spectral analyses targeting epilarynx acoustics (Godoy et. al. 2015) are used to compare clear versus casual speech.

Methods

The clear-casual speech dataset analyzed in this study is from the LUCID corpus (Hazan and Baker, 2010) and is referenced in (Godoy et. al. 2014). Eight speakers were asked to read sentences using both clear and casual styles. Methods follow those in (Godoy et. al. 2015), applying full-resolution (e.g. three pitch periods) and glottal closed-open phase spectral analyses. Focus is on a spectral peak in the expected frequency region of the epilarynx resonance (e.g. 3-4kHz) that exhibits the cyclicity property described in (Honda et. al. 2010), i.e. a significant energy difference between the glottal closed and open phases. The *cyclic peak* thus provides an acoustic representation of the epilarynx resonance. Epilarynx narrowing then translates to an amplification of this spectral peak in a frequency region showing a corresponding energy enhancement in the closed-open difference.

Results

Figure 1 illustrates the results of the clear-casual spectral analyses for a representative selection of speakers. For each speaker, the frequency of the estimated cyclic peak matches that of maximal spectral difference between the clear and casual speech. Moreover, the closed-open phase spectral differences show a redistribution of energy towards the cyclic

peak frequency for the clear speech, consistent with acoustics of epilarynx narrowing. These trends were observed for 7 of the 8 speakers examined. The cyclic peak amplitude gains for the clear speech over the casual ranged from +1.4 to 3.7 dB, with +2dB on average.

Conclusions

Analyses illustrate that acoustics reflecting epilarynx narrowing account for the largest spectral differences observed between clear versus casual speech. Results thus indicate that the majority of individuals adopted a manner of speaking ala loudness projection via epilarynx narrowing in their clear speech strategies. A topic for future investigation is to isolate the impact of this production strategy on speech intelligibility.
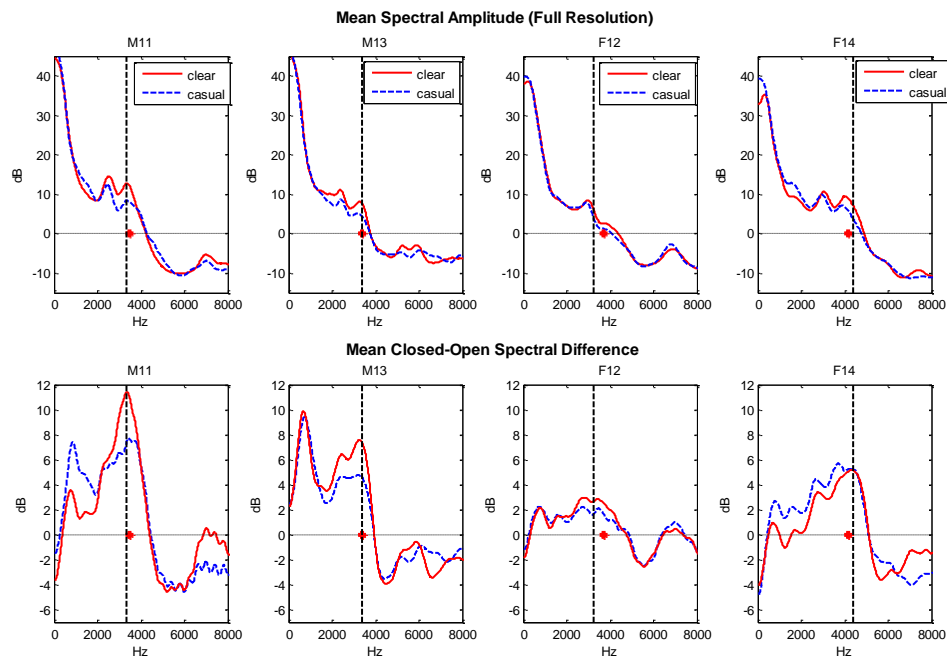
Figure 1. Analysis results for 4 speakers. The estimated cyclic peak frequency for the clear speech of each speaker is indicated with a star in each plot. The black vertical dashed-lines mark the frequency of maximal difference in the clear-casual spectra.

References

Titze, I., & Story, B. (1996). Acoustic interactions of the voice source with the lower vocal tract. J. Acoust. Soc. Am., 101(4), 2234–2243.

Honda, K., Kitamura, T., Takemoto, H., Adachi, S., Mokhtari, P., Takano, S., Nota, Y., Hirata, H., Shimada, Y., Fujimoto, I., Masaki, S., Fujita, S., & Dang, J. (2010). Visualisation of hypopharyngeal cavities and vocal-tract acoustic modelling. Comp. Methods in Biomech. & Biomed. Engineering, 13(4), 443–453.

Godoy, E., Malyska, N., & Quatieri, T. F. (2015). Estimating lower vocal tract features with closed-open phase spectral analyses. Interspeech, 771–775.

Hazan, V., & Baker, R., Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. J. Acoust. Soc. Am., 130(4), 2139–2152.

Godoy, E., Koutsogiannaki, M., & Stylianou, Y. (2014). Approaching speech intelligibility enhancement with inspiration from lombard and clear speaking styles. Computer Speech & Language, 28(2), 629–647.

53. Vocal fold contact pressure estimation over laryngeal high speed videoendoscopy based on a vocal edge tracking method using Kalman Filter and Hertzian Impact model

Manuel Díaz-Cádiz [1], Gabriel E. Galindo [1], Víctor M. Espinoza [1], Matías Zañartu [1]

[1] Department of Electronic Engineering, Universidad Técnica Federíco Santa María, Chile

manuel.diazc@alumnos.usm.cl

Introduction

Finding objective methods to improve the analysis and clinical diagnosis is critical in the assessment of vocal function to take advantage of recent imaging approaches, including laryngeal high-speed videoendoscopy (HSV). Having detailed kinematics allow for studying the mechanical behavior of the vocal folds. The vocal fold collision pressure (or force) could present useful information related to vocal abuse and the development of benign lesions such as nodules. Estimating the contact pressure directly from HSV recordings is the focus of this work.

Methods

In this study, we estimate the contact pressure using the Hertz law (Stronge 2000), which describes the pressure between two elastic bodies in terms of the effective contact surface and apparent penetration. Previous studies have used this approach in numerical models (Gunter 2003), and physical models (Chen et al 2011). Correlation analysis has also been used to estimate the mechanical stresses on physical models (Spencer 2008). However, in this work the area of contact and penetration are obtained from endoscopic videos using computer vision techniques.

The proposed method consists in a vocal edge detection plus tracking with occlusion handling during the collision phase. A polynomial fit is done to register each edge. The tracking module uses a Kalman filter applied over the polynomial coefficients providing the overlap behavior needed during the collision stage. A lumped spring-mass-damper model is used for predicting the edge trajectory during this phase. Finally, this edge description is analyzed to extract estimations of contact area and apparent penetration, and then predict force and contact pressure with Hertz.

Two types of videos are tested: synthetic videos obtained from a numerical model of voice production (Zañartu et al., 2014) with simulated cases of normal and hyperfunctional phonation; and real HSV recordings of some vocal gestures with increasing loudness in a normal subject. The propose of synthetic data is to use as benchmark test for the method, using simulated results from numerical models as reference.

Results

Preliminary results for a real and simulated endoscopic HSV recordings are shown in Fig.1, wherein kymograms and collision predictions are quantified for a sustained vowel. The method can track folds further the contact instants, allowing penetration estimation, as can be see it in kymograms. Similar force and pressure are found in synthetic case. We presume potential predictions on real cases are reachable, according calibration. Sampling rate and resolution are 8750fps, 301x301 pixels for synthetic case and 8000fps, 255x140 pixels for this HSV record. The results in this particular HSV recording can be differ, because calibration parameters for this case are pending.
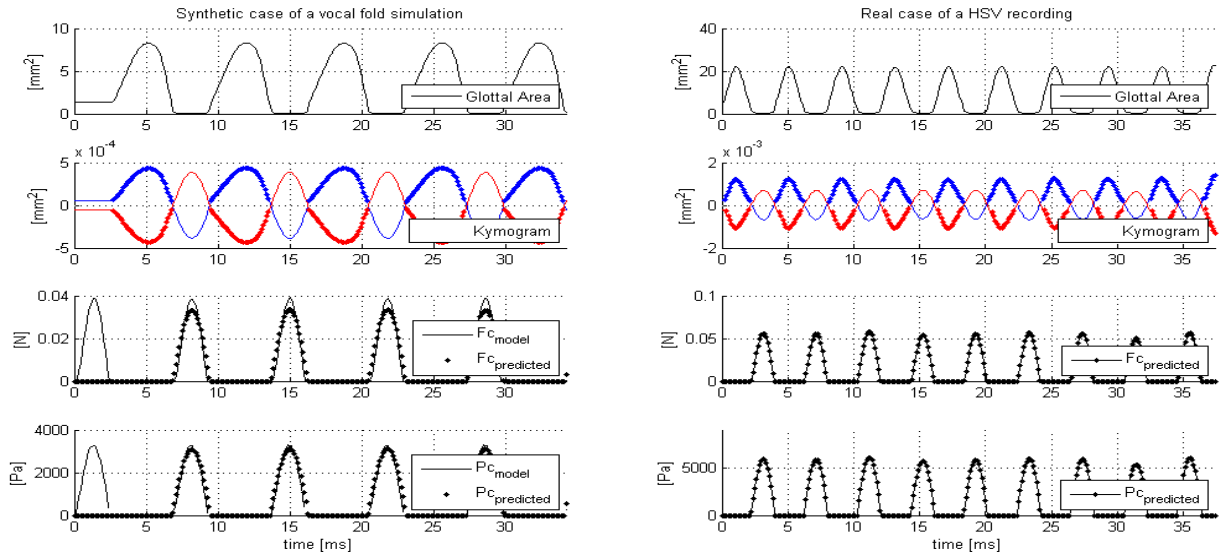
Fig. 1 Preliminary results of the method applied to synthetic and real videoendoscopies.

Conclusions

Current preliminary results of the collision force magnitude are in agreement with previous studies (Gunter, 2003; Zañartu et al., 2014). This indicates that the Kalman filtering with lumped-model approach is feasible for this purpose. Analysis of the synthetic videos and further validation is pending. The proposed method is promising for enhancing the objective assessment of vocal function, and potentially useful for stroboscopic systems.

References

Stronge, W. L. (2000). "Impact Mechanics". Cambridge UK: University Press. Chap. 6, pp. 116-127.

Gunter, H. E. (2003). "A mechanical model of vocal-fold collision with high spatial and temporal resolution". J. Acoust. Soc. Am., 113(2), 994-1000.

Chen, L.J., Mongeau, L. (2011) "Verification of two minimally invasive methods for the estimation of the contact pressure in human vocal folds during phonation". J. Acoust. Soc. Am., 130 (3), 1618-1627.

Spencer, M., Siegmind, T., Mongeau, L. (2008) "Determination of superior surface strains and stresses, and vocal fold contact pressure in a synthetic larynx model using digital image correlation". J. Acoust. Soc. Am. 123(2), 1089-1103.

Zañartu, M., Galindo, G., Erath, B. D., Peterson, S. D., Wodicka, G. R., Hillman R. E. (2014) "Modeling the effects of a posterior glottal opening on voca

Phonation energy utilization and efficiency

Michael H. Krane[1,2,3]

[1] Applied Research Laboratory, Pennsylvania State University,[2] Department of Biomedical Engineering, Pennsylvania State University, [3] Department of Acoustics, Pennsylvania State University

mhk5@psu.edu

Introduction

The need to define and characterize improved measures of phonation energy utilization and efficiency has long been recognized (see, e..g, Titze, 1992; Berry *et al*., 2001).   This presentation gives the beginning of a theoretical foundation for definitions of aeroacoustic and vibration efficiencies, and a path forward to quantifying the approximations needed to perform measurements of efficiency in the clinic.

A control volume analysis of energy utilization in phonation is presented. Conversion of subglottal airstream potential energy into work done vibrating the vocal folds, air flow through the glottis, and radiating sound are described. An approximate numerical model is used to compute the contributions of each of these mechanisms, as a function of subglottal pressure, for normal phonation. An efficiency measure for each energy conversion mechanism is proposed.

Methods

A simplified mechanical energy equation for fluid motion in the glottal control volume, shown in Figure 1, is derived by a scaling analysis and suitable assumptions (see, e.g., Krane and Wei, 2006; Krane, 2015). The simplified equation delineates the principal effects that convert energy of the subglottal airstream into work or energy loss.

A numerical model (Krane, 2015), which computes glottal flow and sound production for prescribed vocal fold motion in an infinite duct, is used to estimate the relevant velocity and pressure distributions needed to compute the terms in the energy equation.

Results

The theoretical analysis yields the following equation describing the rate of change of air kinetic energy in the laryngeal control volume:

$$\underbrace{2\left(p_A{}^+ Q_A - p_D{}^- Q_D\right)}_{\substack{\text{work done by driving pressure force} \\ \text{on fluid in CV} \\ \text{(input)}}} \approx \underbrace{\frac{\rho c}{S} Q_A{}^2}_{\substack{\text{work done by fluid in CV} \\ \text{on trachea acoustic field (loss)}}} + \underbrace{\frac{\rho c}{S} Q_D{}^2}_{\substack{\text{work done by fluid in CV} \\ \text{on vocal tract acoustic field} \\ \text{(acoustic output)}}} + \underbrace{\wp_{VF}}_{\substack{\text{work done by fluid in CV} \\ \text{on vocal folds} \\ \text{(non-acoustic output)}}}$$

The role of each term is indicated underneath, both in terms of its function in the equation, and its role as an input, output, or loss. The latter designation is driven by the perspective that the useful work includes both acoustic output into the vocal tract, and vocal fold vibration, whereas acoustic output back into the lungs is considered an energy loss because it contributes, at best, indirectly to the chosen outputs. This perspective also motivates the following definitions of efficiency:

$$\eta_{acoustic} = \frac{\frac{\rho c}{S} Q_D{}^2}{2\left(p_A{}^+ Q_A - p_D{}^- Q_D\right)}, \qquad \eta_{VF} = \frac{\wp_{VF}}{2\left(p_A{}^+ Q_A - p_D{}^- Q_D\right)}$$

Waveforms of the power terms are shown in Figure 2, over the course of a cycle.
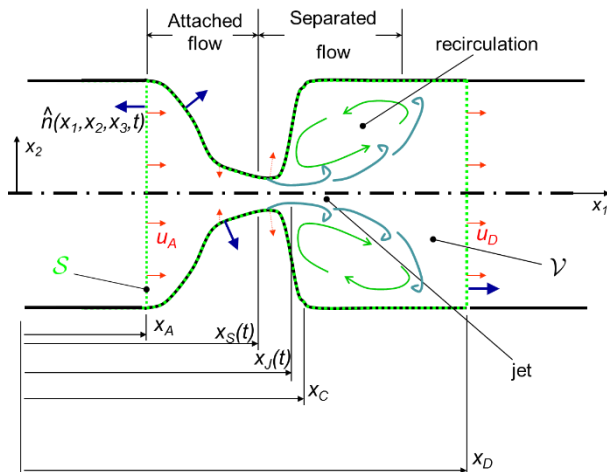
Figure 1. Control volume containing air in larynx, after Krane, 2015.
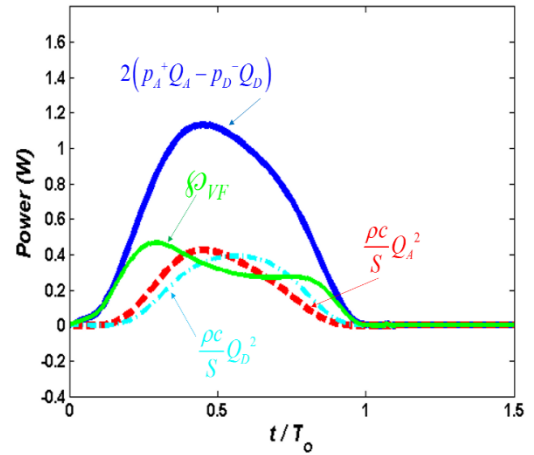
Figure 2. Waveforms of terms in energy equation: Blue: power supplied by driving pressure force, Red: acoustic power radiated to lungs, Cyan: acoustic power radiated to vocal tract, Green: power input to vocal folds

Conclusions

The integral energy equation for air in the laryngeal control volume was simplified to four dominant pressure work terms, and the roles of each were identified. Definitions of acoustic and vibration efficiency were proposed. Evolution of each term was computed using an approximate glottal aerodynamic-aeroacoustic model.

References

Berry, D. A., Verdolini, K., Montequin, D. W., Hess, M. M., Chan, R. W., & Titze, I. R. (2001). A quantitative output-cost ratio in voice production. *Journal of Speech, Language, and Hearing Research*, *44*(1), 29-37.

Krane, M. H., & Wei, T. (2006). Theoretical assessment of unsteady aerodynamic effects in phonation. *The Journal of the Acoustical Society of America*, *120*(3), 1578-1588.

Krane, M.H., (2015) Aeroacoustic sources in phonation – a control volume approach. Submitted to *The Journal of the Acoustical Society of America*.

Poster Presentation

Electroglottographic assessment of
In vivo Japanese Macaque sound production

Christian T. Herbst[1,2], Hiroki koda[3], Takumi Kunieda[3],

Juri Suzuki[3], Takeshi NIshimura[3]

[1] Bioacoustics Laboratory, Department of Cognitive Biology, University Vienna,

[2] Voice Research Lab, Department of Biophysics, Palacký University,

(3) Primate Research Institute, Kyoto University

herbst@ccrma.stanford.edu

Introduction

While the call repertoire of Japanese Macaques (*Macaca fuscata*) has been described based on acoustic evidence (Green, 1975), little is known about the underlying laryngeal function, mostly due to experimental difficulties in vivo. As an alternative to direct laryngeal observation, vocal fold vibration can be assessed non-invasively with electroglottography (EGG). A low intensity, high-frequency current is passed between two electrodes placed on each side of the larynx. The admittance variations resulting from vocal fold (de)contacting during laryngeal sound production are largely proportional to the time-varying relative vocal fold contact area (Hampala et al., 2015).

Methods

Here, we present the results of a pilot study with a female Japanese Macaque who was trained to vocalize upon a visual stimulus. Calls were classified as suggested by Green (1975). A total of 369 "coo" calls, 31 "growls", 5 "chirps", and 26 transitions between "coos" and "grunts" were documented with SPL-calibrated microphone signals and simultaneous EGG recordings.

Results

The "coo" calls had a mean fundamental frequency (F0) of 580.5 Hz ($\pm$44.7, only periodic signals considered, n = 368), a mean dominant frequency (DF) of 677.4 Hz ($\pm$202.2), and a mean sound pressure level (SPL) of 83.0 dB(C) at 10 cm ($\pm$5.2). The respective mean F0, DF and SPL values for the other two call types were: 304.5 Hz ($\pm$117.9, n = 17), 419.7 Hz ($\pm$291.1), and 80.2 dB ($\pm$4.6) for growls; and 1001.2 Hz ($\pm$88.6, n = 3), 2591.4 Hz ($\pm$775.2), and 92.3 dB ($\pm$3.4) for chirps – see Figure 1 A and B. In 83.5% and 11.9% of all cases, respectively, the DF was found at F0 and at twice F0 ($\pm$ 10%) – see Figure 1 C, suggesting that the vocal tract was modified inbetween phonations. In the coos and the grunts, an EGG signal with cyclic content corresponding to the microphone signal was found. The absence

of an EGG trace for the high-frequency chirps might have been caused by a low-pass filter in the EGG device hardware. For the 26 coo/grunt phonations, and the EGG evidence suggests that the transitions between the individual call types occurred over as little as one to five vibratory cycles – see Figure 2 for an example.

Conclusions

The collected evidence suggests that the coos and the grunts might constitute distinct laryngeal mechanisms (comparable to "registers" in human singing), potentially generated with the same vibrating structures. Excised larynx experiments are warranted to test this hypothesis, also investigating the potential influence of the species' vocal membranes.

References

Green, S. (**1975**). "Variation of Vocal Pattern with Social Situation in the Japanese Moneky (*Macaca fuscata*): A FieldStudy," in *Primate Behaviour. Developments in Field and Laboratory Research*, edited by L. A. Rosenblum (Academic Press, New York), pp. 1-102.

Hampala, V., Garcia, M., Svec, J. G., Scherer, R. C., and Herbst, C. T. (**2015**). "Relationship between the Electroglottographic Signal and Vocal Fold Contact Area," Journal of Voice **in press**.
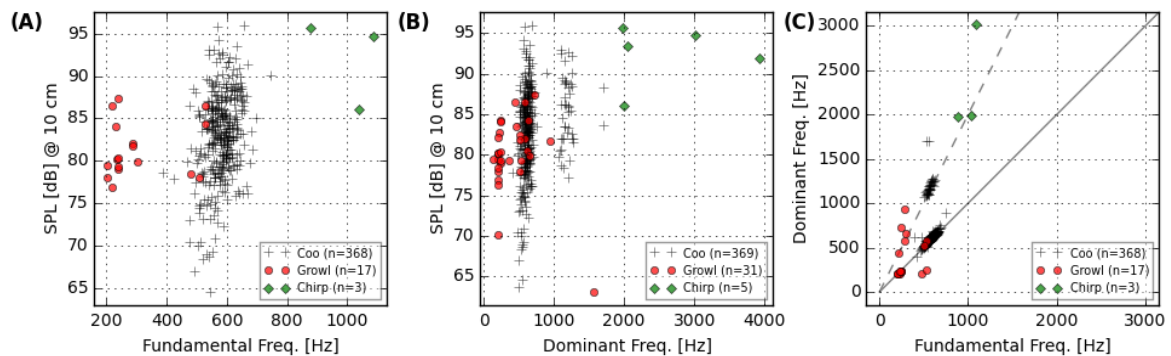
Figure 7: Summary plots of recorded phonations: (A) Fundamental frequency (F0; only periodic signals considered) vs. sound pressure level (SPL); (B) Dominant frequency (DF) vs. SPL; (C) F0 vs. DF (For data points along the solid and dashed lines, the dominant frequency matches the first and second harmonic, respectively).
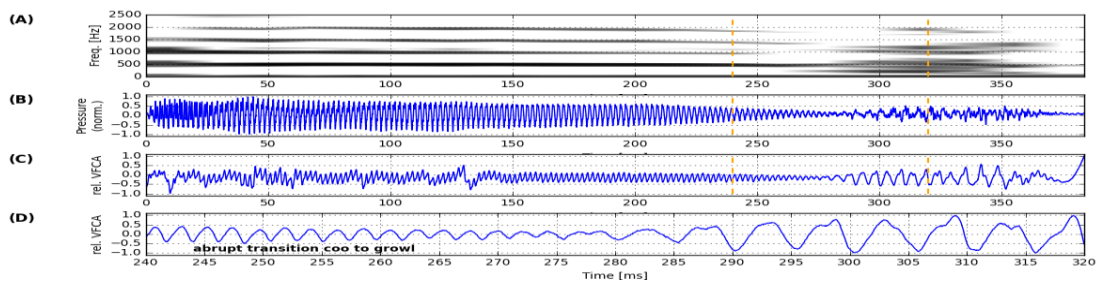
Figure 8: Abrupt transition from "coo" to "growl". (A) Spectrogram of acoustic signal; (B) acoustic signal; (C) EGG signal; (D) portion of the EGG signal, extrated at 240 – 320 ms - see dashed vertical lines in panels (A) – (C).

Cross-cultural adaptation of the Chilean version of the Voice Symptom Scale – VoiSS

Francisco Contreras Ruston[1,4], Felipe Moreti[1], Martin Vivero[2], Celina Malebran[3], Mara Behlau[1]

[1]Centro de Estudos da Voz – CEV – São Paulo (SP), Brasil.,[2]Hospital del Salvador, [3]Universidad Santo Tomás, [4]Universidad de Valparaíso

fco.contreras.ruston@gmail.com

Introduction

The concern with the quality of life of individuals, described at the beginning of the 90 decade[WHO1997], has excelled with dysphonic patients in the last two decades with self-assessment protocols impact of dysphonia[1]despite advances in voice evaluation, none of the tests used both auditory and acoustic indicates the real dimension of behavior that the individual mentioned regarding the quality of life[2].For further information on the development of patients in therapy they were created widespread self-assessment protocols in across the world[3]. This research aims to accomplish the cultural equivalence of the Chilean version of VoiSS protocol, through their cultural and linguistic adaptation.

Methods

A translation was done VoiSS to Chilean Spanish by two bilingual speech therapists and the back translation for English, a comparison was conducted of the items of the original instrument with the previous translated version, the existing discrepancies were modified by a consensus committee of five speech therapist in which it was called "Escala de Sintomas Vocales" – ESV, with 30 questions and five answers: "Never", "Occasionally", "Some of the time", "Most of the time", "Always". For the cultural equivalence, it was applied to 15 individuals with vocal problems. In each question it was added the option of "Not applicable" in the answer choices for identification of the questions not comprehended or not appropriate for the concerned population. Two of the individuals had difficulty at the time of answering two questions, which it was only necessary to adapt the translation of one of them. The modified ESV was applied to three individuals with vocal problems, not finding incomprehensible questions nor inappropriate for the Chilean culture.

Results

The ESV reflects the original English version, both in the number of questions or relating to the domains in which it contains impairment, emotional and physical.

Conclusions

There is cultural equivalence of VoiSS for the Chilean Spanish, titled ESV. The validation of the ESV for Chilean Spanish is already in process of conclusion.

References

[1] Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS et al. The Voice Handicap Index (VHI): development and validation. Am J Speech Lang Pathol. 1997;6(3):66-70.

[2] Branski RC, Cukier-Blaj S, Pusic A, Cano SJ, Klassen A, Mener D, Patel S, Kraus DH. Measuring quality of life in dysphonic patients: a systematic review of content development in patient-reported outcomes measures. J Voice. 2010;24(2):193-8.

[3] Behlau M, Madazio G, Feijó D, Pontes P. Avaliação da Voz. In: Behlau M (org.). Voz: O Livro do Especialista. Rio de Janeiro: Revinter; 2001. v.1. p.83-245.

[4] Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). J Voice. 1999;13(4):557-69.

Phonatory characteristics and Voice quality evaluation of Laryngeal dystonia before and after Botulinum toxin treatment. A case study

Manfred Pützer [(1)] Wolfgang Wokurek [(2)]

[(1)]Neurophonetik & Klinische Phonetik, Universität des Saarlandes,

[(2)]Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, Stuttgart, Deutschland

puetzer@coli.uni-saarland.de

Introduction

Adductor spasmodic dysphonia (laryngeal dystonia) is characterized by irregular hyperadduction of the vocal folds leading to a strangled and hoarse voice quality with breaks in pitch and phonation. Accordingly, the acoustic output is characterized by amplitude and frequency perturbation and diplophonia in parts of the signal [e.g., 1, 2]. Diplophonia is a beat frequency phenomenon caused by the independent vibration of the two vocal folds at slightly different frequencies. The pathophysiology of adductor spasmodic dysphonia is largely unknown. fMRI studies show reduced activities in the sensorimotor cortex and supplementary and pre-supplementary motor areas, respectively [e.g., 3, 4]. Firstly, the aim of the present study is to visualize the glottal excitation for phonation of an [a:] produced by one female patient (adductor type) with laryngeal dystonia before and after botulinum toxin (BTX) treatment, and secondly, to assess the observed phonatory characteristics using a new parametrization of voice quality properties in the acoustic signal.

Methods

Voice quality is parameterized in this study by the bandwidth of the first formant and by spectral gradient features of the source spectrum [5, 6]. The more complete compensation of the vocal tract filter uses the first four formant frequencies F1-F4 and formant bandwidths B1-B4 from LPC estimation. All four formants are compensated. The spectrum is calculated using a 25ms Hamming window, long enough to show the harmonic structure.
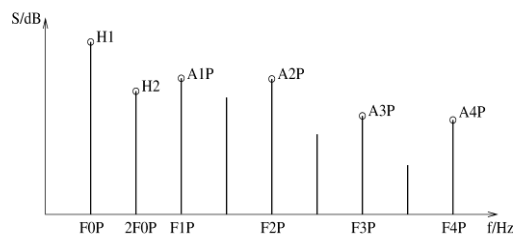
Fig. 9. Schematic harmonic speech spectrum.
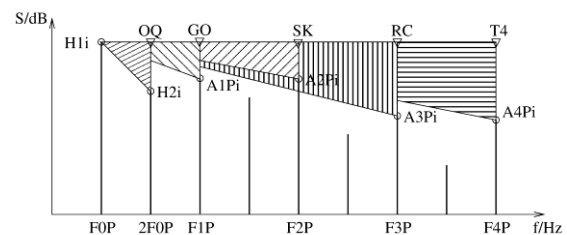
Fig. 2. Schematic harmonic voice spectrum and voice quality parameter definitions.

Figure 1 shows a schematic harmonic speech spectrum. Now the gain of the four formants is subtracted (in decibel amplitude scale) to estimate the source spectrum shown schematically in Figure 2. This inverse filter operation is denoted by the suffix i in all amplitudes. Figure 2 also shows the triangle slopes which are our voice quality parameters OQG opening quotient gradient, GOG glottal opening gradient, SKG skewness gradient, RC rate of closure gradient, and the normalized bandwidth of the first formant IC=B1/F1 incompleteness of closure.

Voice recording was carried out at the Institute of Phonetics of the Saarland University. One female patient (aged 35) with adductor spasmodic dysphonia produced the vowel [a:] at a normal pitch. Electroglottographic (EGG, Portable Laryngograph) and microphone signals (microphone NEM 192.15 Beyerdynamic) were recorded simultaneously (digitized with a 16-bit amplitude resolution and 50 kHz sampling rate) using a Computerised speech Lab (CSL) station.

Results & Conclusions

1. Phonatory characteristics: The acoustic signals (upper) and the EGG signals (lower) are shown in the Figures representing the voice quality (a) before BTX-treatment (hoarseness, diplophonia), (b) five days after BTX-treatment (breathiness), (c) two months after BTX-treatment (modal), and (d) six months after BTX-treatment (state of relapse: hoarseness, diplophonia).

(a)   (b)   (c)   (d)
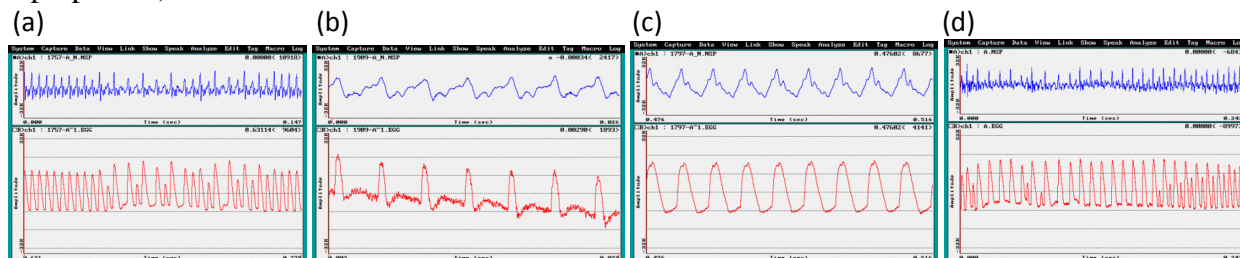
2. Voice quality evaluation: The above mentioned parameters were used for evaluation. Data analysis: ANOVA with posthoc-tests (Scheffé alpha-adjustment) for the effect of the four states.

Analysis approach	Parameter	(a) Before treatment	(b – d) After treatment		
			(b) 5 days	(c) 2 months	(d) 6 months
Glottal openin	GOG	1.83 SD 1.1	4.34 SD 0.6	-3.65 SD 4.52	1.86 SD 0.80
Skewness of glottal pulse	SKG	1.95 SD 0.9	3.77 SD 0.4	0.25 SD 0.38	2.01 SD 0.33

(a)&(d)=Signific. differences compared to (b)&(c); (b)=Signific. differences compared to (c); Nosignific. differences between (a) before treatment and (d) 6 months after treatment ($p<0.001$). The observations and the instrumentally based evaluation of the acoustic signals seem to closely reflect the (patho)physiological behaviour of vocal fold vibration before, and after treatment with BTX. Therefore, they may help to determine the actual voice quality status during this treatment.

Terminology of voice phenomena related to diplophonia

P. Aichinger[1], B. Schneider-Stickler[1], M. Hagmüller[2]

[1] Division of Phoniatrics-Logopedics, Department of Otorhinolaryngology, Medical University of Vienna, [2]Signal Processing and Speech Communication Laboratory, Graz University of Technology

philipp.aichinger@meduniwien.ac.at

Introduction

Terminology with regard to voice quality typing is subject to an ongoing discussion and definitions still remain fuzzy. Of particular interest are nonmodal phonation and diplophonia (Aichinger, 2015, Gerrat and Kreiman, 2001, Titze, 1995). To improve the quality of voice typing a more accurate and standardized terminology is needed. This presentation concerns a proposal for refined terminology of voice phenomena related to diplophonia.

Methods

The proposed terminology is based on observations and experience from a three-years interdisciplinary research engagement, during which audio recordings of 120 subjects had been obtained simultaneously with laryngeal high-speed videos (Aichinger, 2015). Established strategies for successful terminology work are pursued, which include concept analysis and harmonization of concepts and terms (Suonuuti, 2001).

Results

Conflicting definitions have been recognized in the established terminology systems. The term "diplophonia" had been used for two different mental concepts in the literature, namely for a period-two up-down pattern of an arbitrary cyclic parameter (Titze, 1995) and for two simultaneous pitches (Dejonckere, 1983). Additional mental concepts refer to the number of sound sources, the number of glottal oscillators and the number of harmonic series. The concepts have been integrated into the proposed concept system.

The proposed terminology and the relations of its underlying concepts are depicted as a five-dimensional generic concept system diagram (figure 1). We distinguish voice with only one independent sound source ("monophonia") from voice with more than one ("multiphonia"). In particular, voice with two and three independent sound sources is refered to as "biphonia" and "triphonia", which is consistent with Titze's system. It must be noted that the terms may also be used for extralaryngeal sound sources, which contribute to speech sound rather than voice sound. The number of spatially distant glottal oscillators with different fundamental frequencies distinguishes "glottal monophonia", "glottal

multiplophonia", "glottal diplophonia" and "glottal triplophonia". The terms "monoharmonia", "multiharmonia", "biharmonia" and "triharmonia" may be used with regard to the number of harmonic series with different fundamental frequencies. The number of simultaneous pitches with above-threshold salience determines the presence of "auditory monophonia", "auditory multiplophonia, "auditory diplophonia" and "auditory triplophonia". Metacycles occur if an arbitrary cyclic parameter fluctuates in a pattern that repeats itself after an arbitrary number of cycles. We refer to the presence of metacycles as "multicyclia" and to its absence as "monocyclia". In order to establish a clear distinction between the double pitch phenomenon and a metacyclic pattern, we suggest to use the terms "bicyclia" and "tricyclia" for concepts that had been refered to as "diplophonia" and "triplophonia" in Titze's system.
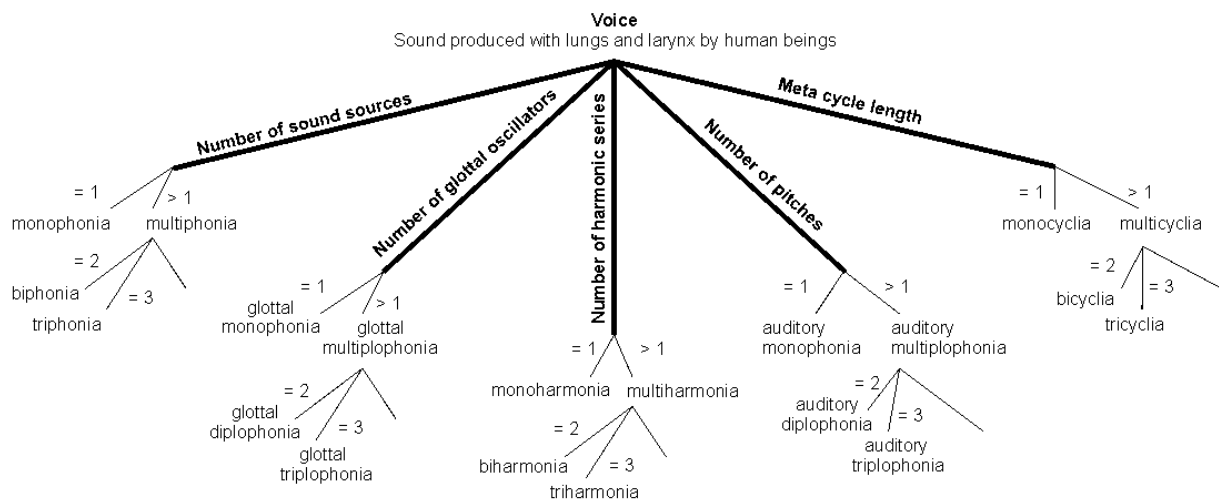
Figure 1: Five-dimensional concept system for voice phenomena related to diplophonia.

Conclusions

A refined terminology for diplophonia related voice phenomena has been proposed. An utmost degree of specificity with regard to the different dimensions of the concept system has been applied. The concepts and terms are harmonized with regard to internal and external consistency. In the future the newly proposed terminology must be evaluated with regard to unambiguousness and efficiacy.

References

Aichinger, P. (2015). Diplophonic Voice - Definitions, models, and detection. Dissertation, Graz University of Technology.

Dejonckere, P. H., & Lebacq, J. (1983). An analysis of the diplophonia phenomenon. Speech Communication, 2(1), 47–56.

Gerratt, B. R., & Kreiman, J. (2001). Toward a taxonomy of nonmodal phonation. Journal of Phonetics, 29(4), 365–381.

Titze, I. (1995). Workshop on acoustic voice analysis: Summary statement. Iowa: National Center for Voice and Speech.

Suonuuti, H. (2001). Guide to Terminology. Helsinki: Nordterm.

Influence on Spectral Energy Distribution of Emotional Expression

Marco Guzman [1], Soledad Correa[2], Daniel Muñoz[1], and Ross Mayerhoff

[1] Universidad de Chile, [2] Universidad de Valparaiso, [3] Department of Otolaryngology-Head and Neck Surgery, Wayne State University

soledad.correa.f@gmail.com

Introduction

In addition to the linguistic messages, the speech acoustic signal also carries information about the identity, age, geographic origin, attitude, and emotional state of the speaker.

Most studies related to emotions in speech have focused on the role of F_0, sound pressure level (SPL), speech rate, segment duration, and overall prosody. The role of voice quality (or timbre) in conveying emotions has been studied to a lesser extent.

Several studies have shown that anger and happiness/joy are generally acoustically characterized by high mean F_0, wider pitch range, high speech rate, increases in high-frequency energy, and usually increases in the rate of articulation. Sadness is characterized by decrease in mean F_0, slightly narrow pitch range, and slower speaking rate.

Therefore, the present study is focused on how emotions are encoded in the speech acoustic signal determined in spectral energy distribution in professional theater actors.

Methods

Thirtyseven actors, native Spanish speakers, were included. All subjects had at least 3 years of professional experience as a theater actor and no history of vocal pathology for the last 5 years. Participants were recorded during a read-aloud task of a 230-word passage, expressing six different emotions (happiness, sadness, fear, anger, tenderness, and eroticism) and without emotion (neutral state). Acoustical analysis with long-term average spectrum included three variables: the energy level difference between the F_1 and fundamental frequency (F_0) regions, ratio between 1–5 kHz and 5–8 kHz, and alpha ratio. The study design is a quasi-experimental study.

Results

All the different emotions differ significantly from the neutral state for alpha ratio and 1–5/5–8 kHz ratio. Only significant differences between ''joy,'' ''anger,'' and ''eroticism'' were found for L1–L0 ratio. Statistically signif- icant differences between genders for the three acoustical variables were also found.

Conclusions

The expression of emotion impacts the spectral energy distribution. On the one hand emotional states characterized by a breathy voice quality such as tenderness, sadness, and eroticism present a low harmonic energy above 1 kHz, high glottal noise energy, and more energy on F_0 than overtones. On the other hand, emotional states such as joy, anger, and fear are characterized by high harmonic energy greater than 1 kHz (less steep spectral slope declination), low glottal noise energy, and more energy on the F_1 than F_0 region.

References

.   Scherer KR. Vocal affect expression: a review and a model for future research. Psychol Bull. 1986;99:143–165.

.   Banse R, Scherer KR. Acoustic profiles in vocal emotion expression. J Pers   Soc Psychol. 1996;70:614–636.

.   Laukkanen A-M, Vilkman E, Alku P, Oksanen H. On the perception of emotions in speech: the role of voice quality. Scand J Logoped Phoniatr Vocol. 1997;22:157–168.

Relationship between sociodemographic variables and parameters teachers Members in Chillan

Dra. Jazmín Pérez Serey [1,2] Dra. Virna Ortiz Araya [3]

[1]Universidad Pedro de Valdivia, [2] Universidad Adventista de Chile, [3] Universidad del Bío Bío.

jperez@upv.cl

Introduction

For the teacher, the voice is the main instrument used to express a message full of content, due to it allows persuading to others and gets their attention(1). Therefore, the risk factors that alter the vocal health have to be clear and they have been of great interest to professionals who support the prevention and rehabilitation of dysphonia in teachers (2-5). The aim of this study is to relate the sociodemographic variables (sex, years of professional practice, weekly hours of working and age) and vocal parameters (maximum phonation time (MFT), Grade Roughness Breathiness Asthenia Strain (GRBAS), Vocal Disability Index (VHI) and harmonic noise ratio (HNR) in teachers from Chillan in order to identify risk factors that alter the parameters of the voice.

Methods

This is a quantitative study with a cross relational descriptive design in 92 teachers from 4 subsidized schools from Chillan. The inclusion criteria that were adopted are active exercise of the profession, complete speech phonological evaluation and sign of the informed consent document. The exclusion criteria were to be out of the exercise of the profession and to have an incomplete speech evaluation. The obtained data were analyzed using Chi- square in 22 SPSS version. Speech evaluation was made considering the Speech Therapy committee basic protocol of European laryngological Society (SEORL) (6), to evaluate the perception it was used the Vocal Handicap Index (VHI-30) (7), the perception of clinician GRBAS was obtained with the scale (8.9), the vocal efficiency was obtained with the TMF in vowel / a / and HNR was obtained recording a sustained vowel / a / using the software PRAAT with a unidirectional microphone Shure brand. The evaluated person was located 10 cm away from de microphone which was connected to a notebook Toshiba Satellite L635-SP3003L with soundcard Mobile Pré Brand M-audio. The data obtained were analyzed using SPSS version 22, using a statistical analysis with chi square. This study was approved by the ethics committee of the Adventist University of Chile (No. 01-2013).

Results

The frequency of women with less than 15 seconds of TMF is greater than men (p = 0.02). Teachers with less than 15 years of professional experience are more likely to have a lower

TMF of 15 seconds (p = 0.03). The VHI is perceived altered in teachers under 40 years of age (p = 0.03). Whereas in the same range, the acoustic parameters shimmer (dB) (p = 0.03) and HNR (0.05) are altered.

The following table summarizes the relationship to the studied vocal parameters versus the sociodemographic variables, comparing the results to others studies founded in the specialized literacy.

Table N°1. Relationship between vocal parameters and demographic variables, discussion of authors in contrast to the results found in this study.

Vocal parameters vs sociodemographic variables	Specialized Literacy	This study
TMF v/s sex	It is more common to find vocal disorders in women than in men (10). The TMF lower than 15 seconds shows insufficiency in the vocal tract (11).	To be a woman increases risk of having a TMF <15.
TMF v/s years of profession practice	It is more common to find voice disorders in early years of the teaching practice (10).	To have less than fifteen years of service have a higher risk of having a TMF <15 seconds.
VHI v/s age	Age has no direct relationship with the presence of hoarseness, but the teachers perceived discomfort and presence of dysphonia (12). There is a weak positive correlation between VHI and teachers´s age (0,063) (13)	To have less than 40 years is a risk factor to perceive an altered VHI.
Shimmer v/s age	Barreto-Munévar et al (2011) found that age has not a direct correlation with the presence of hoarseness, but it is related to the early teaching practice and it increases as working years increases too.	With an error rate of 3% in the shimmer dB, a higher percentage of alteration is observed in the group of 40 years or more. There is a positive correlation between age and the shimmer dB (p = 0.03).
HNR v/s age	An altered HNR denotes a asthenic voice or dysphonic (14)	HNR is observed with a 5% of error compared with 73% of 40 aging people or older who have altered the harmonic noise level.

Conclusion

we conclude that the FMT parameter is significantly related to age and years of professional experience in teaching. Whereas, the VHI Shimmer (dB) and HNR is perceived altered in relation to the age of teachers.

References

(1)Le Huche, Francois., Allalli, André. La Voz, Patología vocal: Semiología y disfonías disfuncionales. Primera Edición Tomo 1. Barcelona: Ed. Masson S.A.; 1993.

(2)Araújo TM, Reis EJFB, Carvalho FM, Porto LA, Reis IC, Andrade JM. Fatores associados a alterações vocais em professoras; Factors associated with voice disorders among women teachers. Cad Saúde Pública. 2008; 24(6):1229-38.

(3)Bruck CC, Peres MA. Alteração vocal auto-referida em professores: prevalência e fatores associados. Rev Saúde Pública. 2011; 45(3):503-11.

(4)Choi-Cardim K, Behlau M, Zambon F. Sintomas vocais e perfil de professores em um programa de saúde vocal. Revista CEFAC. 2010; 12(5):811-9.

(5)Barreto-Munévar DP, Cháux-Ramos OM, Estrada-Rangel MA, Sánchez-Morales J, Moreno-Angarita M, Camargo-Mendoza M. Factores ambientales y hábitos vocales en docentes y funcionarios de pre-escolar con alteraciones de voz. Revista de Salud Pública. 2011; 13(3):410-20.

Overtone measures in the clinic

Mette Pedersen MD [1], Anders Jønsson. [1]

[1]Voice Unit, Medical Centre, Østergade

m.f.pedersen@dadlnet.dk

Introduction

So far voice analysis has not produced any sufficient evidence based clinical diagnostic effect (1, 2, 3) and many voice measures have little significance to pathology. Voice measures are difficult to quantify but overtone analysis software seem promising. The goal of this presentation is to evaluate the possibilities of overtone analysis in pathology.

Methods

Twelve normal persons had overtone analysis carried out with "Sygyt ltd." and "Vocevista" (2008). These methods were compared with the acustical measures of high speed films by "Glottis Analysis Tools", Michael Döllinger, Erlangen (Videos made with Wolf Endocam 5562 Setup from Erlangen, Germany). Both overtone analysers are on the market. The problems in pathology was solved in the way that the fundamental frquency F0 was used for overtone measures.

Results

Normal material for use in our clinic has been made. Comparable results were found with the three methods for formants from 1000 Hz to 5000 Hz, up to 25% variation. Interestingly when "Sygyt Ltd" was compared with "Vocevista" in the same millisecond the variation shrinked to 0% and 7%, which suggests use of the same formulas. Video analysis software measures of the vocal fold movements does not seem to correlate well with the overtone measures.

Conclusions

Overtone analysis softwares proved to be comparable. Further video analysis measures might be needed to yield better results. The implication in pathology of the results are discussed in the oral presentation and future perspectives are included. Evidence based diagnostics of the human voice is needed. Overtones might be useable.

References

1. Pedersen M, McGlashan J (2012) Surgical versus non-surgical interventions for vocal cord nodules (Review) The Cochrane Library. pp. 1-13

2. Leong, K. Hawkshaw, M. Dentchev, D. Gupta, R. Lurie, D. Sataloff, R. Reliability of Objective Voice Measures of Normal Speaking Voices. 2013. Journal of Voice. Vol. 27, No.2. pp. 170-176

3. Roy, N. Barkmeier-Kraemer, J. Eadie, T. Sivasankar, M. Mehta, D. Paul, D. Hillman, R. Evidence-Based Clinical Voice Assessment: A systematic Review. 2012. American Journal of Speech-Language pathology. Vol. 22. pp. 212-226.

Voice type discrimination by analyzing neck surface skin acceleration signals: preliminary results on single vowels.

Zhengdong Lei[1] and Luc Mongeau[1,2]

[1] Department of Mechanical Engineering in McGill Univeristy, [2] CIRMMT lab, 527 Sherbrooke St. West Montreal

zhengdong.lei@mail.mcgill.ca

Introduction

The glottal notch accelerometer (GNA) is a device that can monitor human subjects' long-term voice use [1][2]. The GNA has advantages over microphones and airflow meters, it is less sensitive to airborne acoustic interference and noninvasive manipulation [1]. To investigate features of damaged voice (breathy and pressed) quantitatively, a GNA system was utilized to collect acoustic data for breathy and pressed voices as well as normal voice on one volunteer. Different vocal metrics, such as jitter, shimmer, and short time energy, were calculated by applying advanced digital signal processing algorithms. Differences between these three voice types were identified by comparing these vocal metrics. The results may be useful to investigate the relationship between vocal use and chronic vocal damage in future studies.

Methods

One subject (male, 26 years old) was required to wear the GNA system and perform some phonation exercises, including five single english vowel tokens. Each vowel was produced with three types of voices for a constant period of three seconds and three tokens were done. All the data were recorded by the GNA system and a portable data acquisition (DAQ, NI company) module. A Matlab script was used to detect voice activity and perform data truncation by removing unvoicing segments from the raw data. The truncated data was sequently analyzed by Glottis Analysis Tools (v15.2.1, which is general acoustic analysis tool) to calculate the vocal metrics.

Results

The statistical results are shown in Table 1. In general, the breathy voice type has the lowest Mean Short Time Energy (MSTE, which is defined as the sum of squared amplitude of framed signal) due to the longer period of nonvibration of vocal folds and higher fundamental frequency (f0), shimmer and jitter due to its noise-like feature. Pressed voice has higher

shimmer when compared with normal voice because its vocal fold vibration amplitude is less stable. Normal and pressed voice have similar jitters.

Conclusions

This study was limited to single sustained vowels. Both pressed and breathy voices are less stable than normal voice in terms of f0 and vibration magnitude during phonation. Voice type discrimination in sentences and running speech is much harder to achieve than single vowels due to the complicated factors, such as tones, lexics and artificial phonations. Further work is needed for a larger amount of subjects.

Table 1. Results of data analysis

Vowels	Type	MSTE	Mean f0	Shimmer (%)	Jitter (%)
/a/	normal	77.98	121.35	1.34	1.18
	breathy	0.09	476.63	5.42	28.01
	pressed	114.62	149.39	2.59	0.74
/o/	normal	101.81	110.32	1.36	2.30
	breathy	0.47	382.05	9.48	23.50
	pressed	121.58	123.11	1.79	0.96
/i/	normal	101.02	118.41	2.11	0.91
	breathy	0.11	531.86	6.18	21.30
	pressed	340.99	139.73	3.60	0.91
/u/	normal	64.90	115.39	1.08	1.14
	breathy	0.36	375.16	14.11	20.82
	pressed	73.82	142.07	1.94	1.07
/ə/	normal	70.37	116.59	0.95	0.69
	breathy	0.23	520.89	4.89	19.86
	pressed	72.85	123.11	1.79	0.96

References

[1] Cheyne H. A. (2002). Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck (doctoral dissertation).

[2] Zanartu M., Ho J. C., Mehta D. D., Hillman R. E., and Wodicka G. R. (2013). Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration. IEEE trans. On Audi., Spee., and Lang. Proc., Vol. 21, No. 9, 1929-1939.

Synthetic Multi-Line Kymographic Analysis of High-Speed Videoendoscopy Feature Data

Takeshi Ikuma[1], Melda Kunduk[1,2,3], Daniel Fink[1,3], and Andrew J. McWhorter[1,3]

[1]Department of Otolaryngology–Head and Neck Surgery,
Louisiana State University Health Sciences Center
[2]Department of Communication Sciences and Disorders, Louisiana State University
[3]Our Lady of the Lake Regional Medical Center- Voice Center

mkunduk@lsu.edu

Introduction

Digital kymography (DKG)[1] is perhaps the most commonly utilized technique to visualize vibratory behavior which is captured by a high-speed videoendoscope (HSV). It takes a slice of the video data at a chosen sagittal position and presenting a resulting spatiotemporal data as a two-dimensional image. An inherent weakness of DKG is that it throws away information from all other slices. Multi-line DKG[2] ameliorates this weakness by simultaneously observing multiple sagittal positions of an HSV data. While DKGs are techniques to observe imaging data directly, DKG concept has been extended to the analysis of spatiotemporal feature data.[3] However, such data collected over a dense sagittal grid poses an interpretational challenge with continuous variability along the glottis[4], and outcome measures tend to be holistic such as entropy[3,5] and correlation length[5]. The proposed method synthetically produces a small set of one-dimensional kymographic feature waveforms, which characterize the most dissimilar behaviors across the glottis, enabling subsequent analyses of each waveform and across waveforms.

Methods

The Karhunen-Loève transformation is first applied to spatiotemporal feature data matrix to find their empirical basis functions and corresponding singular values. After eliminating the bases of noise subspaces, the spatial coefficients of the temporal bases are clustered using the *k*-means algorithm to minimize the squared Euclidian distance. The synthetic feature waveforms are obtained by transforming the centroids of the clusters back to the original spatiotemporal space.

Results

The synthesizing process is illustrated with two examples of kymographic glottal width measurements of a normal speaker (Fig. 1) and a speaker with a polyp (Fig. 2). The glottal widths are measured on every HSV frame on every kymographic line separated by one pixel. The kymographic lines are automatically grouped to four groups with distinctive synthetic waveforms. The subharmonic behavior present in the polyp case is present mostly in the posterior, and this irregular behavior is correctly captured only in the strongest synthetic waveform representing the majority of the posterior portion.

Conclusions

The proposed method is shown to transform a dense set of kymographic feature measurements to a handful of synthetic kymographic waveforms to summarize different behaviors observed along the vocal folds. This technique may have an immense benefit in assessing pathological data, in which a highly distinctive behavior may occur only in a localized area of the vocal folds. There are two technical caveats to be aware of. First, it is advised not to use the power of the synthetic waveforms as an outcome measure. The observed power is also a function of the cluster size. Second, uniqueness of these waveforms is not guaranteed as the clustering algorithm may return a suboptimal solution. Multiple clustering runs may be necessary to guarantee the optimal solution.
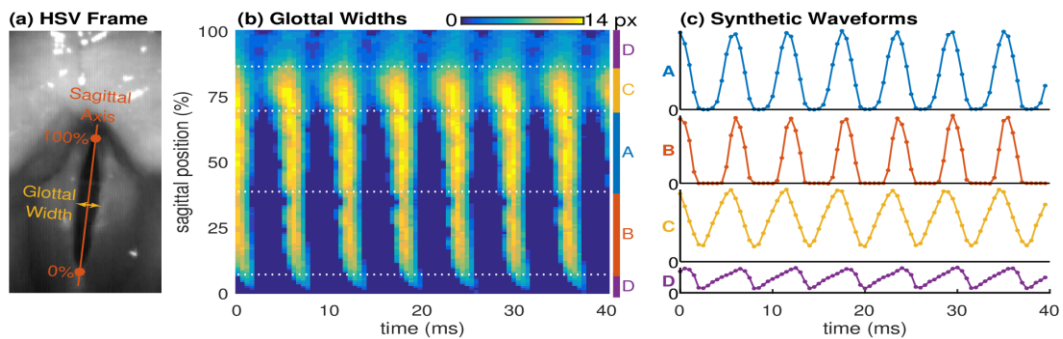
Fig. 2: Normal young male speaker. (a) HSV frame example with sagittal axis annotation, (b) glottal width measurements with four clusters based on glottal width similarity, (c) synthetic glottal width waveforms representing clusters.
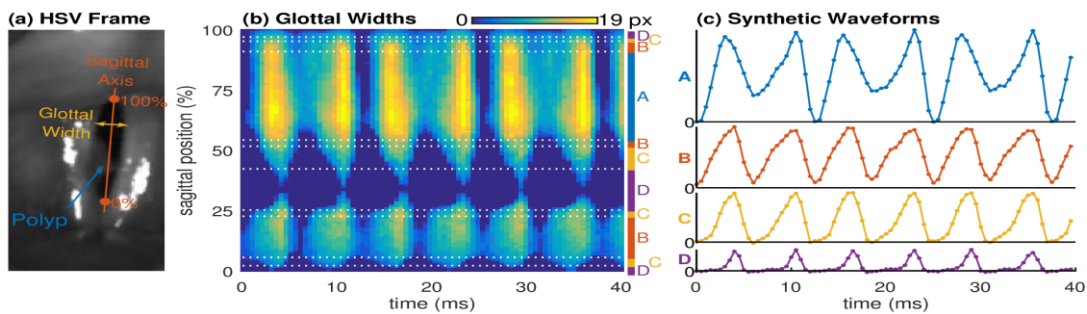
Fig. 3: Male speaker with a polyp on right vocal fold. (a) HSV frame example with sagittal axis annotation, (b) glottal width measurements with four clusters based on glottal width similarity, (c) synthetic glottal width waveforms representing clusters.

References

1. Wittenberg T, Moser M, Tigges M, Eysholdt U. Recording, processing, and analysis of digital high-speed sequences in glottography. *Machine Vision and Applications.* 1995;8(6):399-404.
2. Tigges M, Wittenberg T, Mergell P, Eysholdt U. Imaging of vocal fold vibration by digital multi-plane kymography. *Computerized Medical Imaging and Graphics.* 1999;23(6):323-330.
3. Neubauer J, Mergell P, Eysholdt U, Herzel H. Spatio-temporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes. *J Acoust Soc Am.* 2001;110(6):3179-3192.
4. Lohscheller J, Švec JG, Döllinger M. Vocal fold vibration amplitude, open quotient, speed quotient and their variability along glottal length: Kymographic data from normal subjects. *Logopedics Phoniatrics Vocology.* 2013;38:182-192.
5. Zhang Y, Jiang JJ. Spatiotemporal chaos in excised larynx vibrations. *Physical Review E.* 2005;72(3):035201.

Effects of Volume, Pitch and Phonation Type on Oscillation Initiation and Termination Phases

Melda Kunduk[1,2,3],Takeshi Ikuma[1], Daniel Fink[1,3],and Andrew J. McWhorter[1,3]

[1] Department of Otolaryngology–Head and Neck Surgery,
Louisiana State University Health Sciences Center
[2] Department of Communication Sciences and Disorders, Louisiana State University
[3] Our Lady of the Lake Regional Medical Center- Voice Center

mkunduk@lsu.edu

Introduction

Clinicians and musicians have been long interested in the voice initiation patterns in their clients from the assessment and therapeutics points of vocal function1,2. The initiation and termination of vocal fold (VF) vibratory patterns might be sensitive to changes caused by a voice disorder and expose the limitations of a phonatory system, thus allowing more accurate assessment of vocal function in a disease process. The goal of this work was to investigate the effects of varying volume, pitch, and phonation types (imitated pressed and breathy) on the initiation and termination phases of VF oscillation.

Methods

The high-speed videoendoscopy (HSV) data was recorded at 8000 frames/second using a KayPENTAX HSV system (Model 97109, Montvale, NJ) while a female subject produced pressed and breathy voices and the voices with varying pitch and loudness, including high pitch and loud (HPL), high pitch and comfortable loudness (HPNL), comfortable pitch and loud (NPL), comfortable pitch and comfortable loudness (NPNL). Each voice type was collected six times using block randomization. The vibratory events that are investigated during the initiation and termination of oscillation phases are (Event I) onset/offset of oscillation, (Event II) gain/loss of full-length oscillation, and (Event III) gain/loss of full-amplitude of oscillation. Events I and II were detected based on the glottal vibrating length3 and Event III was detected based on the glottal area root-mean-square (RMS) as illustrated in Fig. 1. Three outcome variables were assessed: the duration between Events I and II (T_{len}), the duration between Events I and III (T_{area}); and the duration between Events II and III (T_{Δ}). Mixed model analyses were conducted in SAS (SAS Institute Inc. v.9.4).

Results

The conditional mean values for the variables during initiation and termination of oscillation are presented in Table 1. Voice type effect was present for all variables during the oscillation initiation phase ($p < 0.05$). Specifically, significant pitch-by-volume interaction effect ($p < 0.01$) was found for T_{area} and T_{Δ}, and the phonation effect (NPNL vs. Breathy vs. Pressed) was found significant ($p < 0.05$) for T_{len} and T_{area}. No significant effect was found in the termination phase. Full length oscillation was always achieved earlier than full-amplitude of oscillation for all voice types during both the initiation and termination of

oscillation phases as demonstrated by the conditional means of outcome parameter measurements.

Conclusions

The study results demonstrated that phonation type significantly affects the parameters during the initiation phase but not during the termination phase. This finding warrants the investigation of whether these parameters during the oscillation termination phase can be used to investigate the effect of pliability of vocal folds and change in vocal function pre-post behavioral, medical, surgical interventions for voice disorders.
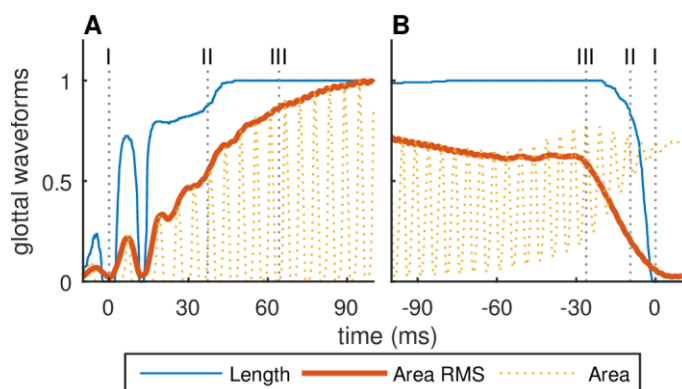
Fig. 1. Illustration of the glottal waveform behaviors during initiation (A) and termination (B). Detected oscillatory events are marked on each phases: Event I-onset/offset, Event II-gain/loss of full-length oscillation, Event III-gain/loss of full-amplitude.

Table 1. Conditional Means of Outcome Parameter Measurements

Segment	Variable	Units	Voice Type[†]					
			NPNL	NPL	HPNL	HPL	Breathy	Pressed
Initiation	T_{len}	ms	29	33	21	35	15	79
	T_{area}	ms	66	55	43	91	33	82
	T_{Δ}	ms	37	22	22	57	18	3
Termination	T_{len}	ms	11	10	16	10	16	36
	T_{area}	ms	33	49	39	39	44	53
	T_{Δ}	ms	22	38	23	29	28	17

† Voice Types: NP - Normal Pitch, HP - High Pitch; NL - Normal Loudness, L – Loud

References

1. Eysholdt U, Tigges M, Wittenberg T, Pröschel U. Direct evaluation of high-speed recordings of vocal fold vibrations. *Folia Phoniatr Logop.* 1996;48(4):163-170.
2. Freeman E, Woo P, Saxman JH, Murry T. A comparison of sung and spoken phonation onset gestures using high-speed digital imaging. *Journal of Voice.* 2012;26(2):226-238.
3. Ikuma T, Melda K, Fink D, McWhorter AJ. A Spatiotemporal approach to the objective analysis of initiation and termination of vocal-fold oscillation with high-speed videoendoscopy. *Journal of Voice.* In Press.

Use of the PROMPT System® in voice training

Ma. Eugenia Ferrando [1], Paola Silberman [1]

[1]Universidad Católica de Santa Fe

meferrando@fibertel.com.ar

Introduction

The Prompt System® (Prompts for restructuring oral muscular phonetic targets)(1) has been used by speech therapists since mid-1980s with abounding research which verifies its effectiveness in speech movement disorders rehabilitation. There is no research related to interventions in voice pathologies or education. This case study shows that the use of the technique produces a better conformation of the vocal tract in case of normal voices. It remains to delve into its effectiveness as regards vocal disorders.

Methods

Fifteen second-year students from the broadcaster course were selected and sound samples from a soundproof recording booth were taken, using one-way dynamic microphones and PRAAT software. They were asked to produce the sound "E" for as long as possible. Bandwidth graphics of the more stable eight seconds were extracted. They were required to put their middle finger in the Prompt® point corresponding to that phoneme (the middle part of the floor of the mouth) applying the counterforce with the mylohyoid muscle while emitting the sound "E", repeating this action until it reaches one minute. Subsequently, they were asked to emit "E" in the same way as before and it was compared with the initial sound. The same methodology was performed with the vocalic sound "U".

Results

The results showed an increase of energy in the formants for both sounds and a larger definition of the glotic pulses after performing the maneuver for one minute as it can be observed in the examples. This would be translated into a better amplification of acoustic energy on the part of the vocal tract. Concurrently, the presence of a greater definition of the glotic pulses would show a better regularity with which the vocal cords vibrate, as Nuñez Batalla says: "the regularity of the phonation or the regularity with which the vocal cords open and close can be valued for the existence of vertical grooves in the spectrogram. The regularity in these grooves is related to the synchrony with which the periodic opening and closure of the vocal cords is produced. It is said that every vertical groove coincides with a glotic pulse."(2)

Sound "E" prior to Prompt®   Sound "E" post Prompt®
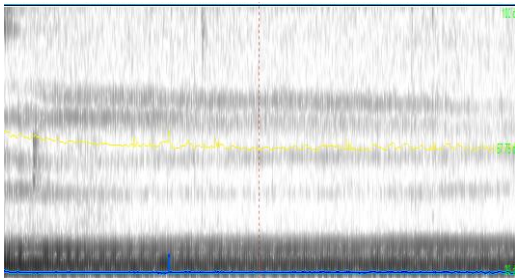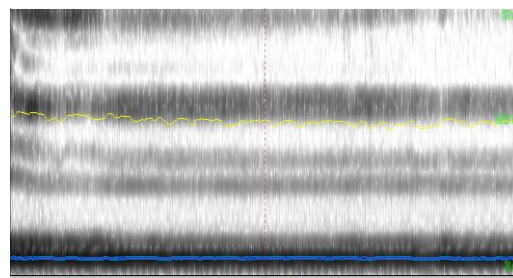
Conclusions

The speech forms an acoustic continuous produced by an uninterrupted movement of the phonatory system organs, going from one configuration to the other.(3) The vocal tract works as an amplifying filter of the sound produced by the vocal cords. The clear presence of the formants would indicate objectively the position adopted by the articulators. Because of the difficulty that represents for a person to identify the position adopted by the middle and rear areas of the mouth, it is important to provide with tactile external information which produces the changes required in the vocal tract. The use of the Prompt System® in the vocal sounds constitutes a valid tool as regards vocal training when it is applied to people with no vocal disorders. Its effectiveness applied to vocal disorders remains to be tested.

References

(1)Chumpelik Débora, M. A. (1984) The Prompt System of Therapy: theoretical framework and applications for developmental apraxia of aspeech .Seminars in Speech and Languge. Vol. 5 Num. 2 Pag. 95

(2)Cobeta, I. Nuñez Batalla, F. Fernandez , S. Patologías de la Voz.(2013) Marge Médica Books. Barcelona.

(3)Farias, Patricia . Ejercicios para restaurar la función vocal.(2007) Ed. Akadia. Bs. As.

Jakcson Menaldi, Ma. C. La voz patológica (2002) Ed. Médica Panamericana. Madrid

Vocal Tremor in Parkinson and Control Speakers

Christophe Mertens[1], Francis Grenez[1], Alain Ghio[2], François Viallet[3], Sabine Skodda[4], Jean Schoentgen[1,5]

[1]Laboratoires d'Images, Signaux et Acoustique, Université Libre de Bruxelles, Belgium, [2]Laboratoire Parole et Langage, Aix-Marseille University, France, [3]Service de Neurologie, Pays d'Aix Hospital, France, [4]Department of Neurology, Knappschaftskrankenhaus, Ruhr-University of Bochum, Germany, [5]National Fund for Scientific Research, Belgium

jschoent@ulb.ac.be

Introduction

The objective of the presentation is to report the size of vocal cycle length jitter and the depth and frequency of vocal cycle length tremor in Parkinson and control speakers. Vocal tremor here designates perturbations the typical frequency of which is larger than 2 Hz and smaller than 15 Hz. The motivations are that few studies have been devoted to vocal tremor in dysphonic or euphonic speakers, most of the corpora that have been studied are small and few studies have addressed jointly vocal jitter and vocal tremor.

Methods

The total number of Parkinson and control speaker recordings has been equal to 661 and 197 respectively in the two corpora (A & B) of sustained vowels [a] that have been analyzed.

Vocal cycle length analysis rests on the sample saliences of sustained vowels. The salience is the duration over which a signal sample is a local maximum. The saliences of vocal cycle peaks enable detecting candidate peaks reporting the instants of glottal excitation loose from the evolving amplitude of the speech sounds (Mertens et al., 2009). The tracking of the vocal cycle lengths involves candidate peak triplets among which a subset is selected via dynamic programming so as to obtain the smallest overall cycle length perturbations (Mertens et al., 2010). Each cycle length time series is decomposed into empirical modes, which are clustered via their instantaneous amplitudes and frequencies into vocal jitter and neurological and physiological tremor (Mertens et al., 2013). The perturbation cues are the relative size of vocal jitter, the relative depths in % of the neurological and physiological vocal tremor and the typical neurological tremor frequency in Hz. The latter is the average of the instantaneous frequencies of the empirical modes assigned to neurological tremor. The instantaneous frequencies are weighted by the complex instantaneous amplitudes.

Results

The neurological tremor depth is statistically significantly higher for female Parkinson speakers than for female control speakers. The neurological tremor frequency differs statistically significantly between male and female speakers and increases statistically significantly for male Parkinson speakers compared to male control speakers. The vocal

frequency increases for male and decreases for female Parkinson speakers. Medication-related (Dopa ON versus OFF) differences for tremor depth and frequency are small and their statistical significance is test-dependent. No significant differences are observed for vocal jitter. Figure 1 reports the (neurological) tremor depth and frequency for male and female Parkinson and control speakers. The histograms of the perturbation cues of Parkinson and control speakers overlap strongly.
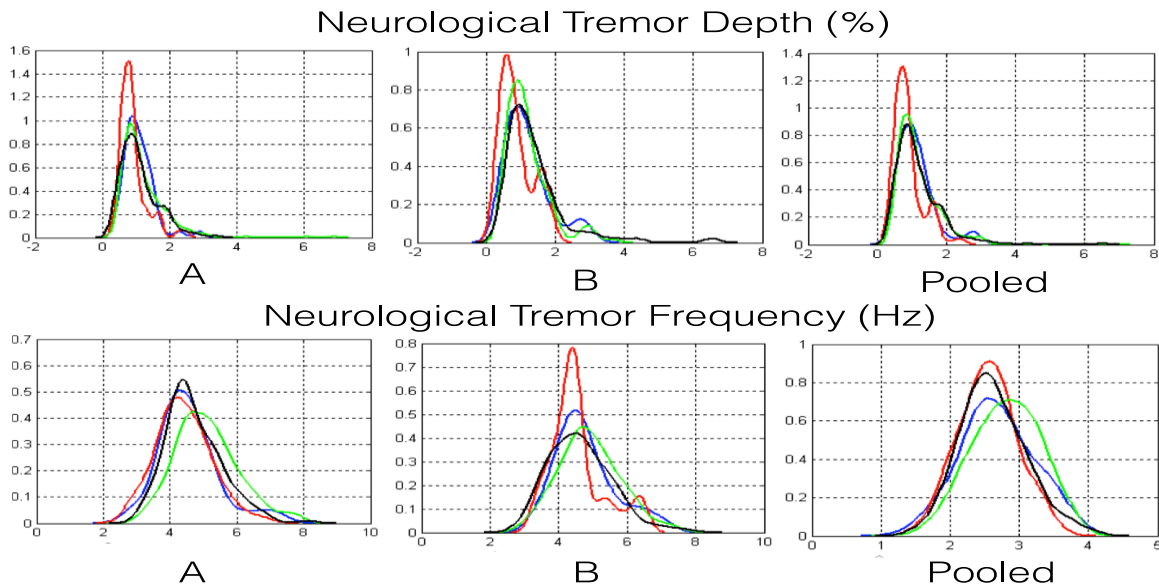
Figure 1 : Normalized histograms of neurological tremor depth in % and frequency in Hz for corpora A, B, individual and pooled. Male and female control speaker histograms are in dark blue and red and male and female Parkinson speaker histograms are in green and black respectively.

Conclusions

The study involves a substantial number of recordings of control and Parkinson speakers. The results report statistically significant differences between control and Parkinson speakers as well as between male and female speakers for tremor depth, tremor frequency and vocal frequency. Statistically significant interactions between pathology and gender are observed for tremor depth and vocal frequency. The strong overlap between the histograms of control and Parkinson speakers suggests that the statistically significant trends are clinically unreliable predictors of the evolution of vocal tremor depth and frequency in individual patients.

References

Christophe Mertens, Francis Grenez, and Jean Schoentgen. "Speech sample salience analysis for speech cycle detection". In: *Proceedings Interspeech - Brighton, U.K.* 2009, pp. 939–942.

Christophe Mertens, Francis Grenez, L. Crevier-Buchman, and Jean Schoentgen. "Reliable tracking based on speech sample salience of vocal cycle length perturbations". In: *Proceedings Interspeech - Makuhari, Japan.* 2010, pp. 2566–2569.

Christophe Mertens, Jean Schoentgen, Francis Grenez, and Sabine Skodda. "Acoustic analysis of vocal tremor in Parkinson speakers". In: *Proceedings 8th International Workshop Models and Analysis of Vocal Emissions for Biomedical Applications - Firenze, Italy.* 2013, pp. 19–22.

A setup to study physical principles underlying speech production for articulation-like movement

A. Van Hirtum[(1)], R. Blandin[(1)], X. Pelorson[(1)]

[(1)] Gipsa-lab, UMR CNRS 5216, Grenoble University, France

annemie.vanhirtum@grenoble-inp.fr

Introduction

The vocal tract duct's shape varies during speech production due to articulation. Sound is generated either aerodynamically somewhere in the vocal tract or is due to a fluid-structure interaction resulting in vocal folds auto-oscillation in the larynx. The detailed geometry of a time-varying human vocal tract is extremely complicated and subject to intra- and inter-speaker variability. In addition measurement of geometrical, flow and acoustic properties on a human speaker is restricted. Therefore, the objective of the current work is to present a setup suitable for physical studies of sound production and propagation through tubes of a finite length subjected to a time-varying forced movement. In particular, it is strived to come up with a setup allowing quantitative measurements in relation to the time-varying variation. Such a setup is pertinent to study human speech production. In the following, the setup is detailed and characterised. Next, examples of quantitative measurements are discussed in order to illustrate phenomena which can be reproduced and hence studied (characterisation or/and model validation).

Methods

The setup (Fig. 1) is inspired on recent work on the deformation of an a-priori uniform circular (diameter $D = 2b_0$, with b_0 the radius of the undeformed circular cross-section) elastic tube by pinching it between two parallel bars [1]. An elegant quasi-analytical geometrical model is proposed to describe the pinched tube portion based on the assumption of a constant perimeter ($P \simeq 2\pi b_0$) and the assumption that each cross-section can be described as a stadium ring. The internal shape of the tube is modelled as a function of the imposed pinching effort $P = 1 - b_{x_c}/b_0$ with $b_{x_c} \leq b_0$ corresponding to half the minimum radius of the compressed tube. Concretely, two pincers are used for which the pinching effort (0%-100%), the speed of pinching (<300mm/s) as well as the pinching position (-L+l<x<0) and movement of pinching position (<3.5mm/s) is computer controlled and can be varied. The setup can be attached to an aero-acoustic facility in order to study the impact of varying the geometry on speech production mechanisms. At the other end different exit conditions can be assessed. In the current work a flat baffle is used. The main variables and flow chart of the geometrical model are illustrated in Fig. 1. Sensors can be added to quantify acoustic and flow variables of interest such as acoustic and flow pressure sensors.
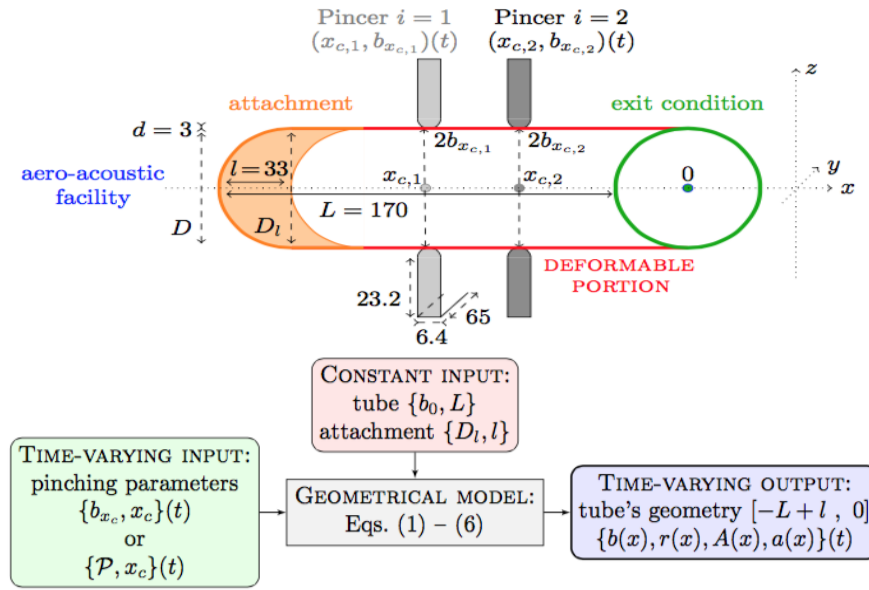
Fig 1: Overview of the deformable tube and main parameters used in the geometrical model. Dimensions are in [mm].

Results

Experiments are performed using either a sound source (controlled excitation features) or a deformable (fluid-structure interaction) replica mimicking vocal folds behaviour at the tube's inlet (x=-L). The tube's resonances are investigated as a function of constriction position and speed of geometry variation.

Conclusions An original setup is presented in order to study physical principles underlying speech production for articulation-like movement. The tubes's deformation can be fully expressed as a function of well defined input parameters favouring an analytical model approach at low computational cost. The variation of tube's resonances is presented for different sound sources at the tubes' inlet as well as onset and offset of voicing as a function of the tube's geometry. In particular transient phenomena can be investigated using the presented setup. More research is required to relate the observations to speech production.

References

[1] Van Hirtum, A. (2015). Deformation of a circular elastic tube between two parallel bars: quasi-analytical geometrical ring models. Mathematical Problems in Engineering, 2015, 1–15.

Perceptual Error Identification of Human and Synthesized Voices

Glaucya Madazio[1], Marina Englert[1,2], Ingrid Gielow[1], Jorge Lucero[3], Mara Behlau[1,2]

[1] CEV – "Centro de Estudos da Voz", [2] Unifesp – Federal University of Sao Paulo
[3] UnB - University of Brasilia

gmadazio@cevfono.com

Introduction
Perceptual auditory analysis is the gold standard to evaluate dysphonia (Gerrat et al., 1993). Auditory training gives this analysis a greater reliability (Gerrat et al., 1993; Awan, Lawson, 2009; Schaeffer, 2013). Synthesized stimuli are advantageous, the vocal characteristics are known and can be manipulated (Chan, Yiu, 2006), but they must sound natural (Lucero et al., 2013).

Objectives
To create a high quality synthetized voice and test voice specialists (VS), general speech language pathologist (SLP) and laymen ability to discriminate human and synthesized voice samples.
Methods: 70 subjects (20 VS, 20 SLP, 30 laymen) preformed a listening task to classify 54 voices: 18 human samples and 18 synthesized vowels, male and female (9 each), with different type and degree of deviation, total of 36 voices. 50% repetition was used for intra-rater reliability. Voice disorders were simulated by: vocal frequency perturbations (jitter - roughness) by the formula $\Delta K = a\varepsilon K$, where K is a vocal fold stiffness coefficient, a is a parameter scale and e is a random variable; additive noise (breathiness) by the formula $\Delta u = b\varepsilon u$, where u is the glottal flow rate, b is a scale parameter and e is a random variable; and by increasing tension and subglottal pressure and decreasing vocal folds' separation (strain). Human voices were collected from a vocal clinic database.

Results
Intra-rater reliability was high (72.22%). The average amount of error considering all groups was 37.77% Voice specialists presented less error than the others groups, 31.9%, SLP presented 39.30% of error and layman 40.83%. The three voices with greater perceptual confusion were human male with severe breathiness, synthesized female with mild breathiness and human female with severe roughness. For the VS group: human male with moderate roughness, human male with moderate strain and with severe breathiness, and human female with severe roughness. For the SLP group: synthesized female with mild breathiness, human male with severe breathiness, and synthesized male with mild strain. For the laymen group: human male with severe breathiness, human female with severe roughness and breathiness, synthesized male with severe breathiness, and synthesized female with mild breathiness (Table 1). Voices with breathiness had greater perceptual confusion, 42.9% for

human and 53.3% for synthesized breathiness. There was no difference in the amount of error considering all groups between human (36.8%) and synthesized (38.9%) voices (p=0,452).

Tables

Table 1. Voices with greater auditory-perceptual confusion for all groups and each group

Group	Voice	Error %
All	Human Male Severe Breathiness	81.4
	Synthesized Female Mild Breathiness	68.6*
	Human Female Severe Roughness	67.1
VS	Human Male Moderate Roughness	80.0
	Human Male Severe Breathiness	70.0
	Human Male Moderate Strain	70.0
	Human Female Severe Roughness	65.0
SLP	Synthesized Female Mild Breathiness	95.0
	Human Male Severe Breathiness	85.0
	Synthesized Male Mild Strain	75.0
Laymen	Human Male Severe Breathiness	86.6
	Human Female Severe Roughness	70.0
	Human Female Severe Breathiness	70.0
	Synthesized Male Severe Breathiness	70.0
	Synthesized Female Mild Breathiness	66.6

Equality of two proportion test

Conclusion

The quality of synthetized voice samples was very high. Human voices with moderate to severe deviation and synthesized voices with mild deviation had greater perceptual confusion. VS presented lower amount of error, which allows us to infer that auditory training assists on vocal analysis tasks.

References

Awan SN, Lawson LL. The effect of anchor modality on the reliability of vocal severity ratings. J Voice. 2009;23(3);341-.

Chan KMK, Yiu EML. A comparison of two perceptual voice evaluation training programs for naive listeners. J Voice. 2006;20(2):229-41.

Gerratt BR, Kreiman J, Antonanzas-Barroso N, Berke GS. Comparing internal and external standards in voice quality judgments. J Speech Hear Res. 1993;36:14-20.

Lucero JC, Schoentgen J, Behlau M. Physics-based synthesis of disordered voices. Proceedings of Interspeech 14th Annual Conference of the International Speech Communication Association; Aug 25-29; Lyon, France: 2013. 587-91.

Divergent or Convergent Glottal Angles: Which Gives Greater Flow?

Ronald C. Scherer[1], Lewis F. Fulcher[1]

[1] Bowling Green State University

ronalds@bgsu.edu

Introduction

During normal phonation a frontal view shows that the glottal angle changes continually, alternating between convergent and divergent angles. For the same angle value, diameter, and transglottal pressure, which angle, divergent or convergent, results in greater flow?

Methods

The symmetric glottal angles of the physical static model M5 were used. Characteristics (life-size) of the model were: axial glottal length 0.30 cm; convergent and divergent angles of 5, 10, 20, and 40 degrees; minimal diameters of 0.005, 0.01, 0.02, 0.04, 0.08, 0.16, and 0.32 cm; transglottal pressures from 1 to 25 cm H2O; and flows from 2.7 to 1536 cm^3/s and Reynolds number from 29.4 to 13,058.

Results

(1) For diameters of 0.04, 0.08 and 0.16 cm, the divergent angle always gave more flow than the convergent angle (about 5-25%) for the same minimal diameter and transglottal pressure.

(2) For the smallest diameter (0.005 cm) and the largest diameter (0.32 cm), the divergent angles always gave less flow (10-30%).

(3) For diameters of 0.01 and 0.02 cm, flow was greater for divergent 5 and 10 degrees, and slightly less for divergent 20 and 40 degrees.

(4) The ratio *(flow for divergent angle) / (flow for convergent angle)* is relatively independent of transglottal pressure except for diam = 0.005 cm (where there is a rise in value from about 0.7 to 0.9 as Ps rises from 3 to 25 cm H2O for 5, 10, and 20 degree angles).

Conclusions

If the quasi-steady assumption is made, these results suggest that the divergent glottal angle will tend to increase the glottal flow for midrange glottal diameters (which could help skew the glottal flow further "to the right"), and create less flow at very small diameters (increasing energy in the higher harmonics when flow nearly reaches baseline as flow decreases during the cycle).

Control of glottal channel geometry by intrinsic laryngeal muscle activation

Dinesh Chhetri [1], Zhaoyan Zhang[1], Andrew Vahabzadeh[1]

[1] University of California, Los Angeles

dchhetri@mednet.ucla.edu

Introduction

The glottal channel shape is considered to critically affect vocal fold vibration and acoustics. In this investigation the shape and posture changes of the medial vocal fold surface upon activation of intrinsic laryngeal muscles (ILMs) were quantitatively evaluated.

Methods

An *in vivo* hemilarynx model was used. The thyroarytenoid (TA), lateral cricoarytenoid (LCA), posterior cricoarytenoid (PCA), and cricothyroid (CT) muscles were individually activated by neuromuscular stimulation of distal nerve branches. Hemilaryngectomy was perfomed and a right angle prism was placed with the hypotenuse along the glottal midline. The prism provided two distinct high-speed camera views of the vocal fold medial surface, from which the three-dimensional motion of the medial surface was calculated using the image-processing software package DaVis (LaVision Inc.). Mapping functions were calculated and were then used to calculate the surface height of all points along the medial surface of the vocal fold. The medial surface contour in 3-dimensional space was reconstructed.

Results

The temporal sequence of medial vocal fold bulging upon TA activation was captured. The bulging started in the inferior portion of the voca fold then expanded to involve the entire vocal fold. LCA activation closed the posterior glottis but could not completely close the anterior and inferior glottis. PCA activation led to superolateral motion at the vocal process. CT activation elongated and anteriorly rotated the vocal fold towards the cricoid cartilage.

Conclusions

Investigations of the vocal fold posture from the medial view reveals some posture changes not seen from a superior view. As the glottal channel shape is convergent at a non-stimulated state, the superior view misses the bulging effect until the degree of medial bulging reaches beyond the superior medial edge of the vocal fold. This bulging effect of the TA muscle may play an important role in F0 and register control.

Acoustic detection of diplophonia among other types of dysphonia

P. Aichinger[1], B. Schneider-Stickler[1], M. Hagmüller[2], J. Schoentgen[3]

[1] Division of Phoniatrics-Logopedics, Department of Otorhinolaryngology, Medical University of Vienna, Austria [2] Signal Processing and Speech Communication Laboratory, Graz University of Technology, Austria [3] Department of Signals, Images and Acoustics, Faculty of Applied Sciences, Université Libre de Bruxelles, Belgium
philipp.aichinger@meduniwien.ac.at

Introduction

Valid descriptions of voice phenomena are required for the clinical care of disordered voice. Known problems with regard to voice quality description include subjectivity and variability of auditory and visual assessment of recorded signals. In addition, non-standardized definitions of voice phenomena impede clinical and scientific communication. This presentation concerns the acoustic model-based detection of diplophonia among other types of non-diplophonic dysphonia.

Methods

The tested audio samples are part of a database of 80 dysphonic subjects and 40 euphonic subjects (Aichinger, 2015). Sustained vowels have been recorded during rigid endoscope high-speed videolaryngoscopy. 55 diplophonic dysphonic and 77 non-diplophonic dysphonic sound fragments of temporally homogeneous voice quality were selected for analysis. The samples stem from 36 dysphonic subjects (31 female and 5 male) and from 9 euphonic subjects (4 female and 5 male). 23 of the dysphonic subjects were clinically categorized as diplophonic (20 female and 3 male).

The samples were analyzed by means of the Diplophonia Diagram (Aichinger, 2015), which is an acoustic analyzer based on an explicit signal model of diplophonic waveforms. The Diplophonia Diagram automatically detects additive secondary harmonic oscillators, the fundamental frequencies of which differ from those of the primary oscillators. A voice sound fragment is tested diplophonic if two fundamental frequencies coincide in time. For comparison, the samples have been analyzed with six conventional features, which were jitter, shimmer, the Göttingen's irregularity and noise, Praat's Harmonic-to-Noise Ratio and the Degree of subharmonics. Göttingen's irregularity is compound of jitter, shimmer and the mean waveform matching coefficient (Michaelis et al., 1998). Göttingen's noise reports correlation between the sub-band envelopes of the pre-whitened signals (Michaelis et al., 1997, Michaelis et al., 1998).

Results

Figure 1 shows the Diplophonia Diagram, which reports the goodness of fit via the maximal quantitative similarity in dB between the analyzed signal and an automatically fitted model. The goodness of fit is visualized for one and two oscillators respectively (horizontal and vertical axes). Diplophonic signals cluster at the upper left, because they are more appropriately represented by a two oscillator model than by a one oscillator model. Table 1 reports the classification accuracy of the remaining acoustic features, obtained via cut-off threshold classification and binomial logistic regression. The Diplophonia Diagram assigns correctly 87.2 % of all samples, while the best conventional feature is the Degree of subharmonics (67.6 %).

Figure 1: The Diplophonia Diagram distinguishes diplophonic dysphonic from non-diplophonic dysphonic voice sound samples.
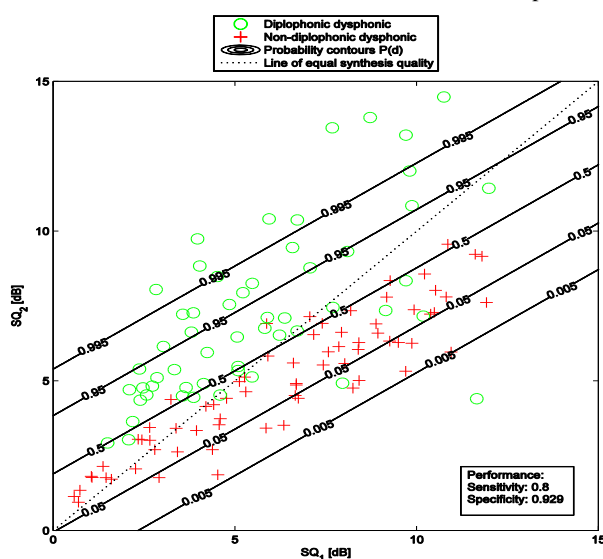
Table 1: Accuracy (Acc) of all features under test. Diplophonia Diagram (DD), Göttingen's irregularity (Irr), Göttingen's noise (Noise), Praat's Harmonics-to-Noise Ratio (HNR), Degree of subharmonics (DSH), 95 % confidence intervals (CI).

	Acc, CI [%]	p-value
DD	87.2, [80.0, 92.5]	-
Jitter	58.4, [46.9, 67.3]	< 0.001
Shimmer	66.2, [56.8, 76.5]	= 0.002
Irr	57.7, [49.2, 66.1]	< 0.002
Noise	64.9, [58.1, 69.7]	< 0.001
HNR	54.0, [48.0, 60.0]	< 0.001
DSH	67.6, [61.8, 75.2]	< 0.001

Conclusions

Six conventional hoarseness features and the Diplophonia Diagram have been evaluated with regard to their ability to detect diplophonia among other types of dysphonia. Jitter, shimmer and the Göttingen's irregularity are valid only if one oscillator is active at any time. The Harmonics-to-Noise Ratio and the Göttingen's noise may increase for both diplophonia and additive noise, which may lead to clinical misinterpretations. The Degree of subharmonics is the only feature that considers that two oscillators may exist simultaneously, but cannot distinguish between diplophonic beating and non-diplophonic amplitude modulation. The Diplophonia Diagram is based on an explicit model of diplophonic signals and provides physiologically interpretable results by means of detecting two fundamental frequencies coinciding in time.

Chromoendoscopy Associated with Endoscopic Laryngeal Surgery for Treatment of Recurrent Respiratory Papillomatosis (Phase II)

Daniel G. Rey Caro MD[1], Enrique P. Rey Caro MD[1] Enrique A. Rey Caro MD[1],

[1] Staff member of the ENT Service of Chutro Clinic in Cordoba, Argentina

gonzaloreycaro@hotmail.com

Introduction

Chromoendoscopy is an endoscopic technique which uses a contrast stain to paint the aerodigestive tract mucosal lining followed by an optical assessment to highlighting any epithelial abnormalities. Detailed and high-definition magnified views achieved with the aid of rigid endoscopes can often allow for identification of the tissue type or pathology based upon the pattern uncovered. According to the literature we reviewed, we may have been the first ones to use indigo carmine in the field of otolaryngology. Tiny lesions that usually go overlooked with conventional microlaryngoscopy become visible upon the instillation of indigo carmine and further decreasing the chances of an early lesion postoperative recurrence.

Chromoendoscopy, in recurrent respiratory papillomatosis (RRP), helps identify unsuspected intraoperative lesions by clearly enhancing the view of their boundaries and surface type. It is also suitable to assess the presence of residual lesions, if any, after their surgical removal.

Objectives: To demonstrate the usefulness of chromoendoscopy in RRP in laryngotracheal surgery.

Methods:

 We used indigo carmine associated with endoscopic laryngeal surgery. Before staining, the mucosa may need to be treated with a mucolytic agent to get rid of excess mucus to boost staining. Rigid suspension laryngoscopes of different proximal and distal diameters were used with chromoendoscopy. Patients underwent chromoendoscopy associated with endoscopic laryngeal surgery under general anesthesia in the O.R

Results:

In this second phase of our research work, this diagnostic technique was applied to eighteen patients with recurrent laryngeal papillomatosis and two patients with suspected carcinoma of the larynx. We were able to optimize the intraoperative diagnosis and reduce the likelihood of the relapse risk in all patients.

Conclusion:

Chromoendoscopy associated with endoscopic laryngeal surgery is an excellent low-cost intraoperative diagnostic method for the treatment of invasive diseases of the larynx such as laryngeal papillomatosis.
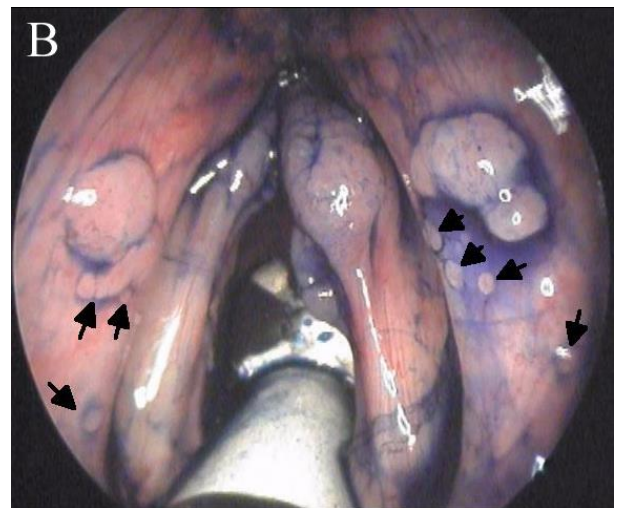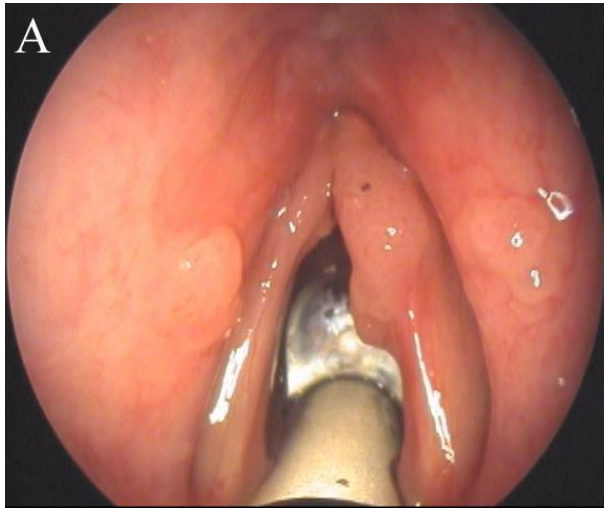
References

1. Kudo S, et al. Colorectal tumors and pitt pattern. J Clin pathol. 1994;(47:880).
2. Sharma p, Weston AP, Topalovski M, Cherian R, Bhattacharyya A, Samplinier RR. Magnification chromoendoscopy for the detection of intestinal metaplasia and dysplasia in Barrett's esophagus. Gut. 2003; 52(24-27).
3. Sharma p, Weston AP, Topalovski M, Cherian R, Bhattacharyya A, Samplinier RR. Magnification chromoendoscopy for the detection of intestinal metaplasia and dysplasia in Barrett's esophagus. Gut. 2003; 52(24-27).
4. Yasushi Sano , Yutaka Saito , Luang-I Fu , Takahisa Matsuda , Toshio Uraoka. Efficacy of magnifying chromoendoscopy for the differential diagnosis of colorectal lesions. Digestive Endoscopy. 2005; 17(105-116).
5. Ribeiro MD. Chromoendoscopy for early diagnosis of gastric cancer. European Journal of Gastroenterology & Hepatology. 2006; 18(831-838).
6. Andrea MD, Dias O, Santos. Contact endoscopy of the vocal cord: normal and pathological patterns. Acta Otolaryngol. 1995; 115(2)(314-6).

.
.
.
.
.

The Tongue- Hyoid- Larynx Complex and its influence on Phonation

Ana Gloria Ortega, SP, PhD[1], Adriana Elina Ortega, Engineer, MSc, PhD[2]

[1]Facultad de Artes y Diseño Universidad Nacional de Cuyo

[2]Instituto Nacional de Tecnología Agropecuaria (INTA)

ortega.adriana@inta.gob.ar

Introduction

Vocal production is primarily regulated by the intrinsic laryngeal muscles. However, the extrinsic laryngeal muscles have an important influence on phonation, given that they provide the base posture from which the intrinsic muscles can operate. Therefore, it is through the extrinsic muscles that the larynx relates with the surrounding organs such as the hyoid, tongue, jaw, sternum, spinal column and skull. It should also be considered that these muscles operate in the voice production through "articulatory gestures". The action of the complex on vertical and horizontal tension of the vocal folds is well known.

In this research, we studied the action of the tongue-hyoid-larynx complex on the variations in horizontal and vertical tensions of the glottal source produced by the phonation of two vowels considered opposing by the action of this complex: the vowel /i/ as in the word "bee" and the vowel /o/ as in the word "board".

Methods

Retrospective investigation; 161 female subjects, aged 18 to 55 all of them having a laryngeal examination carried out:control group of 22 healthy voiced women; 139 dysphonic voiced women with different laryngeal diagnosis.

Perceptual analysis: spontaneous speech; counting from 1 to 20; long vowels /i/-/o/ modal register, habitual pitch and loudness. (GRBAS scale and Vocal profile Examination University of Edinburg).

Acoustic examination vowels /i/-/o/: jitter (%); shimmer (%); Mean Fo (Hz); SD Fo (Hz), and SNR (dB) Absolute and relative measures derived from the comparison of both vowels.

Statistical analysis: descriptive; non parametric (principal component test); Pearson's correlation coefficient test; T test to find significant differences among four groups of phonatory behavior: hypofunctional, hyperfunctional, phonasthenia and normal voices.

Results:

In normal voices the behavior of vowels /i/ /o/ is similar.

In slight dysphonia, vowel /o/ is first affected.

In severe dysphonias, both vowels are affected, with the production of the vowel /o/ having lower pathologic values.

A t-test shows significant differences between hypofunctional- hyperfunctional phonation: jitter i-o (Pr> ItI 0.0334); stdvf i-o (Pr> ItI 0.0552); SNRI- SNRO (Pr> ItI 0.0527); shimmer /i/ (Pr> ItI 0.0346) and between slight vocal fold hypofunction and normal voices: jitter /o/ (Pr> ItI 0.0292), shimmer /o/ (Pr> ItI 0.0320).

Conclusions

The acoustic analysis of Spanish vowels /i/- /o/ adds stability/ instability to the vocal folds vibration by increasing/ decreasing the vertical and horizontal tension of the vocal folds. The diagnosis of phonatory behavior is deduced from the analysis of the acoustic variation produced between these two vowels.

References

Hoole, P.H., (1999) On the lingual organization of the German Vowel System. J. Acoust. Soc. Am. 106 (2) pp 1020
Honda, K. , (1983) Relationship between pitch control and vowel articulation. Vocal Folf Physiology. Edited by D. Bless and J. Abbs. College Hill. San Diego. California.
Honda- Hirai, (1994) Contributions of vocal tract shape to voice quality. MRI data and articulatory modeling. Vocal Fold Physiology. Voice Quality Control. Fujimura- Hirano. Sing. Publ. Group.
Honda- Hirai, (1995) Contributions of vocal tract shape to voice quality: MRI data and articulatory modeling. Vocal Fold Physiology. Edited by Fujimura-Hirano. Sing. Publ. Group
Zheng, Y. & Hasejawa-Johnson, (2003 Analysis of three dimensional tongue shapes using a three-index factor analysis model. J.Acoust.Soc. Am 113 January pp 478

Intrinsic Frequency of Vowels with respect to Phonatory Behaviour.

Ana Gloria Ortega, SP, PhD[1], Adriana Elina Ortega, Engineer, MSc, PhD[2]

[1]Facultad de Artes y Diseño Universidad Nacional de Cuyo Mendoza, Argentina

[2]Instituto Nacional de Tecnología Agropecuaria (INTA) Salta, Argentina

ortega.adriana@inta.gob.ar

Introduction

 Hypothesis
- Healthy larynxes show independence of glottis behavior with respect to articulatory patterns.
- Unhealthy larynxes (loss of phonational synergy) show dependency of glottis behavior with respect to articulatory patterns.

Methods

Retrospective study of voice samples of 188 women: 100 dysphonic, 48 normal with no vocal training and 40 singers with healthy voice. Measuring variables: intrinsic frequency of the vowels /i/ /o/. Perceptual examination of the voices performed by two voice specialists with over 15 years of experience.

The research was conducted using absolute values of the intrinsic frequency, standard deviation of the frequency and relative values corresponding to the difference of frequency and standard deviation of the frequency between both vowels /i/-/o/.

Statistical analysis of the data: Descriptive statistics of each variable and in each group studied, T test and Pearson's correlation coefficient test.

Results

1- Intrinsic frequency of vowels /i/ /o/
T Test: significant differences ($P > 0.05$) in the responses of dysphonic voices with respect to the singers and normal voices without training with respect to singers. Differences between dysphonic and normal voices were not significant.

2- Standard deviation of the frequency of the vowels /i/ /o/
T Test: Very significant differences were obtained ($P > 0.001$) in the responses of the dysphonic voices respect to the singers and of the normal voices without training respect to singers.
A Pearson correlation test applied to the difference of the intrinsic frequency respect to the difference of the standard deviation of the frequency produces a correlation coefficient very significant with a correlation of 0.0014 in the dysphonic group (probability > 1 R 1 under Ho; Rho=0 / N=88).

Conclusions

- The pattern of intrinsic frequency of the studied vowels is maintained across the groups.
- Singers show less dependency of the glottal generator with respect to articulatory mechanisms.
- The dysphonic group shows greater dependency of the glottal generator with respect to articulatory mechanisms.
- People with the most unstable phonatory behavior show a greater dependency of the glottal generator with respect to tongue movements (measured as intrinsic frequency of vowels /i/ /o/).

The selection of the sounds to be evaluated is of capital importance in every situation of perceptual, acoustic or electroglottographic examination.

References

Colton. R. (1988) Physiological mechanisms of vocal frequency control: the role of tension. Journal of Voice. Vol 2 n°3 pp 208-220
Ericson. D. (2002) Articulation of extreme formant patterns for emphasized vowels. Phonetica 59: 134-149
Estella.P. & Edwin. ML. Yu. (2005) Suitability of acoustic perturbation measures in analyzing periodic and nearly perioidc voice signals. Folia Phoniatr Logop 57: 38-47
Honda.k. & Fujimura.o. (1991) Intrinsic vowel fo phrase-final f0 lowering: phonological vs. biological explanations. Vocal Fold Physiology. Edited by Jan Gauffin and Britta Hammarberg. Singular Publishing Group San Diego California Cap 19. 149-157
Honda.K. (1983) Relationship between pitch control and vowel articulation. Vocal Fold Physiology. Contemporary Research and clinical issues. Edited by Diane Bless. James Abbs College Hill San Diego California pp 286-289

Vocal Tract Geometry from a Biomechanical Model

Saeed Dabbaghchian[1], Olov Engwall[1]

[1] Department of Speech, Music, and Hearing, KTH Royal Institute of Technology, Stockholm, Sweden

saeedd@kth.se

Introduction

Recently, 3D biomechanical models (Anderson, et al. 2015) have been developed to study voice production. In these models, a set of muscle activation parameters, moves the different articulators directly and hence deforms the vocal tract indirectly. In order to analyse the acoustics of the vocal tract, a tube representation of it is required. This tube does not explicitly exist in a biomechanical model since the vocal tract is not a physical object, but a cavity that is formed by a large number of structures. The cavity shape may nevertheless be extracted manually for a small number of simulations, but for large-scale simulations, we need an automatic or at least a semi-automatic method with the least user intervention possible. This work addresses the coupling of biomechanics and acoustics by proposing an automatic cavity extraction method.

Methods

In a biomechanical model, the vocal tract tube is created by joining the surface geometries of the all bordering structures, including larynx, pharynx, soft palate, tongue, jaw, and maxilla. Two types of problems may appear when attempting to create the vocal tract tube from the surface geometries of the models. Firstly, the geometries of some connecting structures or tissues may be missing, which results in gaps when forming the vocal tract tube. Secondly, the surface meshes of articulator geometries that are positioned close to each other may intersect each other several times, since they do not share a common boundary. In order to decrease the complexity of the vocal tract tube to be formed, a semi-polar grid is used to intersect all the articulator geometries and create a set of planar cross-sections for the vocal tract. This technique allows reducing the data from 3D geometries to 2D polygons in the semi-polar planes (Dabbaghchian, et al. 2015). Each planar cross-section includes zero, one or several polygons for each structure. A preprocessing step detects gaps between polygons and fills them; it also detects polygons where their edges almost overlap and then snap them, i.e. it creates a common edge for those polygons. Then one airway polygon for each cross-section is extracted using union operation. A post-processing smooths the airway polygons and converts these 2D planar polygons into a 3D vocal tract geometry. Figure 1 illustrates the process by showing the semi-polar grid, one example of the cross-sections and 3D geometry.
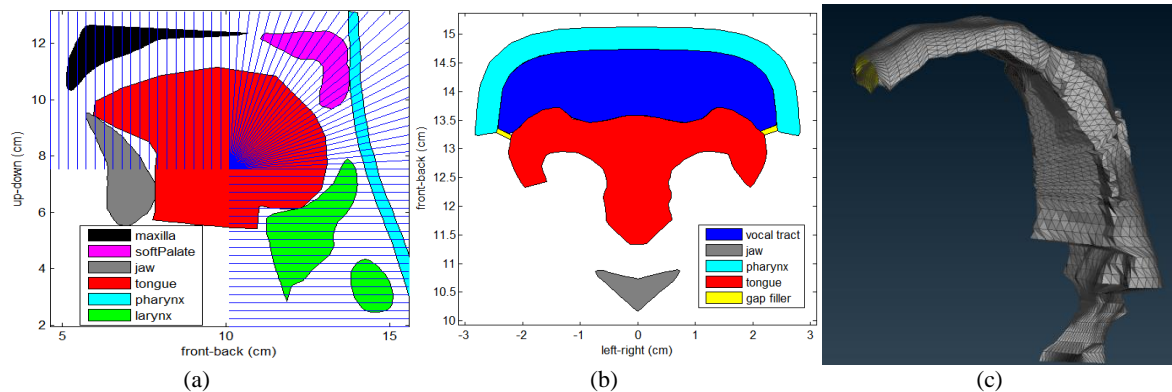
Figure 1. (a) Mid-sagittal view of articulators superimposed on semi-polar grid, (b) One cross-section in the pharyngeal cavity, (c) vocal tract geometry.

Results

We examined the proposed method using a recently developed biomechanical model in ArtiSynth (Lloyd, et al. 2012). A large deformation of the vocal tract, moving from /ɑ:/ to /i:/ was tested while producing a geometry in each time step (10 ms). Figure 1.c shows one of the resulting geometries. There are no self-intersections in the generated mesh and it forms a regular, closed tube, which may be used to study the vocal tract acoustics. This method should work for other similar deformations such as vowel-vowel utterances but for more complex deformations it needs further examination.

Conclusions

We proposed an automatic method to extract the vocal tract geometry using the structures that are surrounding this cavity. This method may be used to couple the biomechanical or geometrical models (Wei, et al. 2015) to acoustics simulations.

As future work, we will improve the method to include the side branches such as the piriformis fossae and small cavities, such as the vallecula, which appears or disappears by the movement of the articulators (tongue root, in the case of the vallecula).

Acknowledgement

References

Anderson, P., Harandi, N. M., Moisik, S., Stavness, I. and Fels, S. (2015). A comprehensive 3D biomechanically-driven vocal tract model including inverse dynamics for speech research, 16th annual conference of the International Speech Communication Association (Interspeech), Dressden, Germany.

Arnela, M., Dabbaghchian, S., Blandin, R., Guasch O., Engwall, O., Pelorson, X., and Van Hirtum, A. (2015). Effects of vocal tract geometry simplifications on the numerical simulation of vowels. 11th Pan-European Voice Conference (PEVOC), 177, Firenz, Italy.

Dabbaghchian, S., Arnela, M., and Engwall, O. (2015). SIMPLIFICATION OF VOCAL TRACT SHAPES WITH DIFFERENT LEVELS OF DETAIL. 18th International Congress of Phonetic Science, Glasgow.

Lloyd, J. E., Stavness, I., and Fels, S., (2012). ArtiSynth: a fast interactive biomechanical modeling toolkit combining multibody and finite element simulation. Soft tissue biomechanical modeling for computer assisted surgery, Springer, 355-394.

Transcriptome Profiling and Molecular Pathway Analysis Using RNA-Seq to Understand VF Development at E15.5 in a Mice Model

Vidisha Mohad[1], Vlasta Lungova, Jeea Choi[1], Christina Kendziorski[1], Xin Sun[2], Susan L. Thibeault[3]

[1] Division of Otolaryngology-Head & Neck Surgery, University of Wisconsin-Madison
[2]Department of Genetics, University of Wisconsin-Madison [3]Department of Biostatistics and Informatics, University of Wisconsin-Madison

mohad@surgery.wisc.edu

Introduction

Our study aims at identifying early gene expression signatures that distinguish vocal fold (VF) lineage from that of trachea and esophagus lineage, and to determine dynamic changes of gene expression in VF epithelium as it progresses through specification, stratification and maintenance. The goal of our study is to build the first comprehensive transcriptome profiling that will serve as a blue print to direct efficient differentiation of pluripotent stem cells into functional VF epithelium.

Methods

Twenty-six FVB/N mice were used. Pregnant females were sacrificed at E15.5, and postnatal day (P0). Mouse larynges, trachea and esophagus were dissected out of each animal and separated at E15.5, P0 and adult. Following dissection total RNA was extracted and DNA digestion was performed. RNA quality was measured and RNAseq analysis was completed with an Illumina HiSeq2000. Two sets of data were compared --VF, esophagus and trachea at E15.5 stage and VF tissue at E15.5, P0 and adult. Gene expression between VF, trachea and esophagus were calculated using EBseq (1) and for between E15.5, P0 and adults, EBSeqHMM (2) was used. We defined a gene as differentially expressed (DE) if 1-posterior probability is greater than 0.95, controlling overall FDR at 5%. Gene ontology (GO) enrichment analyses and KEGG enrichment analysis was used to identify predominant biological themes and pathways.

Results

Between tissue at E15.5 we found a total of 1933 DE genes. Between E15.5VF=E.15.5Trachea≠E.15.5Esophag we found 885 DE genes, with GO terms related to extracellular remodeling, cell adhesion and growth, and protein modification. Between E15.5VF≠E.15.5Esophag=E.15.5Trachea we found 167 genes, with GO terms related to cell adhesion, cytoskeleton, contractile fibers for cell contraction, motility and cell signaling. For E15.5VF≠E.15.5Trachea≠E.15.5Esophag we found 64 DE genes, with GO terms related to ECM growth and remodeling, and protein modification. For E15.5VF≠E.15.5Trachea=E.15.5Esophag only 9 DE genes were found. For VF tissue at

E15.5, P0 and adult, we identified a total of 10289 DE genes. 2552 genes were downregulated, these genes were upregulated during E15.5 and their expression decreased during P0 and adult stages. Downregulated genes showed major common pathways for hedgehog signaling, glycosaminoglycan biosynthesis, ECM receptor and MAPK signaling. 2399 genes were upregulated; these genes were downregulated during E15.5 but an increase in expression was noted during P0 and adulthood. Up-regulated genes showed a major common pathway for protein processing, ECM, and cell adhesion. The constantly changing category consisted of down-up and up-down genes. 1463 genes were found to be DE in the down-up category and 1634 were found to be DE in the up-down category.

Conclusions

Specific molecular pathways that are involved in the development of VF's are significantly different from those involved in maturation of esophagus and trachea. When compared to esophagus, ECM processes peaks at E15.5 during VF development. We have discovered that at E15.5, the majority of genes in the up-regulated pathway are responsible for tissue differentiation. The pathways in the later stages of morphogenesis related toward imparting tissue specific unique properties such as protein processing for VF tissue and muscular characterization for the esophagus.

References

1. Leng, N., Dawson, JA., Stewart, RM., Ruotti, V., Rissman, AI., Smits, BMG., Haag, JD., Gould, MN., Thomson, JA., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics 29.8 1035-1043.
2. Leng, N, and C. Kendziorski, (2014). EBSeqHMM: An R package for identifying gene-expression changes in ordered RNA-seq experiments.

Psychoacoustic parameters present in the vocal prosody as predictors of emotional identification

Carla Badani Schoneweg[(1)], Luis Romero Romero[(1)], Solange Duran Elicer[(1)]

[(1)] Universidad de Chile

cbadani@u.uchile.cl

Introduction

Can listeners infer, through decofication of certain psychoacoustic parameters in the voice, the emotion vocally expressed? This question has been studied by numerous authors[1,2,3,4] and methodological approaches without a conclusive answer. The fact that listeners are able to recognize different emotions in a good shape based on the voice heard, implies that vocal expression of emotion has a discernible pattern[4]. This question has led us to consider what psychoacoustic parameters present in the vocal prosody are good predictors of emotion vocally expressed[1]. Psychoacoustic parameters present in this paper are halftone and loudness. Halftone is the average of the fundamental frecuency f0 emitted during spontaneous speech. It correlates with the perceived pitch of a voice. Loudness is a magnitude of the perceived intensity of the heard voice, correlated with physical sound intensity[2,4]. Another quantity studied corresponds to the duration of the speech signal, which correlates with the perceived speed of elocution.
The importance of this finding will lead to identify vocal effector patterns of emotion, which could in the future be included in the Alba Emoting[TM]

Methods

The subjects participants are people of feminine and masculine gender, with normal hearing, adults between 18 and 60 years, covering two sections of adulthood: early and middle. The subjects have no formal musical training or actoral training before this study.

We proceed to present to normal hearing subjects and native Spanish speakers from Chile, audio recordings containing the voices of actors saying a text in 6 basic emotions according to the Alba Emoting[5,6,7] method (AE). 6 emotions were presented: joy, anger, sadness, erotic love, tenderness and fear. These are being presented in a random manner to the subjects. These recordings were preselected by expert judgment, under the criteria of good acting performance in Alba Emoting method, in which the actors were trained for 4 months. In the selected audio, the variation of prosodic elements in each register of each respective emotion were analyzed, determining the following parameters: Halftone, Halftone range, Loudness, Loudness Range, total length, segment length, speed of elocution. Listeners must perform two tasks of identification:

- Indicate which emotions identified as present in the audio heard regardless of the excitement generated by the actor.

- Indicate which level (high or low) emotional intensity identified as present in the audio heard.

Positive identification results and negative identification and also false positives (misidentification) were noted. Bivariate correlation was calculated between the results of identification tasks and psychoacoustic parameters measured in the recordings, by Pearson correlation coeficients, considering all values of both loudness and tone of successes of emotional identification for men and women and average number of successes of emotional intensity for men and women.

Results

Preliminary results indicate that at a significance level of 0.05, the average loudness is generally a good predictor of emotional identification, especially in the case of men, where the Pearson correlation coefficient reaches a positive value equal to 0.63, whereas in women does not reach the value 0.4 (Figure 1).
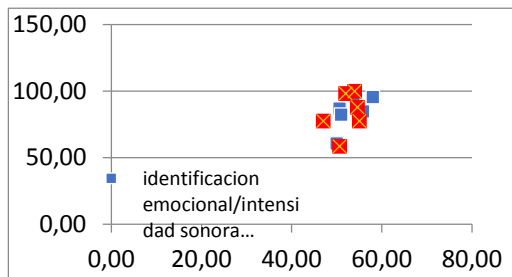
Figure 1 Discrimination of emotion (%) through Loudness (dB), in men and women Comparison of average values.
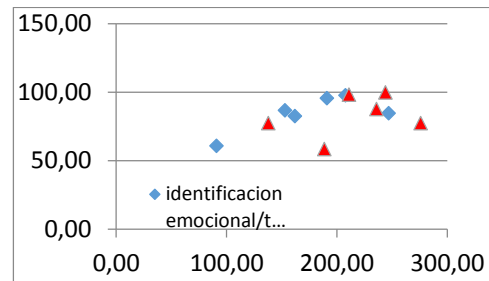
Figure 2 Discrimination of emotion (%) through Halftone (Hz) . Comparison of average values.

Figure 2 shows the dispersion relation between the Halftone spoken for both men and women and emotional identification. Again it is shown that for males, the correlation between these two variables is positive and significantly higher than for women, reaching the value of 0.73. Women only reached a correlation coefficient of 0.33.

Regarding the emotional intensity and its correlation with both loudness and halftone, opposite results were obtained earlier between men and women. The correlation to loudness and emotional intensity in women reached the value 0.56, while males only reached 0.23. In the case of half-tone and emotional intensity, these values were 0.44 for women and 0.29 for men.

The correlation observed between the total duration of each emotion recording and identification of the respective emotion have a negative value for women, inversely correlated, -0.5, and for males is also negative and -0.1 (Figure 3).
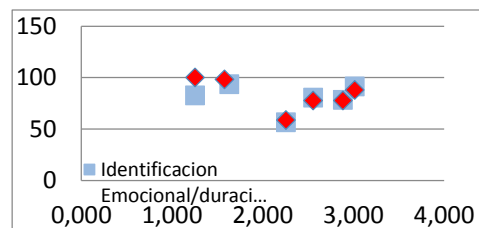
Figure 3 Discrimination of emotion (%) through Duration (s) . Comparison of average values.

Perturbation Measures on the Degree of Naturalness of Synthesized Vowels

Rosiane Yamasaki[1] , Arlindo Montagnoli[1], Adriana Hachiya, Eloísa Gebrim, [1]Jorge Vicente Lopes da Silva[1], Emi Murano, Mara Behlau, Domingos Tsuji[1]

[1]Dep. Otolaryngology, School of Medicine, University of São Paulo, Brazil

r.yamasaki@uol.com.br

Introduction

The naturalness of speech is an essential element in the development of synthesized voices. The human voice holds small perturbations in the acoustic wave that are considered normal variations of the voice.[1] Jitter and shimmer are short-term perturbation measurements[2] that have been widely used both in clinical and in research voice assessments. Adding up these perturbation measurements might be an important way to increase the naturalness of synthesized voices. Therefore, the goal of this study was to verify perceptually the degree of naturalness of synthesized vowels produced with and without the addition of jitter and shimmer. The vowels were developed by acoustical simulations using a mathematical model in order to generate glottal pulses[3] (GP) and a solid model of the vocal tract (SMVT) with the shape of the vowel /ɛ/. The SMVT represents the real shape of the vocal tract and it was previously used in acoustical simulations.[4]

Methods

Synthesized vowels were produced using GP with and without jitter and shimmer, an SMVT and a driver unit. Eighty GP were generated using Rosenberg mathematical model with four perturbation conditions: 20 with addition of jitter; 20 with shimmer; 20 with jitter+shimmer; and 20 without perturbation. The SMVT was based on MRI from a young woman during the phonation of /ɛ/ and using rapid prototyping technology. The driver was used both to amplify and to transfer the GP to SMVT. With that, eighty vowels were produced and two auditory-perceptual evaluation were performed. Experiment 1: three judges rated 120 vowels (20 humans+80 synthesized+20% repetition) as "human" or "synthesized". Experiment 2: three judges rated only the synthesized vowels. Twenty slide sequences were produced plus 20% repetition. Each slide had the four vowels produced with the four perturbation condition. The judges were asked to rank each four-vowel/sequence presented in a 4 number scale, being (1) the most natural and (4) the most artificial.

Results

The inter and intra-rater reliability was high in both perceptual experiments. Experiment 1: all the human vowels were classified as human; and 27 out of eighty synthesized vowels were rated as human. Analyzing these 27 vowels, we observed that: 15 were produced with jitter and shimmer; 10 with jitter; 2 without perturbation; and none with shimmer only. Experiment

2: two judges considered the vowels produced with jitter and shimmer as the most natural, followed by those with jitter; one judge considered the vowels with jitter as the most natural, followed by both jitter and shimmer. All the judges rated the vowels with shimmer and without perturbation as the most artificial (Table 1).

Table 1 - Distribution of the classification of synthesized vowels produced with four perturbation conditions, according to the degree of naturalness

	Vowels	+ Natural	- Natural	- Artificial	+ Artificial	Total
Judge 1	Jitter + Shimmer	12	8	0	0	
	Jitter	8	12	0	0	80
	Shimmer	0	0	11	9	
	No perturbation	0	0	9	11	
Judge 2	Jitter + Shimmer	7	13	0	0	
	Jitter	13	7	0	0	80
	Shimmer	0	0	11	9	
	No perturbation	0	0	9	11	
Judge 3	Jitter + Shimmer	11	9	0	0	
	Jitter	9	11	0	0	80
	Shimmer	0	0	12	8	
	No perturbation	0	0	8	12	

Conclusions

The association of the jitter and shimmer perturbation measurements increased the degree of naturalness of synthesized vowels. The addition of jitter seems to provide naturalness however, the vowels generated with the addition of shimmer only had negative impact. This study suggests that the acoustical simulation performed with GP generated by means of the mathematical model and using an SMVT is an interesting method to test the effect of the perturbation measurements on synthesized voices.

References

1. Brockmann, M.; Drinnan, M.J.; Storck, C.; Carding, P.N. 2011. Reliable jitter and shimmer measurements in voice clinics: the relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. J Voice. 25(1):44-53.

2. Titze, I.R. 2000. Principles of voice production. Iowa City. NCVS. 2nd ed.

3. Rosenberg, A.E. 1971. Effect of glottal pulse shape on the quality of natural vowels. J Acoust Soc Am. 49(2):Suppl 2:583-90.
4. Fujita, S.; Honda, K. 2005. An experimental study of acoustic characteristics of hypopharyngeal cavities using vocal tract solid   models. Acoust. Sci. & Tech. 26(4), 353-7.

Effects of vocal tract inertance on the glottal flow

Charles Farbos de Luzan[1], Liran Oren [2], Ephraim Gutmark[1], and Sid Khosla[2]

[1] Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati

[2] Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati

farboscs@mail.uc.edu

Introduction

According to the classic source-filter theory, sound is produced by the changes in the flow from glottal opening and closing during true vocal folds (TFV) vibration. The greatest rate of change in the flow occurs during the latter part of closing, when the velocity rapidly decreases. This rapid deceleration is quantified by the maximum flow declination rate (MFDR), which has been shown to correlate with acoustic intensity (i.e. loudness).

Several experimental[1] and computational[2] studies have shown that vortices can form near the superior aspect of the glottis during the closing phase of the folds vibration. Our preliminary work highly suggests that these intraglottal vortices can directly affect MFDR[1] and that the glottal flow attributes can vary with the vocal tract inertance[2].

In the current study we develop a computational model based on empirical data.

Methods

The geometry of the computational model was constructed from images taken in a canine larynx model using particle image velocimetry (PIV). The canine larynx was fitted with a vocal tract and inserts for the false vocal folds (FVF). Compressible large eddy simulation was employed to numerically investigate the laryngeal flow. The vocal tract inertance was varied by changing the gap between the FVF. Predicted results were compared with the experimental data.

Changes in laryngeal flow as a function of vocal tract inertance were also investigated for cases with parallel medial walls of the TFV (i.e. no divergence angle of the glottis during closing) and for cases with exaggerated divergence angle of the glottis during closing.

Results

The characteristic of the glottal flow predicted by the computational model were in good agreement with the experimental data. The results showed that the gap between the FVFs stretches the flow structure, reduces the jet curvature, and increases circulation in the flow. The presence of FVFs has a significant effect on the laryngeal flow resistance.

Having parallel TFV during closing showed that the aerodynamic forces near the superior aspect of the glottis did not change as the vocal tract inertance was increased. This observation is important because Titze[3] predicted change in pressure (near the superior aspect) as the vocal tract inertance is increased. Intraglottal flow separation did not occurred due to the parallel folds.

The study also showed that up to a certain angle value, the strength of the intraglottal vortices increased as the divergence angle was increased.

Conclusions

The vocal tract inertance has a significant impact on the glottal flow. The increase in inertance strengthens vortical structure in the flow, which in turn affects aerodynamic forces, which can affect MFDR.

References

[1]Oren, L., Khosla, S., Dembinski, D., Ying, J., & Gutmark, E. (2015). Direct measurement of planar flow rate in an excised canine larynx model. *The Laryngoscope*, *125*(2), 383-388.

[2]de Luzan, C. F., Chen, J., Mihaescu, M., Khosla, S. M., & Gutmark, E. (2015). Computational study of false vocal folds effects on unsteady airflows through static models of the human larynx. *Journal of biomechanics*, *48*(7), 1248-1257.

[3]Titze, I. R. (1988). The physics of small-amplitude oscillation of the vocal folds. *The Journal of the Acoustical Society of America*, *83*(4), 1536-1552.

The Influence of Men and Women Sexual Orientation on the Acoustic Features of Spanish Vowels.

Crisosto, Jaime[1] ; Pérez, Hernán[1]; Dagnino, Franco[1].

[1] Universidad del Desarrollo, Universidad de Concepción

jaimecrisosto@gmail.com

Introduction

The aim of this study is to determine in which extent some acoustic features present on the five spanish vowels (/i/, /e/, /a/, /o/, /u/) can be correlated significantly with men and women's sexual orientation. The investigation dialogues with some data founded on Smyth & Rogers (2008), Pierrehumbert et al. (2004) and Waksler (2001) related to the vowel triangle opening, the segments duration, the fundamental frequency variations and other acoustic parameters that may be relevant to discuss. One of the main study design characteristics is the analyses of the vowels with no linguistic context, to avoid the intonational contours effects.

Methods

20 heterosexuals and homosexuals participants of both sexes should read aloud 15 monosyllabic words. These words have the 5 spanish vowels in a consonant – vowel – consonant phonetic context, with no nasals to minimize the coarticulation effects. The data was recorded on the TASCAM DR-100mkll digital recorder in 41.1 ksps WAV format connected to a unidirectional dynamic SHURE microphone model SM48, with a sampling rate of 44100 Hz. The data was collected and analysed with the software PRAAT. The measured variables were F0, F1, F2, F1 and F2 normalized through Nearey 1 procedure, duration, minimum F0, maximum F0, F0 variation, intensity and the spectral tilt. For analysis, different stadistical tests were used depending on the behavior of the data.

Results

For men, of all measured variables, only vowel /i/ duration and vowel /u/ not normalized F1 shown higher values for the homosexual speakers compared to their heterosexual counterparts.

For women, it was found that vowel /i/ not normalized F1, normalized F1 and spectral tilt were higher for the heterosexual speakers. Vowel /e/ F0, minimum F0 and maximum F0 shown higher values for the homosexual speakers. Vowel /a/ maximum F0 was higher for the homosexual women.

Conclusions

The results do not show an organization around a specific variable or a clear pattern, suggesting that the differences between this population (if any) are not *inside* the vowels and

that it is necessary to complement the studies with suprasegmental information and consonantal data to stablish possible profiles associated with the sexual orientation.

References

Pierrehumbert, J., Bent, T., Munson, B., Bradlow, A., Bailey, JM. (2004). The influence of sexual orientation on vowel production (L). J. Acoust. Soc. Am. 116:4, 1905 – 1908.

Smyth, R., Rogers, H. (2008) Do gay-sounding men speak like women?. En S. Cummins, B. Janowski, P. Shaw (Eds.), Toronto Working Papers in Linguistics 27, 129 – 144. Special Issue: All the things you are: a festschrift in honor of Jack Chambers.

Waksler, R. (2001). Pitch range and women's sexual orientation, WORD, 52,

A low-order vocal-fold model with an aerodynamic degree of freedom: theory and experiments

D. Sciamarella [1,3], G. Artana [2,3]

[1] LIMSI-CNRS, [2] LFD-FIUBA, [3] LIA PMF

denisse.sciamarella@limsi.fr

Introduction

Low-order models for vocal-fold oscillation use simplified airflow models, ruling out interesting phenomena of aerodynamic origin [1]. Intra-glottal flow unsteadiness is often neglected, in spite of its importance during the closing phase of the vibration cycle - known to generate the higher harmonics that are perceptually relevant for speech [2]. A low-model model accounting for intra-glottal flow unsteadiness is built and tested using numerical simulations and experiments.

Methods

Flow unsteadiness and non-uniformity along the glottal channel can be simultaneously included into a low-order system following the stream filament theory [3], where the flow quantities are a function of both, time and position along the streamline. This flow description leads to a dynamical system with a supplementary degree of freedom representing intra-glottal flow behavior. The performance of the approach is evaluated using numerical simulations reproducing measurements from previous works [4] as well as experiments on a self-oscillating replica inspired in [5].

Results

Numerical simulations with the stream-filament model predict high/low glottal pressure peaks at the opening/closing instants (Fig.1a), together with a non-uniform sensitivity of these fluctuations along the glottal length – (Fig. 1b). Pressure measurements with driven oscillations in straight-channel vocal-fold replicas show similar high-/low-pressure peaks [4]. These double peak distributions are retrieved when the stream filament approach is applied to the case of a straight glottal channel. The low-order model also provides satisfactory predictions for self-oscillating experimental results.

Conclusions

An intra-glottal flow model based on the stream filament theory is proposed to improve the prediction of pressure fluctuations along the glottis using low-order vocal-fold models. The aerodynamic degree of freedom allows for strong pressure fluctuations at the opening/closure

instants of the glottal cycle. The model predictions are assessed through numerical and experimental essays.

References

[1] D. Sciamarella, G. Artana. Relaxation to one-dimensional postglottal flow in a vocal fold model. Speech Commmunication, 66, 176–181 (2015)

[2] M. Krane, M. Barry, T. Wei, Dynamics of temporal variations in phonatory flow. J. Acoust. Soc. Am. 128 (1) 372-383 (2010)

[3] J. H. Spurk, Fluid Mechanics, Berlin Heidelberg, Chapter 9: Stream Filament Theory, pp. 261-276 (1997)

[4] Deverge M1, Pelorson X, Vilain C, Lagrée PY, Chentouf F, Willems J, Hirschberg A. Influence of collision on the flow through in-vitro rigid models of the vocal folds. J. Acoust. Soc. Am. 114 (6):3354-62 (2003)

[5] Petr Šidlof , Olivier Doaré, Olivier Cadot, Antoine Chaigne. Measurement of flow separation in a human vocal folds model. Experiments in Fluids 51:123–136 (2011)

Implementation of the vocal function exercises program in teachers: acoustic-perceptives effects.

Zavala Paulina[1], Hinojosa Consuelo[1], Pizarro Nayaret[1], Ramírez Makarena[1], Valenzuela Loreto[1], Yáñez Carolina[1]

[1]Escuela de Fonoaudiología - Universidad de Playa Ancha Convenio Mineduc/Upla

paulina.zavala@upla.cl

Introduction

In our country, teachers do not have any specific training on the use of their professional voice, which includes vocal techniques and hygiene measures for its care. Because of this, there is a risk of acquiring vocal pathologies, added to the inadequate acoustic, noisy environment and long hours in their daily practice. Considering that in these cases voice not only has personal, social and communicational implications, but also professional and economic ones[i], a vocal disturbance impacts in every dimension in an individual's life. A study by UNESCO/OREALC about work and health conditions in Latin American teachers, refers to that in Chile, there is a high percentage of vocal alterations in this group (46,2%). Another investigation about work conditions and teacher's wellness/discomfort in Chile, evidenced ergonomic requirements perceived as excessive by teachers, especially when it comes to standing during work hours (86%) and having to strain their voice (76%). These are distinguished in a high significant level by public establishments teachers[ii]. This study rates and describes the effects of vocal function excercises program (VFE)[iii] in teachers of a Viña del Mar municipal corporation establishment, which was raised on the need to balance the voice subsystems, seeking harmony between high vocal demand in teachers and a physiological and friendly voice production with the vocal body schema.

Method

Quantitative research, quasi-experimental, descriptive/cross-cutting nature. The target population corresponds to teachers that impart elementary education lessons in "Libertador Bernardo O'higgins" school. The non-probabilistic sample counted with 8 teachers, men and woman, with ages between 25 and 59 years old. In the process, the presence/absence of vocal pathology through nasolaryngofibroscopy was valued, a vocal hygiene awareness and misuse/vocal abuse guide[iv]; a protocol to rate life quality voice[v]; vocal/clinical history and phonetic acoustic analysis with PRAAT software were applied. To ensure noise control and interferences in the samples, a "m Fast Track pro" interface and a condenser-cardioid microphone Senheiser e835 were used. Later, the VFE program was introduced to finally apply a vocal perception scale and analyze the post program results.

Results

Initially, the sample showed ignorance in measures of voice care and misuse/vocal abuse, some of them even presented organic and functional dysphonia. After the implementation of

the program, they all improved vocal parameters (f0, intensity and HNR) and awareness of self care measures, confirming hypothesis 1 which determines that this program generates positive effects in this professional group.

SUBJECTS	f0 (Hz)		INTENSITY (dB)		JITTER RAP (%)		LOCAL SHIMMER (dB)		HNR	
VFE	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
TEACHER 1	209	219	53	61	0.146	0.163	0.183	0.186	20.9	22.4
TEACHER 2	163	234	49	53	0.286	0.128	0.264	0.138	18.6	24.8
TEACHER 3	198	242	65	52	0.152	0.111	0.105	0.183	25.9	25.1
TEACHER 4	242	251	63	65	0.329	0.204	0.166	0.195	21.3	19
TEACHER 5	238	248	65	61	0.185	0.149	0.203	0.129	26	27
TEACHER 6	227	238	64	61	0.392	0.241	0.202	0.157	20	21
TEACHER 7	128	142	60	62	0.298	0.232	0.126	0.203	20	21
TEACHER	231	243	73	65	0.159	0.122	0.159	0.137	26	26.6
NORM (Farias, 2012; Boersma & Weenink, 2008).	Fem:205 / Masc: 127 (Aronson y Cols, 2001).		>60 dB y <70 dB Conversational voice		Normal 0,35 / Lìmite 0,65		Normal 0,17 / Lìmite 0,35		Normal 16.5 / Lìmite 12	

Table 1: Habitual tone sustained /a/ acoustic parameters, pre and post VFE implementation.

Conclusions

It is necessary for the group of teachers to perform a daily warm up and warm down vocal exercises routine, since these prevent possible pathologies and improve their vocal productions in the classroom. The Vocal Function Exercises program is an excellent resource to be considered because of its effectiveness, proven based on the evidence and simplicity of its application.However the acoustic measures are not sensitve enough to detect eventual vocal changes pre and post VFE, being in the sample higher sensoperceptive benefits.

Reference

Walker, MB, Líbera, MV. (2013). "Vos y tu voz". Herramientas para docentes, estudiantes, locutores, actores, cantantes, sacerdotes, abogados, martilleros y profesionales que viven con su voz. Editorial Dunken. 1ra edición. Buenos Aires, Argentina.

[1] Cornejo, R. (2009). Condiciones de trabajo y bienestar/malestar docente en profesores de enseñanza media de Santiago de Chile. *Educ. Soc.*, Campinas, vol 30, n. 107, may/ago., p. 409-426

[1] Stemple, JC., Glaze, LE., Gerdeman, BK. (2000). Clinical voice pathology: Theory and management. San Diego: Singular Pub. Group.

[1] Fletcher HM, Drinnan MJ, Carding PN.(2006). "Voice Care Knowledge Among Clinicians and People with Healthy Voices or Dysphonia". J Voice. 2007 Jan;21(1):80-91.

[1] Hogikyan ND, Sethuraman G. (1999) Validation of an instrument to measure voice-related quality of life (V-RQOL). J Voice. 1999;13(4):557-69.

Phonatory aeroacoustics – a Lagrangian perspective

Michael J. McPhail[1,2] and Michael H. Krane[1,2,3]

[1] Applied Research Laboratory, Pennsylvania State University, [2] Department of Biomedical Engineering, Pennsylvania State University, [3] Department of Acoustics, Pennsylvania State University

mjm75@psu.edu

Introduction

As far back as Teager (1980) it has been hypothesized that features of glottal jet topology (e.g. asymmetric jet, Coanda effect) have an effect on sound production. Recent work demonstrates that voice disorders can modify glottal jet topology-- see Mittal *et al.* (2013) for a review. Yet it remains difficult to causally relate airflow topology to sound production, primarily because of the ambiguous definitions of such flow features. Here we use a Lagrangian description of the air motion in the glottis to organize airflow topology, and estimate sound sources in terms of that topology.

Methods

This work is a post-processing of fluid-structure interaction (FSI) simulations conducted at Rensselaer Polytechnic Institute (RPI) (see Wang & Zhang (2013) for a description of their modified immersed boundary method). Simulations consisted of a two-dimensional duct with compliant two-layer model vocal folds using the M5 profile (Scherer & Shinwari 2001). Lagrangian coherent structures (LCS) were identified from the velocity fields of the simulations. LCS are fluid-fluid boundaries that reveal the flow topology by delineating regions of fluid with qualitatively different dynamics. In phonation, LCS delineate the region of air pushed through the glottis during a cycle of vibration. LCS were computed with the techniques of Haller & Yuan (2000).

Results

The LCS, shown in figure 1 (right), delineate the boundaries of the region of air that is pushed through the glottis during a cycle of vibration. The temporal location of the region corresponds to the colored points on the volume flow waveform in figure 1 (left). The region starts as a compact shape in the subglottal space. The region deforms as it is extruded through the glottis by lung pressure. The bulbous portion at the front of the region, enclosing a vortex pair, grows as it is fed by the glottal jet. A secondary vortex pair is observed forming roughly halfway through the cycle. Near the end of the cycle, the region returns to a compact shape as the secondary vortex pair merges with the first.

This analysis allows for a pictorial interpretation of aeroacoustic sources, akin to classical acoustic theory. The rate of change of volume of the region produces a monopole, its net motion through the glottis produces a dipole, and the coalescing vortex pairs produce a quadrupole source. Further analysis, to be presented, allows estimation of these source terms.

Summary

Lagrangian coherent structures (LCS) were identified in a fluid-structure interaction simulation of phonation. LCS marked the boundary of air mass pushed through the glottis during a single cycle of vocal fold vibration, revealing the basic topological organization of the flow. This provides a rational basis to address the relationship between airflow topology and sound production.
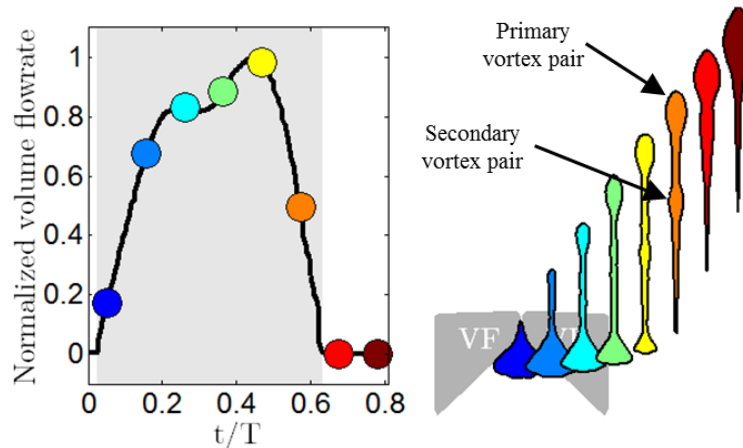
Figure 1. Results from FSI simulations. (Left) Normalized volume flow rate measured at the glottis versus a normalized time. Background colored gray when vocal folds are open. (Right) Series of regions identified with LCS through a cycle of vibration. Vocal folds are shown for orientation with first colored region; others are displaced horizontally for clarity. Color of region corresponds to colored points on the volume flow plot.

References

Haller, G. & Yuang, G. (2000) Lagrangian coherent structures and mixing in two-dimensional turbulence, *Physica D.* **147** (3), 352-370.

Mittal, R., Erath, B. D., & Plesniak, M. W. (2013). Fluid dynamics of human phonation and speech. *Annual Review of Fluid Mechanics*, *45*, 437-467.

Scherer R.C. & Shinwari D. (2001) Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. *J. Acoust. Soc. Am.* **109** (4) 1616-1630.

Teager, H. M. (1980). Some observations on oral air flow during phonation. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, *28*(5), 599-601.

Wang, X. & Zhang L. (2013) Modified immersed finite element for fully-coupled fluid-structure interactions, *Comp. Meth. Appl. Mech. Eng.* **267** 150-169.

Fluid-dynamical double-cavity properties of the laryngeal ventricle

F. Tuerke [1,2,3], D. Sciamarella [1,3], G. Artana [2,3]

[1] LIMSI-CNRS, [2] LFD-FIUBA, [3] LIA PMF

denisse.sciamarella@limsi.fr

Introduction

In the voice production process, the glottal jet encounters, upon exit, a geometry in the form of two fusiform fossa. The space defined by the two facing cavities is known as laryngeal ventricle or Morgagni's sinus.

The fluid dynamical problem of a single cavity flow has received much attention in the last decades [1,2] showing non-harmonic modes in the shear layer spectrum at the impinging edge of the cavity for velocities similar to the ones found in the laryngeal ventricle. These incommensurable frequency peaks were explained in terms of a linear stability analysis in a finite domain [3]. The double-cavity case for a starting jet was studied numerically and with flow visualization techniques [4].

The present work fills the gap between the two previous works [1,2] and [4]. We consider a double-cavity geometry in the stationary regime and investigate its fluid dynamical properties. The double cavity is chosen such that its dimensions match the values for the characteristic parameters inspired in the case of the larynx.

Methods

A numerical investigation was implemented to simulate a jet that exits upon a schematic deep double cavity. The 2D and 3D cases were studied using an incompressible finite volume code for a time-independent channel inflow velocity profile.

Direct numerical simulations (DNS) were carried out on the super-computing cluster at LIMSI-CNRS (Orsay, Paris) using open MP parallel computing. The in-compressible, viscous Navier-Stokes equations were solved in a computational domain comprised of 256x128 grid points using *SUNFLUIDH*, an in-house non-commercial finite volume code developed by Yann Freigneau of LIMSI-CNRS.

Results

Fourier analysis of time series recorded at discrete probe points in the geometry was used to analyze the frequency response of the system as a function of the Reynolds number and the length to depth ratio. Dynamical mode decomposition analysis of snapshots of the entire flow field recorded at roughly 1000Hz was used to detect zones of activity for given frequency peaks.

Conclusions

The fluid-dynamical properties of the laryngeal ventricle were investigated numerically in 2D and 3D schematic geometries. DMD results show zones where specific frequency content is produced. The double-cavity system is found to behave aerodynamically as a coupled system formed by two facing single cavities. Results provide a new framework for the description of the laryngeal ventricle flow dynamics.

References

[1] Basley et al. Experimental investigation of global structures in an incompressible cavity flow using time-resolved PIV Exp. Fluids 50, 905 (2010)

[2] Lusseyran, et al. Dynamical analysis of an intermittency in an open cavity. Phys. Fluids 20, 114101 (2008)

[3] Tuerke et al. Frequency-selection mechanism in incompressible open-cavity flows via reflected instability waves. PHYSICAL REVIEW E 91, 013005 (2015)

[4] Chisari et al. Vortex dipolar structures in a rigid model of the larynx at flow onset. Exp Fluids 50:397–406 (2011)

Acoustic voice measurements after voice warm-up and cool-down in choir singers

Lílian Aguiar-Ricz[1], Fernanda Onofre[1], Gleidy Vannesa E. Rojas[1], Deny Marco Garcia[1], Yuka de Almeida Prado[2]

[1] Department of Ophthalmology, Otorhinolaryngology and Head and Neck Surgery, School of Medicine of Ribeirão Preto, University of São Paulo.

[2] Department of Music, Faculty of Philosophy, Science and Letters of Ribeirão Preto, University of São Paulo.

liricz@fmrp.usp.br

Introduction

Singing is one the functions that most depend on the voice. Refined singing is a multivariate task that depends on extensive formation and training. This vocal competence for the production of singing voice is based on the presence of a functional, healthy and aesthetically acceptable voice, as well as on arduous training. It is known that an exaggerate vocal demand favors vocal fatigue. After years of treaining with classical music, a singer is expected to be able to produce a beautiful voice without effort regardless of the requirements of his performance. However, the voice rest on the day after the performance in opera singers can facilitate voice recovery. On this basis, it should be kept in mind that, after the use of the singing voice it is necessary to performe voice cool-down with a time of rest in order to reestablish vocal tension, preventing the persistence of the stress incorporated into singing. The objective of this study was to evaluate the acoustic measurements of the vowel "a" in modal recording before and after a singing voice test and after 30 minutes of absolute rest in female choir singers.

Methods

Thirteen soprano choir singers with experience in choir singing were evaluated by analysis of acoustic voice parameters at three time points: before continuous use of the voice, after voice warm-up and a singing test of 60 minutes duration respecting the breathing pauses and after voice cool-down and an absolute voice rest for 30 minutes.

Results

The fundamental frequency increased after the voice resistance test (p= 0.012) and after the voice rest (p= 0.01). The jitter decreased after the voice resistance test (p= 0.02) and after the 30 minutes of voice rest. A significant difference was detected for the acoustic voice

parameters relative average perturbation (RAP), p=0.05, and pitch perturbation quotient (PPQ), p= 0.04, compared to the initial time point.

Conclusions

The fundamental frequency increased after 60 minutes of singing and continued to be elevated after vocal cool-down and absolute rest for 30 minutes, proving to be an efficient parameter for the identification of changes inherent to voice demand during singing.

Speech Therapy Treatment of Benign Organic Lesions of the Vocal Fold Mucosa: Case Studies.

Alicia C. Hernandez [1], Valeria I. Pereyra [1]

[1] Universidad Nacional de cordoba

cursofono@hotmail.com

Introduction.

Benign organic lesions of the vocal fold mucosa (BOLM), which occur frequently among voice professionals, are typically resolved through surgery. This study aims at showing the resolution of such pathologies through the application of the inspiratory phonation (IP) technique in cases of early development among voice professionals. In our experience, this technique contributes to the reduction and/or elimination of such lesions as retention cysts (RC) and vascular alterations (VA) of the free edge. In addition, it constitutes a beneficial alternative to non-invasive procedures.

Methods.

The IP technique causes glottis fixation in the midline, decrease in subglottic pressure, high-pitched vocalization, and vocal fold vibration without the intervention of any other laryngeal structure. Besides, it produces the involuntary contraction of the cricothyroid muscle, reducing the approximation strength of the vocal folds. This technique was applied to patients diagnosed with RC and VA through rigid teleendoscopy with stroboscope light (RESL). An initial treatment session was conducted in which the technique was applied for five minutes six times a day during a month's time. Perceptual control (PC) was conducted on a weekly basis and medical control was performed monthly. Positive results were obtained between weeks 4 and 16.

Results.

Favorable changes were reported with PC 3 weeks after the beginning of the treatment and resolution of RCs was observed with RESL after a 4-week period. In addition, favorable changes were reported with PC at weeks 13-14 and resolution of VAs was observed with RESL at week 16.
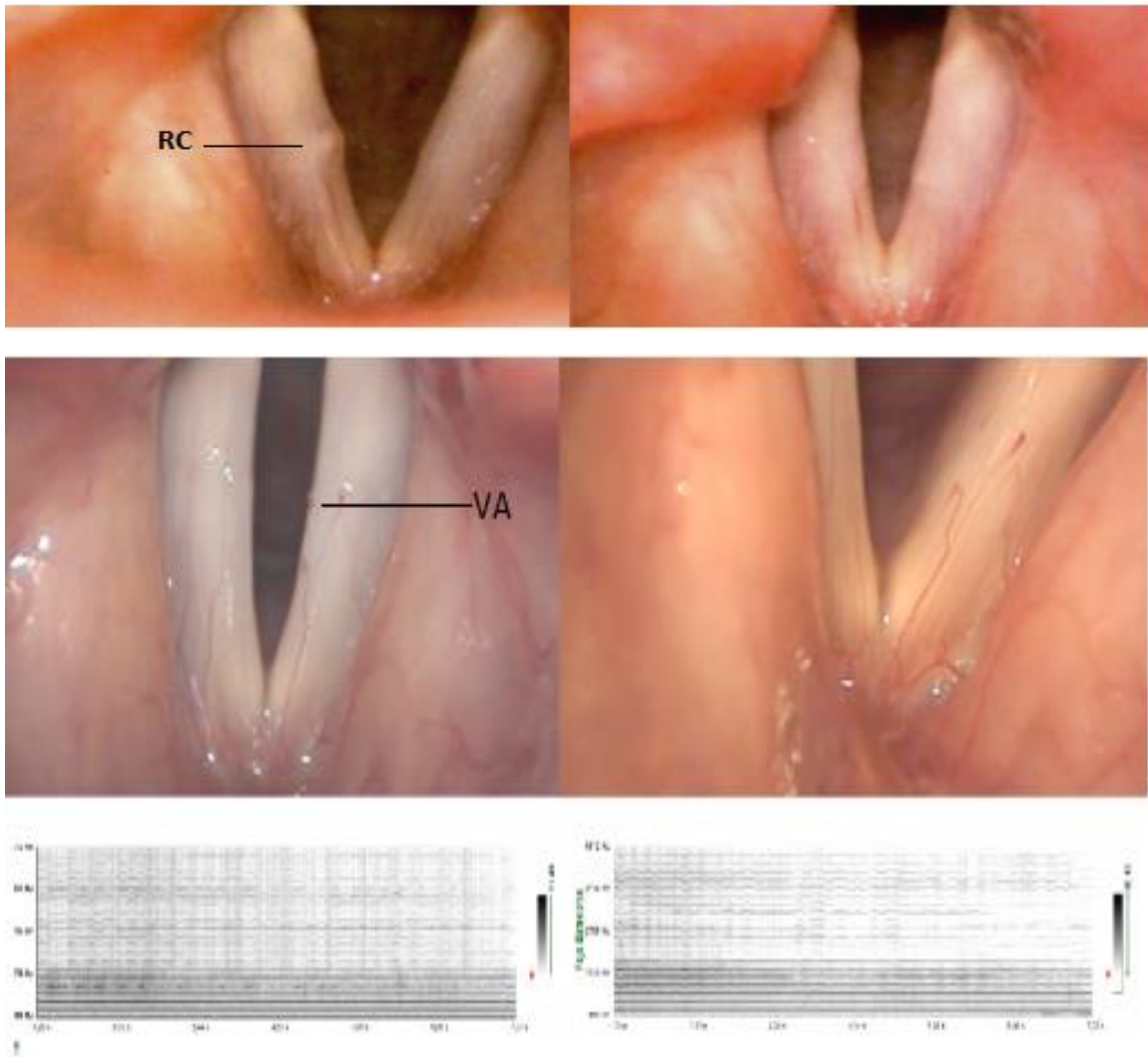
Conclusion.

This study shows that BOLM conditions were resolved successfully through the IP technique, thus avoiding surgical procedures. This alternative therapy becomes a valuable option for voice professionals in the preservation of their quality of life.

References.

Behlau, M.Cobeta, I. 2001. O melhor que vi e ouvi III Actualizacao em Laringe e Voz. Rio de Janeiro: Livraria e Editora Revinter Ltda.

Colton, R.H. Casper, J.K. Leonard, R. 2010.Compreendendo os problemas da voz. Uma Perspectiva Fisiológica no Diagnóstico e Tratamento das Disfonías. Rio de Janeiro: Livraria e Editora Revinter Ltda.

Harrison, G.A. Davis, P.J. Troughrt. R. Winkworth, A. 1992. Inspiratory speech as a management option for spastic disphonia-case study. Annals of Otollogy, Rhinollogy and Laringollogy, 101, 375-382.

Schindler, A. Mozzanica, F. Ginocchio, D. Maruzzi, P. Atac, M. Ottaviani, F. 2012. Vocal improvement after voice therapy in the treatment of benign vocal fold lesions. Acta Otorhinolaringol Ital v 32(5): 304-308.

Preliminary normative data of the Phonatory Aerodynamic System model 6600 for the adult Brazilian population

Marcela Lopes[1], Marília Lugli[1], Telma Kioko Takeshita-Monaretti[1], Patricia Moraes Pereira[1], Gleidy Vannesa E. Rojas[1], Hilton Ricz[1] Lílian Aguiar-Ricz[1]

[1]Department of Ophthalmology, Otorhinolaryngology and Head and Neck Surgery, School of Medicine of Ribeirão Preto, University of São Paulo

liricz@fmrp.usp.br

Introduction

It is believed that normatization of the aerodynamic and vocal parameters should be obtained based on the characteristics of each population, a fact that requires the standardization of the measures of the Phonatory aerodynamic system Model 6600 (Kaypentax®). The objective of the present study was to establish the preliminary normative values of the aerodynamic and phonatory parameters for a Brazilian population of adult men and women.

Methods

Twenty-five individuals, 6 men and 19 women aged on average 36 years, were invited to participate. Subjects with any complaint or disease or subjects previously treated for phonatory function or with harmful voice habits were excluded. The Phonatory Aerodynamic System (PAS) *Model 6600* (Kaypentax®) was used to obtain the aerodynamic and phonatory measures. Each participant performed the following tests: Vital capacity: the subject was instructed to inhale deeply and then exhale into the PAS mask sealed on his face; Maximum phonation time: each subject was asked to produce the sustained vowel "a" at habitual frequency and intensity for the longest possible time and with the highest phonatory efficiency and then to repeat the /pa/ sound several times. To guarantee equality of rhythm, the subjects were trained for this task.

Results

Regarding Vital capacity, two of the measurements obtained in the present study were found to be similar to those obtained by Zraick et al. (2012), who reported normality data for adult individuals divided by age range. Peak Expiratory Airflow was lower for both men and women. Regarding the Maximun Sustained Phonation protocol, all measures used were similar to those obtained by Zraick et al. (2012). For the Voicing Efficiency protocol there was a difference between the values obtained here and those obtained by Zraick et al. (2012), as observed for Expiratory Airflow Duration, which was lower for both men and women and for Peak Air Pressureue, which was higher for both sexes. The table 1 describes the results.

Conclusion

The present study is a preliminary analysis that intends to contribute reference values for the adult Brazilian population regarding aerodynamic and phonatory measures using the phonatory aerodynamic system model, favoring the development of new investigations and applicability to clinical practice.

Reference

Zraick RI, Smith-Olinde L, Shotts LL. Adult normative data for the KayPENTAX Phonatory Aerodynamic System Model 6600. J Voice. 2012 Mar;26:164-76.

Table 1. Phonatory Aerodynamic System Parameters

Parameter	Women		Men	
	Vital Capacity			
	Range	**Mean**	**Range**	**Mean**
Expiratory Airflow Duration (s)	2.42- 13.03	5.88	2.91- 13.94	7.17
Peak Expiratory Airflow (L/s)	0.28- 1.87	0.77	0.48- 5.04	1.49
Expiratory Volume (L)	1.22-3.64	1.981	1.9-3.64	2.85
	Maximun Sustained Phonation			
Maximum SPL (dB)	71.04-99.62	82.80	75.18-99.53	83.26
Minimum SPL (dB)	36.28 - 78.39	59.11	41.25-56.9	51.56
Mean SPL (dB)	64.95-95.92	76.65	64.95-89.68	75.15
Mean SPL During Voicing (dB)	64.95- 95.92	76.90	62.25- 91.98	75.62
Mean Pitch (Hz)	135.06 - 248.56	207.99	106.9 - 214.52	144.09
	Voicing Efficiency			
Mean SPL During Voicing (dB)	73.07 - 99.85	81.25	69.8 - 94.47	82.16
Mean Pitch (Hz)	151.91 - 242.35	213.38	100.8-167.17	134.41
Expiratory Airflow Duration (s)	0.22-2.14	0.68	0.26-1.66	0.64
Peak Air Pressure (cmH$_2$O)	2.19-19.06	8.81	2.05-19.3	9.23
Peak Expiratory Airflow (L/s)	0.02-0.6	0.19	0.1-0.42	0.27
Target Airflow (L/s)	0-0.48	0.12	0.05-0.14	0.15
Mean Airflow During Voicing (L/s)	0-0.47	0.11	0.19-0.47	0.25
Aerodynamic Power	0.002- 0.457	0.11	0.048-2.217	0.12

Preliminary normative data of the Phonatory Aerodynamic System model 6600 for the elderly Brazilian population

Marília Lugli[1], Marcela Lopes[1], Gleidy Vannesa E. Rojas[1], Telma Kioko Takeshita-Monaretti[1], Patricia Moraes Pereira, Hilton Ricz Lílian Aguiar-Ricz

Department of Ophthalmology, Otorhinolaryngology and Head and Neck Surgery, School of Medicine of Ribeirão Preto, University of São Paulo

liricz@fmrp.usp.br

Introduction

It is believed that normatization of the aerodynamic and vocal parameters should be obtained based on the characteristics of each population, a fact that requires the standardization of the measures of the Phonatory aerodynamic system Model 6600 (Kaypentax®). Modifications of the pattern and skill of communication occur physiologically with aging. Changes in voice representing presbyphonia, in turn, may be due to changes in the phonatory apparatus or in the auditory self-monitoring, influencing the psychosocial aspects of elderly subjects by interfering with their social functioning. The objective of the present study was to establish the preliminary normative values of the aerodynamic and phonatory parameters for a Brazilian population of elderly men and women.

Methods

Fourteen subjects, 7 men and 7 women older than 60 years and aged on average 70 years, were invited to participate. Subjects with any complaint or disease or subjects previously treated for phonatory function or with harmful voice habits were excluded. The Phonatory Aerodynamic System (PAS) *Model 6600* (Kaypentax®) was used to obtain the aerodynamic and phonatory measures. Each participant performed the following tests: Vital capacity: the subject was instructed to inhale deeply and then exhale into the PAS mask sealed on his face; Maximum phonation time: each subject was asked to produce the sustained vowel "a" at habitual frequency and intensity for the longest possible time and with the highest phonatory

efficiency and then to repeat the /pa/ sound several times. To guarantee equality of rhythm, the subjects were trained for this task.

Results

Regarding Vital capacity, the values of expiratory airflow duration were higher for men and similar for women when compared to the measures obtained by Zraick et al. (2012), who reported normality data divided by age range (Elderly aged 60-89 yesr). The present peak expiratory airflow and expiratory volume values were lower for both sexes. Regarding the Maximun Sustained Phonation protocol, all measures obtained here were similar to those obtained by Zraick et al. (2012) for women, but were reduced for men. Regarding the Voicing Efficiency protocol, there was a difference between the values obtained here for women and those obtained by Zraick et al. (2012), as observed for mean pitch, which was higher, and for Expiratory Airflow Duration and Expiratory Volume which were lower. For men, most parameters differed from the reference values of Zraick et al. (2012). The table 1 describes the results.

Conclusion

The present study is a preliminary analysis that intends to contribute reference values for the elderly Brazilian population regarding aerodynamic and phonatory measures, so far absent in the literature, using the phonatory aerodynamic system model, favoring the development of new investigations and applicability to clinical practice.

Table 1. Phonatory Aerodynamic System Parameters

Parameter	Women		Men	
	Vital Capacity			
	Range	Mean	Range	Mean
Expiratory Airflow Duration (s)	1.88 - 12.35	5.11	4.31 - 10.36	5.62
Peak Expiratory Airflow (L/s)	0.35 - 1.59	0.93	0.49 - 1.25	0.78
Expiratory Volume (L)	0.32 - 1.9	1.38	0.54 - 3.09	1.58
	Maximun Sustained Phonation			
Maximum SPL (dB)	81.33 - 96.83	87.32	67.43 - 86.92	78.37
Minimum SPL (dB)	41.28 - 78.16	59.64	28.09 - 60.71	44.77
Mean SPL (dB)	73.43 - 89.13	80.82	63.71 - 76.9	69.35
Mean SPL During Voicing (dB)	75.26 - 89.13	81.19	63.84 - 77.8	71.62
Mean Pitch (Hz)	138.59 - 230.31	192.79	110.69 - 162.87	144.41
	Voicing Efficiency			
Mean SPL During Voicing (dB)	80.72 - 95.56	86.3	71.98 - 80.18	77.86
Mean Pitch (Hz)	137.13 - 243.55	205.3	96.36 - 165.31	135.01
Expiratory Airflow Duration (s)	0.38 - 0.94	0.67	0.5 - 0.9	0.68
Peak Air Pressure (cmH$_2$O)	3.55 - 16.06	10.76	4.39 - 12.11	9.68
Peak Expiratory Airflow (L/s)	0.03 - 0.32	0.15	0.09 - 0.13	0.13
Target Airflow (L/s)	0.01 - 0.27	0.11	0.04 - 0.2	0.08
Mean Airflow During Voicing (L/s)	0.01 - 0.25	0.10	0.04 - 0.19	0.08
Aerodynamic Power	0.014 - 0.391	0.11	0.026 - 0.117	0.06

Vocal fold activity detection in a neck surface acceleration signal

Felipe Acevedo [(1)], Agustín González [(1)], Matías Zañartu [(1)]

[(1)] Universidad Técnica Federico Santa María

felipe.acevedo.lr@gmail.com

Introduction

Current approaches for ambulatory monitoring of vocal function use a neck surface accelerometer (ACC) for continuously assessing the activity of the vocal folds, and ultimately providing real time feedback to patients undergoing voice therapy [1]. This approach generates a big volume of data; although, only a small percentage corresponds to the actual activity of the vocal folds. In this study, we analyze features within the accelerometer signal that can robustly detect vocal fold activity, which would help to facilitate subsequent data analysis, storage, processing, transmission, and its associated energy consumption.

There are various algorithms to extract voiced segments from microphone signals (Voice Activity Detector, VAD). These schemes do not work well with accelerometer signals due to differences in this signal source, wherein for instance, unvoiced-sound-like segments should be discarded. Current methods for VAD using ACC [1] are based on heuristic approaches and have not been consistently examined. Thus, a comprehensive analysis of current methods for both microphone and ACC signals is pursued to then develop a specific algorithm for Vocal fold Activity Detection (VfAD).

Methods

We analyzed 8 features from different VAD schemes, i.e., Root Mean Square (RMS), fundamental frequency, periodicity, normalized peak, Cepstral Peak Prominence (CPP), Low-High spectral ratio, spectral tilt, and zero crossing rate.

We implemented these schemes based on one method at the time, and applied each to 11 samples of accelerometer signals. These signals were recorded along with microphone signals during a rainbow passage and an expert speech therapist marked ground truth segments as voiced, unvoiced, and silence.

To compare each of the 8 methods, we registered the best performance of each approach by tuning their detection parameters to minimize the average error among all subjects. The unvoiced segments are considered as silence in the accelerometer results, as these signals are not produced by the vocal folds. We initially opted for a 100 [ms] window size to allow for capturing a minimum number of cycles for an accurate classification in low-frequency voices.

Results

Table 1 shows the percent errors of each method as the sum of type I and II errors when running at its best performance and after taking the average among all subjects.

	Accelerometer		Microphone	
	Avg Error %	Stdev Error %	Avg Error %	Stdev Error %
RMS	12.41	7.96	13.42	2.60
Fund Freq	19.24	15.91	13.99	3.40
Periodicity	36.41	10.58	40.70	7.38
Norm Peak	28.24	6.16	23.41	8.14
CPP	19.98	5.82	21.35	6.25
L-H ratio	11.03	2.45	25.34	6.52
Spectral Tilt	14.01	10.59	22.48	10.61
ZCR	23.44	16.47	11.23	8.53

Table ¡Error! Secuencia no especificada. **Percent error or each VAD scheme**

Note that by using non-overlapped windows for classification, there is around 9% of discretization error in average for the number of transitions in our samples.

Conclusions

The results show that low-high spectral ratio, energy estimation, and spectral tilt are better suited for accelerometer signals. Although zero crossing rate showed to be the best detection scheme for microphone signals, it is not a good detection technique for vocal fold activity detection.

For subsequent steps, we want to study how these features work in presence of the noise that can appear in the accelerometer signal during daily activities and propose an algorithm for Vocal fold Activity Detection. Overlapping windows and other window sizes will be explored.

References

(1) Daryush D. Mehta, Matías Zañartu, Shengran W. Feng, Harold A. Cheyne II, and Robert E, Hillman, et al. "Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform." Biomedical Engineering, IEEE Transactions on 59.11 (2012): 3090-3096.

Voice therapy program for presbyphonia: preliminary results

Gleidy Vannesa Espitia Rojas [(1)], Patrícia Pereira Morais[(1)], Lilian Neto Aguiar- Ricz[(1)]

[(1)] Department of Ophthalmology, Otorhinolaryngology and Head and Neck Surgery, School of Medicine of Ribeirão Preto, University of São Paulo
vannesaespitia08@hotmail.com

Introduction

Presbyphonia is the result of the relation between age and the degenerative alterations of the larynx and breathing which leads to glottic[1] insufficiency and to the decrease of inspiratory and expiratory pressure[2]. These alterations cause a series of voice signals, such as: hoarseness, breathiness, reduction of voice intensity and vocal flexibility, vocal tremor and poor vocal projection[3], causing not only discomfort for the elderly but also frustration and anxiety, which could considerably worsen their quality of life. Considering the elderly's needs and the few pieces of scientific evidence available regarding this topic, it was proposed to develop a therapy program for presbyphonia, aiming at standardizing the duration of treatment, exercises, frequency and dosage, promoting better voice quality of individuals, restoring their lifestyle and improving their quality of life. Therefore, the objective of this study was to verify the effect of a voice therapy program on individuals presenting presbyphonia.

Methods

Pre-test / post-test experimental study, with the participation of 13 adults aging over 60 years (the mean age was 70 years), among which there were 7 female and 5 male with a diagnosis of presbylarynx, randomly divided into two study groups, with the proposal of two different forms of semi-ocluded vocal tract exercises with the use of tubes. Group I (GI) performed the voice exercise with the tube keeping one end open for air flow, making prolonged emission of the sound /uuuuuu/, and Group II (GII) performed the technique Lax Vox, with the prolonged emission of the sound /juuubuuu/. All of them received vocal health orientation during treatment. Participants recorded the sustained emission of the vowel /a/ in the first evaluation and in each of the eleven weekly sessions programmed. To analyze the acoustic measurements the Multi-Dimensional Voice Program (MDVP), from the Computerized Speech Lab/ Model 6103, by Kay Pentax® was used. The acoustic measurements analyzed were: Mean Fundamental Frequency, Jitter Percentual, Relative Average Perturbation, Pitch Perturbation Quotient, Shimmer in dB, Noise to Harmonic Ratio.

Results

All participants, from GI and GII, received 11 weekly sessions of voice therapy, each lasting 45 minutes, and voice exercises were prescribed to be performed daily at home, increasing frequency progressively during sessions, from 3 to 9 times a day, lasting from 2 to 4 minutes, considering as a reference point for each individual their maximum phonation time. GI and GII participants performed the voice exercise with normal frequency and intensity for 6

sessions and other 4 sessions with frequency modulation from low to high sounds. GII participants did the exercise with tube water depth ranging from 2 to 12 centimeters according to their control of the technique.

The assessed acoustic parameters are distributed in table 1, for Groups I and II, and a tendency can be observed in the reduction of vocal measurements in the third and sixth sessions, with result stability in the final assessment.

Table 1. Distribution of the voice acoustic parameters: *Multidimensional Voice Program* (n=13).

Voice acoustic parameters	Voice Initial assessment		3 Sesssion Voice therapy		9 Sesssion Voice therapy		Voice final assessment	
	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD
Mean Fundamental Frequency (Hz)	GI Male (n=3)							
	145,6-164,0	153,6±9,3	124,4-195,2	159,9±35,3	131,4-188,0	157,8±28,4	138,2-142,1	140,2±2,80
	GII Male (n=3)							
	110,9-126,9	116,4±9	115,9-145,3	128,8±15	131,1-141,9	136,3±5,4	127,4-143,2	137,0±8,4
	GI Female (n=4)							
	121,3-193,4	223,8±30,1	178,4-253,7	204,7±34,4	178,4-253,7	204,7±34,4	182,5-257,8	220,2±53,2
	GII Female (n=3)							
	170,5-229,0	192,4±31,9	164,9-182,5	171,4±9,6	163,3-176,6	168,3±7,1	174,3-184,3	178,3±5,2
***Jitter* percentual Jitt (%)**	GI (n=7)							
	0,6-8,2	2,8±2,8	1,9-0,9	0,5±3,5	0,6-1,4	0,8±0,3	0,3-1,5	1,0±0,6
	GII (n=6)							
	0,54-5,0	2,5±1,9	0,5-6,1	1,8±2,1	0,3-4,2	1,42±0,8	0,3-3,4	1,72±1,3
Relative Average Perturbation RAP	GI (n=7)							
	0,42-4,65	1,6±1,5	0,29-2,0	1,15±0,54	0,21-0,90	0,47±0,23	0,1-0,91	0,58±0,36
	GII (n=6)							
	0,3-4,6	1,4±1,1	0,2-2	1,1±0,5	0,19-2,42	0,88±0,84	0,2-2,07	1,0±0,83
Pitch Perturbation Quotient (PPQ)	GI (n=7)							
	0,39-5,7	1,9±2,0	0,30-2,3	1,19±0,62	0,20-0,83	0,46±0,21	0,21-0,97	0,60±0,35
	GII (n=6)							
	0,31-3,2	1,5±1,2	0,31-4,06	1,2±1,4	0,19-2,75	0,92±0,9	0,19-1,88	0,96±0,74
Shimmer in dB	GI (n=7)							
	0,32-2,0	0,95±0,71	0,15-1,27	0,65±0,35	0,17-0,56	0,31±0,13	0,24-0,89	0,59±0,32
	GII (n=6)							
	0,10-3,42	0,57±0,39	0,21-0,82	0,41±0,23	0,10-0,85	0,36±0,28	0,18-1,0	0,51±0,35
Noise to Harmonic Ratio	GI (n=7)							
	0,12-0,66	0,32±0,24	0,11-0,45	0,22±0,11	0,11-0,17	0,13±0,02	0,08-0,17	0,13±0,05
	GII (n=6)							
	0,10-0,55	0,26±0,18	0,10-0,61	0,23±0,19	0,10-0,27	0,15±0,06	0,11-0,24	0,17±0,05

SD: standard deviation, GI: group I, GII: group II.

Conclusions

The therapeutical program proposed may be an option to improve the acoustic parameters and stabilize the voice of participating elderly. However, the sample must be enlarged and comparative statistic tests must be made.

Relationship between laryngeal morpho-functional diagnosis and acoustic parameters of voice

María Esperanza Sastoque H [1], Martha L. Rodríguez P [1],
Marco A. Guzmán N [2], Antonio J. Reyes [3].

[1] University of Valle. Colombia. [2] University of Chile. [3] Centro Médico Imbanaco. Colombia.
maria.sastoque@correounivalle.edu.co

Introduction

Methods used for voice assessment have been strengthened through the implementation of protocols of clinical assessment, perceptual scales and the use of technology for objective evaluation. However, they have a high degree of variability due to the subjectivity in the assessment process and intrarater and interrater variability. This study aimed to determine the relationship between morpho-functional diagnosis of larynx and acoustic parameters of the voice in a group of patients who attended a voice center in the city of Cali, Colombia. Possible correlations of the parameters could provide reference data for assessment, treatment and prevention of voice disorders.

Methods

This is a descriptive, observational and cross-section study. 38 patients were assessed using laryngeal stroboscopy. They were ranked according to the morpho-functional diagnosis. Acoustic voice analysis was performed using the software Anagraf. Fundamental frequency and energy contours, perturbation measures, noise measurements and the Integrated vocal Perturbation Index (IPI), were carried out. Acoustic parameters and laryngeal diagnostics were compared through ANOVA and Kruskal-Wallis test. Multivariate analysis was also performed through Principal Component Analysis (PCA) analysis.

Subjects were classified into three different groups: 55.3% of the subjects had structural damage to vocal folds, 28.9% chronic laryngitis, and 15.8% functional dysphonia. Glottal closure and mucosal wave amplitude in structural lesions associated with the presence of cysts, nodules, or polyps in the vocal folds were detected. Acoustic parameters were significantly altered: shimmer ($P = 0.001$), HNR ($P = 0.002$) and amplitude of the cepstrum ($P = 0.003$). All these values together determined an Integrated vocal Perturbation Index (IPI= Higher than 3) with a p-value = 0.000 in the case of structural lesion. Values of these variables were found be increased and directly related to the amplitude of mucosa wave and glottal closure. The integrated voice index perturbation showed a high degree of relation to these variables in subjects with structural laryngeal dosirders detected through stroboscopy. The quality of measurements and standardization of the procedure among the evaluators resulted in a better reliability in the assessment.

Conclusions

Laryngeal stroboscopy allowed to detect vocal fold vibratory asymmetries, structural and glottal closure alterations or and other conditions that are not visible with continuous light to establish the morpho-functional diagnosis. Acoustic parameters showed consistency in perturbation measurements used to discriminate groups with different degrees of voice alteration. A high relationship between acoustic parameters and morphological and functional diagnosis of the larynx was found, specifically on the amplitude of vibration and glottal closure.

References

1. Jiang JJ, Zhang Y, MacCallum J, Sprecher A, Zhoul L. 2009. Objective acoustic analysis of pathological voices from patients with vocal nodules and polyps. Folia Phoniatr Logop. 61(6):342-9.
2. Kelley, RT; Colton, RH; Casper, J; Paseman, A; Brewer, D. 2011. Evaluation of Stroboscopic Signs. Journal of Voice, Vol. 25, No. 4. pp. 490-495.
3. Preciado, J; Calzada, M. Preciado, P. 2005. Examen de la función vocal y análisis acústico de 905 docentes de La Rioja. Acta Otorrinolaringol Esp; 56: 261-272.
4. Nawka, T; Konerding, U. 2012. The Interrater Reliability of Stroboscopy Evaluations. Journal of Voice, Vol. 26, No. 6. pp. 812.e1-812.e10.
5. Sastoque H, M.E; Bravo G, X. 2013. Acoustical parameters for the normal voice among a young adult population from an institution of higher education in Santiago de Cali, Colombia. Rev Gastrohnup, Vol 15 (3) (July-Dec): pp.181.

The utilization of the vocal biography for the comprehensive approach to the vocal symptomatology

Paola Silberman[1], Ma. Eugenia Ferrando[1]

[1]Universidad Católica de Santa Fe

paosilber@hotmail.com

Introduction

The vocal biography is a tool for serving the personal vocal inquiry. I consider the Voice as one of the most intimate elements,constituent of one´s identity and at the same time, paradoxically, one of the most public ones. Undoubtedly, our voice may sound according to our mood. It can be seen that there is a direct and indissoluble relation between the voice and the mood.

We know that the existing bibliography is very rich, coming from different currents of psychology that explain and support the elaboration of the patient´s Biography with the aim of revising and rewriting (ACÁ ME PARECIÓ QUE QUERÍAS DECIR reescribir. SI ES reinscribir LO BUSCO OTRA VEZ)their own history and its correlate in the present moment, to explain a conflict, emotional disturbances, symptoms. However, regarding vocal symptomatology we are used to hearing "he has nothing … it is emotional" or ,on the contrary, any vocal symptom is said to be of "emotional cause".

I have used this tool for more than 10 years now, trying to hear "the voice of the symptom, when the symptom is in the voice", considering the emotional symptom as integral without underestimating or overstating it. Thus, I believe it is useful to think all together how to help the patient to be part os their own process, changing their own self-image even in deeper levels.

Methods

20 patients were chosen and they attended the office during the year 201. They were between 25 and 60 years old, with vocal symptoms without differentiating the origin or the vocal pathology.

A questionnaire was implemented with questions about the vocal history that help reflect and take a greater register of the voice and communication in general. The questions are always the same and there is a graph that the patient draws about their own voice.

This inquiry is carried out during the one-hour session and it takes between one and three sessions to complete the information

Results

The results showed that in all the cases that were evaluated there were no recurrences in the last three years. In all the cases they could skillfully recognise their voice, as well as acoustic parameters and alterations. They developed a critical and reflexive glance at their vocal

performance with a higher and better register of their own voice and the multiple causes, including the emotional one that affected its optimisation.

Conclusions

The voice is the only manifestation. We are used to saying "I know it thoroughly"but in fact, it is not so.The utilization of this tool allows a more comprehensive insight from both sides (patient and therapist) of both the vocal history and its evolution in a symptom that the patient accuses in the voice. Although the questions are thought to provoke that personal vocal search and growth, there are no expected answers that are correct or "normal". Each answer is unique, as each person is, as each voice. The proposal of the Biography is to give the opportunity for having the courage to see both honestly and creatively a certain and unique scenario.

References

Adriana Schnake. La voz del Síntoma. (2004) Ed. Cuatro Vientos

Laura Gutman. La Biografia Humana.(2013) Ed.Planeta

Rüdiger Dahlke y Thorwald Dethlefsen. La enfermedad como camino. (2004)Ed.Debolsillo

Effect of the electromyographic biofeedback combined with vocal exercises in vocal quality of dysphonic women: a randomized controlled pilot trial

Vanessa Veis Ribeiro [1], Pamela Aparecida Medeiros Moreira [1], Eduardo Carvalho de Andrade[1], Alcione Ghedini Brasolotto [1], Kelly Cristina Silverio [1]

[1] Faculdade de Odontologia de Bauru/Universidade de São Paulo

vanessaribeirooo@hotmail.com

Introduction

One procedure that can be applied in the treatment of functional dysphonia in adult subjects is electromyographic biofeedback[1,2].

The literature reports that some therapeutic programs that applied the electromyographic biofeedback in dysphonic subjects showed improvement in vocal quality[1,3]. However, studies that verified the effectiveness of electromyographic biofeedback in cases of dysphonia differ on the application time and number of sessions, and do not describe the activities and the form of biofeedback control, impeding their replication.

From this information, it is observed that there is need to verify the effects of electromyographic biofeedback associated with a vocal exercise program, randomized and controlled, in the voice quality of dysphonic women, showing scientific evidence on the clinical efficiency this procedure.

Thus, the present study investigated the effect of electromyographic biofeedback associated with vocal exercises in vocal quality of dysphonic women.

Methods

This is a pilot clinical trial, randomized and controlled. This study was approved by the Institutional Ethics Committee (nº 1.235.463). Participated six women with functional dysphonia (ages 19-44 years, mean=37 years) that were divided randomly in: Experimental Group (EG) - three women who received eight sessions of the application of eletromyographic biofeedback (BioTrainer-Biotec software) with electrodes in suprahyodeus and sternocleidomastoid muscles associated with vocal exercises (vibrant sounds, nasal and fricative) for 30 minutes; Placebo Group (PG) - three women who received eight sessions of the placebo biofeedback electromyographic associated with the same vocal exercises, in the same conditions as the EG. There was no control of muscle electrical activity during exercise in PG. The individuals of PG were executed with the computer screen with a software used for electromyography (Miotool 200®), which does not provide visual feedback to the participant. Before and immediately after the therapeutic program, all participants underwent the perceptual-auditory and acoustic vocal assessment. The perceptual-auditory assessment

was performed by three speech therapists, from a visual analog scale. The parameters analyzed were: overall degree of vocal deviation, roughness, breathiness, tension and instability. For the analysis of each parameter was considered the mean between the ratings of the three judges. The acoustic analysis was performed with the Multi Dimension Voice Program (MDVP) of KayPentax®, based on the vowel /a/, extracting the parameters: fundamental frequency, jitter, shimmer and harmonic-noise ratio. Data were statistically analyzed using the Paired T-Test (p=0.05).

Results

There was a significant decrease in the overall degree of vocal deviation in EG after the procedure (p=0.005), while in the PG was observed a decrease in roughness (p=0.013).

Significant changes in acoustic parameters were found after the procedure in both groups. The sample size should be considered as a limitation of the study and its continuation may appoint new results.

Conclusions

We conclude that in dysphonic women studied, the vocal exercises program improved the voice quality parameters such as roughness, however, the application associated with the electromyographic biofeedback promoted improvements in the overall degree of vocal deviation of the participants.

References

1. Watson, T. S., Allen, S. J., & Allen, K. D. (1993). Ventricular fold dysphonia: application of biofeedback technology to a rare voice disorder. Behav Ther, 24(1), 439-46.

2. Wong, A. Y. H., Ma, E. P. T., & Yiu, E. M. L. (2011). Effects os practice variability on learning of relaxed phonation in vocally hyperfunctional speakers. J Voice, 25(3), e103-e113.

3. Allen, K. D., Bernstein, B., & Chait, D. H. (1991). EMG biofeedback treatment of pediatric hyperfunctional dysphonia. J Behav Ther Exp Psychiatry, 22(2), 97-101.

Acoustic and muscular changes in the voice of a singer after respiratory support exercises with semi-occluded vocal tract exercises

Alondra Castillo [1], Paulina Zavala [2], Alba Testart[2]

[1] Universidad de Valparaíso [2] Universidad de Playa Ancha

[2] alondra.castillo@uv.cl

Introduction

Respiratory support required balance between respiratory and laryngeal muscles (2). The appropriate use of these muscles will prevent, during the emission, the presence of stress and/or laryngeal constriction (1), it is proven that the high subglottic pressure affects the glottic closure showing breath in hyperfunctional voices (2). In the clinical vocal therapy we observed alteration of acoustic parameters, fatigue, discomfort and modification of the vocal quality after exercises of respiratory support in singers. The repertoire of vocal techniques includes the *semi-occluded vocal tract exercises* (SOVT), that involve a series of postures with extended and/or occluded vocal tract, producing a change in the vibration pattern of the vocal folds, increasing vocal tract inertance and reducing the phonotrauma (3). We have considered to describe the mismatches muscle in glottic or supraglottic structure of a singer who practices Non Voiced Routine Exercises (NVRE), Voiced Routine Exercise (VRE) and VRE with SOVT, through nasolaringofibroscopy exploration and vocal acoustic analysis.

Methods

We refer the case of a 21 year-old singer, with 3 years of formal training, which demonstrates vocal and sensoperceptive changes undesired after respiratory training. In this investigation of mixed approach, descriptive/transversal and quasi-experimental design, acoustic parameters and glottic/supraglottic adjustments were evaluated pre and post implementation of two workouts of respiratory support exercises (Table 1), by AFA (PRAAT; Interface:Focusrite-PRO24; Mic:Audiotechnica-AT2020; PopFilter and nasolaringofibroscopy (Pentax10RP3). Then the evaluation procedure was replicated after Voiced Routine modified with SOVT. We analyzed and discussed the results and conclusions were presented.

NVRE		VRE		VRE with SOVT	
Exercise, Number of sets	Tipe	Exercise, Number of sets	Tipe	Exercise, Number of sets	Tipe
/s/ (x3)	Legato/Stacatto	/ai/ (x3)	Legato, musical intervals: 3rd, 5th &8th	B buzz (x3)	Legato, musical intervals: 3rd, 5th &8th
/f/ (x3)	Legato/Stacatto	/oi/ (x3)	Stacatto, musical intervals: 3rd, 5th &8th	B buzz (x3)	Stacatto, musical intervals: 3rd, 5th &8th
/sh/ (x3)	Legato/Stacatto			Knoll (x3)	Serie Vocal Function Exercises (VFE) (4)
/s/+/f/+/sh/ (x3)	Legato/Stacatto				

Results

The results evidence that Non Voiced Routine Exercises (NVRE) worsens some acoustic parameters, it does not generate qualitative improvements in acoustic analysis and maintains muscle tension pattern (PTM) type III. Voiced Routine Exercise (VRE) generates minimal positive changes in acoustic parameters, being the fundamental frequency the only significantly increased, it does not generate improvements in acoustic analysis and qualitative analysis and the nasolaringofibroscopy shows: PTM I, II and III in various phonatory tasks, main contact in the middle third, bowing, hypertonic vocal onset and it does not modificate the laryngeal height. The VRE with SOVT improves all quantitative and qualitative acoustic parameters, eliminates PTM I, II and III, reduces bowing triangular, softens the vocal onset and down the larynx on neck.

Table 2: Acoustic qualitative and quantitative data.

| | Análisis Cuantitativo | | | | | | Análisis Cualitativo | | |
	F0 (Hz)	Intensity (dB)	Jitter RAP (%)	Local Shimmer (dB)	NHR	HNR (dB)	Onset	Contour Hz	Contour dB
Initial Evaluation	232 Hz	76,87 dB	0,141%	0,237	0,015	23,075 dB	Isotonic ++	Stable +	Unstable +
Post NVRE	240,3 Hz	78,85 dB	0,193%	0,260 dB	0,011	21,050 dB	Isotonic ++	Stable ++	Unstable ++
Post VRE	267,4 Hz	81,01 dB	0,173 %	0,304 dB	0,014	21,71 dB	Isotonic ++	Stable ++	Unstable ++
Post VRE with SOVT	279 Hz	84,9 dB	0,11 %	0,1 dB	0,004	28 dB	Isotonic +++	Stable +++	Stable +++

Table 3: Nasolaringofibroscopy results.

	PTM I	PTM II	PTM III	Glotal Contact	Glotal Onset	Larynx height
Initial Evaluation	Present	Trend	Trend	Bowing Triangular +++	Hard	----
Post NVRE	-----	-----	Present	-----	-----	Alta
Post VRE	Present	Present	Present	Exclusive contact in the middle third	Hard	Alta
Post VRE with SOVT	Absent	Absent	Absent	Bowing Triangular +++	Iso	Baja

Conclusions

The voice and the vocal system of the singer was according the normal parameters, however some behaviors and symptoms founded were associated with vocal hyperfunction. The NVRE and VRE only improved acoustic parameters Hz and dB, but muscle mismatches indicates laryngeal hyperfunction. Moreover the VRE with SOVT evidence improving in all parameters acoustics, facilitated a harmonious muscular behavior, decreased stress impact and promoted vocal economy.

Finally, the evidence taken from this case opens up lines of research to strengthen the work between singing teachers and vocal therapists, allowing design training programs with SOVT to improve perceptual, acoustic and physiological vocal parameters.

Reference

(1) Tulon, C. (2005). Cantar y hablar. Ed. Paidotribo, Barcelona.
(2) Farias, P. (2012). La Disfonía Ocupacional. Ed. Akadia, Buenos Aires.
(3) Guzmán, M. (2013). Terapia con tracto vocal semi-ocluído. Un estudio de Caso. Revista Chilena de Fonoaudiología. Volumen 11, pp 87-97. ISSN 0717-4659

Voice related quality of life and musculoskeletal pain after use of TENS associated with vocal exercises: a pilot study

Larissa Thais Donalonso Siqueira[1], Camila de Castro Corrêa[1], Pamela Moreira[1], Rinaldo Roberto de Jesus Guirro[1], Alcione Ghedini Brasolotto[1], Kelly Cristina Alves Silveiro[1]

[1] Faculdade de Odontologia de Bauru – Universidade de São Paulo

larisiqueira@usp.br

Introduction

The hyperfunctional muscular pattern associated with dysphonia can or can not be associated with benign lesions in the larynx, as vocal nodules, more common in women. In these cases, the vocal therapy is based on the rebalancing of the extrinsic muscles of the larynx, besides vocal techniques to improve phonation. The use of Transcutaneous Electrical Nerve Stimulation (TENS) has been used to treat these dysphonia, showing positive results[1,2]. However, there are no studies that verified the effectiveness of this feature when combined with vocal exercises. Thus, the aim of this study was to evaluate the therapeutic effects of TENS associated with vocal exercises in musculoskeletal pain, vocal self-perception and voice related quality of life in women with vocal nodules.

Methods

Design: clinical trial, controlled, randomized and blind. After approval by the Ethics Committee (556.273), eight women with vocal nodules participated, 18-45 years of age (mean 33 years), divided, randomly, into two groups: experimental group (EG) - four women who received 12 sessions of the TENS application (pulse: 200 microseconds, frequency: 10 Hz, in motor threshold, with electrodes placed on the trapezius muscle – upper fibers - and submandibular aerea, bilaterally, for 20 minutes followed by 30 minutes of the orientation about vocal health and vocal exercises (nasal sounds, vibrant, fricative and sounded blowing exercise); and Control Group (CG) - four women who received 12 sessions of the application placebo TENS (same conditions EG, including positioning of the electrodes, but without receiving the stimulus in the form of electrical current) for 20 minutes followed by 30 minutes of vocal therapy traditional. All volunteers were assessed about frequency and intensity of the musculoeskeletal pain and quality of life in voice throught Voice-related quality of life protocol (VQROL)[3]. The evaluations were performed before, immediately after treatment, after one and three months. Data from all times of assessments, both groups, were compared by using the paired t-test (p≤0.05).

Results

There were no changes in the frequency and intensity of musculoskeletal pain after treatment in both groups, but it was found that only volunteers from EG reported improvement of vocal self-perception immediately after treatment (p=0.014), considering their voices as "good"

and this result remained after three months of treatment. About VRQoL domains, it was observed that when comparing the "immediate post" with "after three months," there was worsening of the total domain for EG, while for the CG, it was found highest scores for the physical and total domains after three months of treatment (p=0.024 and p=0.035).

Conclusions

Until the present moment, the voice related quality of life of the group was stimulated with TENS worsened after three months, which did not happen with the women in the group who performed only the vocal exercises. But, in general, it is observed that both treatments appear to produce the same therapeutic effect. Just emphasizes that the use of low frequency TENS associated with vocal exercises has been shown to improve and maintain the voice quality, according to the self-perception of women with vocal nodules.

References

1. Guirro RRJ, Bigaton DR, Silverio KCA, Berni KCS, Distéfano G, Santos FL, et al. Transcutaneous electrical nerve stimulation in dysphonic women. Pró-Fono R Atual Cient. 2008;20(3):189-194.

2. Silverio KCA, Brasolotto AG, Siqueira LTD, Carneiro CG, Fukushiro AP, Guirro RRJ. Effect of application of Transcutaneous electrical nerve stimulation and laryngeal manual therapy in dysphonic women: clinical trial. J Voice. 2015;29:200-08.

3. Gasparini G, Behlau M. Quality of life: validation of the brazilian version of the Voice-Related Quality of Life (V-RQOL) measure. J Voice. 2009;23(1):76-81.

Effects of the Transcutaneous Electrical Nervous Stimulation associated with vocal therapy in the vocal and laryngeal symptoms in dysphonic women

Kelly Cristina Alves Silverio[1], Larissa Thaís Donalonso Siqueira[1], Pamela Moreira[1], Rinaldo Roberto de Jesus Guirro[1], Alcione Ghedini Brasolotto[1]

[1] Faculdade de Odontologia de Bauru – Universidade de São Paulo

kellysilverio@usp.br

Introduction

Hyperfunctional voice disorders associated with prolonged, strong contraction of the larynx muscles are commonly associated with high laryngeal position in voices that have a strong component of muscle tension1. In these cases some laryngeal or vocal symptoms are presents because of the muscle tension and vocal misuse. A literature review indicates treatment options for hyperfunctional voice disorders such as indirect and direct therapy; medical treatment; and surgery for secondary organic lesions. In recommending laryngeal relaxation in individuals with hyperfunctional dysphonia the application of transcutaneous electrical nerve stimulation (TENS) may be an effective therapeutic tool1,2. Thus, the objective of this study was to evaluate the effects of TENS followed by vocal exercises in the vocal and laryngeal symptoms in women with vocal nodules.

Methods

Participated of this pilot study: eight women with vocal nodules, 18-45 years of age (mean 33 years), divided, randomly, into two groups: experimental group (EG) and control group (CG). The EG received 12 sessions of TENS for 20 minutes (pulse: 200 microseconds, frequency: 10 Hz, the motor threshold). The electrodes were placed on the trapezius muscle – upper fibers and submandibular area bilaterally. After, the EG group received 30 minutes of vocal exercises (nasal sounds, vibrant, fricative and semioccluded vocal tract exercises). The CG received 12 sessions of the TENS placebo for 20 minutes (same conditions of the EG, but without receiving the stimulus from electrical current) followed by 30 minutes of vocal exercises (same GE). All volunteers were evaluated about the frequency of vocal and laryngeal symptoms using a protocol described by Guirardi et al3. The frequency of symptoms was classified in "never", "sometimes", "almost always" and "always" by the individuals. A likert scale with four points was applied on the responses to do statistical analyses: "never" = 0; "sometimes" = 1; "almost always" = 2; "always" = 3. The symptoms were evaluated before, immediately after treatment, after one and three months of the treatment. The data were compared by using the paired t test (p < 0.05).

Results

The results are shown in Tables 1 and 2, considering the time of the assessments before the treatment, immediately after, after one month and after three months of the treatment. There was observed so little change, considering the treatment vocal applied. Maybe the individuals

had some difficulties to perceive the vocal and laryngeal symptoms. This pilot study contains few subjects per group and we can observe that the standard deviation is high considering the size of the sample. Others analyses are necessary to understand the results.

Conclusion

TENS followed by vocal exercises didn't reduce the frequency of laryngeal and vocal symptoms in women with vocal nodules. By the way, the vocal exercises provided improved the frequency of laryngeal symptom "pain when speak".

Table 1. Frequency of the vocal and laryngeal symptoms of the experimental group and different follow-up phases: before treatment, immediately after treatment, after one and three months after the treatment (N=4)

Symptoms	Before Mean (SD)	After Mean (SD)	p	After Mean (SD)	After 1 Mean (SD)	p	After Mean (SD)	After 3 Mean (SD)	P
Hoarsen	2,50 (1,83)	1,83 (0,98)	0,141	1,83 (0,98)	1,67 (1,03)	0,718	1,83 (0,98)	1,25	0,718
Voice loss	1,83 (0,75)	1,17 (0,75)	0,252	1,17 (0,75)	0,83 (0,75)	0,718	1,17 (0,75)	1,00	0,718
Vocal	2,50 (0,58)	1,67 (0,52)	0,092	1,67 (0,52)	1,50 (0,84)	0,718	1,67 (0,52)	1,33	0,495
Low-	2,00 (0,89)	1,67 (0,82)	1,000	1,67 (0,82)	1,50 (0,84)	0,638	1,67 (0,82)	0,75	0,001
Tightness	1,67 (1,03)	1,67 (1,21)	0,638	1,67 (1,21)	2,17 (1,17)	0,215	1,67 (1,21)	1,00	1,000
Dry cough	2,00 (0,89)	0,83 (0,75)	0,080	0,83 (0,75)	0,83 (0,98)	1,000	0,83 (0,75)	0,00	0,182
Cough	1,00 (0,63)	0,83 (1,17)	0,319	0,83 (1,17)	1,00 (1,26)	0,391	0,83 (1,17)	0,25	1,000
Pain when	2,00 (0,63)	0,67 (0,82)	0,058	0,67 (0,82)	0,83 (0,75)	0,638	0,67 (0,82)	0,75	0,638
Pain when	1,67 (1,03)	0,83 (0,98)	0,194	0,83 (0,98)	1,17 (0,98)	0,215	0,83 (0,98)	0,50	0,638
Secretio	1,83 (1,33)	1,67 (1,51)	0,824	1,67 (1,51)	1,83 (1,47)	0,718	1,67 (1,51)	1,50	0,391
Dry	1,67 (1,03)	0,83 (0,98)	0,182	0,83 (0,98)	1,33 (1,03)	0,215	0,83 (0,98)	1,25	0,092
Effort to	2,33 (0,82)	0,67 (0,82)	0,252	0,67 (0,82)	0,60 (0,55)	0,391	0,67 (0,82)	1,25	1,000

Table 2. Frequency of the vocal and laryngeal symptoms of the control group and different follow-up phases: pre-treatment, immediately after treatment, after one and three months (N=4)

Symptoms	Before Mean (SD)	After Mean (SD)	P	After Mean (SD)	After 1 Mean (SD)	p	After 1 Mean (SD)	After 3 Mean (SD)	p
Hoarseness	2,6	2,00 (0,71)	0,058	2,00 (0,71)	2,2 (0,84)	0,391	2,00	1,75 (1,26)	1,00
Voice loss	1,8 (1,1)	1,2 (0,45)	0,638	1,2 (0,45)	1,8 (0,84)	0,391	1,2 (0,45)	1,25 (0,96)	1,00
Vocal failure	2,00	1,6 (1,34)	0,058	1,6 (1,34)	1,8 (0,84)	0,182	1,6 (1,34)	1,6 (1,14)	1,00
Low-pitched voice	2,00	1,2 (1,3)	0,182	1,2 (1,3)	1,2 (1,30)	1,000	1,2 (1,3)	0,5 (1,00)	0,71
Tightness	2,6	1,8 (1,3)	0,182	1,8 (1,3)	1,8 (1,3)	1,000	1,8 (1,3)	1 (1,41)	0,39
Dry cough	2 (1,00)	1,6 (1,14)	0,182	1,6 (1,14)	1,4 (0,55)	1,000	1,6 (1,14)	1,5 (1,00)	0,39
Cough with	1,8	1 (0,71)	1,000	1 (0,71)	1 (0,71)	0,391	1 (0,71)	1,5 (0,58)	1,00
Pain when speak	1,2	1 (0,71)	0,001*	1 (0,71)	1 (0,71)	1,000	1 (0,71)	1 (0,82)	0,39
Pain when	2,2 (1,1)	1,4 (1,34)	0,391	1,4 (1,34)	1 (1,00)	1,000	1,4 (1,34)	0,5 (1,00)	0,71
Secretion in the	1,4	0,6 (0,89)	0,638	0,6 (0,89)	0,8 (0,84)	0,391	0,6 (0,89)	0,5 (0,58)	0,71
Dry throat	1,8	1,6 (0,55)	0,058	1,6 (0,55)	1,6 (0,89)	0,391	1,6 (0,55)	1,25 (0,96)	0,71
Effort to speak	2,4	1,2 (1,3)	0,391	1,2 (1,3)	1,4 (1,52)	0,391	1,2 (1,3)	0,5 (0,58)	0,25

Adipose Tissue as a Lateral Vocal Fold Boundary Condition

Hiba Saidi[1], Cameron D. Jones [1], Byron D. Erath[1],

[1] Clarkson University

Introduction

Computational and physical models of the vocal folds are used extensively in scientific investigations of both normal and pathological voiced phonation. The most widely implemented boundary condition involves applying a fixed constraint on the anterior, posterior, and lateral surfaces, which is justified by the attachment of the tissue at these points to the surrounding cartilages. However, physiologically, the lateral face of the thyroarytenoid muscle is bounded by the paraglottic space, an interstitial region between the vocal folds and the thyroarytenoid cartilage that is comprised of soft, fatty, adipose tissue (AT). Experimental measures of porcine AT stiffness identify a viscoelastic behavior with stiffness, based on strain rates found during normal phonation, of $\sim 1-8$ kPa. This work employs numerical and experimental models of the vocal folds to determine how altering the lateral boundary condition to include AT influences the vocal fold dynamics.

Methods

Two modeling modalities are implemented. The first utilizes synthetic, self-oscillating silicone models vocal fold models. An additional substrate layer of adipose tissue with stiffness ranging from $1-8$ kPa is added to the lateral face. The dynamics of the model are quantified by acquiring the mean flow rate, fundamental frequency, subglottal pressure, and glottal area.

The silicone synthetic models are isotropic, whereas the actual vocal folds are transversally isotropic. To extend the experimental investigation to provide a more physiologically-representative tissue representation, a numerical finite element model of the vocal folds utilizing a transversally-isotropic assumption was also investigated for the same fixed and AT boundary conditions as the experimental configuration.

Results and Conclusions

Numerical results are shown in Figure 1 with the first 6 modes presented for the case of a model with no AT layer (Figure 1a) and with a layer of AT with a stiffness of 8 kPa (Figure 1b). The first two modes are the same, with the first mode representing a z-10 mode. However, the modes quickly become more complicated due to the three-dimensional nature of the structure.

Figure 1: The first 6 modes of the numerical vocal fold model. (a) M5 geometry with fixed lateral boundary condtion. (b) M5 geometry with a substrate layer of adipose tissue.
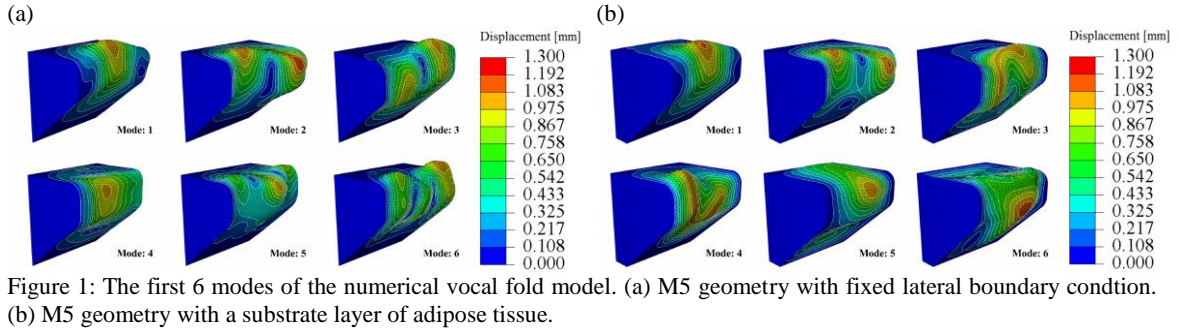
Table 1 presents the fundamental frequency for the first 6 modes as a function of AT stiffness. The fundamental frequency is only minimally influenced by stiffness values greater than 100 kPa. While minor variations are observed when it decreases to 8 kPa, significant differences are observed over the range of 1 − 8 kPa, the range of values estimated to be encountered in the human vocal folds. This is largely because at this value, the AT stiffness is on order with the stiffness of the vocal fold tissue.

Table 1: Fundamental frequency as a function of mode shape and adipose tissue stiffness.

Young's Modulus for AT [kPa]	Mode 1 [Hz]	Mode 2 [Hz]	Mode 3 [Hz]	Mode 4 [Hz]	Mode 5 [Hz]	Mode 6 [Hz]
No AT/ Fixed Case	111.9	116.7	134.2	140.7	161.9	181.7
1000	112.8	120.3	136.3	143.9	164.8	186.5
100	108.0	120.0	135.9	142.8	164.4	185.3
10	93.7	117.6	133.0	135.8	147.2	158.9
8	92.0	117.0	131.9	135.1	140.1	154.2
1	73.5	91.0	93.5	103.4	119.0	123.8
0.1	41.6	43.0	44.1	46.3	46.6	47.9

Results from the experimental investigations (not shown here) , which also include fluid loading, exhibit the same tendency in the fundamental frequency of oscillation while also revealing that the onset pressure decreases in tandem with the fundamental frequency.

The results provide insight into how vocal fold dynamics are dependent upon the prescription of boundary conditions. The addition of a layer of AT addresses common complaints about synthetic, silicone, self-oscillating models; namely, the onset pressures are elevated above what are considered "normal" values, and the fundamental frequency is also elevated. It is emphasized that the current work models the viscoelastic AT as an isotropic material (as necessitated by limitations in the material selection in the physical model), and the effective stiffness is chosen based on approximations of the strain rate. More physiologically-realistic specifications of the AT properties will likely introduce additional complexities in the model behavior.

Laryngoscopic and Spectral Analysis of Laryngeal and Pharyngeal Configuration in Non-Classical Singing Styles

Marco Guzmán[1], Andrés Lanas[3], Christian Olavarria[2], Maria Josefina Azocar[1], Daniel Muñoz[2], Sofía Madrid[1], Sebastián Monsalve[1], Francisca Martínez[1], Sindy Vargas[1], Pedro Cortéz[2], Ross M. Mayerhoff[4]

[1] University of Chile, Department of communication Sciences and Disorders. [2] Department of Otolaryngology, Voice Center, University of Chile Hospital [3] Department of Otolaryngology, Voice Center, Las Condes Clinic [4] Department of Otolaryngology- Head and Neck Surgery, Wayne State University

smadrid @ug.uchile.cl

Introduction

Earlier studies have suggested that supraglottic activity may not necessarily be a sign of vocal hyperfunction or harmful behavior to vocal folds, but rather a normal and even desirable muscle activity[.1-2]. Titze[3] states that the source-filter interaction and the vocal tract inertance may be increased by narrowing the epilarynx tube in an anterior-posterior (A-P) direction.

There is evidence showing supraglottic activity may not be pathologic during classical and nonclassical singing, although they have not compared different singing styles produced by the same subjects. In addition, most of the studies have only evaluated laryngeal compression, not other features such as VLP or pharyngeal compression.

Present study aimed to assess three different singing styles (pop, rock, and jazz) with laryngoscopic, acoustic, and perceptual analysis in healthy singers at different loudness levels. Special emphasis was given to the degree of anterior-posterior (A-P) laryngeal compression, medial laryngeal compression, vertical laryngeal position (VLP), and pharyngeal compression.

Methods

Twelve female trained singers conservatories with at least 5 years of voice training and absence of any voice pathology were included. Flexible and rigid laryngeal endoscopic examinations were performed. During the transnasal endoscopic examination, each participant was instructed to sing the song ''Happy Birthday'' in three different styles (pop, rock, and jazz) and three different loudness levels (medium, high, and low). Voice recording performing the same singing phonatory tasks was also carried out. Four blinded judges were asked to assess laryngoscopic and auditory perceptual variables using a visual analog scale.

Results

All laryngoscopic parameters showed significant differences for all singing styles. Rock showed the greatest degree for all of them. Overall A-P laryngeal compression scores demonstrated significantly higher values than overall medial compression and VLP. High loudness level produced the highest degree of A-P compression, medial compression, pharyngeal compression, and the lowest VLP for all singing styles. Additionally, rock demonstrated the highest values for alpha ratio (less steep spectral slope), L1-L0 ratio (more glottal adduction), and Leq (more vocal intensity). Statistically significant differences between the three loudness levels were also found for these acoustic parameters.

Conclusions

Rock singing seems to be the style with the highest degree of both laryngeal and pharyngeal activity in healthy singers. Although, supraglottic activity during singing could be labeled as hyperfunctional vocal behavior, it may not necessarily be harmful, but a strategy to avoid vocal fold damage

References

1. Lawrence, V. (1979). Laryngological observations on belting. J Res Singing, 2, 26-28.

2. Stager, S., Bielamowicz, S., Regnell, J., Gupta, A. & Brakmeier, J. (2000). Supraglottic activity: evidence of hyperfunction or laryngeal articulation?. J Speech Hear Res.,43, 229-238.

3. Titze, I & Story, B(1997). Acoustic interactions of the voice source with the lower vocal tract. J Acoust Soc Am.,101, 2234-2243

Air pressure and contact quotient measures during different semi-occluded postures in subjects with different voice conditions

Marco Guzman[1], Christian Castro[2], Sofia Madrid[1], Christian Olavarria, Levia[1,3,4]
Daniel Muñoz[1], Elisabeth Jaramillo[2], Anne-Maria Laukkanen[5]

[1]University of Chile, Department of communication Sciences and Disorders, [2] University of Valparaiso. Department of communication Sciences, [3]University of Chile Hospital, Santiago, [4] Barros Luco-Trudeau Hospital, Department of Network Management. Santiago, [5]University of Tampere, School of Education, Speech and Voice Research Laboratory, Tampere

guzmann.marcoa@gmail.com

Introduction

Semi-occluded vocal tract exercises are widely used in voice therapy and training. It is well known that longer and narrower tubes offer more resistance to the airflow due to frictional losses. Moreover, when a tube is submerged in water, an extra resistance is added due to hydrostatic pressure, which is dependent mainly on the depth of immersion. The purpose of this study was to investigate the effect of different artificial lengthening of the vocal tract (with the free end in water and in air) on air pressure variables and vocal fold adduction in subjects with different voice conditions.

Methods

Forty-five participants were assessed. They were divided in four groups: normal voice without voice training (n=12), normal with voice training (n=9), functional dysphonia (n=14), and subjects with unilateral vocal fold paralysis (n=10). Participants were asked to produce a series of six semi-occluded vocal tract postures: base line, drinking straw, stirring straw, silicon tube with the free end in air (10 mm of inner diameter and 55 cm in length) silicon tube with the free end submerged 3 cm below the water surface, and silicon tube with 10 cm below the water surface.

For each task was obtain the mean glottal contact quotient (CQ), mean fundamental frequency, mean subglottic pressure (Psub), mean oral pressure (Poral), and mean transglottal pressure. Maximum and minimum Poral were also measured during both silicon tube submerged 3 and 10 cm below the water surface.

Numerical variables were described by median and interquartile range (IQR), and compared by phonatory task and vocal status separately using Kruskal-Wallis test. A generalized multivariable linear model to observe the joint influence of phonatory task and vocal status in vocal parameters was fitted. Separate subgroup analysis for minimal and maximal oral

pressure was also performed. Finally, linear correlation analysis using Pearson coefficient for overall correlation was used.

Results

All semi-occluded postures produced an increase in Psub and Poral compared to the baseline condition (vowel phonation). Phonation with tube into the water (10 and 3 cm below the surface) and phonation into a stirring straw produced the three highest values for Poral. Interestingly, the same three phonatory tasks caused the highest values for Psub. The correlation analysis demonstrated a high correlation between these two dependent variables. No significant differences were observed between voice conditions.

Multivariate linear regression model showed that all voice condition behaved similarly regarding air pressure variables. Tube 10 cm, tube 3 cm, and stirring straw presented the highest values in both, Psub and Psupra for all voice conditions. It seems that these variables are in general more dependent on the degree of airflow resistance than vocal status of participants.

Results from the multivariate linear regression model including F0, CQ, Psub, Poral, and Ptrans as outcomes as well as phonatory task as predictive variable are shown in Tables 1.

	F0 (Hz)		CQ (%)	Psub (cm H^{2}O)	Poral (cm H^{2}O)	Ptrans (cm H^{2}O)
Baseline	Men	Women	54.14 (10.20)	7.28 (2.53)	0.22 (0.14)	6.90 (2.42)
	128.27 (17.50)	215.13 (32.45)				
Drinking Straw	130.37 (9.14)	208.38 (23.46)	56.43 (9.66)	10.62 (4.73)	2.84 (2.54)	7.20 (3.11)
Stirring Straw	127.63 (7.94)	160.55 (90.49)	57.50 (8.56)	14.47 (5.57)	6.75 (4.44)	6.58 (2.93)
Silicon Tube in Air	129.45 (18.20)	215.97 (50.60)	54.61 (13.26)	10.32 (4.33)	0.55 (0.82)	9.38 (3.85)
Silicon Tube in Water (3 cm)	134.91 (19.12)	222.04 (62.15)	54.99 (13.40)	12.05 (3.60)	32.94 (7.67)	8.31 (3.56)
Silicon Tube in Water (10 cm)	143.91 (42.08)	207.64 (50.03)	59.44 (11.09)	17.82 (3.42)	38.03 (7.01)	9.07 (3.49)
P value	0.27622	0.30380	0.01519	0.0001	0.0001	0.0001

Conclusion

During semi-occlusion exercises most variables behaved in a way regardless of the vocal status of the participants. This means that if different outcome is expected in different patients. Avoidance of stirring straw and tube 10 cm in water may be safest for patients with hyperfunctional voice disorder, while they may be a good option for subjects with hypofunctional dysphonia.

Acoustic effects of implementing a Minimum Vocal Warming Plan (PCVM) in Radio Newscasters

Mª Celina Malebran B.[1], Patricio Saldívar O.[2], Paulina Zavala[2]

.[1] Universidad Santo Tomás, [2] Universidad de las Américas

celinamalebranbmello@gmail.com

<u>Introduction</u>

Certain work-related activities require a clear, projected, and persuasive voice. People who work with the voice can be classified into two groups: Artistic Voice Professionals (AVP) and Non-Artistic Voice Professionals (NAVP). Radio newscasters and commentators are in the AVP group according to Vilkman's definition (2000); these professionals must preserve the integrity of their phonatory apparatus and thus take care of their voice. However, they lack specific training in the correct use of the voice and the knowledge of voice techniques exercises and vocal warm-ups (Malebran & Saldívar, 2010).
This assertion agrees with what was detected by Cielo & Bazo (2008) who state that this group of professionals do not know the techiques to improve vocal resistance and optimize phonation.
AVP is a group who requires the handling of several vocal resources that give the speaker the power to express him/herself; among the requirements the application of warm-up vocal exercises are fundamental to the care of their vocal health.

Therefore, the concern arises to verify and characterize the voice of broadcasters. The study proposed to apply to this group of VPA an Plan of Minimum Vocal Warming-up (PCVM) developed by Behlau (2005). The plan consisted of ten sessions with PCVM, accompanied by vocal hygiene guidelines.

<u>Methods</u>

The subjects of this study were 2 women and 6 men, broadcasters, with diverse requirements in terms of vocal quality and use, without prior vocal training. Before applying the PCVM, a nasofibroscopy was performed to discard vocal pathology.

The study is quantitative, descriptive, quasi-experimental and transversal, with a sample who was non-probabilistic and selected by convenience. To ensure environment without noise or interference, the sound pressure level was evaluated using a sound level meter IEC 8810 CE-651 type II. The recordings were made using model AKG® condenser microphone attached to Perception 120 audio interface M- Audio® v.1.03.

PRAAT® software was used to capture and voice recording, with a sampling frequency of 44.1 kHz and 16 bit al a recording studio. For this study, it was analyzed the vowel /a/, which is characterized by the absence of audible noise in the open air passage, with its clear glottal

pulse and contains all the physiological information system (Russo & Behlau, 1993).

Results

Speakers showed a modification in a majority of the vocal parameters analyzed.
The results before and after application of the PCM are shown in the following table.

SUJETOS	f0 (Hz)		INTENSIDAD (dB)		JITTER RAP (%)		SHIMMER LOCAL (dB)		NHR		HNR (dB)		INICIO VOCAL	
PCVM	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
LOCUTOR 1	213	205	63.27	68.09	0.198	0.424	0.226	0.437	0.012	0.039	19.74	15.05	hipertónico	isotónico
LOCUTOR 2	97	90	75.52	61.77	0.25	0.538	0.356	0.375	0.945	0.08	0.877	13.770	isotónico	isotónico
LOCUTOR 3	93	90	62.9	62.02	0.57	0.218	0.263	0.247	0.086	0.028	12.61	17.478	hipertónico	isotónico
LOCUTOR 4	176	152	72.37	70.6	0.08	0.068	0.09	0.136	0.003	0.006	28.56	29.627	hipertónico	isotónico
LOCUTOR 5	229	245	69.35	76.25	0.425	0.348	0.305	0.333	0.216	0.012	17.9	20.65	isotónico	isotónico
LOCUTOR 6	93	108	72.83	74.08	0.142	0.126	0.256	0.181	0.017	0.01	19.37	23.992	isotónico	isotónico
LOCUTOR 7	119	126	68.2	57.18	0.223	0.111	0.234	0.158	0.014	0.015	22.83	22.123	hipertónico	hipertónico
LOCUTOR 8	96	142	67.87	68.61	0.295	0.114	0.289	0.154	0.022	0.003	18.77	26.902	hipertónico	isotónico
NORMA (Farias, 2012; Broersma & Weenink, 2001)	Fem: 205 Masc:127 (Aronson y Cols. 2001)		>60 dB y <70 dB para voz conversacional		normal 0.35 límite 068		normal 0.17 límite 0,35		normal 0.11 límite 0,19		normal 16.5 límite 12		isotónico	

Tabla 1: parámetros acústicos /a/ sostenida y tono habitual, pre y post implementación del PCVM

Conclusions

The application of PCVM has shown a positive difference in the voices, regulating the intensity of the signal and thus the vocal projection, in addition to a tendency to increase the F0.
The warm-up vocal exercises as those of the PCVM should be incorporated as a daily practice by speakers since they help bring about maximum performance with minimum use, guaranteeing protection of vocal health.

Greater investigation needs to be done by this group of voice professionals regarding vocal warm-ups in order to find ways of preventing future pathologies and to optimize vocal resources.

References

Behlau, M., Pontes, P. (2005) A voz que ensina. Rio de Janeiro: Revinter. 68p.

Cielo, C. A. e Bazo, G. (2008) Pefil vocal de locutores de rádio. *Salusvita*. Bauru, v. 27, n. 2, p. 167-181.

Farghaly, S.F., Furquim, C.R. (2008) *Voice training program for radio presenters.* Rev Soc Bras Fonoaudiol. 13(4):316-24

Guzmán, M., Malebrán, MªC., Zavala, P., Saldívar, P., Muñoz, D. (2013) Cambios acústicos de la voz como signos de fatiga vocal en locutores de radio: resultados preliminares. *Acta Otorrinolaringológica Española.* Vol 64 Nº3

Hitomi, K, Zambuze, L., Bittante, I. (2008) *25 years of professional voice care: analyzing the actions* Rev CEFAC, São Paulo, v.10, n.4, 557-565

Malebran, Mª C.; Saldívar, P. (2010) "Elaboración del perfil vocal de un grupo de locutores radiales chilenos"; Iº Congreso Iberoamericano de voz cantada y hablada, Memorias. CD Room. ISBN 978-987-26209-0-5.Buenos Aires, Argentina

Russo, I., Behlau, M. (1993) Percepciones del habla: análisis acústico del Portugués Brasileño. Sao Paulo: Lovise

Effects of resonance tube phonation in water in the asymmetric oscillatory entrainment of unilateral vocal fold paralysis

Christian Castro[1,2], Alba Testart[1], Gabriel Galindo[2], Matías Zañartu[2]

[1] Universidad de Valparaíso, Department of Speech, Language, and Hearing Sciences.. [2] Universidad Técnica Federico Santa María. Department of Electronic Engineering.

christian.castro@uv.cl

Introduction

The semi-occluded vocal tract exercises are widely used in voice therapy and vocal training. Within this category, the resonance tube phonation in water (RTPW) uses a silicon tube with its free end immersed in varying depths of water. Previous studies have reported that the water immersion increases oral pressure (Poral), subglottal pressure (Psub), and open quotient (OQ), also, the bubbling created with exercise, creating a beneficial massaging effect that improves voice quality that can endure after the exercise[1]. Therefore, RTPW exercises may be useful for pathologies such as unilateral vocal fold paralysis. However, the biomechanical and kinematic effects that this technique has over vibration of the vocal folds are still largely unknown. The purpose of this study is to assess the kinematic effects immediately after RTPW exercises, by means of laryngeal high-speed videoendoscopy in subjects with unilateral vocal fold paralysis.

Methods

Rigid HSV was recorded at 8000 fps for two female patients with unilateral vocal fold paralysis uttering a sustained vowel /e/ at confortable loudness. Synchronous recordings of microphone, electroglottography, and neck surface acceleration were also obtained. Two different recording sessions were performed for each patient, i.e., pre and post RTPW exercises. A 15 cm silicone tube with 1cm of diameter was immersed into 10 cm of water. The duration of each RTPW exercise was 1 minute with breathing pauses.

Digital videokymography (DVK) analysis was performed for every recording, and glottal area, vocal fold displacements and signal spectra were also obtained.

Results.

Both patients exhibited an increase in the mean displacement and maximum displacement of the vibratory amplitude of the non-paralyzed vocal fold. In the analysis of glottal area and DVK, changes in the entrainments pre and post exercised were observed. For subject 1, pre-RTPW entrainment was 2:1 and post-RTPW entrainment became 1:1. For subject 2, pre-RTPW entrainment was 3:1 and post-RTPW entrainment became 3:2, as shown in Fig 1. Both changes in entrainment are considered beneficial, as they significantly enhanced the energy of the fundamental frequency in the resulting spectra.
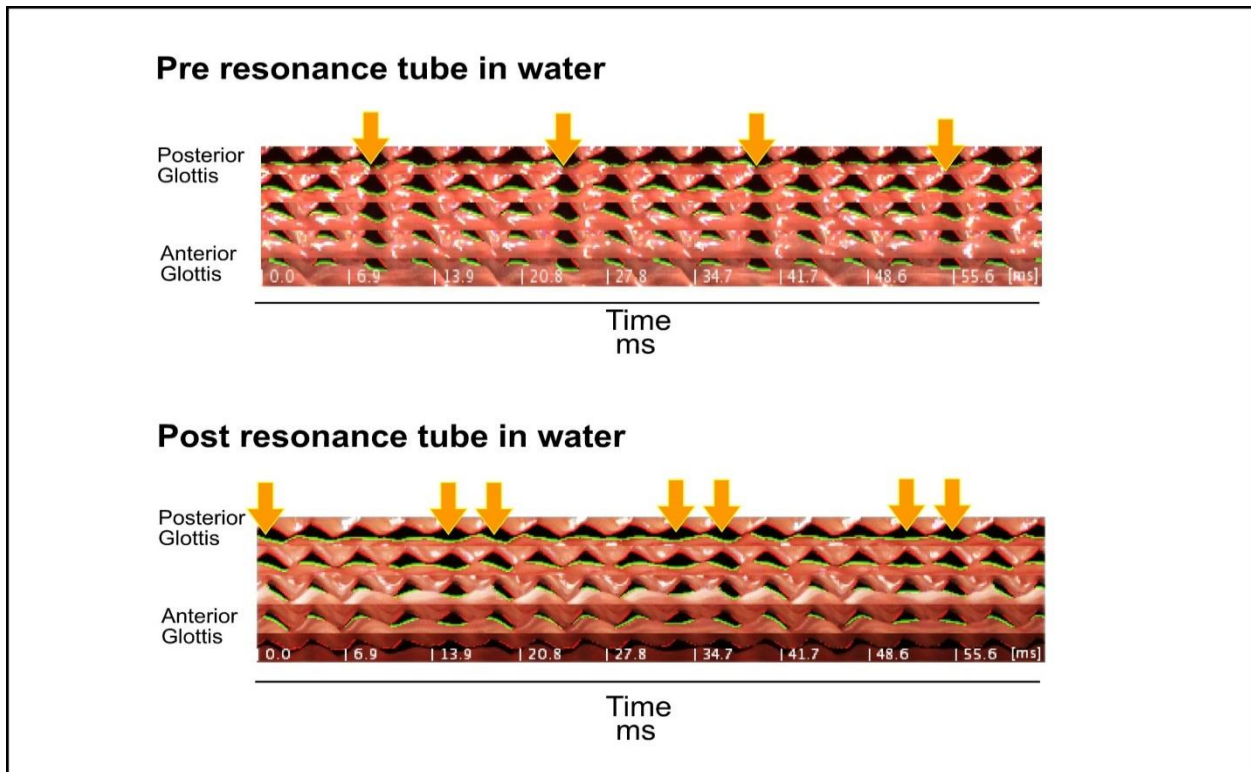
Figure 1: Videokymography of subject N2. Pre and Post resonance tube phonation in water at 10 cm of immersion. The arrows indicate the oscillatory entrainment for each case

Conclusion

Changes in asymmetric oscillatory entrainment have been noted previous studies with asymmetric numerical models[2,3], which have been shown to be dependent upon degree of asymmetry and subglottal pressure. Following this idea, we hypothesize that the observed changes in entrainment in this study are related to changes in Psub and Poral product of the semi-occluded vocal tract exercises.

The beneficial modification of oscillatory entrainment vocal fold oscillation and the increase in the amplitude of vibration of the non-paralytic fold suggest that the tube resonance phonation in water may be useful tool for therapy in patients with unilateral vocal fold paralysis.

References

1. Granqvist S, Simmbeg Hetegard S, Holmwvist S, Larson H, Lindestad P, Sodersten M, Hammaberg B, Resonance tube phonation in water: High-speed imaging, electroglottographic and oral pressure observations of vocal fold vibrations - a pilot study
2. Steinecke, I. and Herzel, H. (1995). "Bifurcations in an asymmetric vocal-fold model". J. Acoust. Soc. Am., 97(3):1874–84.
3. Sommer DE, Erath BD, Zañartu M, and Peterson SD (2013), "The Impact of glottal area discontinuities on block-type vocal fold models with asymmetric tissue properties", J. Acoust. Soc. Am., 133(3):EL214-EL22

Differentiating between females with vocal hyperfunction and matched-controls using inverse filtered aerodynamic measures

Víctor M. Espinoza[1,2], Matías Zañartu[1],

Jarrad H. Van Stan[3,4], Daryush D. Mehta[3,4,5], Robert E. Hillman[3,4,5]

[1] Department of Electronic Engineering, Universidad Técnica Federico Santa María,[2] Department of Music and Sonology, Faculty of Arts, Universidad de Chile, [3] Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital,[4] Department of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, [5] Department of Surgery, Harvard Medical Schoo

vespinoza@u.uchile.cl

Introduction

Many of the most common voice disorders (e.g., nodules, polyps, muscle tension dysphonia) are chronic or recurring conditions that are believed to result from repeated detrimental patterns of vocal behavior, referred to as vocal hyperfunction (VH). Subjects with these disorders often attempt to compensate for their vocal deficits, which has been shown to produce increased aerodynamic energy transfer and high collision forces. Normal and Regressed z-scores were used as statistical tools to discriminate VH subjects against a normative set of vocally healthy controls (Hillman et al., 1989). However, this framework considered primarily male subjects, low number of patients, reduced statistical significance, and no group analysis. With this in mind, the goal of the study is to extend early approaches for the objective detection of VH using inverse filtered aerodynamic measures and robust statistical methods in a database of female voices.

Methods

Indirect measurements of vocal function (i.e., peak to peak glottal airflow, maximum flow declination rate, among others) were obtained through noninvasive inverse filtered approximations to glottal airflow (Perkell et al., 1991). Multiple /pae/ gestures were recorded at three levels of vocal intensity in sound-treated room, for 16 pairs of patients diagnosed with nodules (non-singers associated with phonotraumatic VH) and muscle tension dysphonia (non-phonotraumatic VH) and corresponding matched controls (age and occupation). Robust multiple regression (Maronna et al., 2006) is executed using each parameter covariate with SPL and fundamental frequency (f0) to estimate a model for the matched-control group. If an acceptable model fit is achieved (Hastie et al., 2009),

independent-samples t-test for means are performed between normal and pathological groups, in both regressed and non-regressed versions (without covariate SPL and f0).

Results

Using the regressed data, the preliminary results show statistically significant differences ($p<0.05$) between the normative and phonotraumatic group of female voices in 5 of 8 parameters. However, only 1 of 8 parameters was significantly different between the normative and non-phonotraumatic group. All differences represented moderate and large effect sizes using Cohen's approach. The non-regressed data yielded minimal findings when comparing the normative group to the phonotraumatic (1 of 8 were significantly different) and non-phonotraumatic (0 of 8 were significantly different) groups.

Conclusions

We have presented a simple approach to apply a t-test to detect VH in female voices. Using regression for correcting/adjusting for SPL and f0 increase the sensitivity of the inverse filtered measures in the independent-samples t-test. The regressed approach tends to reduce the variance and separate mean values, instead of the non-regressed version. Further research will be conducted using multivariate models, discriminant analysis and machine learning techniques with the aim to improve discrimination.

References

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective Assessment of Vocal Hyperfunction: An Experimental Framework and Initial Results. J Speech Hear. Res., 32, 373–392.

Maronna, R., Martin, R., & Yohai, V. (2006). Robust Statistics: Theory and Methods. John Wiley & Sons, Ltda.

Perkell, J. S., Holmberg, E. B., & Hillman, R. E. (1991). A system for signal processing and data extraction from aerodynamic, acoustic, and electroglottographic signals in the study of voice production. The Journal of the Acoustical Society of America, 89(4), 1777–1781. http://doi.org/http://dx.doi.org/10.1121/1.401011

Hastie, T. and Tibshirani, R. and Friedman, J. (2009), The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer.

Mehta DD, Van Stan JH, Zañartu M, Ghassemi M, Guttag JV, Espinoza VM, Cortés JP, Cheyne HA and Hillman RE. (2015).Using ambulatory voice monitoring to investigate common voice disorders: Research update. Front. Bioeng. Biotechnol. 3:155.doi:10.3389/fbioe.2015.00155

Hard glottal attack from a physiological and behavioural review

Felipe Cerda Sandoval [1]

[1] Asociación chilena de Seguridad

Lanfal@achs.cl Felipecerda@euvoz.cl

Introduction

The hard Glottal attack is one of the most investigated phonotraumatic behaviours in the phonoaudiologic evaluation process of a dysphonic patient. In this context, corrective measures must take an ethiologic-physiologic approach into consideration rather than a facilitator sound which is not coincident with what was previously mentioned

Therefore, the objective of this investigation is to identify the physiological cause of the hard Glottal attack, thus, show the variations of phonatory learning in the clinical consultation of the dysphonic patient to tackle this behaviour from a propioceptive learning model.

Methods

Using a monograph design with qualitative data, 4 phonoaudiology students – participants were recruited since they are able to produce and understand the difference between the hard and normal glottal attack. They were given an acoustic model of a vowel /a/ with or without a vocal cord hit to make sure the sound is produced appropriately. Two patterns were used: isolated and non isolated hit. Then, they were given sequences of /a/ which were produced with and without a hit to describe and observe the behaviours of the hard glottal attack using videostroboscopy and spectrographic measurements objectively. The materials used for the measurement were: nasofibroscopy (Atmos), narrowband spectogram and oscillogram (praat) a channel interface and microphone.

Results

We found a modification of the hyperfunctional configuration of glotic as much as supraglotic patterns before producing a hard glottal attack. This is considered apnea. As the procedure was carried out, the participants develop apnea prior to the sound they produced when they were asked to make a hit of vocal cords. The participants closed the larynx developing apnea prior to the sound they made. The existance of this hyperfunctional pattern produces an increase of the resistance before phonation which is supported by a cordal hit. All of this is correlated with the oscillogram variation which is observed in an increase of the peak at the start phonatory as well as "nasofibroscopicos" findings. When there's no existence of vocal cords hit, there's neither alteration in the glotic and superglotic pattern in the image diagnosis nor oscillogram variations.

Conclusions

The modification of the structures when producing a glotic and supraglotic hyperfunction before phonation produces a change upon glotic, supraglotic and transglottic pressures. The relation is proportional in terms of subglotic pressure and the hard glottal attack or adduction speed at the beginning of phonation. That is to say, if the vocal folds resistance increases, there will be an increase of the subglotic pressure. Therefore, if resistance is overcome, glotis will be abducted, the transglottic flow will increase. If the flow is higher, the negative intercordal pressure will increase generating a quick approximation of vocal folds to a middle position. This is considered as impact stress or phonotrauma.

Finally, in terms of projections for this investigation it would be accurate to measure the specified behaviour from other visions or using objective complementary tools to support what has been presented.

Referencias

Cobeta, I. Núñez, F. Fernández, S. (2013) Patología de la voz. (1° ed). España. Editorial Marge medica books.

Farías, P. (2007).Ejercicios para restaurar la función vocal, observaciones clínicas. (1° ed). Argentina. Editorial Akadia.

Prater, R. Swift, R. (1986) Terapéutica de la voz. (2° ed). España. Editorial Salvat.

Case presentation: Riedel's Thyroiditis and Cordal Paralysis

Gonzalo Inostroza Moreno[1], Macarena Manríquez Muñoz[1]

[1] Universidad Mayor

gonzalo.inostroza.moreno@gmail.com

Introduction

Nowadays, cases of thyroid gland pathologies, such as Riedel's thyroiditis - which is chronic thyroid gland inflammation where the fibrous tissue replaces the glandular tissue and extends to adjacent structures, are published less frequently. Unfortunately, these patients usually present voice alterations due to the proximity of its structures to the larynx. This case is presented with the aim of showing the development of the vocal intervention on a patient who has undergone thyroid surgery. The patient has been diagnosed with aphonia.

Methods

The process starts with vocal rehabilitation, but ends with no significant improvement. After the first six sessions, bilateral cordal paralysis in abductor position is diagnosed, but presents no response to treatment. Cordal contact is persued, with the objective of improving the patient's life quality in the family and social sphere.

Results

After six months of treatment - this is required by the Chilean health norms - there was no significant improvement. Therefore, the patient got a disability retirement pension with the aim of improving his life quality. Concerning the patient's voice, there were minimal changes, mainly in laryngeal manipulation. However, the vocal folds passed over the medium line.

Conclusion

It is of great importance to develop these topics, especially in our region, since many patients suffer from Riedel's thyroiditis. Therefore, it is important to know how to proceed and what the best treatment to follow is, in order to improve the patients' quality of life, and to be aware, from a physiological and biomechanical point of view, of the best and most suitable procedures for the patients.

References

Díaz Pérez, J. A., García Vera, J. A., & Herrera, A. (2008). Riedel´s thyroiditis. REV ESP PATOL, Vol 4º (nº4), 297 - 302.

González López, S., Suárez Núñez, E. (2005). Tiroiditis de riedel. Presentación de un caso. Acta Médica Grupo Ángeles, Vol 3, (nº3), 99 – 101.

Morpho-functional, perceptual and acoustic parameters of the voice of people with functional dysphonia

Alejandro Rodríguez Campo[1,] María Esperanza Sastoque H[1] ,Leonardo Fierro Pérez[1] ,Claudia Gómez Perdomo[1]

Universidad del Valle. Colombia.
alejandro.rodca@correounivalle.edu.co

Introduction

The voice is the product of the coordinated function of several body systems, mainly phonatory, resonatory and respiratory systems. Moreover, it is a multidimensional phenomenon that is part of a socio-communicative context. In order to assess the voice a battery of tests that complement each other is required. The morpho-physiology of the larynx refers to an integrated laryngeal structure and function study. Perceptual evaluation investigates the degree of involvement vowel discrimination by ear. Acoustic evaluation identifies physical parameters of the voice which are hardly perceptible in some cases. When functional alterations occur in the voice, some changes at the level of the laryngeal structure may occur requiring speech therapy. In accordance to the previous, this study aimed to characterize the morpho-functional, perceptual and acoustic parameters of the voice in a group of people with functional dysphonia in Santiago de Cali.

Methods

This study of quantitative, observational and descriptive type, was conducted in the framework of the research project entitled: "Effect of voice therapy on the morpho-functional, perceptual and acoustic parameters of the voice in functional dysphonia in Santiago de Cali," supported by the Ethics Committee of the Faculty of Health from Universidad del Valle, Colombia. The study was carried out with 10 people diagnosed with functional dysphonia, aged between 20 and 38 years old. Morpho-functional evaluation was made through fibronasolaryngoscopy, conducted by an experienced otolaryngologist in larynx. The state of the laryngeal mucosa, the glottal closure and the vestibular vocal folds behavior were observed during the phonation. Perceptual evaluation was made by using the scale GRBAS with a phonetic balanced reading by which only the overall degree of dysphonia, roughness and vocal strain were analyzed. And acoustic evaluation was made through Anagraf software from the analysis of the fundamental frequency (Fo), Energy, speech disturbance measures (Jitter and Shimmer) and the integrated index of disturbance voice. Both perceptual and acoustic evaluation were conducted by an experienced speech therapist in voice.

Both otolaryngologist and speech therapist conducted tests of concordance, each one of them with two experts in their same area, through the Intraclass Correlation Coefficient (ICC) producing results of a high level of concordance.

Results

In the morpho-physiology was found: edema interarytenoid (40%), adduction of vestibular folds (10%) and altered glottal closure (10%). In perceptual evaluation was found: overall degree of mild (80 %) and moderate (20 %) dysphonia; a slight (80%) and moderate (20 %) roughness, and a slight (70%) and moderate (30%) tension. The acoustic variables found were: energy altered (60 %), Shimmer altered (70%), Jitter altered (10%), and an integrated index of voice disturbance by finding voice risk (20 %), no voice risk (70%) and altered voice (10 %). The fundamental frequency (Fo) was not found altered.

Conclusions

The morphofunctional state of the larynx is not only affected when there is an organic dysphonia, but also when an inadequate vocal technique which comes to limiting the use of the voice in activities of daily life. A battery of tests is necessary to properly plan the voice therapy.

References

1. Carding, P., Carlson, E., & Epstein, R. (2001). Re: Evaluation of voice quality. *International Journal of Language & Communication Disorders*, *36*, 127-134.

2. Jiang, J. J., Shah, A. G., Hess, M. M., Verdolini, K., Banzali, F. M., Jr., & Hanson, D. G. (2001). Vocal fold impact stress analysis. J Voice, 15(1), 4-14.

3. Khoddami, S. M., Nakhostin Ansari, N., Izadi, F., & Talebian Moghadam, S. (2013). The Assessment Methods of Laryngeal Muscle Activity in Muscle Tension Dysphonia: A Review. *The Scientific World Journal*, *2013*, 507397. http://doi.org/10.1155/2013/507397

4. Rodríguez-Parra, M. J., Adrián, J. A., y Casado, J. C. (2009). Voice therapy used to test a basic protocol for multidimensional assessment of dysphonia. *Journal of Voice, 23*(3), 304-318.

5. Speyer, R., Wieneke, G. H., & Dejonckere, P. H. (2004). Documentation of progress in voice therapy: perceptual, acoustic, and laryngostroboscopic findings pretherapy and posttherapy. *J Voice, 18*(3), 325-340.

Vocal profile using a phonetic acoustic analysis voice of patients with Parkinson disease by stages of evolution.

Mauricio Alfaro[1], Hernan Pérez [2]

[1] University of Bio – Bio, [2]University of Concepción

malfaro@ubiobio.cl

Introduction.

Parkinson disease (PD) is a neurodegenerative disorder, mainly characterized by the degeneration of the nigrostriatal system, specifically for the presence of the Lewy Body in the functional neurons' cytoplasm. Even though literature describes the vocal symptomatology in patients with PD it is important to considerate that the vocal characteristics according to PD's evolution stages have not been described in Chile. The aim of this study is to characterize the voice by an acoustic-phonetic analysis of PD's patients according to stages of evolution by Hoehn y Yahr (H&Y) scale.

Methods.

A non-experimental design of descriptive and cross sectional type was accomplished, in 18 patients, diagnosed with PD, who receive the assistance from Rehabilitation Community Center and "Dr. Federico Puga Bone" Family Health Center in the districts of Chillán and Chillán Viejo, respectively. To prevent the participants to performed out from the norm, and considering that the Multi-Dimentional Voice Program (MDVP) Software is not standardized for the population's target parameters, a calibration of acoustics parameters in senior adults without PD was performed, considering they belong to an age segment of the population with different voice characteristics to the average adult. In order to make this adjustment more reliable, 87 workers of University of Bio-Bio were considered. From this probabilistic sample with 18 people was collected. These people were between 50 and 89 years, male and female, without PD or neurologic complications Patients with PD were assessed according UKPDSBB (UK Parkinson Disease Society Brain Bank) diagnosis criteria and Barthel index to determine the stages of evolution as stated by H&Y scale. A voice evaluation was applied to both groups. A phonation of vowel /a/ was requested for 8 seconds using an audio recording equipment Tascam DR-40, after this, the results were analyzed using the MDVP. Data was analyzed with SPSS (v15). Interval variables were described using the arithmetic mean and standard deviation and qualitative variables with observed frequencies and percentages. For the bivariate analysis were used tests to compare means for independent

samples of Anova, and for two groups with t-Student. The Pearson r coefficient was applied in order to correlate the variables of interest. We considered as significant 5%.

Results.

Intrinsic and extrinsic larynx muscles' movement range is reduced, this is associated to the rigidity, which is the main characteristic of PD's patients; therefore, the movement control of thyroarytenoid and cricothyroid, and the elevate and lower the laryngeal muscles are also modified. Accordingly there will not be control of the subglottic pressure. Resulting in patients with difficulty in phonation's control. For this reason, Stability parameters of frequency (RAP: relative Average Perturbation) and stability parameters of amplitude (Shimmer), parameter of noise (NHR: noise to harmonic ratio) and turbulence in signal (VTI: voice turbulence index), and tremor (FTRI: frequency tremor intensity index; ATRI: amplitude tremor intensity index), were evaluated. Results reveal, only RAP and ATRI measures reach a linear correlation between the disease's development stage and the measure from the parameter (p=0.005, p=0.006 respectively), whereas FTRI measures, even though show alterations, they are not correlated according to the corresponding PD stage.

Conclusions.

Only RAP, ATRI and FTRI, are attributable characteristic to Parkinson Diseases, where RAP and ATRI are progressive alterations. Voice characteristics in people with PD can be useful in assessment and the establishment of PD progression indicators, and allows the future possibilities of creating a Vocal Intervention Program in order to satisfy the specific needs of each stage of the disease.

References

ARONSON A. (2009). Clinical voice disorders. New York: Thieme Inc.

FURTADO L., CÔRTES A., COSTA F., DA CONCEICAO C., BARRETO I (2012). "Idiopathic Parkinson's disease: vocal and quality of life analysis" en Arq Neuropsiquiatr. 79(9), pp.: 674-679.

HOLMES R., OATES J., PHYLAND D., HUGHES A. (2000). "Voice characteristics in the progression of Parkinson's disease" en Journal Language Communication Disorders, 35, pp.: 18-31.

SKODDA S. (2011). "Aspects of speech rate and regularity in Parkinson's disease" en Journal of the Neurological Sciences, 310, pp.: 231–236.

XUE S., DELIYSKI D. (2001). "Effects of aging on selected acoustic voice parameter: preliminary normative data and educational implications" en Educational Gerontology. Universidad de Ohio. 27(2), pp.: 159-168.

Effect of the technique of semi-occluded vocal tract in the acoustic parameters and phonatory ability in the students of choral group.

Mauricio Alfaro [1], Karina Arriagada [1], Alejandra Godoy [1], Patricia San Martin [1], Elinita Stuardo [1], Francisca Carrasco [1] , Angela Contreras[1]

[1] University of Bio – Bio,

Introduction.

Singing voice requires learning a vocal technique in order to allow the proper development of vocal skills for the healthy singing, which will prevent voice disorders. The vocal technique includes, the vocal warming, which prepares the muscles for the future exercise of greater performance. In the same way, semi-occluded vocal tract (SOVT) technique has for aim a phonation with minimum effort and maximum efficiency. Both techniques are useful in preventing vocal alterations. The aim of this study was to evaluate the effect of the SOVT technique in the acoustic parameters and phonatory ability of choral group using students from Purisima Conception school, Chillán.

Method.

This study presents a design of randomized clinical trials (parallel design), whose sample was composed of 52 students aged 14-18 years from choral group of the Purisima Concepción school, which were randomly assigned to two test groups: intervention group and an active control group. Sample was submitted to an assessment of the acoustic parameters. A phonation of vowel /a/ was requested for 8 seconds using an audio recording equipment Tascam DR-40. Jitter, Shimer and NHR, were evaluated by applying the PRAAT software, therefore were also evaluated the phonatory parameters like maximum phonation time and quotient s/z, which pretend relate the pulmonary and laryngeal function. Subsequently those students meeting eligibility criteria are divided into two groups: The first is a control group which consisted by twenty-seven students who continued with the technique of vocal warming and the second group is the intervened group which consisted by twenty-five students which was subjected an intervention with SOVT technique. The intervention was performed twice a day for 6 weeks. Data was analyzed with SPSS (v15). Interval variables were described using the arithmetic mean and standard deviation, using the Kolmogorov-Smirnov test.

It was analyzed with t-student test for related samples in order to compare the means between the two groups at the beginning and end of the study. We considered as significant for this study P value $>0,05$.

Results.

Stability parameters of frequency and amplitude (Jitter, Shimmer), noise, s/z ratio and Maximum Phonation Time (/s/) were evaluated. Results showed a significant improve in Jitter (p=0,048), Shimmer (p<0,001), NHR (p<0,001) and s/z ratio (p=0,050) to intervention group. On the other hand, the control group showed a significant improve in NHR(p=0,002) and Maximum Breath Time (p= 0,012).

Conclusions.

The continuous training using the semi-occluded vocal tract technique improved acoustic parameters and phonatory ability of choral group students. This could be explained for the greatest source-filter interaction, because the exercises increase the intraoral pressure with torax's expansion and decrease of vocal fold contact force. This study allows the future possibilities of creating a Vocal Prevention Program, using semi-occluded vocal tract in order to satisfy the specific needs of choral group.

References

Gish, A., et al. (2012). Vocal Warm-Up Practices and Perceptions in Vocalists: A Pilot Survey.Journal of Voice, vol. 26, n°

Guzmán M. et al. (2012). Efecto terapéutico de los ejercicios con tracto vocal semi-ocluido en pacientes con disfonía músculo tensional tipo I. Revista de Logopedia, Foniatría y Audiología 32, pp.139 -146.

Guzmán, M. et al. (2013).Effect on long-term average spectrum of pop singers' vocal warm-up with vocal function exercises. International Journal of Speech-Language Pathology, 2013; 15(2): 127–135.

Nix J, Simpson C. (2008). Semi-occluded vocal tract postures and their applications in the singing voice studio. J Singing, pp. 339-340

Titze, I. (2006). Voice training and therapy with a semi-occluded vocal tract: rational and scientific underpinnings. J. Speech Lang Hear Res, vol. 49, n° 2: 448-59.