

Meeting Report

The 13th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research (June 2–4, 2019, Montreal, Quebec, Canada)

Luc Mongeau 

Department of Mechanical Engineering, McGill University, Montreal, QC H3A 0G4, Canada;
luc.mongeau@mcgill.ca

Received: 28 May 2019; Accepted: 29 May 2019; Published: 30 June 2019



Abstract: The 13th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research (AQL 2019) will be held in Montreal, Canada, 3–4 June 2019. Pre-conference workshops will be held on 2 June 2019. The conference and workshops provide a unique opportunity for partnership and collaboration in the advancement of quantitative methods for the measurement and modelling of voice and speech. The AQL accomplishes this mandate by facilitating an interprofessional scientific conference and training intended for an international community of otolaryngologists, speech–language pathologists and voice scientists. With a continued drive toward advancements in translational and clinical voice science, the AQL has rapidly expanded over the past 20 years, from a forum of 15 European member laboratories to a globally recognized symposium, connecting over 100 delegates from across the world.

Contents

1	Pre-Conference	4
1.1	Hybrid Aeroacoustic Approach for the Efficient Numerical Simulation of Human Phonation	4
1.2	simVoice—Numerical Computation of the Human Voice Source	6
1.3	Aeroacoustic and Vibroacoustic Mechanisms during Phonation	7
1.4	A Machine-Learning Based Reduced-Order Modeling of Glottal Flow	8
1.5	Updated Rules for Constructing a Triangular Body-Cover Model of the Vocal Folds from Intrinsic Laryngeal Muscle Activation	9
1.6	Synthetic Vocal Fold Model Closed Quotient Optimization	10
1.7	Contact Pressure and Length as a Function of Posterior Glottal Area: Synthetic Vocal Fold Investigations	12
2	Session 1	13
2.1	Vocal-Fold 3D Micro-Architecture and Micro-Mechanics: A Multimodal Imaging Study	13
2.2	Influence of Recording Perspective in Laryngoscopy on Perceived Asymmetry	14
2.3	Extracting Reduced-Order Model Parameters from High-Speed Video of Silicone Vocal Folds Using a Gradient-Based Approach	16
2.4	Segmenter’s Influence on Objective Glottal Area Waveform Measures from High-Speed Laryngoscopy	17
2.5	Vocal Fold Collision Pressure Amplitude and Timing in an Excised Hemilarynx Setup with Dual High-Speed Videoendoscopy	18
2.6	Recent Advancements in Acoustic Analysis for Assessing Laryngeal Function	20

2.7	Optimization of Relative Fundamental Frequency Estimation Algorithms: Accounting for Sample Characteristics and Fundamental Frequency Estimation Method	21
2.8	Acoustic Phonatory Tremor Index: Objective Quantification of Perceived Vocal Tremor Severity	22
2.9	Accelerometer-Based Prediction of Subglottal Pressure in Healthy Speakers Producing Non-Modal Phonation	23
2.10	Classification of Vocal Gestures Extracted from Quasi-Daily Sentences to Detect Vocal Fatigue	25
2.11	Uncertainty of Ambulatory Airflow Estimates and Its Effect on the Classification of Phonotraumatic Vocal Hyperfunction	26
2.12	How Is Vocal Loudness Affected by Spectral Slope	27
3	Poster Session 1	28
3.1	Riedel’s Thyroiditis Cordal Paralysis: A Single Case Study	28
3.2	Influence of Voice Focus Adjustments on Oral-Nasal Balance in Speech and Singing	29
3.3	Immunological Profiling of Vocal Fold Hydrogel Scaffolds	30
3.4	Chemical Receptors of the Larynx: A Comparison of Human and Mouse	32
3.5	An Investigation of Vocal Fatigue Using a Dose-Based Vocal Loading Task	33
3.6	Passive Vowel Devoicing in Osaka Japanese: Case Study Using Electromyography (Emg) and Photoglottography (PGG)	34
3.7	High-Resolution CFD Simulation of Flow in Glottis Using Les	35
3.8	Quantification of the Degree of Vocal Fatigue in Teachers by Means of an Interface That Characterizes Voice Signals	37
3.9	Clinical Practicability of a Newly Developed Real-Time Digital Kymographic System	38
3.10	Functional Changes of Submandibular Gland by Steatosis-Induced Ferroptosis in Ovariectomized Rats	39
3.11	Extracellular Matrix Turnover in Human Larynx	40
3.12	Tissue Hysteresis and Relaxation, Phonation Onset, and Phonation Offset in The Context of the Surface Wave Model	41
3.13	3D Printed Scaffold Design for Vocal Fold Tissue Engineering Application	42
3.14	A Preliminary Study on Pharyngoesophageal Segment Vibration in Tracheoesophageal Speech by Means of a Collapsible Channel Model	43
3.15	Application of Two Different Modalities for the Vibratory Characteristics in Vocal Fold Vibration of Vocal Cord Paralysis before and after Injection Laryngoplasty-Laryngeal Videostroboscopy and Two Dimensional Scanning Videokymography	44
3.16	BiOChemical Alterations in Vocal Fold Tissue in the Production of Decellularized Extracellular Matrix Hydrogels	45
4	Session 2	47
4.1	Vocal Fold Visco-Hyperelastic Properties: Characterization and Multiscale Modeling upon Finite Strains	47
4.2	Investigation of Constrains on Vocal fold Viscoelastic Properties Using an Inverse Mapping Approach	48
4.3	Vocal Fold Contact Pressure in a Three-Dimensional Body-Cover Phonation Model	49
4.4	Numerical Study of the Influence of Vascular Morphology on the Evolution of Vortical Flow Structures through the Blood-Feeding Arteries of the Human Vocal Folds: Application to Drug Delivery for Laryngeal Cancer	50
4.5	Development of a High-Fidelity Voice Simulator—From Muscle Contraction to Running Speech	51
4.6	SpEAR: A Speech Database for the Advancement of Intra-Aural Wearable Technology	52
4.7	High Performance Simulation and Visualization of 3D Vocal Fold Agent-Based Model	53

5	Poster Session 2	54
5.1	Development, Validation and Analysis of Numerical Larynx Models with Regard to Computational Costs	54
5.2	Agent-Based Model of Hyaluronic Acid-Gelatin Scaffold for Vocal Fold Tissue Engineering	56
5.3	Usefulness of Cepstral Peak Prominence (CPP) in Post-Thyroidectomy Dysphonia Evaluation	57
5.4	Decoding Phonation with Artificial Intelligence (D ϵ P AI): Proof of Concept	58
5.5	Glottal Area Waveform Modeling Based Voice Quality Typing	60
5.6	Automated Quantification of Inflection Events in the Electroglottographic Signal	61
5.7	Characteristics of the Pharyngoesophageal Segment: Literature Review	62
5.8	Designing Audible Sound Spots Using Metamaterial Based Phased Array	63
5.9	Characterizing Injury Recovery in Rabbit Vocal Folds with Multimodal Imaging	64
6	Session 3	66
6.1	The Causes and Laryngeal Electromyography Characteristics of Unilateral Vocal Fold Paralysis	66
6.2	Arytenoid Adduction and Type 1 Thyroplasty for Unilateral Vocal Fold Paralysis: Measurements from Six Excised Canine Larynges	67
6.3	Increased Calcium Channel in the Lamina Propria of Aging Rat	68
6.4	Localization of the Tight Junction Proteins Claudin Family in the Laryngeal Glands: A Rat Study	69
6.5	Macrophages in the Vocal Fold	70
6.6	Vocal Fold-Mimetic Environment for the Modulation of Stem Cell Functions	71
6.7	Bioprinting Highly Porous Chitosan-Based Scaffolds With Tunable Stiffness and Viscoelasticity for Vocal Fold Repair	73
6.8	The Effects of Laryngeal Massage and Nebulized Saline on High-Voice Users	74
6.9	Investigating the Pathobiology of Vocal Fold Dehydration and Rehydration	75
6.10	Increased Laryngeal Mucosal Cellular Proliferation in Mice Exposed Short-Term to Cigarette Smoke	76
6.11	Effects of Voice Changes under Testosterone Therapy on Listener Perception of Gender: A Transgender Case Study	78
7	Poster Session 3	79
7.1	The Use of Nasalance for Voice Stabilisation during the Tenors' Passaggio	79
7.2	Numerical Analysis of the Airflow Downstream from a Tracheoesophageal Voice Prosthesis	80
7.3	Beneficial Effects of Choral Singing on Speech and Voice in Normal Aging	81
7.4	Esophageal Wall Compliance and Its Influence on the Driving Pressures of Tracheoesophageal Speech	82
7.5	On the Role of Simultaneous Observations for a Bayesian Estimation of Subglottal Pressure and Laryngeal Muscle Activation	84
7.6	Comparing Accelerometer and Oral Airflow Based Aerodynamic Measures in Patients with Vocal Hyperfunction	85
7.7	Development of a Vocal Warm Up Protocol for Vocal Fatigue Prevention	86
7.8	Evaluation of Anti-Fibrotic Activity of Wound Healing Macrophages in a 3D In Vitro Model for Vocal Fold Scar Treatment	87
7.9	Characterizing Injury Recovery in Rabbit Vocal Folds with Multimodal Imaging	89
7.10	Stress Relaxation in Carbon Nanotube Composite Hydrogels for Vocal Fold Tissue Regeneration	90

7.11	Three-Dimensional Vocal Fold Deformation under Simulated Lateral Cricothyroid Muscle Activation in an Excised Human Larynx	91
7.12	High throughput Drug & Kinase Inhibitor Screening for Idiopathic Subglottic Stenosis	92
7.13	Clinical and Surgical Implications of Intraoperative Optical Coherence Tomography Imaging for Benign Pediatric Vocal Fold Lesions	93
8	Session 4	94
8.1	The Relationship between Speech Rate, Voice Quality and Listeners' Purchase Intentions	94
8.2	Predicting Emphatic Speech: Classification of Non-Literal Utterances	95
8.3	Cortical Mechanisms Controlling the Speech Production during Lombard Effect: An EEG Study	96
8.4	Auditory Acuity to Fundamental Frequency in Children with and without Vocal Fold Nodules	98
8.5	Phonation Type and Amplitude of Voice Source Fundamental	99
8.6	Comparison of Voice Onset Measures with Glottal Pulse Identification in Acoustic Signals: Preliminary Analyses	100
8.7	Differences in Ambulatory Vocal Behavior between Patients with Phonotraumatic Lesions and Matched Healthy Controls	101
8.8	Automatic Voice Signal Typing Using Classic and Nonlinear Dynamics Features	102
8.9	Vocal Tract Shape and Acoustic Adjustments of Children during Phonation into Narrow Flow-Resistant Tubes	103
9	Poster Session 4	105
9.1	Estimating Patient-Specific Contact Pressures Using a Finite Element Model	105
9.2	Methodological Barriers in Building an Audiovideo Database for Automatic Identification of Fatigue Levels through Speech and Facial Expressions in People with a Neurological Condition	106
9.3	Simultaneous Measurements of Glottal Velocities and Vocal Folds Geometry in a Canine Larynx Model	108
9.4	Application of a Promotion of Vocal Health Program (Virtual + Face to Face) for College Professors	109
9.5	Investigation of Vocal Folds Poroelastic Behaviour under Mechanical Loading in Different Bath Concentrations	110
9.6	In Vitro Analysis of Polymeric Microspheres Containing Human Vocal Fold Fibroblasts for Vocal Fold Lamina Propria Regeneration	111
9.7	Laser-Projection System and Method for 3D Calibrated Laryngeal Measurements Using Transnasal Flexible High-Speed Videoendoscopy	112

1. Pre-Conference

1.1. Hybrid Aeroacoustic Approach for the Efficient Numerical Simulation of Human Phonation

Stefan Schoder ¹, Sebastian Falk ², Michael Döllinger ² and Manfred Kaltenbacher ¹

¹ Institute of Mechanis and Mechatronics, TU Wien, Austria

² Division of Phoniatics and Pediatric Audiology, Friedrich-Alexander University Erlangen-Nürnberg, Germany

Keywords: hybrid aeroacoustics; numerical simulation; source term interpolation

Objectives

Our key objective is to develop an aeroacoustic computational model *simVoice* for clinical applicability. The incompressible fluid dynamics equations using a LES (Large Eddy Simulation) turbulence model will be based on prescribed vocal fold oscillations identified first from synthetic and then from in-vivo and ex-vivo high-speed imaging (see, e.g., [1]). In this way the fluid-solid interaction problem, whose accuracy critically depends on reliable geometrical and material parameters of all layers of the vocal folds, is circumvented. According to a perturbation ansatz, the acoustic model is based on the perturbed convective wave equation with the substantial derivative of the incompressible pressure as a source term [2]. Thereby, a main challenge is the interpolation of the aeroacoustic source from the flow to the acoustic grid.

Introduction

Voice research is carried out experimentally by applying physical (synthetic or ex-vivo/in-vivo animal and human larynges) and numerical models. Experimental investigations involve high personnel, material and hence financial costs [3]. Furthermore, the scientific outcome is limited to the parameters that are measured at few discrete positions in the larynx replica. In contrast, numerical models provide a much higher amount of results since parameters from any location in the model can be obtained and analyzed. Furthermore, since the advent of affordable high-performance computing, computer simulation methods based on modeling the fundamental physical phenomena by partial differential equations and solving them numerically have been steadily gaining importance and are therefore a highly promising alternative approaches [4].

Methods

We follow a hybrid aeroacoustic approach, which performs, on a first step, an incompressible flow simulation on a computational grid being capable to resolve all relevant turbulent scales. In a second step, we compute the acoustic source terms on the flow grid and perform a conservative interpolation to the acoustic grid, on which we solve the perturbed convective wave equation to obtain the acoustic field. Thereby, the conservative transformation of the acoustic sources from the flow grid to the acoustic grid is a key step to allow coarse acoustic grids without reducing accuracy. We use an advanced cut-volume-cell approach, which guarantees high accuracy for regions where the flow grid is finer and for regions where the flow grid is coarser as the acoustic grid. In our framework, the radial basis function method was incorporated to compute spatial derivatives of the flow data as necessary in the computation of the acoustic sources.

Results

The application of the cut-volume-cell interpolation method allows for an accurate transfer of the acoustic sources from the flow grid to the acoustic grid. The acoustic result changed only little, even for a reduction of elements in the source region of more than 10 times compared to the initial acoustic grid. Therefore, we obtain now a strong reduction in the elapsed CPU time for the computation of the acoustic field.

Acknowledgments: The authors acknowledge support from the German Research Foundation (DFG) under DO1247/10-1 no. 391215328 and the Austrian Research Council (FWF) under no. I 3702.

References

1. Schützenberger, A.; Kunduk, M.; Döllinger, M.; Alexiou, C.; Dubrovskiy, D.; Seger, A.; Semmler, M.; Bohr, C. Laryngeal high-speed videendoscopy: Sensitivity of objective parameters towards recording frame rate. *BioMed Res. Int.* **2016**, *2016*, 4575437.
2. Kaltenbacher, M.; Hüppe, A.; Reppenhagen, A.; Zenger, F.; Becker, S. Computational Aeroacoustics for Rotating Systems with Application to an Axial Fan. *AIAA J.* **2017**, doi:10.2514/1.J055931.

3. Döllinger, M.; Kobler, J.; Berry, D.A.; Mehta, D.D.; Luegmair, G.; Bohr, C. Experiments on Analysing Voice Production: Excised (Human, Animal) and In Vivo (Animal) Approaches. *Curr. Bioinform.* **2011**, *6*, 286–304.
4. Tian, F.-B.; Dai, H.; Luo, H.; Doyle, J.F.; Rousseau, B. Fluid-structure interaction involving large deformations: 3D simulations and applications to biological systems. *J. Comput. Phys.* **2014**, *258*, 451–469.

1.2. *simVoice*—Numerical Computation of the Human Voice Source

Sebastian Falk¹, Stefan Kniesburges¹, Hossein Sadeghi¹, Stefan Schoder², Manfred Kaltenbacher² and Michael Döllinger¹

¹ Div. of Phoniatrics & Pediatric Audiology, Dep. of Otorhinolaryngology, University Hospital Erlangen, Germany

² Institute of Mechanics and Mechatronics, TU Wien, Austria

Keywords: phonation; computer modelling; CFD

Introduction

The central objective of this project is the development of a three dimensional aero-acoustic numerical model (*simVoice*) for a prospective application in a clinical environment. The larynx model considers the fluid flow through the glottis, the vocal fold motions, and the resulting acoustic signal. Thereby, we solve the partial differential equations of the fluid flow by a Finite Volume (FV) and the acoustic field by a Finite Element (FE) method. The main advantage of this approach compared to experimental investigations (synthetic or ex-vivo/in-vivo animal and human larynges) is the high spatial and temporal access to the flow as well as acoustic quantities and the generated acoustic source terms.

Methods

The *simVoice*-model is a hybrid model. It consists of a fluid dynamic simulation model with an external driven vocal fold motion, based on the 3D FV method [1], and an aero-acoustic model, based on the 3D FE method. The numerical model of *simVoice* considers the vocal folds, the ventricular folds and various vocal tract geometries based on a synthetic vocal fold model [2]. The oscillation of the vocal folds, identified from in-vivo and ex-vivo high-speed imaging (see, e.g., [3]), is externally forced. The fluid dynamic simulation model uses the Large Eddy Simulation (LES) turbulence model to solve the incompressible fluid dynamic equations. *simVoice* is currently optimized concerning computing time and complexity considering the computational grid resolving all relevant turbulent scales. This optimization will achieve the prospective clinical application of the hybrid model *simVoice*.

Results

In a first step, we investigate two different glottis configurations with symmetric and asymmetric motions of the vocal folds. The symmetric vocal folds motion is associated with a healthy voice and the asymmetric one to a disordered voice [4]. Within the closed glottis configuration, the vocal folds are entirely interrupting the fluid flow while in the half-open glottis configuration there is permanent fluid flow through the glottis during vocal fold oscillation.

The different glottis geometries and vocal fold motions allow us to identify the fluid dynamical differences to obtain a better understanding of the disturbed and healthy vocalization. While the glottal jet is symmetrically expanding through the glottis and the ventricular folds within the symmetric vocal folds motion, the glottal jet is asymmetric and just hitting one ventricular fold during asymmetric dynamics. Compared to the half-open glottis configuration with an uninterrupted glottal jet, periodic

interruptions of the glottal jet and consequently periodic pressure fluctuations occur in the closed glottis configuration.

Conclusions

The innovative scientific aspects of *simVoice* are: (1) Analysis of the magnitude of dissolving the time-dependent turbulent fluid flows to obtain physical correct acoustic source terms; (2) To gain insight in the cause and effect of the vocal folds motion, fluid flow and acoustics; (3) An extensive study of the impact of various glottis geometries, vocal fold motions and vocal tracts on the acoustic signal.

Acknowledgments: The authors acknowledge support from the German Research Foundation (DFG) under DO 1247/10-1 (no. 391215328) and the Austrian Research Council (FWF) under no. I 3702.

References

1. Sadeghi, H.; Kniesburges, S.; Kaltenbacher, M.; Schützenberger, A.; Döllinger, M. Computational models of laryngeal aerodynamics: Potentials and numerical costs. *J. Voice* **2018**, doi:10.1016/j.jvoice.2018.01.001.
2. Kniesburges, S. Fluid-Structure-Acoustic Interaction During Phonation in a Synthetic Larynx Model; Shaker: Düren, Germany, 2014.
3. Schützenberger, A.; Kunduk, M.; Döllinger, M.; Alexiou, C.; Dubrovskiy, D.; Seger, A.; Semmler, M.; Bohr, C. Laryngeal high-speed videoendoscopy: Sensitivity of objective parameters towards recording frame rate. *BioMed Res. Int.* **2016**, *2016*, 4575437.
4. Inwald, E.; Döllinger, M.; Schuster, M.; Eysholdt, U.; Bohr, C. Multiparametric analysis of vocal fold vibrations in healthy and disordered voices in high-speed imaging. *J. Voice* **2011**, *25*, 576–590.

1.3. Aeroacoustic and Vibroacoustic Mechanisms during Phonation

A. Lodermeier^{1,2}, E. Bagheri¹, C. Näger¹, K. Nusser¹, S. Becker¹, M. Döllinger³ and S. Kniesburges³

¹ Institute of Process Machinery and Systems Engineering, FAU Erlangen-Nürnberg, Erlangen, Germany

² Erlangen Graduate School in Advanced Optical Technologies (SAOT), FAU Erlangen-Nürnberg, Erlangen, Germany

³ Division for Phoniatics and Pediatric Audiology, Department of Otorhinolaryngology, Head & Neck Surgery, FAU Erlangen-Nürnberg, Erlangen, Germany

Keywords: vibroacoustics; aeroacoustics; Hybrid Acoustic PIV; perturbed convective wave equation

Objectives

The mechanisms of basic sound generation for healthy and disordered voice production is of scientific and clinical interest. In the present study, we investigate sound generation in an experimental model using an approach that covers both the aeroacoustic (i.e., sound induced by the flow field) and the vibroacoustic (i.e., sound induced by the vocal fold surface vibration) sound generation.

Introduction

Howe and McGowan (2007) showed in an analytical model that sound is mainly generated by aeroacoustic sources immediately downstream of the glottis. In an experimental approach, Lodermeier et al. (2018) found that highly intense tonal pressure fluctuations are generated at the glottal exit, while broadband sound is evoked along the full length of the glottal jet. This approach was based on Lighthill's acoustic analogy, which is unable to show purely acoustic characteristics within the flow field, hence it is limited to far field observations. In the present study, we eliminate this drawback and analyze the sound field within the supraglottal region.

Methods

In our Hybrid Acoustic PIV (HAcouPIV) approach, we acquire the instantaneous flow field using time-resolved PIV. The measured flow field data is used to compute the aeroacoustic sources within the supraglottal channel. We apply a state of the art perturbation ansatz known as the perturbed convective wave equation (PCWE). By that, the sound pressure field is fully separated from the fluid-mechanical pressure field. Hence, a direct analysis of the sound sources and its radiation is possible.

Additionally, the vibrational behavior of the vocal fold surface is investigated by a scanning vibrometer.

Results

The simulated sound outside of the supraglottal channel shows good agreement with the experimental microphone measurements up to a frequency of 2 kHz. All harmonics in this range as well as the broadband slope of the spectrum are well reproduced by the PCWE approach. The sound field within the supraglottal channel shows that strong tonal components of the basic voice signal are at the glottal exit. In contrast to a Lighthill-based evaluation, additional tonal sources are situated along the supraglottal channel with similar intensity.

The structural motion of the vocal folds exhibits mostly harmonic content. High harmonic content was found near the medial surface of the vocal folds.

Conclusions

Our experimental method to determine sound generation was further improved by the PCWE approach, as the sound field within the supraglottal region is accessible. The acoustic source field may be evaluated with that separation approach and shows a spatially more extensive source region than claimed in previous studies.

Acknowledgments: We acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) within the framework of the German Excellence Initiative. Additionally, the work was supported by the Else Kröner-Fresenius Stiftung under Grant agreement no. 2016 A78.

References

1. Howe, M.; McGowan, R. Sound generated by aerodynamic sources near a deformable body, with application to voiced speech. *J. Fluid Mech.* **2007**, *592*, 367–392.
2. Lodermeier, A.; Tautz, M.; Becker, S.; Döllinger, M.; Birk, V.; Kniesburges, S. Aeroacoustic analysis of the human phonation process based on a hybrid acoustic PIV approach. *Exp. Fluids* **2018**, *59*, 13.

1.4. A Machine-Learning Based Reduced-Order Modeling of Glottal Flow

Yang Zhang, Xudong Zheng and Qian Xue

Department of Mechanical Engineering, University of Maine, Orono, ME, USA

Keywords: glottal flow model; machine learning; reduced-order modeling

Objectives/Introduction

With the recent rapid growth of computing power, many high-fidelity computational fluid dynamics (CFD) models of phonation have been developed and combined with patient-specific anatomical laryngeal geometries extracted from medical imaging, to provide more detailed information about the underlying biophysics of human phonation. However, the use of this type of model in both fundamental and clinical research is still very limited due to very high computational cost which only

allows explorations of limited behaviors. To this end, we aim to develop a machine-learning based reduced-order glottal flow model that can provide fast and accurate prediction of intraglottal flow pressures in all kinds of regular and irregular glottal shapes. The model is coupled with a continuum mechanics based vocal fold model to predict the three-dimensional dynamics of vocal fold vibrations in near real-time.

Methods

A series of three-dimensional glottal configurations are obtained through symmetric and asymmetric left-right combinations of the five lowest-order normal modes of vocal fold, superimposed on varied vocal fold postural configurations. An in-house Navier-Stokes flow solver is used to simulate the flow rate and pressure in the glottis, based on which, the flow resistance is computed. Then a deep neural network (DNN) machine-learning model based on Keras is trained to develop a reduced-order model of flow resistance. With this trained model, we can predict the flow resistance in any shapes of the glottis, and finally obtain the flow rate and pressure with the modified Bernoulli model. A continuum-mechanics based vocal fold model is then coupled to predict the three-dimensional vibratory dynamics. To demonstrate that the present machine-learning based reduced-order model outperforms in terms of both accuracy and CPU time, the Navier-Stokes and original Bernoulli flow model are respectively coupled with the vocal fold model for comparison.

Results/Conclusions

The neural network model is evaluated by performing the 10-fold cross validation. The result shows that the mean absolute error of the flow resistance on the train-test folds is around 6%. Further results show that the reduced-order model can have a good prediction on the intraglottal pressure for arbitrary shapes of the glottis. Finally, the comparison of the pressure on a continuum-mechanics based vocal fold model obtained by the Navier-Stokes, the original Bernoulli, and the present reduced-order flow models show that, the present model can give a good prediction of the three-dimensional vibratory dynamics with a significant reduction of the CPU time, which holds a great promise for future clinical use.

1.5. Updated Rules for Constructing a Triangular Body-Cover Model of the Vocal Folds from Intrinsic Laryngeal Muscle Activation

Gabriel A. Alzamendi ¹, Sean D. Peterson ², Byron D. Erath ³ and Matías Zañartu ¹

¹ Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

² Department of Mechanical and Mechatronics Engineering, University of Waterloo, On, Canada

³ Department of Mechanical & Aeronautical Engineering, Clarkson University, Potsdam, NY, USA

Keywords: lumped element model; laryngeal muscle control; intrinsic muscles; laryngeal posturing

Objective

To introduce a scheme to construct a triangular body-cover model (TBCM) of the vocal folds [1] by the independent activation of five intrinsic laryngeal muscles, thus providing revised physiologically-relevant rules that yield model parameters and prephonatory configurations.

Introduction

Prior studies have introduced physiologically-inspired rules for mimicking the effects of intrinsic muscles in the myoelastic properties of the vocal folds and in the larynx configuration through a body-cover model of the vocal folds [2]. In this context, activation of the cricothyroid and thyroarytenoid muscles provides independent control of vocal fold posturing, but the remaining intrinsic muscles are disregarded, thus limiting the ability of body-cover models to describe the role of antagonistic muscle

pairs and to study muscle tension dysphonia. Efforts to account for muscle dynamics have also been performed [3], but they have not been related, or applied, to self-sustained models of the vocal folds. The triangular glottal shape is of particular interest because it accounts for a gradual anterior-posterior closure, has a well-defined posterior gap, and has been used to study vocal hyperfunction [1].

Methods

A two-dimensional biomechanical representation of vocal fold posturing [3] was utilized to obtain adduction and elongation variables using all five intrinsic muscles as input parameters. The adduction and elongation variables were input into known rules for controlling low-dimensional vocal fold models with muscle activation [2], while maintaining the remaining assumptions. A quasi-steady scenario was first implemented, where all passive effects in the muscle dynamics were neglected. The proposed scheme was studied by simulating and assessing the resultant glottal signals for different scenarios.

Results

The proposed scheme is capable of generating physiologically-relevant, complex laryngeal postures as a result of the coordinated activation of the five intrinsic laryngeal muscles. The simulations illustrate that even though laryngeal posture is important, additional parameters (e.g., volumetric flow rate, glottal area, vocal fold displacement) are required to predict the dynamic model behavior. For example, whereas similar vocal fold posturing can be obtained with different muscle activation configurations, the dynamic behavior can be significantly different.

Conclusions

The proposed scheme exhibits great potential to describe complex glottal conditions and configurations with reduced-order vocal fold models. Additional efforts are required for exploring the role of passive muscle components and their impact on transients, such as voicing onset and offset, and for exploring the proposed scheme to study the effect of antagonistic muscles and pathological conditions.

Acknowledgments: Research reported in this work was supported by the NIDCD of the NIH under award P50DC015446, CONICYT FONDECYT 1151077, and the Ontario Ministry of Research and Innovation through the Early Researcher Award program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Galindo, G.E.; Peterson, S.D.; Erath, B.D.; Castro, C.; Hillman, R.E.; Zañartu, M. Modeling the Pathophysiology of Phonotraumatic Vocal Hyperfunction With a Triangular Glottal Model of the Vocal Folds. *J. Speech Lang. Hear. Res.* **2017**, *60*, 2452.
2. Titze, I.R.; Story, B.H. Rules for controlling low-dimensional vocal fold models with muscle activation. *J. Acoust. Soc. Am.* **2002**, *112*, 1064–1076
3. Titze, I.R.; Hunter, E.J. A two-dimensional biomechanical model of vocal fold posturing. *J. Acoust. Soc. Am.* **2007**, *121*, 2254–2260.

1.6. Synthetic Vocal Fold Model Closed Quotient Optimization

Cassandra J. Taylor, Austin C. Vaterlaus, Michael S. Farnsworth and Scott L. Thomson

Department of Mechanical Engineering, Brigham Young University, Provo, UT, USA

Keywords: synthetic vocal fold models; computational vocal fold models; glottal closed quotient

Objectives

A parameterized, low-fidelity, two-dimensional finite element model of vocal fold (VF) flow-induced vibration was coupled with a genetic algorithm optimization tool. The objective was to use this coupled numerical platform to identify geometric and stiffness properties of a synthetic VF model that would exhibit a glottal closed quotient within the normal human physiological range while maintaining reasonable geometry, stiffness, onset pressure, and frequency characteristics. Synthetic VF models were then fabricated based on the optimization results and tested to evaluate their closed quotients.

Introduction

Synthetic VF models are often used to study aspects of voice biomechanics. Although synthetic VFs are capable of self-sustained vibration with favorable lifelike characteristics, creating models with adequate closed quotient values has largely been elusive. Results from a computational study by Zhang (2016) suggested that the medial surface length and VF stiffness characteristics could be prescribed to achieve adequate closed quotient. However, the computational model was defined using anisotropic material properties. This feature is observed in human VFs due to fibers that primarily run along the anterior-posterior direction; however, most synthetic silicone VF models are comprised of isotropic materials. It thus remains to be seen whether materially-isotropic synthetic VF models could be developed that achieve adequate closed quotient values.

Methods

A genetic algorithm coupled with a computational four-layer VF model was used to optimize the geometry and stiffness in order to achieve a flow-induced vibratory frequency in the range of 85 to 150 Hz and a closed quotient in the range of approximately 0.35 to 0.65. The genetic algorithm results were tested by conducting high-speed imaging experiments using synthetic VF models that had been fabricated using parameters based on the optimized configuration.

Results

The algorithm was allowed to proceed for 11 generations and 550 simulations. The results predicted that changes in geometry and stiffness would lead to a model that exhibited the desired characteristics. Various physical models based on optimized parameters were then fabricated and the closed quotient was tested. The physical models successfully vibrated with nonzero closed quotient as predicted by the computational models.

Conclusions

The results predicted that medial surface length, cover thickness, and ligament thickness all played important roles in closure and frequency, and that adequate closed quotient response could be achieved for a range of these parameters. The beginnings of a Pareto front were evident in the results, showing a trade-off between frequency and closed quotient. As predicted, synthetic models based on the geometry and of approximately the same material properties as selected computational models exhibited significant, non-zero closed quotient responses.

Acknowledgments: Support of NIH Grant R01DC005788 is gratefully acknowledged.

References

1. Zhang, Z. Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *J. Acoust. Soc. Am.* **2016**, *139*, 1493–1507.

1.7. Contact Pressure and Length as a Function of Posterior Glottal Area: Synthetic Vocal Fold Investigations

Mohsen Motie-Shirazi ¹, Sean D. Peterson ², Matías Zañartu ³, Daryush D. Mehta ⁴, James B. Kobler ⁴, Robert E. Hillman ⁴ and Byron D. Erath ¹

¹ Department of Mechanical and Aeronautical Engineering, Clarkson University, Potsdam, NY, USA

² Department Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON, Canada

³ Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaiso, Chile

⁴ Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital, Boston, MA, USA

Keywords: synthetic vocal fold models; contact pressure; posterior glottal opening; vocal fold contact

Objectives

This study aims to determine the exact position of vocal fold (VF) contact, and investigate the effects of VF geometry, medial compression, and posterior glottal opening (PGO) on contact length and pressure in synthetic self-oscillating VF models using a hemilaryngeal configuration.

Introduction

A PGO causes air leakage through the glottis, which reduces net energy transfer to the VFs and radiated sound pressure level (SPL). It has been hypothesized that increasing subglottal pressure to compensate for reduced SPL may lead to higher contact pressures and phonotraumatic vocal hyperfunction.

Methods

Newly developed synthetic self-oscillating VF models are fabricated using physiologically-based geometry that includes a substrate layer of adipose tissue and an undercut superior surface¹. The location and length of VF contact is found by coating the model with graphite powder to make its surface electrically conductive, while placing a copper strip on a moveable contact plate, and measuring the resistance between the VF and the copper strip in a Wheatstone bridge configuration. The position of the contact plate is controlled by a micro-positioner, allowing precise measurements of the contact location. A Millar Mikro-Cath pressure sensor (O (>4 kHz) frequency response) is flush-mounted in the contact plate to measure the contact pressure. A PGO is modeled with areas of 0.02, 0.05, 0.08, and 0.1 cm². The radiated SPL is recorded, and a Photron high-speed camera synchronizes the vocal fold motion and pressure measurements. For each PGO area, the radiated acoustic output is maintained constant by adjusting the subglottal pressure.

Results

The addition of a substrate layer of soft adipose tissue, a longer medial VF surface, and an undercut in the superior surface, results in a robust mucosal wave with a clear convergent-divergent transition during each cycle. Kinematic measures of the synthetic VF oscillations demonstrate good agreement with excised larynx measures. Intraglottal contact pressure shows high sensitivity to sensor position in the inferior-superior direction, even within the region of contact. Increasing medial compression increases the contact length and onset pressure, and reduces the mean flow rate. Compensating for increased PGO by increasing subglottal pressure greatly increases the magnitude of contact pressure, the flow rate, and the spectral tilt (less steep).

Conclusions

Identifying the exact contact location is crucial for measuring the correct contact pressure. The amount of medial compression also plays an important role on the aerodynamic performance. In addition, increasing subglottal pressure to compensate for reduced radiated acoustic pressure due to the presence of a PGO increases the contact pressure, which increases the risk of vocal hyperfunction.

Acknowledgments: Research reported in this work was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under award P50DC015446. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Hirano, M.; Sato, K. *Histological Color Atlas of the Human Larynx*; Singular: San Diego, CA, USA, 1993.
2. Syndergaard, K.L.; Dushku, S.; Thomson, S.L. Electrically conductive synthetic vocal fold replicas for voice production. *J. Acoust. Soc. Am.* 2017, *142*, EL63.

2. Session 1

2.1. Vocal-Fold 3D Micro-Architecture and Micro-Mechanics: A Multimodal Imaging Study

Thibaud Cochereau ^{1,2}, Hamid Yousefi-Mashouf ^{1,2}, Lucie Bailly ¹, Jérôme Sohier ³, Laurent Orgéas ¹, N. Henrich Bernardoni ², S. Rolland du Roscoat ¹, Anne McLeer-Florin ⁴ and Olivier Guiraud ⁵

¹ Univ. Grenoble Alpes, CNRS, Grenoble INP, 3SR, Grenoble, France

² Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France

³ Univ. Lyon 1, CNRS, LBTI-IBCP, Lyon, France

⁴ Univ. Grenoble Alpes, CHU Grenoble Alpes, Histology Lab, IAB, Grenoble, France

⁵ Novitom, Grenoble, France

Keywords: vocal folds; biomechanics; fibres; synchrotron X-ray imaging; biphotonic confocal microscopy; histology

Objectives

Current understanding of the histological features of the vocal folds is still insufficient to make the link to their vibromechanical performance. In particular, the 3D microscale rearrangement of the loaded tissues is still to be explored. Thus, the aim of this work is to characterize the 3D histological specificities of human vocal folds' fibrous networks and their strain-induced microstructure evolutions under tensile loading, at the scale of the muscular, collagen and elastin microfiber bundles.

Introduction

The 3D *ex vivo* observation of vocal folds at micron scale is a challenging task using micro Magnetic Resonance Imaging (limited spatial resolution), multiphoton nonlinear scanning microscopy (limited depth of field) and X-ray microtomography in absorption mode (low contrast) [1–3]. Recently, the 3D hierarchical architecture of human vocal folds was revealed by means of synchrotron X-ray microtomography with phase retrieval imaging mode [4]. High-resolution (voxel size of $0.65^3 \mu\text{m}^3$) and in depth 3D images of soft tissues (subvolumes of 1.3^3mm^3) were acquired with fast scanning times (1–2 min), to quantify structure descriptors of *lamina propria* and *vocalis* fibrous networks at multiple length scales. In continuation of such developments, this work presents the preliminary *in situ* tensile tests on vocal-fold tissue coupled to synchrotron X-ray imaging.

Methods

10 tissue samples made of *lamina propria* and *vocalis* sublayers were excised from vocal folds of the same human larynx (male donor, 76 yo), to avoid inter-subject variability, and enable imaging at micro-scale resolution. The dissection yielded to 5 samples from each left/right vocal fold [4], of typical dimensions $15 \times 5 \times 3 \text{ mm}^3$. Several conditions of tissue conservation were used before testing (ethanol, gel, cryopreservation). First, a multimodal imaging study was conducted to investigate the tissue fibrous architecture at rest, thanks to three complementary techniques: (1) Synchrotron 3D X-ray imaging at high-resolution ($0.65^3 \mu\text{m}^3$), using the microtomographs of the ESRF's ID19 beamline; (2) Biphotonic 3D microscopy imaging ($0.35 \times 0.35 \times 0.70 \mu\text{m}^3$) using the confocal microscope of Léon Bérard's Cancer Center (LSM 780, Laser 900 nm); (3) Histological stainings (Hematoxylin-Eosin-Saffron), differentiating vocal-fold constituents on 2D micrographs. Then, several samples were subjected to uniaxial tensile tests interrupted by relaxation steps, and combined to 3D X-ray imaging using a mechanical device positioned onto the ESRF microtomographs. Holding conditions were optimized to avoid stick-slip effects and tissue damage.

Results

Micromechanisms of deformation of the vocal fold tissues under tensile loading will be described and quantified by tracking various structure descriptors as a function of the applied strain: shape and size of their layered fibrous architectures within the *lamina propria* and the *vocalis*; orientation, shape and size of muscle fibres as well as collagen and elastin fibre bundles constituting these layers. Based on the multimodal study, a specific focus will be paid to the distinction of the extracellular matrix fibres within the *lamina propria*. Finally, the impact of the contrast agent on both the tissue mechanics and the quality of the acquired images will be assessed, depending on the immersion procedure (diffusion time, concentration).

Conclusions

Based on advanced micro-imaging techniques, this study provides a quantitative database of the 3D and multiscale descriptors of vocal fold tissues, and their evolution during a mechanical loading.

Acknowledgments: This work was supported by the ANR MICROVOICE N° ANR-17-CE19-0015-01 and the LabEx Tec 21 (Investissements d'Avenir—grant agreement no. ANR-11-LABX-0030).

References

1. Kelleher, J.E.; Siegmund, T.; Du, M.; Naseri, E.; Chan, R.W. *J. Acoust. Soc. Am.* **2013**, *133*, 1625–1636.
2. Miri, A.K.; Heris, H.K.; Tripathy, U.; Wiseman, P.W.; Mongeau, L. *Acta Biomater* **2013**, *9*, 7957–7967.
3. Strupler, M.; et al. *Biomedical Optics in Otorhinolaryngology: Head and Neck Surgery*; Springer: Basel, Switzerland, 2016; pp. 511–528.
4. Bailly, L.; Cochereau, T.; Orgéas, L.; Henrich Bernardoni, N.; Rolland du Roscoat, S.; McLeer-Florin, A.; et al. *Sci. Rep.* **2018**, *8*, 14.

2.2. Influence of Recording Perspective in Laryngoscopy on Perceived Asymmetry

Marion Semmler, Sahar Fattoum, Reinhard Veltrup, Stefan Kniesburges, Anne Schützenberger and Michael Döllinger

Dep. of Otorhinolaryngology, Head and Neck Surgery, Div. of Phoniatics, University Hospital Erlangen, Erlangen, Germany

Keywords: laryngoscopy; asymmetry; perspective distortion; high-speed recordings

Introduction

The visual inspection of the larynx and the phonation process is essential for the diagnostic process in laryngology. Besides structural alterations of the vocal fold tissue and neurologic pathologies, the dynamic behavior enables diagnostic conclusions. Insufficient closure, asymmetry and irregularity in the oscillation pattern of the vocal folds are associated with functional dysphonia and chronic hoarseness [1]. However, during oral endoscopy, the positioning of the rigid laryngoscope depends on the recorded test subject (anatomy, compliance, etc.) as well as the recording operator (ability, handedness, etc.). Distance and angle between the glottal plane and the imaging system are hardly controllable and the resulting perspective distortion is not yet considered in the diagnostic evaluation.

Objectives

Our goal is a systematic investigation of the influence of the perspective distortion in 2D high-speed imaging and laser-based 3D reconstruction. We will test the following two hypotheses: (1) Under certain recording conditions, symmetric vocal fold oscillations can be falsely perceived as asymmetrical in the video footage and even in objective asymmetry parameters. (2) Furthermore, we assume that 3D asymmetry parameters are less susceptible to the influence of the recording perspective than 2D parameters.

Methods

A systematic analysis is performed on a validated synthetic vocal folds model from silicone, which is excited to regular and symmetric oscillation patterns by a mass flow generator [2]. Two different oscillation modes, i.e., without (M1) and with (M2) vocal fold contact can be simulated by variation of the applied flow rate. The synthetic vocal fold model is recorded at 4 kHz by a laser-based 3D imaging system under 13 different angles in-between $\pm 30^\circ$ for 3 different distances (50 mm, 65 mm, 80 mm) above the glottal plane. For each perspective, the combination of a laser projection unit (LPU) and a high-speed (HS) camera provides 2D and 3D data of the oscillating surface [3]. As a reference, a second synchronized HS camera is mounted directly above the synthetic model, recording 2D data under an angle of 0° .

Results

By the use of an in-house software (Glottis Analysis Tool), the 2D recordings of reference and perspective view will be segmented and analyzed with respect to their spatial and temporal asymmetry. We determine objective and clinically established parameters based on the glottal area waveform. In analogy to the more frequently used 2D parameters, we will determine the 3D asymmetry on the basis of the reconstructed 3D surface model.

Conclusions

We will be able to quantify the perspective distortion in video recordings from rigid laryngoscopes and discuss the susceptibility to diagnostic misinterpretation in 2D and 3D imaging.

Acknowledgments: The Else-Kröner-Fresenius Stiftung is highly acknowledged for its funding (grant-no. 2016_A78).

References

1. Eysholdt, U.; Rosanowski, F.; Hoppe, U. Vocal fold vibration irregularities caused by different types of laryngeal asymmetry. *Eur. Arch. Otorhinolaryngol.* **2003**, *260*, 412–417.
2. Kniesburges, S.; Hesselmann, C.; Becker, S.; Schlücker, E.; Döllinger, M. Influence of vortical flow structures on the glottal jet location in the supraglottal region. *J. Voice* **2013**, *27*, 531–544.

3. Semmler, M.; Kniesburges, S.; Birk, V.; Ziethe, A.; Patel, R.; Döllinger, M. 3D reconstruction of human laryngeal dynamics based on endoscopic high-speed recordings. *IEEE TMI* **2016**, *35*, 1615–1624.

2.3. Extracting Reduced-Order Model Parameters from High-Speed Video of Silicone Vocal Folds Using a Gradient-Based Approach

Jonathan J. Deng ¹, Paul J. Hadwin ¹, Mohsen Motie-Shirazi ², Byron D. Erath ², Matías Zañartu ³ and Sean D. Peterson ^{1,*}

¹ Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON, Canada

² Mechanical and Aeronautical Engineering, Clarkson University, Potsdam, NY, USA

³ Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

* Correspondence: peterson@uwaterloo.ca

Keywords: inverse analysis; bayesian inference; optimization; patient specific modeling

Objective

The ultimate aim of this work is to develop patient-specific reduced-order vocal fold models capable of providing clinically relevant measures, such as contact forces. Herein, we aim to develop a “patient-specific” reduced-order model of self-oscillating silicone vocal folds and explore the relationship between the estimated model parameters and the actual physical vocal fold properties.

Introduction

Patient specific reduced-order models (ROMs) of vocal fold motion have been successfully generated using inverse analysis, primarily through two approaches, traditional least squares [1] and Bayesian estimation [2]. The Bayesian framework provides a probable range for the estimated parameters in the form of estimated values and their uncertainties, whereas least squares generally provides only the point estimates. While both methods have successfully inferred patient-specific ROM parameters, it remains unclear how these parameters map to real vocal fold (VF) tissue properties and other voice control features (muscle tension, articulation, etc.).

Methods

An adjoint-state gradient-based optimization method is used to infer body-cover model (BCM) parameters from observations of the kinematics of silicone VFs. Specifically, the gradient of the objective function (the squared difference between the measured and model glottal area waveforms) is used to fit the model parameters to observed measurements. This method computes an exact gradient with the cost of just two model evaluations, independent of the number of parameters. The method is first verified by inferring parameters from synthetic data and is then used to estimate parameters from observations of silicone VF kinematics. Parameter uncertainties are estimated using a Gaussian approximation of the posterior of the BCM parameters.

Results

Validation against simulated data demonstrates that the gradient-based approach is a fast and viable method for inference of BCM parameters. The estimation process requires approximately 300 evaluations of the forward model, with the averaged error in glottal area on the order of 10^{-3} cm². Estimated subglottal pressures from the silicone VFs show reasonable agreement with known subglottal pressures, however estimated material properties have a large bias in comparison with the known material properties, a potential artefact of the high-dimensionality of the parameters space. Uncertainty in estimated parameters is also high, with standard deviations exceeding 40% of some parameter values, largely due to the high-dimensionality.

Conclusions

Gradient-based optimization methods are an attractive approach for inferring patient-specific parameters since they scale well to high-dimensional parameter spaces. This usually leads to an increase in parameter uncertainty, however, which could be offset by incorporating additional subject measurements, such as glottal flow, by adding prior information, or by reducing the dimensionality. Estimated material parameters of the BCM show large biases, while the subglottal pressure shows reasonable qualitative agreement. A general approach to map material properties to ROM parameters is the subject of ongoing work.

Acknowledgments: Research reported in this work was supported by the NIDCD of the NIH under award P50DC015446, the Ontario Ministry of Research and Innovation through the Early Researcher Award, and NSERC's CGS-M program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Döllinger, M.; Hoppe, U.; Hettlich, F.; Lohscheller, J.; Schubert, S.; Eysholdt, U. Vibration parameter extraction from endoscopic image series of the vocal folds. *IEEE Trans. Biomed. Eng.* **2002**, *49*, 773–781.
2. Hadwin, P.J.; Galindo, G.E.; Daun, K.J.; Zañartu, M.; Erath, B.D.; Cataldo, E.; Peterson, S.D. Non-stationary Bayesian estimation of parameters from a body cover model of the vocal folds. *J. Acoust. Soc. Am.* **2016**, *139*, 2683–2696.

2.4. Segmenter's Influence on Objective Glottal Area Waveform Measures from High-Speed Laryngoscopy

Youri Maryn ^{1,2,3,4}, Monique Verguts ^{1,5}, Hannelore Demarsin ¹, Pablo Gomez ⁶, Patrick Schlegel ⁶ and Michael Döllinger ⁶

- ¹ European Institute for ORL-HNS, Dep. of Otorhinolaryngology, GZA Sint-Augustinus Hospital, Wilrijk, Belgium
- ² Dep. of Rehabilitation Sciences, University of Ghent, Ghent, Belgium
- ³ Fac. of Education, Health and Social Work, University College Ghent, Ghent, Belgium
- ⁴ Phonanium "Elements of phonatory sound", Lokeren, Belgium
- ⁵ Dep. of Otorhinolaryngology, General Hospital Diest, Diest, Belgium
- ⁶ Div. of Phoniatrics and Pediatric Audiology, Dep. of Otorhinolaryngology, University Hospital Erlangen, Erlangen, Germany

Keywords: glottal area waveform; measures; high-speed laryngoscopy; segmenter reliability

Objectives/Introduction

With sufficient spatio-temporal resolution, high-speed laryngeal imaging has potential to objectively quantify vibratory vocal fold characteristics. This is relevant especially for clinical tracking of vocal folds status and vibration as well as scientific study of vocal fold physiology and pathology. Glottal Analysis Tools (GAT) version 2018 (Friedrich-Alexander Universität Erlangen-Nürnberg, University Hospital Erlangen, Germany), for example, is software to objectively determine various glottal area waveform (GAW) quantities. However, prior to having GAT analyzing high-speed laryngeal images, clinicians/researchers (i.e., 'segmenters') have to define laryngeal region of interest and to segment glottal area across videos in a semi-automatic segmentation algorithm. Such subjective human interventions are hypothesized to induce variability across segmenters and consequently to attenuate GAT measures' reliability. This study therefore explored variability in GAT's GAW measures based on differences in glottis segmentation within and between segmenters.

Methods

Twenty high-speed laryngeal videos from normophonic as well as dysphonic subjects with various laryngeal pathologies were selected for study. Videos were recorded at the ENT department of GZA Sint-Augustinus Hospital (Wilrijk, Belgium) with following equipment: Machida LY-C30 rigid endoscope (10 mm diameter, 70° view) (Chiba, Japan), CUDA Surgical E300 xenon lightsource (300 Watt) (Jacksonville, FL, USA), Photron Fastcam MC2.1 unit with MC2 camera head (Tokyo, Japan), 15-to-36 lens adaptor, and Laryngograph high-speed video recording software (London, UK). Frame rate was 4 kHz in sixteen recordings and 8 kHz in four recordings. There were three segmenters: one ENT trainee (S1), one laryngology consultant (S2), one voice therapist (S3). After onsite training in Erlangen (S2 and S3) and peer-learning in Wilrijk (S1, S2 and S3), they separately delineated glottis areas in the same frame sets of these twenty-five videos. Upon analysis of GAW, GAT automatically offers sixty measures related to fundamental frequency, amplitude/period/energy perturbation, noise, mechanics, GAW quotients/periodicity/derivatives and symmetry. To address GAT's reliability, intra- and inter-segmenter-based variability in these measures was examined with single-measures consistency-type intraclass correlation coefficient (ICC) in SPSS version 20.

Results

In general, ICC behavior of the sixty GAW measures across the three raters was highly acceptable. Inter-segmenter-based variability was acceptably low, with reliability ranging between $ICC_{INTER} = 0.65$ to $ICC_{INTER} = 1.00$, with $ICC_{INTER} > 0.9$ for 50 measures (83.3%), $0.9 \geq ICC_{INTER} > 0.7$ for 9 measures (15.0%), and $0.7 \geq ICC_{INTER}$ for only 1 measure (1.7%).

Conclusions

The found high ICC_{INTER} values for 59 out of 60 parameter confirm the applicability of the GAT software. Naturally, manual user interaction (with for example decisions in grey-value adjustment or glottal middle axis determination) effects the outcome to a certain extent. Next steps will investigate the exact influence of the differences in clinical assessment. However, current results suggest that these small inter-rater differences will not noticeably influence clinical assessment. To further improve such software's performance, training in glottis area segmentation and/or even fully automated segmentation algorithms are desired.

2.5. Vocal Fold Collision Pressure Amplitude and Timing in an Excised Hemilarynx Setup with Dual High-Speed Videoendoscopy

Daryush D. Mehta ¹, James B. Kobler ¹, Matías Zañartu ², Byron D. Erath ³, Mohsen Motie-Shirazi ³, Sean D. Peterson ⁴, Robert H. Petrillo ¹ and Robert E. Hillman ¹

¹ Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital, Boston, MA, USA

² Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

³ Department of Mechanical & Aeronautical Engineering, Clarkson University, Potsdam, NY, USA

⁴ Department of Mechanical and Mechatronics Engineering, University of Waterloo, Ontario, Canada

Keywords: Subglottal Pressure; Intraglottal Pressure; Vocal Fold Collision; Hemilarynx

Objectives

The purpose of this study was to measure vocal fold collision pressures during the self-sustained phonation of an excised human hemilarynx model and compare the results with collision measurements made in silicone vocal fold models that offer convenience and mechanical precision but lack some of the biomechanical complexity of real laryngeal tissue.

Introduction

Knowledge of how vocal fold collision forces are related to characteristics of vocal fold closure is critical to improving physical and computational models of voice production. Developing more physiologically relevant models is also an important step in better understanding the pathophysiology of phonotraumatic voice disorders (e.g., vocal fold nodules). In this study, an excised human hemilarynx model allowed for a detailed understanding of vocal fold collision amplitude and timing using high-speed videoendoscopy of the superior and medial vocal fold surfaces that can inform future *in vivo* measurement.

Methods

Three hemilarynx models were prepared using excised tissue from adult male cadavers. The right vocal fold and associated supraglottal tissues were removed, and the specimens were mounted in a custom apparatus such that the left vocal fold vibrated against a transparent Lucite acrylic window. Two pressure transducers (Mikro-Cath, Millar, Inc.) were embedded in a dovetailed slider that could be advanced through the glottis in a superior-inferior direction. A third pressure transducer served as a reference measure of subglottal pressure below the trachea. High-speed video was recorded at 4000 frames per second using two cameras for top-down and *en face* imaging of the superior and medial vocal fold surfaces, respectively. Synchronization of the cameras and sensor data was achieved using a common external clock. A TTL signal synchronized video data with signals from the pressure sensors, an acoustic microphone, and a high-bandwidth accelerometer mounted externally on the anterior tracheal wall. Trials consisted of aerodynamically driven phonation with systematic variations in pressure sensor positioning (superior-inferior, medial-lateral) and subglottal pressure.

Results

Self-sustained oscillation was achieved at subglottal pressures ranging from 10 to 60 cm H₂O. The location and timing of vocal fold collision were verified using a custom graphical user interface that visualized frame-by-frame high-speed video data along with the time-synchronized sensor signals. The intraglottal pressure signal exhibited an impulse-like peak when vocal fold contact occurred, followed by a broader peak that is theoretically related to intraglottal pressure build-up during the de-contacting phase. As subglottal pressure was increased, the peak amplitude of the collision pressure increased and typically reached a value below that of the average subglottal pressure; this pattern was similar to that seen in previous excised animal and computational models.^{1,2} Pressure measurements made just above or below the glottis were significantly different from those made in the contact zone, in contrast to observations in silicone vocal fold model experiments. As expected, vocal fold collision pressure was highest in the mid-glottis compared with pressures in regions toward the anterior commissure and vocal process.

Conclusions

The excised hemilarynx experimental setup provided important baseline vocal fold collision pressure data with which computational models of voice production can be developed and *in vivo* measurements can be referenced. A long-term goal of this work is to continue developing vocal dose measures that incorporate vocal fold collision information and can be estimated from noninvasive neck-surface vibration signals.

Acknowledgments: Funding provided by the Voice Health Institute and the National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders (Grant P50 DC015446). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Jiang; et al. *J. Voice* **2001**, *15*, 4–14.

2. Chen; Mongeau. *J. Acoust. Soc. Am.* **2001**, *130*, 1618–1627.

2.6. Recent Advancements in Acoustic Analysis for Assessing Laryngeal Function

Jack J. Jiang

Department of Surgery—Division of Otolaryngology, University of Wisconsin School of Medicine, Madison, WI, USA

Keywords: voice; acoustic analysis; chaos

Objective

Review recent developments of acoustic analysis and its role in assessing laryngeal function.

Introduction

Acoustic analysis is a non-invasive and objective method for distinguishing normal from disordered patients. Traditional analysis includes perturbation parameters and nonlinear dynamic methods such as correlation dimension and Lyapunov exponents. While these nonlinear dynamic methods are successful in distinguishing between chaos and periodicity, error frequently occurs in the analysis of highly disordered voice signals or what is now considered type 4 voice. The voice typing paradigm was first introduced by Titze to describe type 1 voice which is nearly periodic, type 2 voice which is primarily periodic with subharmonics and bifurcations, and type 3 voice which is aperiodic and chaotic. Type 4 voice was later introduced to describe chaotic signals that contain a significant portion of stochastic noise, stemming from turbulent air in the vocal tract. More recent analyses have successfully differentiated between type 3 and 4 voices; however, they only produce singular values and simply reflect the degree of aperiodicity or disorder in a voice signal. In reality, periodic elements and complex nonlinear phenomena, such as subharmonic frequencies, bifurcations, deterministic chaos and stochastic noise, are simultaneously present in voice.

Methods

Two parameters, the intrinsic dimension and diffusive chaos methods, were recently developed to analyze the distribution of voice type components (VTCs) that are present in the voice signal. The intrinsic dimension represents the lowest dimension at which the data remains fully intelligible, and this method functions by making multiple estimates of local dimension throughout the signal. Diffusive chaos analysis is a signal processing technique that functions by repeatedly assessing the bounded or unbounded trajectories of two variable parameters within the time series. Both methods output the percentage of voice type components 1, 2, 3, and 4 that are present. To assess the effectiveness of these parameters, 135 disordered voice samples of sustained /a/ vowels were selected from the Disordered Voice Database 4337, classified according to the voice type paradigm using spectrogram analysis, and analyzed with diffusive chaos and intrinsic dimension analyses.

Results

Both methods demonstrate that the distribution of VTCs varies distinctly across traditional voice type groups. Highly disordered type 4 voices contain high proportions of voice type component 4 (VTC4), indicating the presence of noise; however, there are also smaller proportions of VTC1, indicating an underlying periodicity. Notably, the VTCs of type 3 voices are all significantly different from the VTCs of type 4 voices ($p < 0.001$). These results were compared to calculations of correlation dimension and spectrum convergence ratio, which both demonstrated limited effectiveness in differentiating between all four voice types.

Conclusions

These analyses provide a comprehensive description of the signal elements present in the voice, thus, they have the potential to provide a more complete and quantitative evaluation of the voice and treatment efficacy.

Acknowledgments: This work was supported in part by NIH grant R01 DC006019 from the National Institute of Deafness and Other Communication Disorders.

References

1. Liu, B.; Polce, E.; Jiang, J. Application of Local Intrinsic Dimension for Acoustical Analysis of Voice Signal Components. *Ann. Otol. Rhinol. Laryngol.* **2018**, *127*, 588–597.
2. Liu, B.; Polce, E.; Sprott, J.C.; Jiang, J.J. Applied Chaos Level Test for Validation of Signal Conditions Underlying Optimal Performance of Voice Classification Methods. *J. Speech Lang. Hear. Res.* **2018**, *61*, 1130–1139.

2.7. *Optimization of Relative Fundamental Frequency Estimation Algorithms: Accounting for Sample Characteristics and Fundamental Frequency Estimation Method*

Jennifer M. Vojtech^{1,2}, Katharine R. Kolin^{2,3}, Roxanne K. Segina³ and Cara E. Stepp^{1,2,4}

¹ Department of Biomedical Engineering, Boston University, Boston, MA, USA

² Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA

³ Undergraduate Program in Neuroscience, Boston University, Boston, MA, USA

⁴ Department of Otolaryngology–Head and Neck Surgery, Boston University School of Medicine, Boston, MA, USA

Keywords: voice; relative fundamental frequency; laryngeal muscle tension; signal processing

Objectives

Quantitative measures of laryngeal muscle tension are needed to improve assessment and track clinical progress. While relative fundamental frequency (RFF) has shown promise as an acoustic measure for evaluating tension-based voice disorders, it is not yet transferable to the clinic. A primary obstacle is the accuracy of semi-automated algorithms to calculate RFF from a wide range of vocal signals; here, we evaluate the impact of sample characteristics and fundamental frequency (f_0) estimation techniques on the correspondence between automated and gold-standard manual RFF estimates.

Introduction

During a voiced sonorant–voiceless consonant–voiced sonorant (VCV) production, RFF captures instantaneous changes in f_0 corresponding to the transition into and out of the voiceless consonant. Current algorithms [1] rely on autocorrelation-based f_0 estimation; however, this method assumes signal periodicity and f_0 stability. Additionally, the relationship between manually- and algorithmically-extracted RFF values is dependent on sample characteristics (e.g., signal acquisition quality, overall dysphonia severity), but current RFF algorithms have yet to take this into account. Optimizing the method of f_0 estimation and understanding the effects of sample characteristics on the resultant accuracy of RFF estimations is a necessary step towards improving semi-automated RFF estimation for clinical translation.

Methods

Acoustic recordings were collected from individuals with voice disorders (VD; 148 female, 79 male; M = 52.9 years, SD = 17.7 years; overall severity range = 0–100) and healthy controls (C; 152 female,

104 male; $M = 37.6$ years, $SD = 22.3$ years; overall severity range = 0–44.6). Participants produced three sets of three VCV utterances in either a sound-treated room (128 VD, 207 C) or in a quiet room/waiting area (99 VD, 49 C), for a total of 1449 speech samples from 483 independent speakers. Sample characteristics were quantified via pitch strength [2] and signal-to-noise ratio. Common f_0 estimation algorithms (Halcyon, RAPT, A-SWIPE', and YIN) were compared to autocorrelation. Using a training set ($N = 1158$), categories based on pitch strength values were then constructed and RFF algorithm thresholds were tuned to each category. RFF values were then recalculated on a test set ($N = 291$) using category-specific thresholds. Algorithmically-extracted RFF values were evaluated against manually-extracted RFF values using mean error (ME).

Results

The RFF algorithms with Halcyon for f_0 estimation led to the greatest correspondence with manual RFF; thus, Halcyon was used in concert with category-specific thresholds. Optimizing f_0 estimation and accounting for sample characteristics led to increased correspondence with manual RFF (ME = 0.016 ST) when compared to the algorithms without modifications (i.e., autocorrelation for f_0 estimation; ME = 0.088 ST) and only optimizing f_0 estimation method (i.e., Halcyon; ME = 0.071 ST).

Conclusions

Optimizing the f_0 estimation method and accounting for sample characteristics led to improved correspondence between semi-automated RFF and gold-standard manual RFF. These findings highlight the importance of considering f_0 estimation method and sample characteristics for semi-automated RFF computations across a broad range of vocal function.

Acknowledgments: This work was supported by grants DC015570 from the National Institute on Deafness and Other Communication Disorders and DGE-1247312 from the National Science Foundation.

References

1. Lien, Y.S.; et al. Validation of an Algorithm for Semi-automated Estimation of Voice Relative Fundamental Frequency. *Ann. Otol. Rhinol. Laryngol.* **2017**, *126*, 712–716.
2. Kopf, L.M.; et al. Pitch Strength as an Outcome Measure for Treatment of Dysphonia. *J. Voice* **2017**, *31*, 691–696.

2.8. Acoustic Phonatory Tremor Index: Objective Quantification of Perceived Vocal Tremor Severity

Youri Maryn ^{1,2,3,4,5}, Andrzej Zarowski ¹, Marc Leblans ¹ and Julie Barkmeier-Kraemer ⁵

¹ European Institute for ORL-HNS, Dep. of Otorhinolaryngology, GZA Sint-Augustinus Hospital, Wilrijk, Belgium

² Dep. of Rehabilitation Sciences, University of Ghent, Ghent, Belgium

³ Fac. of Education, Health and Social Work, University College Ghent, Ghent, Belgium

⁴ Phonanium “Elements of phonatory sound”, Lokeren, Belgium

⁵ Div. of Otolaryngology and Head & Neck Surgery, University of Utah, Salt Lake City, UT, USA

Keywords: vocal tremor; auditory-perceptual evaluation; acoustic measurement; validity

Objectives/Introduction

Vocal tremor can be described as a “quavering type of speech” due to tremor affecting muscles of speech structures. At the phonatory level, vocal tremor is characterized by modulations in fundamental frequency (f_0) and intensity level (IL). Measurement of the different modulation properties—i.e., rate, extent and perturbation—in f_0 and IL is regarded relevant to diagnostically differentiate between

patient groups and to track the severity of tremor across time and clinical treatment paths. As it typically are tremor characteristics in the speech signal that trigger the speaker's complaint or help/care search, auditory-perceptual evaluation of vocal tremor severity can be regarded as the primary, yet subjective, criterion to be accounted for in clinical assessment of vocal tremor. Acoustic analysis of that speech signal then is the logical choice for objective quantification. This study explored associations between (a) auditory-perceptual vocal tremor severity evaluation (i.e., PVTS), and (b) monivariate acoustic markers as well as a multivariate acoustic model of tremorous vowel signal modulation (i.e., Acoustic Phonatory Tremor Index or APTI) in a multi-centric and clinically representative batch of recordings including various vocal tremor severity levels. PVTS reliability was addressed and APTI's clinical utility was assessed in terms of correlation with PVTS (i.e., assessment across the complete continuum) and diagnostic precision (i.e., assessment between negative and positive cases).

Methods

Fifty-six mid-vowel sustained [a:] recordings were selected to have a convenience sample including the continuum from completely absent to most severe vocal tremor. Four female audiologists were asked to rate each of these sustained vowel samples on 'voice tremor severity' on a continuous 10-cm scale. Prior to this task, however, they were requested to listen to a set of tremor-induced training samples synthesized in the program Praat. At the end of the rating session, fifteen randomly selected recordings were presented a second time for the purpose of intra-rater reliability assessment. Customized audio signal processing in Praat yielded ten acoustic measures of rate, extent and perturbation of f_0 and IL modulation. Enter-type multiple linear regression analysis was applied to weight and combine these acoustic variables into an acoustic model of PVTS.

Results

After removing the PVTS ratings of one of the audiologists because of insufficient intra- and inter-rater reliability, mean single-measures consistency-type intraclass correlation coefficients (i.e., ICC) equalled a reasonable 0.83 within raters and 0.72 between raters. Correlation between mean PVTS and the ten acoustic markers ranged from 0.76 for median extent of f_0 modulation to 0.11 for rate of IL modulation. Correlation between mean PVTS and APTI was 0.88. Analysis of APTI's receiver operating characteristics (i.e., ROC) yielded an area under ROC of 0.93, denoting sensitivity = 0.87 and specificity = 0.91.

Conclusions

This study demonstrated that auditory-perceptual ratings of vocal tremor severity are guided primarily by f_0 and IL modulation extent, less by modulation perturbation, and least by modulation rate. The APTI model covering all these modulation properties yielded very acceptable results in terms of both concurrent and diagnostic validity. However, external cross-validation of APTI is warranted before applying it in clinical voice/speech assessment.

2.9. Accelerometer-Based Prediction of Subglottal Pressure in Healthy Speakers Producing Non-Modal Phonation

Jonathan Z. Lin ¹, Víctor M. Espinoza ^{2,3}, Matías Zañartu ³, Katherine L. Marks ^{1,4} and Daryush D. Mehta ^{1,4,5}

¹ Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, MA, USA

² Department of Sound, Universidad de Chile, Santiago, Chile

³ Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

⁴ MGH Institute of Health Professions, Massachusetts General Hospital, Boston, MA, USA

⁵ Harvard Medical School, Boston, MA, USA

Keywords: subglottal air pressure; ambulatory voice monitoring; accelerometer; inverse filtering

Objectives

Develop a methodology for accelerometer (ACC)-based estimation of subglottal air pressure (Ps) that incorporates ACC-based measures of vocal function to achieve improved prediction of Ps during non-modal phonation.

Introduction

Ps plays a major role in voice production and is a primary factor in controlling voice onset, offset, and intensity, and contributes to variation in fundamental frequency. In terms of clinical voice assessment, Ps alone, or in combination with other parameters (e.g., aerodynamic resistance and vocal efficiency measures), has been shown to differentiate between normal and disordered voice production, providing insight into changes in vocal function associated with treating voice disorders. In particular, the method uses an unobtrusive miniature ACC sensor attached to the anterior base of the neck. Previous work has shown promise for the robust estimation of Ps from ACC signal amplitude during typical modal voice production across multiple pitch and vowel contexts.¹ This study expands on that work to incorporate additional ACC-based measures of vocal function to compensate for non-modal phonation characteristics and achieve better estimation of Ps.

Methods

Subjects with normal voices repeated /p/-vowel syllable strings from loud-to-soft levels in multiple vowel contexts (/pa/, /pi/, and /pu/), pitch conditions (comfortable, lower than comfortable, higher than comfortable), and voice quality type (modal, breathy, strained, and rough). Ps estimates were obtained via intraoral pressure (IOP) recordings during occlusive plosives using an intraoral catheter connected to a pressure sensor. Simultaneously, oral airflow was captured using a circumferentially vented pneumotachograph mask. Ps for each vowel was estimated by taking the average of IOP peaks preceding and following the vowel. Subject-specific, linear regression models were constructed using root-mean-square (RMS) values of the ACC signal (ACC RMS) alone and in combination with additional flow- or ACC-based measures to estimate Ps across vowel, pitch, and voice quality contexts. These additional measures included fundamental frequency, cepstral peak prominence, and glottal airflow parameters from inverse filtering (IF) the oral airflow³ and subglottal impedance-based inverse filtering (IBIF) the ACC signal². Cross-validation assessed the robustness of model performance using the root-mean-square error (RMSE) metric for each regression model.

Results

Each fold of the 5-fold cross-validation exhibited an increase in RMSE when ACC RMS-alone models were used to predict Ps across both modal and non-modal phonation. Improvements to model performance (decreases in RMSE) were found when the following glottal airflow measures of vocal function were included in the model: open quotient, speed quotient, normalized amplitude quotient, maximum flow declination rate, harmonic richness factor, peak-to-peak amplitude of the unsteady glottal airflow, and the difference between first and second harmonic amplitudes. Critically, similar model performance was achieved when the same flow-based IF measures were derived from the ACC signal using IBIF, thus showing promise for ACC-only prediction of Ps for modal and non-modal phonation.

Conclusions

Improved estimation of Ps for non-modal phonation is achievable with additional ACC-based measures, leading to future exploration of subglottal pressure estimation in patients with voice disorders and in ambulatory voice recordings.

Acknowledgments: This work is supported by the NIH National Institute on Deafness and Other Communication Disorders (Grants R21 DC015877 and P50 DC015446). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Fryd, A.S.; et al. Estimating subglottal pressure from neck-surface acceleration during normal voice production. *J. Speech Lang. Hear. Res.* **2016**, *59*, 1335–1345.
2. Zaňartu, M.; et al. Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration. *IEEE Trans. Audio Speech Lang. Process.* **2013**, *21*, 1929–1939.
3. Espinoza, V. M.; et al. Glottal aerodynamic measures in adult females with phonotraumatic and non-phonotraumatic vocal hyperfunction, *J. Speech Lang. Hear. Res.* **2017**, *60*, 2159–2169.

2.10. Classification of Vocal Gestures Extracted from Quasi-Daily Sentences to Detect Vocal Fatigue

Yixiang Gao ¹, Maria Dietrich ², Melinda Pfeiffer ², Allison Walker ² and Guilherme N. DeSouza ¹

¹ Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, USA

² Department of Speech, Language and Hearing Sciences, University of Missouri, Columbia, MO, USA

Keywords: voice; vocal fatigue; surface electromyography; pattern recognition

Objectives

In this study, we used techniques for Natural Language Processing (NLP) to extract vocal gestures—such as the /a/, /i/, and /u/ vowels—from quasi-daily sentences. This extraction allows us to employ our previously developed classifier of surface Electromyography (sEMG) signals to detect vocal dysfunction from subjects with and without self-reported vocal fatigue.

Introduction

Voice disorders pose a significant threat to teachers' careers. However, signs of common symptoms such as vocal fatigue are elusive during assessments. From our previous study [1], we have developed a pattern recognition system for detecting vocal fatigue based on data from surface electrodes placed on the anterior neck by classifying normal and fatigued vowel productions based on the subjects' Vocal Fatigue Index (VFI) index scores on the factor tiredness of voice [2]. In this study, we expanded our system to be no longer limited by isolated vowel gestures, but sentences as a more general form of voice production in teachers' daily life.

Methods

The study was conducted on 61 female subjects including early career teachers within their first ten years of teaching and control subjects. We used Vocal Fatigue Index factor 1 (VFI-1, tiredness of voice) [2] scores to determine whether they had vocal fatigue. Each subject had to perform a series of voice gestures including normal vowel and sentence repetitions. Only the data collected from the sentences were considered in this study. The sentences were "The dew shimmered over my shiny blue shell again" and "Only we feel you do fail in new fallen dew," which were developed for acoustic analysis of relative fundamental frequency [3]. Both sEMG and acoustic signals were collected simultaneously. First, we applied a speech recognition system using Hidden Markov Models on the acoustic signals to locate certain words and vowel-fricative-vowel combinations (e.g., /i/, /u/, /ifi/) from sentences. Next, we extracted regions of interest (ROI) in the sEMG signals associated with the vowel

locations in those sentences. Finally, a Support Vector Machine based pattern recognition system was used to classify extracted sEMG signals between vocally healthy subjects and vocally fatigued subjects.

Results

This study consists of two parts. Data are presented for a matched sample of teachers and controls ($n = 26$). The first part establishes the accuracy of our speech recognition in successfully detecting and locating relevant ROIs within sEMG signals using speech. The second part focuses on comparing the classification of vocally fatigued subjects using sEMG ROI extractions using speech instead of isolated vowel gestures. That is, we present our overall classification accuracy with and without sEMG signal extraction to demonstrate that locating vowels within sentences as opposed to analyzing isolated vowel productions can improve the sensitivity of vocal fatigue detection.

Conclusions

We combined a pattern recognition framework with speech recognition to demonstrate a preliminary daily voice usage monitoring system. The system first detected certain vowels of interests from the speech production of sentences. Then the system extracted sEMG signals from ROIs to be used for classifying whether the subject had vocal fatigue.

Acknowledgments: The study was funded by NIDCD grant R15 DC015335 to MD. Thanks to Ashton Bernskoetter, Taylor Hall, Katherine Johnson, and Haley McCabe for help with data collection.

References

1. Yixiang, G.; et al. Classification of sEMG Signals for the Detection of Vocal Fatigue Based on VFI Scores. In Proceedings of the 2018 IEEE 40th EMBC, Honolulu, HI, USA, 18–21 July 2018; pp. 5014–5017.
2. Nanjundeswaran, C.; et al. Vocal Fatigue Index (VFI): Development and Validation. *J. Voice* **2015**, *29*, 433–440.
3. Lien YA, S.; Gattuccio, C.I.; Stepp, C.E. Effects of Phonetic Context on Relative Fundamental Frequency. *J. Speech Lang. Hear. Res.* **2014**, *57*, 1259–1267.

2.11. Uncertainty of Ambulatory Airflow Estimates and Its Effect on the Classification of Phonotraumatic Vocal Hyperfunction

Juan P. Cortés ¹, Gabriel A. Alzamendi ¹, Alejandro Weinstein ², Juan I. Yuz ¹, Víctor M. Espinoza ¹, Daryush D. Mehta ³, Jarrad H. Van Stan ³, Robert E. Hillman ³ and Matías Zañartu ¹

¹ Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

² School of Biomedical Engineering, Universidad de Valparaíso, Valparaíso, Chile

³ Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, MA, USA

Keywords: phonotraumatic vocal hyperfunction; ambulatory monitoring; machine learning; kalman filter

Objectives

To introduce a method to select the most reliable glottal airflow estimates from the output of an impedance-based inverse filtering method (IBIF) and to assess the effect of the uncertainty of the estimates on differentiating between the daily voice use (classification performance) of subjects with phonotraumatic vocal hyperfunction (PVH) and healthy controls.

Introduction

Ambulatory monitoring using a neck-mounted accelerometer (ACC) and a smartphone platform (referred to as the Voice Health Monitor, VHM) has the potential to provide important information about the role of daily voice use in disorders associated with vocal hyperfunction that cannot be obtained in a clinical setting. The capability to differentiate between healthy and hyperfunctional daily voice use could be greatly enhanced by using IBIF to extract estimates of aerodynamic features from the ACC signal that have been shown to represent physiologically meaningful characteristics of vocal hyperfunction. However, current performance of IBIF estimates in ambulatory scenarios could be affected by a mismatch between laboratory and in-field conditions, thus resulting in an unknown uncertainty in the aerodynamic estimates.

Methods

We used IBIF to estimate the unsteady glottal volume velocity (GVV) airflow waveform from the neck-surface accelerometer and optimized the way its parameters are obtained through a laboratory calibration session to better match ambulatory scenarios. In parallel, we applied a Kalman filter to the accelerometer signal by constructing IBIF into a state-space observation model. The output is a point estimate and uncertainty of the GVV estimated from IBIF (ground truth signal). The root-mean squared error (RMSE) between IBIF- and Kalman-based GVV waveforms is obtained for each 50 ms frame. This scheme is used to select frames with better fit to the IBIF model, thus resulting in more reliable aerodynamic measures, namely AC Flow, maximum flow declination rate, open quotient, speed quotient, and spectral tilt.

Results

Preliminary results using features of the glottal airflow were obtained in an ambulatory study of 5 patients with PVH and 5 healthy matched-control subjects, each wearing the VHM for one week. When we processed all voiced frames (171,220), using a support vector machine (SVM) with Gaussian kernel and Random Forest we obtained AUCs of (0.87, 0.87) and accuracies of (80.4%, 79.0%), respectively. By selecting only those frames with an RMSE lower than the 10th percentile of the total RMSE (best RMSE cases, 15,294 frames), we obtained an AUC of (0.87, 0.87) and an accuracy of (80.3%, 78.8%) with the same SVM and Random Forest, respectively. By using frames with higher than 90th percentile of RMSE (worst RMSE cases, 16,948 frames), we obtained AUCs of (0.81, 0.81) and accuracies (74.5%, 72.5%).

Conclusions

The classification performance is not improved by using the best frames from the Kalman filtering, and the reduction in the number of frames is quite large. The worst adapted frames to the Kalman configuration only degrade the baseline performance by approximately 5 percentage points. These results provide preliminary evidence that relying on only one laboratory calibration procedure does not have a large negative impact on the classification performance (differentiating PVH and controls) when analyzing large amounts of ambulatory data.

Acknowledgments: This research was supported by the National Institute on Deafness and Other Communication Disorders (grants R33 DC011588 and P50DC015446), and CONICYT (grants FONDECYT 1151077 and BASAL FB0008). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

2.12. How Is Vocal Loudness Affected by Spectral Slope

Ingo R. Titze

The National Center for Voice and Speech, University of Utah, Salt Lake City, UT, USA

Keywords: vocal loudness; sound pressure level

Objectives

To determine how perceived vocal loudness is affected by spectral slope and how it differs from SPL.

Introduction

Vocal loudness is often assessed by measuring sound pressure level. It is well-known, however, that loudness in *sones* does not have a direct proportionality to SPL in dB when multiple frequencies are contained in the sound. The auditory system is more sensitive to some harmonics than others. In speech and singing, higher harmonics can contribute more to total loudness than to total SPL.

Methods

The ISO Standard equations for equal loudness level in phons and loudness in sones are used to compute the overall loudness variation that can be obtained with spectral slopes of 0.0, 1.5, 3.0, and 6.0 dB/octave in a vocal output signal when the fundamental frequency varies between 125 Hz and 1000 Hz. The effect of individual harmonics being tuned by vocal tract resonances is also investigated.

Results

Decreasing the spectral slope (more energy in the collective higher harmonic spectrum) is much more effective for increasing loudness than boosting a single harmonic with resonance tuning.

Conclusions

These results explain why talkers tend to press their voices to increase loudness and why singers search for strategies to reinforce multiple harmonics simultaneously with source-filter interaction.

3. Poster Session 1

3.1. Riedel's Thyroiditis Cordal Paralysis: A Single Case Study

Gonzalo Inostroza

Bachelor in Biomedical Sciences of Human Communication, Degree in Speech Therapist, Speech Therapist, Professor, Department of Voice, Mayor's University and Professional Institute of Chile; gonzalo.inostroza.moreno@gmail.com; Tel.: +569-99097684

Objective

To compare rehabilitation from symptomatic and physiological orientation in a patient previously subjected to surgery of Riedel's Thyroiditis—which is chronic thyroid gland inflammation where the fibrous tissue replaces the glandular tissue and extends to adjacent structures. Unfortunately, these patients usually present voice alterations due to the proximity of its structures to the larynx. The present case report aims to show the process of voice therapy on a patient who has undergone thyroid surgery.

Methods

The voice was evaluated by clinical and acoustic procedures, using a long-term average spectrum, f_0 and intensity. The samples were compared before and after performing three months digital laryngeal manipulation exercises and, in the following three months, semi-occluded vocal tract exercises with water resistance, always keeping the same acoustic measurements.

Results

It is determined that with laryngeal manipulation exercises, there are favorable changes in the short term, however, there is no transfer of what has been learned to other contexts. On the other hand, with semi-occluded vocal tract exercises, long-term improvements can be seen, the appearance of a more stable F0 and an intensity that allows the patient to communicate.

Conclusions

It is of great clinical importance to be able to show this case, especially in our country, since no significant improvements have been found in patients suffering from Riedel's thyroiditis. Therefore, it is important to know how to proceed, what is the best treatment to follow to improve the patients' quality of life and the best and most appropriate procedure for patients, from a physiological and symptomatic point of view, that it is proposed to use motor sensory learning with phonatory high pressure tasks to achieve a better degree of contact of the vocal folds.

3.2. Influence of Voice Focus Adjustments on Oral-Nasal Balance in Speech and Singing

Charlene Santoni ¹, Gillian de Boer ², Michael Thaut ³ and Tim Bressmann ⁴

¹ Faculty of Music, University of Toronto, Toronto, ON, Canada

² Department of Speech-Language Pathology, University of Toronto, Toronto, ON, Canada

³ Faculty of Music, University of Toronto, Toronto, ON, Canada

⁴ Department of Speech-Language Pathology, University of Toronto, Toronto, ON, Canada

Keywords: oral-nasal balance; velopharyngeal control; voice focus; nasalance; speech language pathology; singing

Objectives

This study investigated the role of backward and forward voice focus training adjustments on the regulation of oral-nasal balance in speech and singing.

Introduction

Oral-nasal balance is regulated by the degree of opening and closing of the velopharyngeal sphincter, and a competent velopharyngeal sphincter is fundamental to the normal execution of both speech and song. When there is insufficient separation between the oral and nasal cavities, the result is abnormal sound transmission through the vocal tract filter resulting in an oral-nasal balance disorder. Treating velopharyngeal dysfunction using speech therapy alone is difficult because proprioception in the velopharyngeal sphincter is limited, which makes it difficult to control with volition. Kummer (2014) theorized that a yawning maneuver could prompt oral sound redirection. This is indicative of a backward voice focus adjustment. Voice focus describes vocal tract shape and length modifications (forward or backward), which affect voice quality perception (Boone, 1997) and "... produce measurable changes in voice (less perturbation, formant shifts, quality differences)" (Boone et al., 2010, pp. 210). Previous research has indicated that speaking in a forward voice focus with a raised larynx and a shortened vocal tract increases nasal production, whereas speaking in a backward voice focus with a lowered larynx and a lengthened vocal tract reduces nasal production (de Boer & Bressmann, 2015; de Boer et al., 2016). This study expanded on previous research by using singing-voice pedagogy instructional strategies for teaching voice focus adjustments, and by including a singing-task in the study design.

Methods

Twenty participants (10M/10F) aged 24.25 (SD 3.73) read phonetically-balanced, nasal and oral speech stimuli, and sang a phonetically-balanced musical passage in both forward and backward voice focus conditions. A Nasometer 6450 was used to quantify nasalance scores in the different phases of the experiment.

Results

Results indicated that forward voice focus resulted in more nasality ($p < 0.01$) for the oral stimulus and phonetically-balanced song; while backward voice focus prompted a decrease in nasality ($p < 0.01$) for the nasal stimulus, phonetically balanced paragraph and phonetically-balanced song. During production of the phonetically-balanced song, males responded to the forward voice focus condition significantly more than females ($p < 0.01$); and overall the singing task produced the highest numerical score in the forward voice focus condition (41.45%), and the lowest numerical score in the backward voice focus condition (21.95%), in comparison to the phonetically-balanced paragraph.

Conclusions

Voice focus influences oral-nasal balance in normal speakers. These findings have helped to refine an interventional protocol for a follow-up experiment on hypernasal speakers with cleft palate.

Acknowledgments: This research was supported by the *Music and Health Research Collaboratory (MaHRC)* at the University of Toronto.

References

1. Boone, D.R. *Is Your Voice Telling on You?* 2nd ed.; Singular Publishing Group: San Diego, CA, USA, 1997.
2. Boone, D.R.; McFarlane, S.C.; Von Berg, S.L.; Zraick, R.I. *The Voice and Voice Therapy*; Pearson: Boston, MA, USA, 2010.
3. De Boer, G.; Bressmann, T. Influence of Voice Focus on Oral-Nasal Balance in Speech. *J. Voice* **2016**, *30*, 705–710.
4. De Boer, G.; Marino, V.C.; Berti, L.C.; Fabron, E.M.; Bressmann, T. Influence of Voice Focus on Oral-Nasal Balance in Speakers of Brazilian Portuguese. *Folia Phoniatr. Logop.* **2016**, *68*, 152–158.
5. Kummer, A.W. *Cleft Palate and Craniofacial Anomalies—Effects on Speech and Resonance*, 3rd ed.; Delmar Cengage Learning: Clifton Park, NY, USA, 2014.

3.3. Immunological Profiling of Vocal Fold Hydrogel Scaffolds

P. T. Coburn¹, A. Herbay¹ and N. Y. K. Li-Jessen^{1,2,3}

¹ School of Communication Sciences and Disorders, McGill University, Canada

² Department of Biomedical Engineering, McGill University, Canada

³ Department of Otolaryngology, McGill University, Canada

Keywords: hydrogels; tissue engineering; biocompatibility; immunology

Objectives

To evaluate the macrophage inflammatory profile following exposure to glycol-chitosan vocal fold (VF) hydrogels.

Introduction

The foreign body response is a significant obstacle to overcome for VF hydrogels that aim to resolve severe VF scarring. Critical to the success of VF hydrogels is the avoidance of an adverse immune response post-implantation. Rigorous, *in vitro* assessments are required to characterise VF hydrogels prior to *in vivo* pre-clinical testing. In this study, an *in vitro* comprehensive protocol was developed to profile macrophage-mediated inflammatory response to a chitosan-based hydrogel for VF reconstruction.

Methods

Mono- and co-cultures of THP-1 macrophages (M ϕ) and immortalised human vocal fold fibroblasts (HVFF) were exposed to glycol-chitosan hydrogels. The concentration of the crosslinker glyoxal (0.005%, 0.01%, 0.02%) was varied to assess the impact of hydrogel stiffness upon the M ϕ inflammatory response. Rheometry and atomic force microscopy were used to characterise the mechanical properties of hydrogels. Transwells were used to physically separate cell types whilst allowing paracrine signalling. HVFF were embedded in hydrogel 3D culture within the transwell insert whilst M ϕ were seeded on the basal membrane of the insert. At three time-points (3 h, 24 h, 72 h), enzyme-linked immunosorbent assays (tumour necrosis factor (TNF)- α and interleukin (IL)-10) and cell viability were used to analyse the M ϕ inflammatory response. TNF- α and IL-10 were used as pro- and anti-inflammatory markers respectively. Co-cultures were compared to M ϕ monoculture controls and co-culture non-hydrogel controls. Three-way, linear mixed effect regression models were applied for statistical analysis.

In addition, four-colour flow cytometry (CD11b, CD33, CD80, and CD206) investigated the effect of diffusion distance on the ability of HVFF to modulate M ϕ phenotype over a 24 h period. The transwell set-up differed by seeding M ϕ at the base of the lower well instead of the membrane of the insert. This increased diffusion distance for paracrine signalling between M ϕ and HVFF. Co-culture samples were compared with M ϕ monoculture controls, and non-hydrogel, tissue culture plastic, M ϕ monoculture controls.

Results

Mechanical characterisation of the three hydrogels (0.005%, 0.01%, 0.02%) found 0.02% had the highest storage and Young's moduli, representing the stiffest hydrogel among those tested. Secreted protein levels of TNF- α decreased for all glyoxal concentrations in co-culture hydrogels ($p < 0.05$) compared to monoculture hydrogel controls. In parallel, levels of IL-10 increased in co-culture hydrogels ($p < 0.05$) compared to monoculture controls. Additionally, increased hydrogel stiffness produced increased M ϕ viability ($p < 0.05$). Co-culture hydrogels demonstrated higher viability than monoculture equivalents, notably at 72h.

Based on the flow cytometry, both co-cultures and monocultures contained a small M1 M ϕ population (<3% of total cell population). The M2 M ϕ population was more ranged with 0.01% cultures demonstrating a differing response compared to 0.005% and 0.02%. Co-cultures for 0.005% and 0.02% demonstrated a marginal M2 M ϕ population (<4%) compared to equivalent monocultures (~17%). In contrast, 0.01% co-cultures reported an increased M2 M ϕ population (11%) compared to monocultures (<1%).

Conclusions

Co-culture of macrophages with HVFF and increased hydrogel stiffness contributed to an increasingly anti-inflammatory response of macrophages compared to monoculture controls. The paracrine effect of HVFF appeared to be greater than that of hydrogel stiffness for reducing the M ϕ pro-inflammatory response. Further experiments aim to expand the current *in vitro* model to incorporate other relevant cell types of the VF immunological response in addition to a more dynamic culture set-up.

Acknowledgments: The Biomechanics Laboratory, McGill University, for use of the glycol-chitosan hydrogel. Financial support was received from National Institute of Deafness and other Communication Disorders of the National Institutes of Health (R01DC005788 L. Mongeau) and the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-03843 N. Li-Jessen).

3.4. Chemical Receptors of the Larynx: A Comparison of Human and Mouse

Marie E. Jetté ^{1,2}, Matthew S. Clary ¹, Jeremy D. Prager ^{1,3} and Thomas E. Finger ^{2,4}

¹ Department of Otolaryngology, University of Colorado School of Medicine, Aurora, CO, USA

² Rocky Mountain Taste and Smell Center, University of Colorado School of Medicine, Aurora, CO, USA

³ Children's Hospital Colorado, Aurora, CO, USA

⁴ Department of Cell and Developmental Biology, University of Colorado School of Medicine, Aurora, CO, USA

Keywords: chemical receptors; arytenoids; taste buds; innervation

Objectives/Introduction

The larynx is a highly responsive organ exposed to mechanical, thermal, and chemical stimuli. Chemicals elicit responses both in intraepithelial nerve fibers and in specialized chemosensory cells, including scattered solitary cells as well as taste cells organized into taste buds. Activation of both chemosensory cells and taste buds in the larynx elicit cough, swallow, or apnea with exposure to sour or bitter substances, and even by water or sweet-tasting chemicals. In an effort to begin understanding their function, we sought to compare the distribution, density, and types of chemosensory cells and chemoresponsive nerve fibers in laryngeal epithelium of humans and mice.

Methods

Using immunohistochemistry, we identified taste cells and polymodal nociceptive nerve fibers in the arytenoid area of the laryngeal epithelium of the following: (1) infants undergoing supraglottoplasty for laryngomalacia, and (2) a cadaveric specimen procured from a 34-year-old donor. We then compared these findings to both pre-weanling and mature mouse tissue.

Results

Arytenoid tissue from both human and mouse contained many taste buds containing Type II taste cells—bitter, sweet, or umami-sensing—which were innervated by nerve fibers expressing P2X3 type ATP receptors. Type III cells (acid-responsive) were also present, but they were fewer in human tissue than in equivalent tissue from mice. In both species, the epithelium was densely innervated by free-nerve endings.

Conclusions

Our findings suggest that from a standpoint of chemosensation, human and mouse larynges are biologically similar. This suggests that a murine model can be used effectively in laryngeal chemosensory research.

Acknowledgments: This work was supported by the Seymour Cohen Award of the American Laryngological Association granted to MEJ and MSC. This work was also supported by the National Institute on Deafness and other Communication Disorders at the National Institutes of Health [grant numbers T32DC012280, R01DC014728 to TEF, and K23DC014747]. We thank Mei Li for histological preparations, Amanda Ruiz, Emily Jensen, and Christopher Greenlee for assisting with tissue retrieval,

Todd Wine and Melissa Scholes for procuring infant tissue, and Sue Kinnamon and Vijay Ramakrishnan for valuable insights and critical reflection.

Reference

1. Jetté, M.E.; Clary, M.S.; Prager, J.D.; Finger, T.E. Chemical receptors of the arytenoid: A comparison of human and mouse. *Laryngoscope* **2019**, in press.

3.5. An Investigation of Vocal Fatigue Using a Dose-Based Vocal Loading Task

Zhengdong Lei ¹, Laura Fasanella ¹, Nicole Li-Jessen ² and Luc Mongeau ¹

¹ Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

² School of Communication Sciences and Disorders, McGill University, Montreal, QC, Canada

Keywords: vocal fatigue; distance dose; voice quality

Objectives

This study aims at investigating the relationship between voice use and vocal fatigue using a uniquely designed distance-dose-based vocal loading task (VLT).

Introduction

VLTs are often used to investigate the relationship between voice use and vocal fatigue in laboratorial settings. Previous studies on vocal fatigue have reported inconsistent results [1]. One possible cause of inconsistency might be that the VLTs were not standardized and well monitored [1,2]. Very few studies have simultaneously imposed multiple voice parameters, namely, fundamental frequency, loudness, duration, and duty ratio, in the design of the VLTs. Most previous studies have controlled only one or two of these parameters, allowing the other parameters to vary freely [3,4]. The present study investigated a different approach for quantifying vocal loading based on the vocal distance dose in the hope of obtain results that are more consistent.

Methods

A distance-dose-based [5,6] VLT which consisted of six phonation sessions followed by a rest session was designed for the present study. Nine female native Canadian English speakers performed the VLT sessions under the monitoring of an online vocal distance dose calculator. Subjective ratings (CAPE-V and SAVRa) were implemented to evaluate the participants' voice qualities before, between, and after the VLT sessions. Instrumental measures using an acoustic microphone and a neck surface accelerometer (NSA) were done throughout the six VLT sessions. Fatigue-indicative symptoms were manually recorded during the VLT sessions. Across-session variation analyses were performed on subjective ratings, instrumental measures, and symptoms recordings, to evaluate the consistency between different methods.

Results

The overall severity, the roughness, and the strain scores in the CAPE-V showed a similar trend with all three dimensional ratings of the SAVRa. An increase followed by a saturation plateau and a slight decrease formed an "arch-shaped" variation trajectory from the first to the last session. No significant trend was identified for the fatigue symptoms as the discrepancies between certain symptoms for different individuals were too large. Across-session variations of four microphone features (fundamental frequency, SPL, duty ratio, shimmer) and two NSA features (shimmer, spectral tilt) were closely correlated ($r_{\min} > 0.7$) with each other and showed a general trend of a vocal adjustment period followed by a vocal saturation period. This trend was also consistent with that of the subjective ratings.

Conclusions

Under the proposed framework of a distance-dose-based VLT measurement protocol, a vocal adjustment period was consistently observed in subjective ratings and objective measures for the early VLT session, and a vocal rest effect was consistently observed in subjective ratings for the last session. The trends in the SLP-rated CAPE-V results and the self-reported SAVRa results all exhibited a similar arch-shaped variation. Among the fifteen microphone and NSA features considered, six features yielded a consistent voice quality variation.

Acknowledgments: The financial support of the National Institutes of Health (Grant R01 DC-005788) and the National Science and Engineering Research Council is gratefully acknowledged.

References

1. Fujiki, R.B.; Sivasankar, M.P. A Review of Vocal Loading Tasks in the Voice Literature. *J. Voice* **2017**, *31*, 338.e33–388.e39.
2. Welham, N.V.; Maclagan, M.A. Vocal Fatigue: Current Knowledge and Future Directions. *J. Voice* **2003**, *17*, 21–30.
3. Chang, A.; Karnell, M.P. Perceived Phonatory Effort & Phonation Threshold Pressure Across Prolonged Voice Loading Task: A Study of Vocal Fatigue. *J. Voice* **2004**, *18*, 454–466.
4. Laukkanen, A.; Ilomaki, I.; Leppanen, K.; Vilkmann, E. Acoustic Measures and Self-reports of Vocal Fatigue by Female Teachers. *J. Voice* **2008**, *22*, 283–289.
5. Titze, I.R.; Svec, J.G.; Popolo, P.S. Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues. *J. Speech Lang. Hear. Res.* **2003**, *46*, 919–932.
6. Švec, J.G.; Popolo, P.S.; Titze, I.R. Measurement of vocal doses in speech: Experimental procedure and signal processing. *Logop. Phoniatr. Vocol.* **2003**, *28*, 181–192.

3.6. Passive Vowel Devoicing in Osaka Japanese: Case Study Using Electromyography (Emg) and Photoglottography (PGG)

Masako Fujimoto ¹, Ken-Ichi Sakakibara ², Niro Tayama ³ and Kiyoshi Honda ⁴

¹ Advanced Research Center for Human Sciences, Waseda University, Saitama, Japan

² Department of Communication Disorders, Health Sciences University of Hokkaido, Tobetu-cho, Hokkaido, Japan

³ Otolaryngology, Center Hospital of The National Center of Global Health and Medicine, Tokyo, Japan

⁴ College of Intelligence and Computing, Tianjin University, Tianjin, China

Keywords: vowel devoicing; Japanese; dialects; electromyography (EMG), photoglottography (PGG)

Introduction

In the Tokyo (standard) Japanese, high vowels /i/ and /u/ are regularly devoiced when they are preceded and followed by a voiceless consonant. Vowel devoicing in typical environments occurs systematically regardless of speaking rate. Photoglottographic (PGG) observations have shown that the glottis opens continuously with no closing movement for the vowel during the devoiced /CVC/ sequences. Electromyographic (EMG) investigations have revealed that the PCA (post cricoarytenoid) showed a single activation during the voiceless /CVC/. These results suggest that the production of the voiceless vowel is assigned at the stage of speech planning. In other words, vowel devoicing is phonologized and positively controlled at least at the neural level. However, in the Osaka-Kyoto dialects, the phenomenon appears to be less frequent and less systematic. Acoustic studies have revealed that, the accent pattern being equal, the devoicing rate for Osaka speakers varied among

speakers from virtually 0% to 100% in the typical devoicing environment. The glottal opening pattern often showed closing movement for the vowel during voiceless /CVC/ indicating an adduction of glottis for the vowel. Devoicing of Osaka speakers is thus phonetic, occurring at a lower level in speech production. In the present study we carried out an EMG and PGG recording of an Osaka speaker in order to investigate the mechanism of vowel devoicing in the Osaka dialect.

Methods

The subject was a male speaker of the Osaka dialect. EMG signals from the PCA and INT (interarytenoid) muscles were recorded. PGG and audio signals were simultaneously recorded. The laryngeal view was monitored during the session. Test words were /akite/, /agite/, /akise/, /agise/ and /asise/. The vowel /i/ in /akite/, /akise/ and /asise/ is devoiceable since it is placed between voiceless consonants. The subject repeated the test words 20 times in a carrier phrase “Kono XX ga aru (There is this XX.)”.

Results

The devoicing rate of the subject was on average 90%. This rate is as high as the rate of the Tokyo Japanese case. However, PGG signals showed quasi-bimodal or plateau-like patterns, which differs from the monomodal pattern by Tokyo speakers. The PCA activity clearly showed a bimodal pattern. This indicates that the motor neurons for glottal opening were activated twice, each corresponding to each consonant. Hence, while the subject intended to produce a voiced vowel in the /CiC/ sequence, the vowel was devoiced passively. The PGG signal of some of the individual tokens showed a mono-modal pattern. However, the degree of glottal opening of these tokens was comparable to that of a single consonant, which differs from the pattern of Tokyo speakers.

Conclusions

The results clearly indicated that, despite the high devoicing rate which is similar to that of Tokyo Japanese, the subject partially maintains the dialectal rule of laryngeal articulation for vowel voicing (at least at the speech planning level).

Acknowledgments: This study is partly supported by JSPS KAKENHI JP-17K02707.

References

1. Yoshioka, H. Laryngeal Adjustment in the Production of the Fricative Consonants and Devoiced Vowels in Japanese. *Phonetica* **1981**, *38*, 236–235.
2. Yoshioka, H.; Löfqvist, A.; Hirose, H. Laryngeal adjustments in Japanese voiceless sound production. *J. Phon.* **1982**, *10*, 1–10.
3. Fujimoto, M. Vowel devoicing. In *The Handbook of Japanese Phonetics and Phonology*; Haruo, K., Ed.; De Gruyter Mouton: Berlin, Germany, 2015; pp. 167–214.
4. Fujimoto, M.; Kiritani, S. Vowel duration and its effect on the frequency of vowel devoicing in Japanese: A comparison between Tokyo- and Osaka dialect speakers. In *Proceedings of the 15th ICPhS, Barcelona, Spain, 2015*; pp. 3189–3192.

3.7. High-Resolution CFD Simulation of Flow in Glottis Using Les

Petr Šidlof^{1,2} and Martin Lasota²

¹ Institute of Thermomechanics of the Czech Academy of Sciences, Prague, Czech Republic

² Technical University of Liberec, NTI FM, Liberec, Czech Republic

Keywords: voice modelling; numerical simulation; Large Eddy Simulation

Objectives

The objective of the study is to perform a high-resolution CFD simulation of airflow during human phonation. Instead of developing a complex model with two-way fluid-structure interaction, nonlinear mechanical properties of the tissues and MRI-based geometry, the model keeps the 3D channel geometry and vocal fold motion easily parametrizable and focuses on the fluid dynamics. The goal is to assess what level of detail is necessary and how much do the fine turbulent structures influence the aeroacoustic sources and voice generation. In future, the results can be used e.g., for development of voice prostheses.

Introduction

Numerical simulation of the unsteady separated airflow in glottis during phonation, which may be laminar in trachea but undergoes transition to turbulence, is a challenging issue. Largely used laminar model introduces inaccuracy, since it neglects turbulent momentum transfer. Reynolds-Averaged (RANS) models are inappropriate for aeroacoustic simulations, since they provide only mean flow solution with turbulent fluctuations averaged out. Since Direct Numerical Simulation (DNS) is currently unfeasible due to enormous computational cost, the most promising approach seems to be Large Eddy Simulation (LES). In some studies, e.g., (Mihaescu et al. 2011, Sadeghi et al. 2018) LES has been already employed. However, a number of open questions related to optimal boundary layer treatment, turbulence initialization at inlet, subgrid scale model etc. still remain.

Methods

A simplified 3D model of glottis with forced convergent-divergent motion of vocal folds has been developed. The vocal fold geometry is based on Scherer's M5 parametric geometry, the ventricles and ventricular folds are modeled according to data published by Agarwal. The domain was meshed in such a way, that the boundary layer in the convergent glottal channel and intraglottal space is well captured, thus avoiding use of special wall-treatment models. Computational grids used in the simulation range from two to five million elements. The fluid flow is modeled using LES with Smagorinski and one-equation SGS models, and solved by finite volume method using second-order discretization schemes.

Results

The velocity fields, pressure distribution and flow waveforms are compared for different computational meshes. Particular attention is paid to regions, where the major aeroacoustic sources are located, as known from previous studies using a laminar flow model (Šidlof et al. 2015). The results show that the flowrate waveform simulated by LES is significantly lower than the results of the laminar model due to SGS viscosity, which exceeds molecular viscosity by almost one order of magnitude in the intraglottal space.

Conclusions

Large eddy simulation is a promising approach for high-fidelity CFD simulation of airflow during phonation, especially for modeling of fundamental aeroacoustic processes in voice generation. The optimal wall-treatment approach, turbulence initialization, choice of the SGS model and required level of detail for realistic voice generation simulation will be analyzed in further studies.

Acknowledgments: The research was supported by the Czech Science Foundation, project 19-04477S "Modelling and measurements of fluid-structure-acoustic interactions in biomechanics of human voice production".

References

1. Mihaescu, M.; Mylavarapu, G.; Gutmark, E.J.; Powell, N.B. Large Eddy Simulation of the pharyngeal airflow associated with Obstructive Sleep Apnea Syndrome at pre and post-surgical treatment. *J. Biomech.* **2011**, *44*, 2221–2228.
2. Sadeghi, H.; Kniesburges, S.; Kaltenbacher, M.; Schützenberger, A.; Döllinger, M. Computational Models of Laryngeal Aerodynamics: Potentials and Numerical Costs. *J. Voice* **2018**, in press.
3. Šidlof, P.; Zörner, S.; Hüppe, A. A hybrid approach to computational aeroacoustics of human voice production. *Biomech. Model. Mechanobiol.* **2015**, *14*, 473–488.

3.8. Quantification of the Degree of Vocal Fatigue in Teachers by Means of an Interface That Characterizes Voice Signals

Diego Morales ¹, Stephanie Cuellar ¹, Hédrick Robles ¹, Emilio Sánchez ¹ and Lady Catherine Cantor-Cutiva ²

¹ Biomedical Engineering Program, Manuela Beltrán University, Bogotá D.C. (Colombia)

² Speech and Language Pathology Program, Manuela Beltrán University, Bogotá D.C. (Colombia)

Keywords: vocal fatigue; teaching-related factors; signal processing; voice research

Objectives

Determine vocal fatigue levels among college professors in Manuela Beltrán University in Bogotá (Colombia).

Introduction

For some professionals, the voice is the main tool of work. Therefore, voice disorders impact their daily performance and quality of life greater proportion. These workers are called “voice occupational users” [1]. Among those, teachers, singers, speakers, among others have been widely investigated. Generally, voice problems are identified through laryngoscopic examinations, perceptual examinations by speech and language pathologists, self-report or acoustic characterization of voice parameters [2]. In the case of voice occupational users, self-report is one of the most common evaluation methods. Through this has been detected a high occurrence of vocal fatigue in these occupational groups. However, there is no objective measure that quantifies the degree of vocal fatigue [3]. That is the reason why we propose to determine the degree of vocal fatigue in teachers. This will allow voice professionals to identify and recommend healthy vocal use conditions.

Methods

For this study, technological resources will be used to register and characterize the signals obtained. Therefore, differentiate the changes or visible variations in the waves. In order to do this, we will recruit college professors from a private university in Bogota, Colombia. The signals will be obtained at the beginning of the academic day and at the end of the day (repeated measures -intrasubject-). This will show changes that take place during a work day and will allow to detect short-term vocal fatigue. The first step will be to identify parameters that allow to determine variations in the voice signals obtained through digital signal processing protocols. The second step is to quantify variations in the identified variables to obtain vocal fatigue degree. The last step was to define the relation between the identified variables and vocal fatigue control variables (self-report, standard deviation of the fundamental frequency and standard deviation of the level of vocal sound pressure).

Results and Conclusions

At the end of this project, we expect to obtain parameters/variables that help to estimate the degree of vocal fatigue, which will be made by an interface that characterizes the voice signal.

References

1. La voz, huella digital única para el ser humano. *Fundación UNAM*, 2017. Available online: <http://www.fundacionunam.org.mx/unam-al-dia/la-voz-huella-digital-unica-para-el-ser-humano/> (accessed on 11 December 2018).
2. Remacle, A.; Garnier, M.; Gerber, S.; David, C.; Petillon, C. Vocal Change Patterns During a Teaching Day: Inter- and Intra-subject Variability. *J. Voice* **2018**, *32*, 57–63.
3. Hunter, E.; Banks, R. Gender Differences in the Reporting of Vocal Fatigue in Teachers as Quantified by the Vocal Fatigue Index. *Ann. Otol. Rhinol. Laryngol.* **2017**, *126*, 813–818.

3.9. Clinical Practicability of a Newly Developed Real-Time Digital Kymographic System

Jin-Choon Lee ¹, Soo-Geun Wang ¹, Eui-Suk Sung ¹, In-Ho Bae ², Seong-Tae Kim ³ and Yeon-Woo Lee ⁴

¹ Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University School of Medicine, Yangsan, Gyeongsangnam-do, Korea

² Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University Yangsan Hospital, Yangsan, Gyeongsangnam-do, Korea

³ Department of Speech-Language Pathology, Dongshin University, Naju, Jeollanam-do, Korea

⁴ Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University Hospital, Busan, Korea

Keywords: real-time digital kymography; intracordal injection; vocal cord paralysis

Objectives

A digital kymogram shows real images of vocal fold vibration. However, Digital kymography is difficult to use in clinical practice because the recorded image cannot be seen instantaneously after examination, as considerable encoding time is required to visualize a digital kymogram. In addition, frame-by-frame analysis should be implemented to evaluate high-speed videoendoscopy data, but is time- and labor-intensive. The purpose of the study was to validate the clinical practicability of a real-time multislice digital kymographic system developed by the authors. We analyzed the promptness and accuracy of the examination before and after intracordal injections in patients with unilateral vocal fold paralysis.

Methods

To assess the clinical applicability of this system, six patients with unilateral vocal fold paralysis were selected. Real-time DKG was performed before and immediately after intracordal injection. We observed changes in the digital kymogram after the intracordal injection. We could identify the change by checking kymogram before and after procedure.

Results

Using this system, 10 scanning lines and up to five vertical pixel row could be obtained in real time, and the maximum acquisition time for the DKG image was 10 s. A digital kymogram of the patients could be instantaneously acquired, and whether the intracordal injection was appropriate or not.

Conclusions

This article is the first validation study after the development of the real-time multislice digital kymographic system. Our system may be a promising tool in clinical practice for immediate assessment of the vibratory patterns of the vocal cords. More research is necessary for further clinical validation.

3.10. Functional Changes of Submandibular Gland by Steatosis-Induced Ferroptosis in Ovariectomized Rats

Han-Seul Na ¹, Ji Min Kim ², Sung Chan Shin ¹, Jin-Choon Lee ³, Eui-Suk Sung ³ and Byung-Joo Lee ^{1,*}

¹ Department of Otorhinolaryngology, Head and Neck Surgery, Pusan National University Hospital, Busan, Democratic People's Republic of Korea

² Pusan National University Medical Research Institute, Pusan National University School of Medicine, Pusan National University, Busan, Democratic People's Republic of Korea

³ Department of Otorhinolaryngology, Head and Neck Surgery, Pusan National University Yangsan Hospital, Yangsan, Gyeongnam, Democratic People's Republic of Korea

Keywords: submandibular gland; ferroptosis; lipid peroxidation; ovariectomy

Objectives

Our objective was to determine the dysfunction of submandibular gland in ovariectomized rats in order to clarify effects of estrogen deficiency. The aim of this study was to evaluate functional changes of submandibular gland by lipid accumulation in post-menopausal female rats.

Introduction

Xerostomia is representative post-menopausal oral symptoms that may result from decreased salivary secretion. Menopause-related hormonal deficiency may affect the oral conditions by various mechanisms. This study was performed to determine the histological and molecular changes of salivary gland in order to clarify effects of estrogen deficiency on lipid accumulation and cellular dysfunction of submandibular gland in ovariectomized rats.

Methods

Forty-eight female Sprague-Dawley rats aged eight weeks were randomly divided into four groups: group I (1 month after sham-operated rats as CON-OVX1), group II (1 month after ovariectomized rats as OVX1), group III (3 month sham-operated rats as CON-OVX3) and group IV (3 month after ovariectomized rats as OVX3). To investigate whether estradiol effect on lipid accumulation and fibrosis production in submandibular gland, hematoxylin-eosin, masson's trichrome and transmission electron microscope methods were used. Malondealdehyde (MDA) and hydroxyalkens (HAE) absorbance assay was used to evaluate lipid peroxidation levels. Reactive oxygen species (ROS) and glutathione (GSH)/oxidized glutathione (GSSG) assays were used to explore redox imbalance. Real-time qPCR and Western blot method used for mRNA expression of IL-6, TNF α and protein expression of ACC, PPARs and SREBP-1C for investigating lipid metabolism and inflammatory response.

Results

To begin with present study, we confirmed serum estradiol levels were down regulated by surgery and estrogen receptors were well expressed on submandibular gland. MDA and HAE levels increased on serum and cytosolic fraction in OVX group and these results due to upregulated ROS and transformed from GSH to GSSG. Moreover, lipid metabolism-related ACC, PPARs and SREBP-1C mRNA and protein expressions also increased in OVX group compare with CON-OVX group. Eventually, lipid peroxidation-related inflammatory cytokines include IL-6 and TNF α were elevated on submandibular gland. These results occurred fibrosis of submandibular on OVX.

Conclusions

Our finding confirms that estrogen effects on lipid accumulation-induced pathophysiological processes and cellular dysfunction of submandibular gland through lipid peroxidation and its related pro-inflammatory response in estrogen deficiency rats. Furthermore, present results might have clinical important evidence and these results could be used for therapeutic targeting of post-menopausal oral symptoms include dry mouth.

Reference

1. Leimola-Virtanen, R.; Salo, T.; Toikkanen, S.; Pulkkinen, J.; Syrjänen, S. Expression of estrogen receptor (ER) in oral mucosa and salivary gland. *Maturitas* **2000**, *36*, 131–137.

3.11. Extracellular Matrix Turnover in Human Larynx

Yoshitaka Kawai ¹, Brian L. Frey ², Bruce A. Buchholz ³ and Nathan V. Welham ¹

¹ Division of Otolaryngology, Department of Surgery, University of Wisconsin-Madison, Madison, WI, USA

² Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA

³ Center for Accelerator Mass Spectrometry, Lawrence Livermore National Laboratory, Livermore, CA, USA

Keywords: aging; carbon dating; metabolism; regenerative therapy; tissue engineering

Objective

To assess the natural rate of extracellular matrix turnover in normal human larynx throughout the lifespan.

Introduction

There are no data on the natural turnover rates of tissues in the human larynx; however, such information is desirable to better understand development and aging, the biological response to voice- and swallow-induced mechanical forces, and appropriate engineering of biomaterials. The carbon-14 (¹⁴C) bomb-pulse method is a sensitive dating technique that can be applied to measurement of cell and tissue turnover. ¹⁴C, a heavy radioisotope of carbon with a 5730-year half-life, exhibited a sharp increase in concentration in the earth's atmosphere in the 1950s and 1960s due to aboveground nuclear weapons testing, peaked in 1963, and then began to decline following the international Limited Test Ban Treaty. All living organisms incorporate ¹⁴C, either via photosynthesis or dietary intake, and the ¹⁴C concentration in a given biological material is reflective of atmospheric ¹⁴C at the time it was synthesized. Thus, ¹⁴C data can be used to accurately date tissues and estimate their turnover rate. Here, we measured ¹⁴C levels in five subregions of larynges from humans born before, during, and after the nuclear bomb pulse.

Methods

First, we refined a procedure to isolate and purify the core matrisome (structural ECM proteins) via depletion of glycans and cells from laryngeal tissue. Next, 14 adult human larynges were obtained at autopsy and microdissected to obtain 5 subregions: superficial lamina propria, vocal ligament (mid-membranous), macula flava, thyroid cartilage (mid-lamina), and arytenoid cartilage (hyaline core). Tissue samples were purified and subjected to accelerator mass spectrometry to measure the ratio of heavy (¹⁴C) to modern (¹³C) carbon; data were analyzed with respect to human date-of-birth and change in atmospheric ¹⁴C throughout the lifespan.

Results

^{14}C levels in superficial lamina propria, ligament, and macula flava yielded similar curves, with a peak corresponding to a 1948 birth date. Given that atmospheric ^{14}C peaked in 1963, this finding suggests that these laryngeal connective tissues may finish maturing around 15 years of age and have little turnover during adult life. ^{14}C data from thyroid and arytenoid cartilage exhibited a later peak; additionally, thyroid cartilage samples from humans born in the 1960s and 1970s exhibited a sharp drop in ^{14}C levels. This observation might be attributable to matrix remodeling during middle age, perhaps in concert with cartilage mineralization and ossification.

Conclusions

Matrix turnover within the larynx differs by region. Following development, the connective tissues appear to have little turnover during adult life; the thyroid cartilage appears to undergo remodeling during midlife.

3.12. Tissue Hysteresis and Relaxation, Phonation Onset, and Phonation Offset in The Context of the Surface Wave Model

Lewis P. Fulcher ¹ and Ronald C. Scherer ²

¹ Physics and Astronomy, Bowling Green State University, Bowling Green, OH 43403, USA

² Communication Sciences and Disorders, Bowling Green State University, Bowling Green, OH 43403, USA

Keywords: tissue properties; computer modeling; phonation onset; phonation offset

Objectives

To find a method that allows one to determine the relative importance of tissue hysteresis and viscous tissue damping in accounting for phonation thresholds. This work is an extension of earlier work¹ that considered only the effects of tissue hysteresis in determining these thresholds.

Introduction

Alipour and Vigmostad² made a number of measurements of the difference of the force required to stretch vocal fold tissue and the force this tissue produces when allowed to return to its original position. Further, they measured the area of the hysteresis loops for such a cyclic procedure and thus obtained a quantitative measure of the energy loss from hysteresis during the oscillation cycle. They also showed that the elastic parameters of the tissue decreased after a number of cycles, that is, after the tissue reached a preconditioned state.

Methods

The surface wave model provides a convenient framework for incorporating the effects of energy loss due to hysteresis and those due to viscous tissue damping. A careful examination of the equations of motion for the oscillating tissue allows one to obtain an analytic expression for the threshold pressure, which connects the elastic parameters used to describe tissue hysteresis, the parameter used to describe the tissue damping, the geometric properties of the vocal fold, and the threshold pressure. The result from the analytic expression may be checked with a numerical solution of the underlying dynamical equation.

Results

Chan and Titze³ measured threshold pressures for a vocal fold model consisting of a silicone membrane with biomaterials, such as hyaluronic acid and fibronectin, implanted under it. As an

independent variable, they chose the glottal half-width and collected data for 7 or 8 half-widths for each implant. This data set provides a good test of the results of our surface wave model calculations for each biomaterial implant, and in most cases the fits to the data are reasonable. Alipour and Vigmostad's measurements¹ also provide a means for estimating the effects of tissue relaxation. Using their data as a guide, it becomes possible to determine the elastic parameters appropriate for phonation offset, by adjusting one stiffness parameter. Thus, calculations for phonation offset are also compared with Chan and Titze data². Again, in most cases agreement of the calculations and measurements is reasonable.

Conclusions

The surface wave model provides a good context for deciding on the relative importance of tissue hysteresis and viscous tissue damping in explaining phonation thresholds. Adding viscous damping effects to the formalism allows one to explain the Chan and Titze experiments with hysteresis stiffness parameters that are in accord with the Alipour and Vigmostad measurements.

References

1. Fulcher, L.; Scherer, R. Hysteresis and Relaxation of Vocal Fold Tissue and the Difference between Phonation Onset and Offset. In Proceedings of the 10th International Conference on Voice Physiology and Biomechanics, Vina del Mar, Chile, 2016; p. 20
2. Alipour, F.; Vigmostad, S. Measurement of vocal fold elastic properties for continuum modeling. *J. Voice* **2012**, *26*, 816.e21–816.e29.
3. Chan, R.; Titze, I. Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics. *J. Acoust. Soc. Am.* **2006**, *119*, 2351–2362.

3.13. 3D Printed Scaffold Design for Vocal Fold Tissue Engineering Application

Anete Branco ¹, Peter Moua ², Amit J. Nimunkar ² and Susan L. Thibeault ¹

¹ Department of Surgery, University of Wisconsin-Madison, Madison, WI, USA

² Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI, USA

Keywords: vocal fold; extracellular matrix; tissue engineering; scaffolds; 3D printer

Objectives

To develop a three-dimensional (3D) printed scaffold that can be manufactured to match the biomechanical and anatomical properties of the vocal fold lamina propria.

Introduction

Vocal folds are delicate 3D tissues that provide phonatory function. The vocal fold lamina propria extracellular matrix is critical to vocal quality, but its poor healing ability and limitations of surgical repair have motivated tissue engineering based strategies to engineer living tissue replacement. To date tissue-engineering approaches in this arena have included hydrogel injections, tecoflex scaffolds, stem cell therapy and the use of decellularized tissue. However, these methods are far from being utilized clinically¹. Many of the vocal fold regeneration failures can be attributed to the inability to recapitulate the anatomical multilayered 3D lamina propria structure and the failure to withstand the in vivo biophysical forces² that the vocal fold endures.

Methods

Initially, we have created virtual models of the scaffold using specific software in order to produce a multilayered geometry design. For the prototype step of this ongoing research, we used a 3D printer to print thermoplastic polyurethane microscaffolds. Three specific percentage of infills (30, 50 and 70) and

pore sizes were printed oriented in 90 degrees from adjacent layers to form a basket-like architecture. Primary human vocal fold fibroblasts were cultured and encapsulated on hydrogel and injected into the 3D printed scaffolds. Cell morphology, viability and proliferation assays were performed.

Results

We have successfully microengineered 3D thermoplastic polyurethane printed scaffolds. The higher percentage of infills, the easier the cell-encapsulated hydrogel injection. Vocal fold fibroblasts presented intact spindle-shape morphology and size when attached on 50% infill scaffolds. Live cells spread and attached to the scaffold, with the majority of dead cells in the middle of the pores. 3D multilayered printed scaffold allowed vocal fold fibroblasts to be highly viable up to 96 h.

Conclusions

In this research, novel 3D printed scaffolds for vocal fold tissue engineering have been produced, characterized and tested in human vocal fold fibroblasts using partially physically cross-linked hydrogel. 3D printing has high potential for producing scaffolds with high uniformity and multilayered compositions from various materials that until now has been absent in the field of laryngology. Utilizing the present 3D printed scaffold model, we are able to investigate how varying pore geometry, infills and coating affect cellular behavior.

References

1. Li, L.; et al. Tissue engineering-based therapeutic strategies for vocal fold repair and regeneration. *Biomaterials* **2016**, *108*, 91–110.
2. Wrona, E.A.; et al. Extracellular matrix for vocal fold lamina propria replacement: A review. *Tissue Eng.* **2016**, *22*, 421–429.

3.14. *A Preliminary Study on Pharyngoesophageal Segment Vibration in Tracheoesophageal Speech by Means of a Collapsible Channel Model*

André M. C. Tourinho¹, Fernando H. T. Santos¹ and Andrey R. da Silva¹

Department of Mechanical Engineering, Federal University of Santa Catarina, Florianópolis, Santa Catarina, Brazil

Keywords: pharyngoesophageal segment, tracheoesophageal voice; collapsible tubes; collapsible channel

Objectives

To study how the tonicity of the pharyngoesophageal segment (PES) affects its vibration by means of a collapsible channel model.

Introduction

Total laryngectomy is commonly used to treat advanced-stage laryngeal cancer. With the removal of the larynx, the patient will lose the ability to speak. However, different techniques of voice rehabilitation are available to the laryngectomee, among which, the tracheoesophageal voice is the most widely used. In spite of its high success rate, some laryngectomees are not able to satisfactorily produce the tracheoesophageal voice. Such impediment is usually associated with hypertonicity or hypotonicity of the PES. Therefore, it would be desirable to quantitatively assess the influence of the tonicity of the PES on its vibration. In this study it is argued that the vibration of the PES shares several similarities with the problem of the flow-induced vibrations in collapsible tubes (Heil and Hazel, 2011). Given the complexity of the problem, a simplified collapsible channel model will be used to begin evaluating how tonicity might affect the vibration of the PES.

Methods

The model used is an adaptation of that proposed by Stewart (2017), where the governing equations of the flow and membrane are reduced to two partial differential equations by means of a von Kármán-Pohlhausen approximation. The tonicity of the PES is represented as an external pressure applied to the flexible membrane. Solutions of the coupled pair of partial differential equations are obtained by means of the finite difference method.

Results

The behavior of the system is assessed with regards to changes of the external pressure representing tonicity. The inlet parameters necessary for self-excited oscillations of the membrane are determined for different values of the external pressure, and the movement of the membrane is analyzed in each case.

Conclusions

The simplified model adopted provides useful information on the role played by the tonicity of the PES on its vibration. The limitations of the model in representing PES vibration are discussed, and possible solutions are presented.

Acknowledgments: Financial support by CAPES, CNPq and FINEP is gratefully acknowledged. We also thank Luiz Medina from CEPON for the helpful discussions.

References

1. Heil, M.; HAZEL, A.L. Fluid-Structure Interaction in Internal Physiological Flows. *Annu. Rev. Fluid Mech.* **2011**, *43*, 141–162.
2. Stewart, P.S. Instabilities in flexible channel flow with large external pressure. *J. Fluid Mech.* **2017**, *825*, 922–960.

3.15. *Application of Two Different Modalities for the Vibratory Characteristics in Vocal Fold Vibration of Vocal Cord Paralysis before and after Injection Laryngoplasty-Laryngeal Videostroboscopy and Two Dimensional Scanning Videokymography*

Eui-Suk Sung ¹, Soo-Geun Wang ², Byung-Joo Lee ², Han-Seul Na ², In-Ho Bae ¹ and Jin-Choon Lee ¹

¹ Department of Otolaryngology - Head and Neck Surgery, College of Medicine, Pusan National University, Yangsan, Korea

² Department of Otolaryngology - Head and Neck Surgery, College of Medicine, Pusan National University, Pusan, Korea

Keywords: two dimensional scanning videokymography; stroboscopy; vocal fold; vocal cord paralysis

Objectives

Our objective is to analyze of mucosal wave of vocal folds in vocal cord paralysis (VCP) before and after injection laryngoplasty using laryngeal videostroboscopy (LVS) two-dimensional (2D) scanning videokymography (VKG).

Introduction

The voice is produced through vibration of the vocal folds during the exhalation of airflow. Since mucous membrane vibration is an important factor in sound quality, various tools have been developed to evaluate vocal fold vibration. So it is important to analyze the vibration pattern of the vocal fold in phonation. LVS is widely used to study the vibration of the vocal folds in clinical practice. We analyzed vocal fold vibration using 2D scanning VKG in normal person without vocal fold disease in previous article. It is possible to evaluate the whole mucosal wave pattern of the entire vocal folds mucous

membrane both laryngeal videostroboscopy (LVS) and 2D scanning VKG. There was no comparative study of mucosal wave pattern between LVS and 2D scanning VKG in VCP so far.

Methods

Ten patients with unilateral VCP participated in the study, and LVS and 2D VKG were used to assess the vibratory pattern of vocal folds before and after injection laryngoplasty. And the qualitative analysis (glottal gap, amplitude difference, phase difference) using a 100 mm visual analog scale was performed in two modalities, and quantitative analysis was performed in 2D VKG (contact quotient, CQ; phase symmetric index, PSI; amplitude symmetric index, ASI) and LVS (CQ, PSI, and ASI in glottal area waveform).

Results

The results of qualitative (glottal gap and amplitude difference except phase difference) and quantitative (CQ and ASI except PSI) analysis showed significant improvement in both two modality after injection laryngoplasty. And, both quantitative and qualitative analysis showed a significant correlation between the two modality. Two patients with severe dysphonia were unable to be assessed by LVS due to periodicity disruption or insufficient sustained phonation, but mucosal wave could be observed in 2D VKG.

Conclusions

2D scanning VKG can support the analysis of the dynamic status such as degree of atrophy and elasticity of the vocal fold even at aperiodic voice phonation in unilateral VCP before and after injection laryngoplasty. It is useful for assessing the quality of voice by recognizing the vocal fold mucosal vibration pattern at one time before and after injection laryngoplasty in VCP.

References

1. Anastaplo, S.; Karnell, M.P. Synchronized videostroboscopic and electroglottographic examination of glottal opening. *J. Acoust. Soc. Am.* **1988**, *83*, 1883–1890.
2. Trapp TK, Berke GS. Photoelectric measurement of laryngeal paralyses correlated with videostroboscopy. *Laryngoscope* **1988**, *98*, 486–492.
3. Wang, S.G.; Park, H.J.; Lee, B.J.; et al. A new videokymography system for evaluation of the vibration pattern of entire vocal folds. *Auris Nasus Larynx* **2016**, *43*, 315–321.
4. Wang, S.G.; Park, H.J.; Cho, J.K.; et al. The First Application of the Two-Dimensional Scanning Videokymography in Excised Canine Larynx Model. *J. Voice* **2016**, *30*, 1–4.
5. Kim, G.H.; Wang, S.G.; Lee, B.J.; et al. Real-time dual visualization of two different modalities for the evaluation of vocal fold vibration—Laryngeal videoendoscopy and 2D scanning videokymography: Preliminary report. *Auris Nasus Larynx* **2017**, *44*, 174–181.
6. Wang, S.G.; Lee, B.J.; Lee, J.C.; Lim, Y.S.; Park, Y.M.; Park, H.J.; Shin, B.J. Development of Two-Dimensional Scanning Videokymography for Analysis of Vocal Fold Vibration. *Korean J. Laryngol. Phoniatr. Logop.* **2013**, *24*, 107–111.

3.16. BiOChemical Alterations in Vocal Fold Tissue in the Production of Decellularized Extracellular Matrix Hydrogels

M. Brown¹ and N.Y.K. Li-Jessen^{1,2,3}

¹ Biological and Biomedical Engineering

² School of Communication Sciences and Disorders

³ Otolaryngology—Head and Neck Surgery, McGill University, Montréal, QC, Canada

Keywords: vocal folds; tissue engineering; decellularized extracellular matrix; hydrogels

Objectives

To evaluate the morphology, viability and proliferation of human vocal fold fibroblasts (HVFF) encapsulated in hydrogels that consisted of vocal fold (VF) decellularized extracellular matrix (dECM) microparticles.

Introduction

Biomaterials produced from dECM have been suggested for VF tissue engineering due to their unique prospective for replicating the composition of native VF-ECM. Previous work focused on development of whole VF-dECM scaffolds as surgical implants. However, surgical procedures may carry risks of causing further tissue alterations to already damaged VF¹. Injectable hydrogels composed of dECM microparticles have shown potential as alternatives for soft tissue engineering. In previous studies using dECM hydrogels for VF tissue engineering, small intestinal submucosa was used³. However, the efficacy of dECM hydrogels has been suggested to be tissue-specific, with favorable outcomes for dECM derived from the same tissue type². Changes in ECM composition can occur during decellularization, with a demonstrable impact on dECM scaffold outcomes, and it was therefore important to quantify these alterations⁴. We hypothesized that dECM would facilitate HVFF proliferation and functional morphology over a collagen-HA (CHA) control, which has previously been investigated for VF tissue engineering.

Methods

Porcine VF were dissected from larynges and subjected to a decellularization protocol at 37 °C under constant agitation, consisting of: (1) 2 h in 4% sodium deoxycholate, (2) 24 h in 0.75 mg/mL deoxyribonuclease I and 0.1 mg/mL ribonuclease I, (3) 0.1% peracetic acid, (4) three 15 min washes with deionized water (ddH₂O), and (5) repetition of 2-4. Nucleic acid removal was quantified by the Quant-iT™ Pico-Green™ dsDNA Assay (Invitrogen). Homogenization was then conducted on dECM in a bead mill, using 2.8 mm ceramic beads in ddH₂O to produce dECM microparticles. Assays were performed on native VF, whole dECM, and dECM microparticles to determine the effect of decellularization and homogenization on ECM composition. Collagen, elastin, and hyaluronic acid were quantified using the Sircol® Total Collagen Assay (Biocolor), Fastin™ Elastin Assay (Biocolor), and carbazole assay, respectively. Protein results were normalized using the Pierce BCA Total Protein Assay. To produce hydrogels, dECM microparticles were solubilized in 3 mg/mL pepsin and 0.1 M HCl for 48 h at RT, neutralized with sodium hydroxide, lyophilized, and resuspended at 3% dECM in Dulbecco's Modified Eagle Medium (DMEM). The pre-gel solution was mixed with an equal volume of DMEM containing 1.6 million HVFF/mL and incubated for 90 min to produce 1.5% dECM hydrogels. Cellular viability and proliferation was monitored over 7 days using a Live/Dead Viability/Cytotoxicity Staining Kit (Invitrogen). A CHA hydrogel (0.5% collagen, 0.5% HA) was used as a positive control.

Results

The decellularization protocol reduced DNA content by $99.6 \pm 0.1\%$ ($p < 0.05$), reliably $> 95\%$ removal, a measure for successful decellularization¹. Collagen content, 58.1 ± 15.0 (w/w) % of total protein in native VF, was not reduced by decellularization ($69.0 \pm 13.1\%$, $p > 0.05$) or homogenization ($65.3 \pm 12.5\%$, $p > 0.05$). While elastin content, $6.94 \pm 0.58\%$ of total protein in native VF, was not decreased by decellularization ($6.38 \pm 0.63\%$, $p > 0.05$), it was by homogenization ($4.55 \pm 0.48\%$, $p < 0.05$). Hyaluronic acid content was reduced by $75.00 \pm 4.00\%$ ($p < 0.05$) by decellularization, and $86.11 \pm 3.29\%$ ($p < 0.05$) from native VF after homogenization. When encapsulated in CHA hydrogels, HVFF demonstrated elongated morphology typical of fibroblasts, with minor but significant cell death between day 1 and 3 ($p < 0.05$), and a stable population between days 3 and 7 ($p > 0.05$). HVFF encapsulated in 1.5% dECM hydrogels instead exhibited spherical morphology. Although the population remained stable between day 1 and 7 ($p > 0.05$), HVFF displayed signs of dying.

Conclusions

Elastin and hyaluronic acid content was significantly reduced after homogenization. These changes are a possible contributing factor to the lack of elongated HVFF in the 1.5% dECM hydrogel, unlike the CHA control. Further optimization of the decellularization and homogenization protocol is suggested for use of a VF-dECM hydrogel in VF tissue engineering.

Acknowledgments: Financial support was received from the Natural Sciences and Engineering Research Council of Canada RGPIN-2018-03843 and the National Institute on Deafness and Other Communication Disorders R01 DC005788.

References

1. Wrona, E.; et al. *Tissue Eng. Part B Rev.* **2016**, *22*, 421–429.
2. Gibson, M.; et al. *BioMed Res. Int.* **2014**.
3. Huang, D.; et al. *BioMed Res. Int.* **2016**.
4. Keane, T.; et al. *Methods* **2015**, *84*, 25–34.

4. Session 2

4.1. Vocal Fold Visco-Hyperelastic Properties: Characterization and Multiscale Modeling upon Finite Strains

Alberto Terzolo ¹, Thibaud Cochereau ^{1,2}, Lucie Bailly ¹, Laurent Orgéas ¹ and Nathalie Henrich Bernardoni ²

¹ Univ. Grenoble Alpes, CNRS, Grenoble INP, 3SR, Grenoble F-38000, France

² Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble F-38000, France

Keywords: vocal folds fibrous microstructure; multiscale modeling; viscoelasticity; finite strains

Objectives

Analytical models currently simulating vocal folds mechanical behavior at the fiber scale are very promising to better understand their remarkable macroscale performances. In continuity with these developments, this study aims at proposing a 3D multiscale biomechanical model of the vocal tissue able to properly predict its time-dependent properties upon finite strains.

Introduction

Since the 2010s, the investigation and modeling of human vocal-fold fibrous microstructure has opened a new insight into voice biomechanics [1–3]. So far, most of the multiscale formulations have focused on the *lamina propria* anisotropic and hyperelastic properties, predicting the unfolding of crimped collagen fibers and, consequently, the tissue's J-shaped response under large-strains uniaxial tensile loading. A few authors have proposed a 1D viscoelastic model of *lamina propria* in tension, to better investigate the regulation of key phonatory parameters such as the acoustic fundamental frequency [4]. *In vivo*, however, vocal folds are subjected to numerous complex and coupled 3D mechanical solicitations (tension, compression and shear), experienced upon finite strains and at various strain rates. Therefore, this work presents: (i) a set of data completing the available characterization of vocal-fold viscoelasticity under multiple mechanical loadings [4,5]; (ii) a 3D micro-mechanical model of the tissue layers, able to predict their visco-hyperelastic and anisotropic properties.

Methods

Histological descriptors of the fibrous networks within *lamina propria* and *vocalis* layers were firstly collected from 2D/3D human vocal folds images [3]. Available measurements of their time-dependent

response [4,5] were then completed by mechanical tests performed on excised samples subjected to finite strain loadings (cyclic traction, compression, shear, relaxation). These tests were conducted by means of a dedicated micro-press coupled with an optical tracking of the deformed samples. A 3D homogenized micro-mechanical model was also developed, to predict the macroscale anisotropic and hyperelastic properties of the tissues [6]. The formulation was furtherly enriched by adding a Maxwell viscoelastic stress contribution to the hyperelastic stress of each collagen fiber. The model is therefore able to simulate a large variety of loading conditions at various rates.

Results

The model microscale parameters are optimized within a range provided by 2D/3D histo-mechanical data for both *lamina propria* and *vocalis* layers. For each layer, the resulting macroscale predictions reproduce typical stress-strain responses under various loading conditions, from small to finite strains. The influence of the strain rate and strain amplitude on their mechanical behavior is also highlighted. Predictions are in each case critically discussed with regard to the evolution of the collagen and muscular fibrous microstructures. Particular attention is paid to the coupling between fibers rotation and their own deformability as well as to fibers steric interactions, notably critical under compression loadings.

Conclusions

This study points out the finite strain and time-dependent mechanical response of the vocal-fold sublayers under various physiological loadings, which are barely studied experimentally or theoretically so far. These advances are expected to provide a better understanding of the link between the micromechanics of vibrating tissues and their macroscale performances.

Acknowledgments: This work was supported by the ANR MICROVOICE No. ANR-17-CE19-0015-01) and the LabEx Tec 21 (Investissements d’Avenir—grant agreement no. ANR-11-LABX-0030).

References

1. Miri, A.K.; Heris, H.K.; Tripathy, U.; Wiseman, P.W.; Mongeau, L. *Acta Biomater* **2013**, *9*, 7957–7967.
2. Kelleher, J.E.; Siegmund, T.; Du, M.; Naseri, E.; Chan, R.W. *J. Acoust. Soc. Am.* **2013**, *133*, 1625–1636.
3. Bailly, L.; Cochereau, T.; Orgéas, L.; Henrich Bernardoni, N.; Rolland du Roscoat, S.; McLeer-Florin, A.; et al. *Sci. Rep.* **2018**, *8*, 14.
4. Zhang, K.; Siegmund, T.; Chan, R.W. *J. Mech. Behav. Biomed. Mater.* **2008**, 93–104.
5. Chan, R.W. *J. Rheol.* **2018**, *62*, 695–712.
6. Cochereau, T. Ph.D. Thesis, Université Grenoble Alpes, Grenoble, France, 2018.

4.2. Investigation of Constrains on Vocal fold Viscoelastic Properties Using an Inverse Mapping Approach

Ted Mau ¹, Anil Palaparthi ² and Ingo R. Titze ²

¹ Department of Otolaryngology-Head and Neck Surgery, UT Southwestern Medical Center, Dallas, TX, USA

² The National Center for Voice and Speech, University of Utah, Salt Lake City, UT, USA

Keywords: voice simulation; multiobjective optimization; inverse mapping; vocal fold morphology

Objectives

To determine if a single set of acoustic target (defined by a particular range of target F_0 and SPL) can be delivered by vocal fold morphologies that differ in the viscoelastic properties of the tissue layers.

Introduction

Any given vocal acoustic property (e.g., F_0) has a dynamic dependence on multiple physiologic input and vocal fold morphologic parameters. Many simulation approaches have investigated how input parameters determine the vocal output. Here we examine the inverse problem: Given a desired acoustic target, what are possible morphologies that can deliver such a target? This inverse mapping problem is addressed using a previously reported technique that integrates voice simulation with multiobjective optimization. Specifically, we test the hypothesis that the same acoustic output can be achieved with different vocal fold morphologies.

Methods

A 3-layered fiber-gel vocal fold model consisting of the superficial layer of lamina propria (SLLP), vocal ligament, and thyroarytenoid muscle was used (Titze et al., 2017). Self-sustained oscillation was produced with modified Bernoulli airflow. Voice simulation was coupled to the NSGA2 multiobjective genetic algorithm (Palaparathi et al., 2014), with target objective functions matching a male voice with $F_0 = 120 \pm 30$ Hz and $SPL = 70 \pm 10$ dB at 30 cm. Three parameters were allowed to vary: lung pressure (P_L), transverse shear modulus of the SLLP (μ_{SLLP}), and longitudinal shear modulus of the vocal ligament (μ'_{lig}). μ_{SLLP} and μ'_{lig} were allowed to vary over two orders of magnitude. Each simulation was run to produce 200 ms of voice signal. For the genetic algorithm, 200 generations of 20 solutions (simulations) each were produced to arrive at optimized values of P_L , μ_{SLLP} and μ'_{lig} to deliver the target F_0 and SPL . In Experiment 1, the longitudinal shear modulus of the SLLP (μ'_{SLLP}) was fixed at 1 kPa. In Experiment 2, μ'_{SLLP} was fixed at 10 kPa. Each experiment was run twice. The outcome measures were the optimized μ_{SLLP} and μ'_{lig} in each experiment.

Results

With μ'_{SLLP} fixed at 1 kPa, the optimal μ_{SLLP} was around 2 kPa and optimal μ'_{lig} was 26–32 kPa. With μ'_{SLLP} fixed at 10 kPa, the optimal μ_{SLLP} was 2.2 kPa, but optimal μ'_{lig} was around 3–10 kPa. In other words, to achieve the same F_0 and SPL targets, μ_{SLLP} remained essentially unchanged even as μ'_{SLLP} changed by a factor of 10. This particular set of acoustic targets had little dependence on μ'_{lig} , which ranged from 3–32 kPa.

Conclusions

These results are consistent with the notion that phonation in modal register at comfortable loudness is largely determined by the transverse shear property of the SLLP, which must remain within a narrow range for a given acoustic target. The longitudinal shear properties of the SLLP and the ligament are less constrained, with a reciprocal relationship between the two. These findings may have implications for the mechanical design of bioengineered lamina propria replacement constructs. This preliminary study also illustrates the utility of the optimized simulation approach in exploring alternative vocal fold morphologies.

References

1. Titze, I.R.; Alipour, F.; Blake, D.; Palaparathi, A. Comparison of a fiber-gel finite element model of vocal fold vibration to a transversely isotropic stiffness model. *J. Acoust. Soc. Am.* **2017**, *142*, 1376–1383.
2. Palaparathi, A.; Riede, T.; Titze, I.R. Combining multiobjective optimization and cluster analysis to study vocal fold functional morphology. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 2199–2208.

4.3. Vocal Fold Contact Pressure in a Three-Dimensional Body-Cover Phonation Model

Zhaoyan Zhang

Department of Head and Neck Surgery, University of California, Los Angeles, Los Angeles, CA, USA

Keywords: vocal fold contact pressure; vocal fold biomechanics; vocal tract configuration; vocal fold injury

The objective of this study is to identify vocal fold geometric and mechanical properties that can be manipulated to reduce contact pressure between the vocal folds and minimize vocal fold injury during phonation. Using a three-dimensional computational model of phonation, parametric simulations are performed with co-variations in vocal fold geometry and stiffness, with and without a vocal tract. For each simulation, the peak contact pressure and peak contact area are calculated. The results show that the subglottal pressure and transverse stiffness have the most consistent and dominant effect on the peak contact pressure, which decreases with decreasing subglottal pressure or increasing transverse stiffness of the vocal fold. In most vocal fold conditions investigated, the peak contact pressure can also be reduced by decreasing the vertical thickness of vocal fold medial surface. Changes in vocal fold stiffness along the anterior-posterior direction have the least and inconsistent effect on the peak contact pressure. The presence of a vocal tract generally increases the peak contact pressure, with the increase significantly larger for the /a/ vocal tract than the /i/ vocal tract, suggesting the potential usefulness of a constricted vocal tract configuration in voice therapy. While a low degree of vocal fold approximation significantly reduces vocal fold contact pressure, strong interaction among the degree of vocal fold approximation, vocal fold stiffness, and vocal tract configuration is observed for conditions of moderate and tight vocal fold approximation.

Acknowledgments: Research was supported from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health.

4.4. Numerical Study of the Influence of Vascular Morphology on the Evolution of Vortical Flow Structures through the Blood-Feeding Arteries of the Human Vocal Folds: Application to Drug Delivery for Laryngeal Cancer

Mehdi Shamshiri, Rosaire Mongrain and Luc Mongeau

Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

Keywords: computational fluid dynamics; blood flow modelling; vortical structures; drug delivery

Objectives

Computational fluid dynamics methods were used to non-invasively investigate the impact of vascular morphology on the evolution of vortical flow structures through the major blood-feeding arteries of the human vocal folds, and to explore the potential impact of such complex structures on anticancer drug transport to the Larynx.

Introduction

Head and neck cancer represents the sixth most-common type of cancer worldwide, with the larynx and oral cavity being the most common subsites. The anatomy and morphology of the vasculature irrigating the vocal folds varies significantly between individuals, with distinct primary and secondary flow structures through blood-feeding arteries. Little is known about the formation of such blood flow structures and their potential impact on drug delivery for laryngeal cancer patients.

Methods

Three-dimensional (3D) computational models of the blood-supporting vasculature of the human vocal folds were created. Computed tomographic angiography (CTA) data available in the literature, defining the size and the morphology of the arterial branches, was used to create patient-specific

computer models. Blood flow was assumed to be governed by the Navier-Stokes equations of continuity and conservation of momentum with consideration of the non-Newtonian Quemada rheological behavior. The vessel walls were assumed to be rigid and impermeable. The no-slip velocity condition was applied at the solid wall boundaries. Representative values of the cardiac input for one cardiac cycle were prescribed over the inflow boundary. In a verification study, the numerical model was found to yield velocity predictions in good agreement with available experimental data for blood flow through an idealized curved artery. A variety of common arterial morphologies were then considered, and the simulations were repeated for different values of the Reynolds and Dean numbers. The ANSYS software products (Ansys Inc., Canonsburg, PA) were used for the simulations.

Results

For a typical C-shape blood-feeding arterial branch of the vocal folds, at low Dean Numbers (De), steady blood flow featured two symmetric, counter-rotating vortices, namely Dean vortices. As the De is sufficiently increased, multiple pairs of secondary flow vortices (up to three symmetrical vortices, or six-vortex patterns) form under steady inflow conditions. In the next step, the effect of different arterial morphologies and physiological conditions on hemodynamic flow of the human vocal folds, and the potential impact of such morphologies on administration of drug to the Larynx was investigated.

Conclusions

Anatomical factors have considerable effect on the evolution of the vortical flow structures associated with local hemodynamics of the Larynx and Hypopharynx. Targeted drug administration and local drug mixing can be remarkably influenced by the formation and evolution of such complex flow structures. We anticipate this work will significantly contribute to the ongoing research on laryngeal and hypopharyngeal cancer and the concept of organ preservation.

4.5. Development of a High-Fidelity Voice Simulator—From Muscle Contraction to Running Speech

Biao Geng¹, Xudong Zheng¹, Ngoc Hong Pham² and Qian Xue¹

¹ Department of Mechanical Engineering, University of Maine, Orono, ME, USA

² Sarkeys Energy Center, the University of Oklahoma, Norman, OK, USA

Keywords: realistic model; finite element; vocal fold posturing; speech simulation

Objectives

Production of voice in a wide range requires complex coordination between the neuromuscular control of the laryngeal muscles and the respiratory system. However, the role of each of the muscles as well as the internal status of the vocal folds in the full vocal range remain elusive. While experimental studies provide valuable knowledge on various aspects, the observations are generally limited by the inaccessibility of the vocal folds. This study aims to develop a realistic finite element (FE) larynx model that incorporates muscle and joint mechanisms to dynamically posture the vocal fold. This model is then coupled with a flow solver and an acoustic solver to simulate running speech.

Methods

The laryngeal structures, including all the intrinsic muscles and cartilages, are constructed from MRI images. A previously proposed 1-D muscle activation model is integrated into a 3-D finite element code to model the contractile behaviors of the muscles. The behaviors of the major joints are modeled using multi-point constraints. The posturing model is then coupled with a 1-D Bernoulli flow model and a wave-reflection acoustic model for speech simulation. The activation of the muscles is simultaneously and independently controlled to produce simple English phrasal utterance.

Results

The high-fidelity larynx model developed in this study is capable of basic posturing tasks with given activation commands, including adduction, abduction and stretching. Vocal fold strain at different activation levels of a variety of muscle combinations are systematically validated against in-vivo experimental results reported in the literature. Control commands are inversely determined to coordinate all the muscles to produce pitch change and aspiration, two of the key features in running speech. Simulations using the model generate recognizable English phrases.

Conclusions

A high-fidelity voice simulator is developed with various areas of application. By inversely determining the required neuromuscular command input for different voice types, simulations using this model could help reveal the corresponding control mechanisms and vibrational characteristics. The model can also be used to simulate pathological cases where one side of the vocal folds is inactive to a certain extent. Surgical planning simulation is also possible. For example, medialization can be simulated using numerical contact models and the voice quality can be checked for different insertion configurations, based on which an optimal solution can be selected.

4.6. SpEAR: A Speech Database for the Advancement of Intra-Aural Wearable Technology

Rachel E. Bouserhal ^{1,2} and Jérémie Voix ^{1,2}

¹ École de technologie supérieure, Montréal, QC, Canada

² Centre for Interdisciplinary Research in Music Media and Technology, Montréal, QC, Canada

Keywords: speech; lombard effect; database; wearables; in-ear microphone

Introduction

The use of in-ear microphones with intelligent intra-aural devices is growing. This is due to the rise in hearable technology and the advantages to communication offered by using an in-ear microphone. In noisy conditions, the in-ear microphone captures a speech signal with a relatively high signal-to-noise ratio since it is usually placed past the passive attenuation of the earplug. However, speech captured inside the occluded ear is limited in its frequency bandwidth and has an amplified low frequency content. In addition, occluding the ear canal with an intra-aural device affects speech production, which could have detrimental effects on speech processing algorithms. An in-ear microphone speech database in noise and in quiet is therefore essential to the advancement of intra-aural technologies utilizing an in-ear microphone. Yet, to the authors' knowledge no such database exists. This work presents a Speech-in-EAR (SpEAR) database in various conditions of the audio-phonation loop.

Methods

Twenty-four participants (11 in English and 13 in French) took part in this study. Speech was collected in an audiometric booth using a reference microphone placed in front of the mouth as well as an intra-aural device equipped with in-ear microphones, outer-ear microphones, and miniature loudspeakers. Hearing-in-noise test sentences in French and English were read in four conditions: (1) quiet open-ear, (2) quiet occluded, (3) occluded in noise (ambient), and (4) occluded in noise (regenerated inside the ear). The first and second condition serve as a baseline reference for comparison when understanding the effects on speech production caused by occluding the ear in quiet and in noise. For the third condition, factory noise is played at 95 dBA in the room while the participants read each sentence. This condition is a realistic condition meant to aid researchers and developers working on denoising algorithms for intra-aural wearable devices. In the fourth condition noise is regenerated inside the occluded ear, causing the Lombard effect to be triggered but leaving the

outer-ear microphones and the microphone in front of the mouth free of noise. In this case, the changes invoked by the Lombard effect and their interaction with occluding the ear can be studied without the nuisance of noise.

Results

Preliminary acoustical analysis showed that the presence of noise and occlusion have a significant ($p < 0.001$) effect on the speech level of participants. Once occluded and in quiet, on average participants raise their speech level by 2.6 dBA compared to the open-ear condition. At the introduction of 95 dBA of factory noise, the average speech level increases on average by 6.5 dB. Analysis showed that males and females do not speak at different average speech levels and that language spoken has no significant effect on speech level over all conditions.

Conclusions

SpEAR aims to aid researchers and developers working with intra-aural wearable technology to develop speech algorithms for adverse realistic conditions. It is meant to respond to a lack of in-ear speech database and can deepen the understanding of changes in speech production caused by noise and occluding the ear.

Acknowledgments: The authors would like to acknowledge the funding received from the Centre for Interdisciplinary Research in Music Media and Technology, the Fonds de recherche du Québec-Nature et technologies, the Natural Sciences and Engineering Research Council of Canada, and the NSERC-EERS Industrial Research Chair in In-Ear Technologies.

4.7. High Performance Simulation and Visualization of 3D Vocal Fold Agent-Based Model

Nuttiya Seekhao¹, Grace Yu², Samson Yuen⁴, Joseph Jaja¹, Luc Mongeau³ and Nicole Y.K. Li-Jessen⁴

¹ Department of Electrical and Computer Engineering, University of Maryland, College Park, VA, USA

² Department of Physiology, McGill University, Montreal, QC, Canada

³ Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

⁴ School of Communication Sciences and Disorders, McGill University, Montreal, QC, Canada

Keywords: vocal folds; computer modeling; 3d agent-based modeling; data visualization

Objectives

Personalized or precision medicine remains an open challenge in voice care. One biggest challenge is the complexity of vocal disease mechanism and unpredictable patient response to voice treatments. At present, clinicians have to make the best clinical decision based on their own experience. The overarching goal of this research is to harness a computational approach to forecast individual response to voice treatments and tailor the treatment to individual needs.

Introduction

Computational medicine is a growing field in personalized medicine. Computational models with high fidelity are necessary to accurately mimic complex diseases. At the same time, high fidelity models often require expensive computational resources and generate large amounts of numerical data. To address these challenges, we developed a unique parallel simulation platform that takes advantage of both multicore CPUs and GPUs to enhance the performance of large multi-scale models [1]. We also developed a comprehensive data visualization protocol to make the data analytic more intuitive and straightforward.

Methods

We present a highly-interactive remote simulation and visualization framework for vocal fold (VF) agent-based modeling (ABM) [1,2]. The physiologically representative human VF ABM consists of more than 15 million mobile biological cells including neutrophils, macrophages and fibroblasts. The model maintained and generated 1.7 billion signaling and extracellular matrix (ECM) protein data points in each iteration. The VF ABM employs high-performance computing (HPC) techniques to optimize its performance by concurrently utilizing multi-core CPU and multiple GPUs. Multiple GPUs were used to visualize protein contents using ray-casting to perform direct volume rendering on 3D time-varying output data.

Results

The HPC version of VF ABM achieved a speedup of 35 times over the optimized sequential version. Further, data footprint and data transfer reduction techniques were used to achieve real-time visualization. This fast visualization allows the users to observe the evolution of the signaling and ECM protein distribution with average framerate of 42.8 fps.

Conclusions

The optimizations developed for HPC VF ABM framework have made simulations of large high-fidelity models feasible. A five-day long biological event of a full-size VF surgical repair would take less than 30 min to simulate using our framework. This computing performance is necessary to perform a full-scale sensitivity analysis and model calibration in a reasonable amount of time. Furthermore, by coupling with the aforesaid interactive visualization tool, users can explore the simulated objects such as cell-ECM interactions in real-time. The real-time data exploration can lead to better understanding of the underlying mechanisms of the model. Additionally, real-time visualization also lets the user steer the course of the computation. Computational steering can save the modeler time and resources while debugging and verifying the model. The performance and functionality of the HPC VF ABM framework accommodates modelers in developing predictive tools applicable to in silico clinical trials. This development can potentially lead to faster-than-real-time predictive tools with the ability to design optimal personalized treatments.

Acknowledgments: Canadian Institutes of Health Research Project Grants 388583 (Li-Jessen), Natural Sciences and Engineering Research Council of Canada RGPIN-2018-03843 (Li-Jessen), Compute Canada (Li-Jessen), National Institute of Deafness R03DC012112 (Li-Jessen), Communication Disorder of the National Institutes R01DC005788 (Mongeau).

References

1. Seekhao, N.; Shung, C.; JaJa, J.; Mongeau, L.; Li-Jessen, N.Y. High-performance agent-based modeling applied to vocal fold inflammation and repair. *Front. Physiol.* **2018**, *9*, 304.
2. Seekhao, N.; JaJa, J.; Mongeau, L.; Li-Jessen, N.Y. In situ visualization for 3d agent-based vocal fold inflammation and repair simulation. *Supercomput. Front. Innov.* **2017**, *4*, 68.

5. Poster Session 2

5.1. Development, Validation and Analysis of Numerical Larynx Models with Regard to Computational Costs

S. Kniesburges¹, Hossein Sadeghi¹, Sebastian Falk¹, Manfred Kaltenbacher² and Michael Döllinger¹

¹ Div. of Phoniatrics & Pediatric Audiology, Dep. of Otorhinolaryngology, University Hospital Erlangen

² Institute of Mechanis and Mechatronics, TU Wien, Austria

Keywords: phonation; computer modelling; CFD; computational costs

Introduction

The clinical application of computational larynx models would commence new strategies for planning and controlling subject specific therapeutic concepts for the treatment of voice disorders. Besides the visualization and segmentation of the relevant anatomical dimensions in the larynx, the initialization, the execution and the evaluation of the simulations demand large time intervals and computational hardware resources especially for an accurate simulation of the laryngeal aerodynamics [1].

Objectives

The first aim of the current work is to evaluate the computational costs that are necessary to simulate the laryngeal aerodynamics. In a second step, we analyze strategies for reducing the size of the numerical model with regard to the validity of the simulation results. The goal is to minimize the wall time and hardware resources for a future clinical application.

Methods

Large eddy simulations of the laryngeal flow have been performed with externally impressed vocal fold oscillations. The geometry of the simulation model corresponds to an experimental synthetic larynx model that includes silicone M5 models of the vocal folds [2,3]. The vocal folds motion and the inlet/outlet boundary conditions were adapted from the experimental model. A mesh independence study was performed to find the appropriate spatial resolution of the mesh and time step size for performing a large eddy simulation. After the validation of this base model (BM) with experimental data, we gradually reduced the mesh resolution and adjusted the time step size generating three additional models M1 to M3. In the following, we measured the wall time for one oscillation cycle and compared the aerodynamic parameters with the model BM.

Results

The model BM with the highest resolution consists of 2.4 million control volumes (CV) and a time step size of 1.0×10^{-6} s. Based on pressure data acquired in the experimental model [4], the average L^2 error of the numerical results was 0.256. The flow velocity field in the mid coronal plane showed also a high consistency with the phase resolved flow field measured by particle image velocimetry [5]. By reducing the resolution, the number of CVs could be decreased to 1.1 million with a time step of 1.36×10^{-6} s for the model M3. The comparison between BM and the three models M1-M3 showed only small deviations. The mean L^2 error is between 0.012 and 0.163 computed with regard to the flow rate, the static pressure and the flow velocity. Thus, all models with reduced resolution produced valid results compared to BM. The wall time for one oscillation cycle could be reduced by 75% from 79 h for BM to 19 h for M3. By further optimizing the parallel resources of our high performance computer (70 physical cores), the wall time could be further decreased to close to 7h for simulating one oscillation cycle in M3.

Conclusions

The results show that the wall time for valid laryngeal flow simulations can be decreased to values relevant for a clinical application. Moreover, the results also show further potential in reducing the model size for further speed-up.

Acknowledgments: This work was funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) in the framework of Project No. DFG DO 1247/10-1 and the Austrian Research Council (FWF) under No. I 3702.

References

1. Sadeghi, H.; Kniesburges, S.; Kaltenbacher, M.; Schützenberger, A.; Döllinger, M. Computational models of laryngeal aerodynamics: Potentials and numerical costs. *J. Voice* **2018**, in press.
2. Scherer, R.; Shinwari, D.; Witt, K.D.; Zhang, C.; Kucinski, B.; Afjeh, A. Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. *J. Acoust. Soc. Am.* **2001**, *109*, 1616–1630.
3. Thomson, S.; Mongeau, L.; Frankel, S. Aerodynamic transfer of energy to the vocal folds. *J. Acoust. Soc. Am.* **2005**, *118*, 1689–1700.
4. Kniesburges, S.; Hesselmann, C.; Becker, S.; Schlücker, E.; Döllinger, M. Influence of vortical flow structures on the glottal jet location in the supraglottal region. *J. Voice* **2013**, *27*, 531–544.
5. Lodermeier, A.; Becker, S.; Döllinger, M.; Kniesburges, S. Phase-locked flow field analysis in a synthetic human larynx model. *Exp. Fluids* **2015**, *56*, 77.1–77.13.

5.2. Agent-Based Model of Hyaluronic Acid-Gelatin Scaffold for Vocal Fold Tissue Engineering

Grace Yu ¹, Nuttitiya Seekhao ², Caroline Shung ³, Luc Mongeau ³ and Nicole Y. K. Li-Jessen ⁴

¹ Department of Physiology, McGill University, Montreal, QC, Canada

² Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA

³ Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

⁴ School of Communication Sciences and Disorders, McGill University, Montreal, QC, Canada

Keywords: voice; tissue engineering; computer modelling

Objectives

The primary objective of this study was to develop an agent-based computational model for the design of scaffold biomaterial for vocal fold tissue engineering.

Introduction

Biomaterial scaffolds have been engineered to support regeneration of defected or diseased vocal folds for the restoration of voice quality. During the design and evaluation process, it is necessary to identify scaffold compositions that are optimal to tissue-specific anatomical and mechanical requirements. Costly in vitro and in vivo experiments have been used to optimize the large number of composition parameters, such as porosity and compliance. In this study, stochastic-based computer simulations, namely agent-based models (ABMs), were used to narrow down the parameter space with reduced cost, time, and labour. ABMs have been used to accelerate engineering of biomaterials for bone [1], cartilage [2], blood vessels [3], and generic tissues [4], but not for vocal folds. Moreover, these models are limited by specificity, scalability, capacity for personalization, and accuracy. As such, there is a need to explore in silico methods to aid in designing biomaterials for vocal folds.

Methods

Hyaluronan-gelatin (HA-Gtn) hydrogels were used as a test case for the biomaterial ABM simulation. Chemical and cellular compositions of the biomaterial, such as HA-Gtn ratios, were used as inputs to the model. Agent rules were derived from known relationships in structure and function found in literature. For example, an increased total protein concentration increased the elastic modulus of the simulated scaffold, which decreased vocal fold fibroblast migration speed. The model outputs included cell population, cytokine and extracellular matrix concentrations, and biomechanical scaffold properties. The HA-Gtn ABM was implemented in C++ and parallelized with OpenMP and

CUDA. Concurrently, in vitro experiments were performed for calibration and validation. Human vocal fold fibroblasts were seeded in HA-Gtn scaffolds of varying HA-Gtn ratios, and the total protein levels, collagen content, and cell count were measured using biochemical assays. Morris parameter screening and Sobol sensitivity analysis were used to select the most influential model parameters to the model outputs.

Results

The most influential parameters were related to swelling ratio, degradation rate, cell proliferation, extracellular matrix synthesis, and cytokine synthesis. The mechanical dynamics of the biomaterial properties were similar to those reported in literature. However, while the model captured cell population trends, it did not accurately predict cell population magnitudes.

Conclusions

The ABM has shown promising results for parameter evaluation. It can be further developed into a computing framework for evaluating biological and mechanical outcomes of bioengineered scaffolds for vocal fold reconstruction.

Acknowledgments: Canadian Institutes of Health Research Project Grants 388583 (Li-Jessen), Natural Sciences and Engineering Research Council of Canada RGPIN-2018-03843 (Li-Jessen), Compute Canada (Li-Jessen), National Institute of Deafness R03DC012112 (Li-Jessen), National Institute of Communication Disorders R01DC005788 (Mongeau).

References

1. Murphy, J.T.; et al. Simulating 3-D bone tissue growth using repast HPC: Initial simulation design and performance results. In Proceedings of the 2016 Winter Simulation Conference (WSC), Washington, DC, USA, 11–14 December 2016.
2. Bryant, J.S.; Vernerey, F.J. Programmable Hydrogels for Cell Encapsulation and Neo-Tissue Growth to Enable Personalized Tissue Engineering. *Adv. Healthc. Mater.* **2018**, *7*, 1700605.
3. Zahedmanesh, H; Lally, C. A multiscale mechanobiological modelling framework using agent-based models and finite element analysis: application to vascular tissue engineering. *Biomech. Model. Mechanobiol.* **2012**, *11*, 363–377.
4. Artel, A.; et al. An Agent-Based Model for the Investigation of Neovascularization within Porous Scaffolds. *Tissue Eng. Part A* **2011**, *17*, 2133–2141.

5.3. Usefulness of Cepstral Peak Prominence (CPP) in Post-Thyroidectomy Dysphonia Evaluation

Hee Young Son

Department of Otorhinolaryngology, Dongnam Institute of Radiological & Medical Sciences, Busan, Korea

Keywords: voice; cepstrum; thyroid

Objectives

The purpose of this study was to compare the usefulness of Cepstral peak prominence (CPP) with parameter of Multi-Dimensional Voice Program (MDVP) in evaluating patients with subjective voice impairment after thyroidectomy.

Introduction

Thyroid disease are very common and 30% of that will receive thyroidectomy as a treatment. But the 40–80% of that patients complained for dysphonia after surgery. The possibility of nerve damage

after thyroidectomy is considerably low. Therefore, the possible causes of voice changes other than nerve damage are studied variously. Cepstrum is a function of quefrequency, the reciprocal of frequency, and magnitude, the concept of intensity. CPP is the distance from cepstral peak in regression line which is average of sound energy in cepstrum. CPP is sensitive to the patient 's worse subjective symptoms.

Methods

Eighty patients who underwent thyroidectomy were enrolled in this study. We measured vowel extension and paragraph reading before thyroid surgery, 2 weeks and 3 months after surgery. Speech tools were used to analyze the speech cepstrum, and CPPs and mean CPP F0 were measured. Independent sample t-tests were performed to see the differences between the two groups. One-way ANOVA was used to determine the differences between the time of test measurements.

Results

There was a significant difference between CPPs and mean CPP F0 for prolonged speech and paragraph reading between the two groups. There was no difference according to the measurement time. MDVP parameter such as Jitter, shimmer, and Noise to Harmony Ratio (NHR) showed similar results to CPP. It was confirmed that the threshold value for estimating the degree of voice impairment in patients who had subjective voice symptoms.

Conclusions

The Cepstral peak prominence (CPP) was found to be a useful index in assessing post-thyroidectomy dysphonia compared with MDVP parameters.

Acknowledgments: There are no acknowledgments to be made.

References

1. Fraile, R.; Godino-Llorente, J.I. Cepstral peak prominence: A comprehensive analysis. *J. Biomed. Signal Process. Control* **2014**, *14*, 1–24.
2. Heman-Ackah, Y.D.; Michael DDGoding, G.S., Jr. The relationship between cepstral peak prominence and selected parameters of dysphonia. *J. Voice* **2002**, *1*, 20–27.
3. Park, M.C.; Mun, M.K.; Lee, S.H.; Jin, S.M. Clinical usefulness of cepstral analysis in dysphonia evaluation. *Korean J. Otorhinolaryngol.-Head Neck Surg.* **2013**, *56*, 574–578.
4. Kim, T.H.; Choi, J.I.; Lee, S.H.; Jin, S.M. Comparison of vowel and text-based cepstral analysis in dysphonia evaluation. *J. Korean Soc. Laryngol. Phoniatr. Logop.* **2015**, *26*, 117–121.

5.4. Decoding Phonation with Artificial Intelligence (DEP AI): Proof of Concept

Maria E. Powell¹, Marcelino Rodriguez Cancio², David Young¹, William Nock³, Beshoy Abdelmessih¹, Amy Zeller¹, Irvin Perez Morales^{4,5}, Peng Zhang³, C Gaelyn Garrett¹, Douglas Schmidt³, Jules White³ and Alexander Gelbard¹

¹ Vanderbilt Bill Wilkerson Center for Otolaryngology, Vanderbilt University Medical Center, Nashville, TN, USA

² Department of Information Technology, Vanderbilt University, Nashville, TN, USA

³ Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA

⁴ Center for Computational & Numerical Methods in Engineering, Central Univ Marta Abreu of Las Villas, Santa Clara, Cuba

⁵ Infralab, University of Brasilia, Brasilia, Brazil

Keywords: voice disorders; detection; acoustic analysis; convolutional neural network; classification

Objectives

The purpose of this study is to provide a proof-of-concept that embedded data within human phonation can be accurately and efficiently decoded with deep learning neural network analysis to differentiate normal from disordered voices.

Introduction

Acoustic analysis of voice has the potential to expedite detection and diagnosis of voice disorders. Despite its widespread use for screening and progress monitoring, intrinsic limitations have prevented its effective application for automated detection and diagnosis of voice disorders.¹ Acoustic analysis has traditionally relied on the characterization of limited numbers of acoustic parameters; however, the mechanism of human speech production is highly complex, and any given pathology affects multiple acoustic parameters simultaneously. Applying an image-based, convolutional neural-network (CNN) approach to analyzing the acoustic signal based on spectral content may be an effective means for detecting and differentially diagnosing voice disorders. Not only do the spectrograms maintain the full frequency resolution of the acoustic signal, but they also have the unique characteristic of being data-rich *images* that can be analyzed via image analysis techniques.

Methods

Acoustic recordings of the Rainbow Passage were acquired from 10 vocally-healthy speakers and 70 patients with one of seven voice disorders (n = 10 per group). Diagnoses included adductor spasmodic dysphonia (AdSD), vocal fold polyp (Polyp), polypoid corditis (PCord), recurrent respiratory papillomatosis (RRP), muscle tension dysphonia (MTD), unilateral vocal fold paralysis (UVFP), and essential tremor of voice (ETV). Acoustic signals were cut into 3-s segments and converted into grayscale, wide-band, linear spectrograms using the short-time Fourier transform with a block-size of 2048 data points and overlap of 1536 (75%). Spectrograms were filtered above 8kHz and scaled by an 8/3 ratio at 1 inch-per-second to create standardized images. To augment the sample size of the training set, each organic spectrogram was randomly divided with a single, vertical splice and the subsequent pieces were reversed in order to create a synthetic 3-s spectrogram. This process was repeated 10 times for each spectrogram, rendering a synthetic training set of 4510 images, which were used to train a CNN developed with the Keras library. The network architecture was trained separately for each of the seven diagnostic categories, and binary classification tasks (normal vs disordered) were performed for each diagnosis. All models were validated using the 10-fold cross validation technique, where all spectrograms for an individual speaker were included in a single fold to prevent information leakage from the training set to the validation set. Outcome measures for each model were accuracy and loss.

Results

Five models achieved at least 80% accuracy in the binary classification task: AdSD (90%), UVFP (87%), Polyp (86%), PCord (84%), and RRP (80%). While the MTD (58%) and ETV (78%) models were comparably less robust, results coincided with other studies that employ other neural network techniques, such as the Mel frequency cepstral coefficient or wavelet-based artificial intelligence models.² Overfitting was noted in all models, likely due to the small sample size.

Conclusions

Despite the small size of the available dataset, satisfactory results were obtained for the AdSD, UVFP, Polyp, PCord, and RRP diagnostic groups, with accuracy in the validation set substantially higher than the baseline accuracy of the naïve algorithm. Larger data sets are needed to optimize these neural networks; however, these preliminary results support further study of deep neural networks for clinical detection and diagnosis of human voice disorders.

References

1. Saenz-Lechon, N.; Godino-Llorente, J.I.; Osma-Ruiz, V.; Gómez-Vilda, P. Methodological issues in the development of automatic systems for voice pathology detection. *Biomed. Signal Process. Control* **2006**, *1*, 120–128.
2. Schönweiler, R.; Hess, M.; Wübbelt, P.; Ptok, M. Novel approach to acoustical voice analysis using artificial neural networks. *J. Assoc. Res. Otolaryngol.* **2000**, *1*, 270–282.

5.5. Glottal Area Waveform Modeling Based Voice Quality Typing

Philipp Aichinger ¹, Imme Roesner ¹, Franz Pernkopf ² and Jean Schoentgen ^{1,3}

- ¹ Division of Phoniatics-Logopedics, Department of Otorhinolaryngology, Medical University of Vienna and Austria
- ² Signal Processing and Speech Communication Lab, Graz University of Technology, Austria
- ³ Department of Bio-, Electro- And Mechanical Systems, Faculty of Applied Sciences, Université libre de Bruxelles, Belgium

Keywords: high-speed videoendoscopy; glottal area waveform; voice source modeling; voice quality typing

Objectives

The aim of the study is to automatically distinguish four types of voice qualities by means of glottal area waveform (GAW) modeling.

Introduction

Voice quality typing is pivotal to the clinical care of voice disorders, because it is important for the indication, selection, evaluation, and optimization of clinical treatment techniques. Three of the selected voicing types are typical for pathological voices, i.e., (1) diplophonia, (2) voices with random extra pulses during quasi-closed glottal cycle phases, and (3) voices with random phase differences between the left and the right vocal folds. The fourth voice type is a control type, i.e., (4) normophonia.

Methods

To enable experimentation under controlled conditions, per voicing type 20 GAWs are generated synthetically. Synthesizer structures and parameters are identified based on qualitative observations made on an available database comprising laryngeal high-speed videos of 120 subjects. Automatic 4-way classification is proposed. It is based on analysis-by-synthesis, and uses as predictors the four root mean square model errors obtained with each of the synthesis structures. The predictors are input to a 4-way logistic regression model.

Results

The overall accuracy of the proposed 4-way logistic regression model is 56.5%. The classifier performs best for the random phase difference GAW type (74.1%), and worst for the extrapulsed GAW type (45.8%). The GAW type is predicted correctly in the majority of the signals. No confusions of the extrapulsed type with the random phase differences type are observed, although they are not easily distinguishable auditorily (Table 1).

Table 1. Confusion matrix regarding logistic regression based automatic classification of glottal area waveform (GAW) types. GAW types are shown in columns; predicted types are shown in rows. The diplophonic type (D), the extrapulsed type (EP), the normophonic type (N), and the random phase differences type (PD) are shown.

		GAW Type			
		D	EP	N	PD
Predicted Type	D	56.1%	25.6%	30.6%	23.9%
	EP	30.0%	45.8%	13.5%	0.0%
	N	0.0%	28.6%	54.2%	1.96%
	PD	13.9%	0.0%	1.7%	74.1%

Discussion and Conclusions

The four proposed types of GAWs are often distinguishable automatically. Classification accuracy is better than guessing. A challenge may be the observation that the proposed GAW types are not mutually exclusive, e.g., diplophonic GAWs that contain extra pulses were observed. The analysis is limited to GAWs only, because the model complexity of this class of signals is lower than the model complexity of spatially resolved vocal fold vibration signals. Also, it is hypothesized that the GAW contains all information relevant to vocal fold vibration induced voice quality features.

Acknowledgments: This work was supported by the Austrian Science Fund (FWF): KLI 722-B30.

5.6. Automated Quantification of Inflection Events in the Electroglottographic Signal

Juliana Codino ¹, María Eugenia Torres ^{2,3}, Adam Rubin ^{1,4,5} and María Cristina Jackson Menaldi ^{1,6}

¹ Lakeshore Professional Voice Center, Lakeshore Ear, Nose and Throat Center, St. Clair Shores, MI, USA

² Laboratorio de señales y dinámicas no lineales, Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Argentina

³ National Council for Scientific and Technical Research, CONICET, Argentina

⁴ Department of Surgery, Oakland University, William Beaumont School of Medicine, Rochester, MI, USA

⁵ Department of Surgery, Michigan State University, College of Osteopathic Medicine, Lansing, MI, USA

⁶ Department of Otolaryngology, School of Medicine, Wayne State University, Detroit, MI, USA

Keywords: electroglottography; inflection events; contact phase; automated report

Objectives

The goal of this study is to further investigate an automated tool (ATool), developed by our team, which can report and classify inflection events in the electroglottographic signal. Its performance for different thresholds is here studied.

Introduction

Electroglottography (EGG) has been widely used in the clinic and in voice research. Its main quantitative parameters include: contact quotient, open quotient, and F0 extraction; Titze¹ described four distinct electroglottographic signal shapes, which are used during voice assessment. Many publications refer to the knee at the beginning of the closing phase or during the opening phase. However, little has been said about quantifying these inflection events (IE). This study is built on previously published research². With the goal of improving the automated performance of the ATool, we ran a statistical analysis including larger database and different threshold settings.

Methods

We conducted a retrospective study of patients evaluated for dysphonia at a private voice center, where electroglottographic measurements are routinely performed. Sixty-one EGG signals demonstrating “peak-skewing” were selected. Based on the waves’ mathematical properties, we developed an ATool to be run on Matlab©. The IE within the signal are reported and classified by ATool with 7 different detection threshold settings. Blinded to the automatic output, 4 voice specialists manually evaluated the signals (by sight). Agreement between the manual analysis and ATool was determined.

Results

The agreement between the 4 blinded raters and the ATool was calculated using Cohen’s Kappa coefficient³ for the 7 threshold settings. As expected, some of the threshold settings exhibited better agreement with the blind raters than others. This would indicate that the user could maximize performance of the ATool by adjusting thresholds. We also observed that the ATool did not have the same performance in the ascending and descending portions of the signal, with slightly lesser agreement in the descending portion. It was found to be “good” for the former and “moderate” and “fair” for the latter.

Conclusions

EGG signals can display IE in the ascending and/or descending portion of the signal that may be valuable to report. The ATool provides a detailed quantified analysis of IE in the EGG signal throughout a voice sample. Agreement was acceptable compared to observation performed by clinicians, obviating the need for time consuming manual analysis. The ATool also includes the option of adjusting threshold settings for this quantification, to enhance performance levels.

References

1. Titze, I. Interpretation of the electroglottographic signal. *J. Voice* **1990**, *4*, 1–9.
2. Codino, J.; Torres, M.E.; Rubin, A.; Jackson-Menaldi, M.C. Automated electroglottographic inflection events detection. A pilot study. *J. Voice* **2016**, *30*, 768
3. Fleiss, J.; Levin, B.; Cho Paik, M. *Statistical Methods for Rates and Proportions*; Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003.

5.7. Characteristics of the Pharyngoesophageal Segment: Literature Review

Ana Carolina A. M. Ghirardi ¹, Andrey Ricardo da Silva ², Thaiana Volkmann Nakandakari ¹ and Rayane Délcia da Silva ¹

¹ Department of Speech-Language Pathology and Audiology, Federal University of Santa Catarina, Florianópolis, SC, Brazil

² Department of Mechanical Engineering/Laboratory of Vibrations and Acoustics, Federal University of Santa Catarina, Florianópolis, SC, Brazil

Keywords: voice; voice acoustics; alaryngeal speech; pharyngoesophageal segment

Objectives

To conduct a literature review regarding the dynamic properties of the pharyngoesophageal segment (PES), aiming to better understand the mechanisms involved in (tracheo)esophageal speech production.

Introduction

The vibratory behavior of the vocal folds is widely studied, but little is known about the vibratory aspects of the PES and of the fibers of the pharyngoesophageal segment, that provide the voice source for totally laryngectomized individuals. Understanding this structure's vibratory behavior may aid in developing assistive technologies that can enhance rehabilitation of these patients' oral communication.

Methods

This is a literature review study, conducted on the following data bases: *Web of Science*, *Scopus*, *Medline*, *SciELO*, *Lilacs* and *Google Scholar* between March and August, 2018. The guideline question was: "what are the anatomic, geometric and dynamic characteristics of the pharyngoesophageal segment?". Key terms were selected and entered into the search fields, interacting through the Boolean operators 'and', 'or' and 'not'. Studies were screened by title, and those that did not concern the guideline question, and/or found in more than one data base and/or were not available for reading were excluded. Thereafter, all abstracts of the selected papers were read and those that did not specifically address the guideline question or those not written in the previously selected languages were excluded. Finally, the papers remaining after the previous screenings were fully read and were included in the study when providing answers to any part of the guideline question.

Results

This review was composed of 16 papers published between 1967 and 2017, concerning the characteristics of the PSE, using esophageal manometry, videostroboscopy, high-speed imaging, videofluoroscopy and EMG to assess different aspects of this structure. The PES is reportedly located between the C3 and C6 vertebrae and, during speech, it may be slightly elevated. Its shape may vary but the most commonly reported shapes are circular, triangular or linear. Studies generally report that patients with circular PES have better vocal/speech outcomes but its radius or the length of linear segments have no significant correlation with voice quality. The location of the vibration is also found to vary among patients, who may have a PES that vibrates entirely throughout its circumference or there may be one or two vibratory segments within the structure. When the pressure along the esophagus and at the PES during phonation is higher than at rest, around 77mmHg, there is generally better voice quality than when pressure is very low or has high peaks during phonation. Studies discuss that better control of the cricopharyngeal muscle and, ultimately, of the PES leads to a more regular vibration pattern of this structure, which may also aid in wider dynamic extension, better fundamental frequency control and voice inflections during speech.

Conclusions

The shape of the PES varies according to different aspects and, therefore, has different mechanical properties in each patient. Positive vocal outcomes are associated with a circular structure, well-distributed pressure along the esophagus during phonation and relative patient control of the contraction of the cricopharyngeal muscle.

Acknowledgments: FINEP/Brazil for grant number 01.16.0044.00(0346/15) and PIBIC/CNPq/Brazil.

5.8. Designing Audible Sound Spots Using Metamaterial Based Phased Array

Mahdi Derayatifar ^{1,2}, Mohsen Habibi ^{1,2}, Muthukumaran Packirisamy ^{1,2} and Rama Bhat ¹

¹ Department of Mechanical Engineering, Concordia University, Montreal, QC, Canada

² Optical BioMEMS Laboratory, Concordia University, Montreal, QC, Canada

Keywords: ultrasound wave; acoustic holography; audio spot; directional loudspeaker

Objectives

The purpose of the present study is to investigate the capability of combining the passive acoustic hologram in super-directional loudspeakers which emit focused ultrasound wave at single or multiple spots, while are inexpensive and has less complexity than active ultrasound phased arrays.

Introduction

Traditional loudspeakers mostly generate sound uniformly in all directions which makes them to have broad radiation pattern. This is usually an unfavorable feature which cause noise for non-listeners in public places [1]. In recent technology, highly directional speakers has been generated by employing a phased array ultrasound transducer. They are capable of generating a single or multiple focused ultrasound [2]. However, since they require complex phase calibration circuit and careful tuning, they can only be used in small numbers of elements which hinders their degree of freedom in generating the more complex wavefronts. Acoustic holography is one way to overcome these obstacles, which requires much less complexity with higher degree of freedom than ultrasound phased arrays [3].

Methods

Similar to optical holography, an iterative phase retrieval algorithm is used to generate the acoustic hologram. In this method, the two amplitudes on the hologram surface and plane of the desired focus spot are known. The phase distribution in these planes are reconstructed by forward and backward propagation of the wavefront along with updating the amplitudes at each iteration with the already known amplitudes. Finally, the phase distribution on the hologram surface can be retrieved.

Results

By employing acoustic holography, single or sparse multi focal ultrasound wave can be achieved. The results will be discussed in details, since it has not been specifically optimized in the existing literature.

Conclusions

This study highlights an important role of acoustic holography in super-directional loudspeakers. The major advantages of the present study is that acoustic holograms are superior in flexibility and performance in comparison with ultrasound phased arrays.

References

1. Ibaraki, T. Holographic Whisper: Rendering audible sound spots in three-dimensional space by focusing ultrasonic waves. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017.
2. Shi, C.; Kajikawa, Y.; Gan, W.S. An overview of directivity control methods of the parametric array loudspeaker. *APSIPA Trans. Signal Inf. Process.* **2014**, *3*, 1–12.
3. Xie, Y.; et al. Acoustic Holographic Rendering with Two-dimensional Metamaterial-based Passive Phased Array. *Sci. Rep.* **2016**, *6*, 1–6.

5.9. Characterizing Injury Recovery in Rabbit Vocal Folds with Multimodal Imaging

Ksenia Kolosova ¹, Marius Tuznik ², Qiman Gao ³, Sarah Bouhabel ⁴, Huijie Wang ⁵, Luc Mongeau ⁵ and Paul W. Wiseman ^{1,6}

¹ Department of Physics, McGill University, Montreal, QC, Canada

² McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, QC, Canada

³ Department of Dentistry, McGill University, Montreal, QC, Canada

⁴ Department of Otolaryngology, Head and Neck Surgery, McGill University, Montreal, QC, Canada

⁵ Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

⁶ Department of Chemistry, McGill University, Montreal, QC, Canada

Keywords: vocal fold injury; nonlinear microscopy; microcomputed tomography; magnetic resonance imaging

Objectives

We visualized vocal fold injury recovery in a rabbit model using three imaging modalities. These techniques can be applied to characterize the injury recovery process and test treatments for vocal fold scarring.

Introduction

Rabbits are frequently used for studying vocal fold injury and testing injectable biomaterial treatments for vocal fold scarring. In these studies, imaging-based evaluation is most often conducted by tissue slicing and histological staining. To obtain three-dimensional information without physical slicing, we recently used nonlinear laser-scanning microscopy and nanoscale computed tomography to visualize a dissected rabbit vocal fold specimen¹. As they require a small specimen size, these techniques oblige precise injury localization prior to dissection. We sequentially applied magnetic resonance imaging (MRI), microscale computed tomography (CT), and nonlinear laser-scanning microscopy (NLSM) to visualize injury with and without labelling.

Methods

A unilateral injury was created using microcup forceps in the left vocal fold of three New Zealand White rabbits. Animals were sacrificed at 3, 10, and 39 days post injury, and the larynx was excised and dissected. Three imaging methods were sequentially applied to each specimen. MRI was performed using the 7-T Bruker Pharmascan at the McConnell Brain Imaging Center. CT was performed using the Bruker SkyScan 1172 at the McGill Institute for Advanced Materials. NLSM was performed with optical clearing sample preparation following methods described in our previous study¹. Images were analyzed using Dragonfly (Object Research Systems) and a Python script.

Results

The MRI modality allowed visualization of the injury location label-free with 100 μm resolution. The CT modality achieved finer resolution down to the micrometer scale. However, the intrinsic density contrast of CT was insufficient to visualize features of the injury beyond the contour, necessitating heavy metal staining for contrast enhancement. The NLSM modality provided simultaneous specific visualization of second harmonic generation from fibrillar collagen and two-photon autofluorescence of elastin with near diffraction-limited spatial resolution, allowing clear resolution of collagen fibers in the vocal fold lamina propria, muscle, and surrounding cartilages at submicrometer scales. It allowed quantitative evaluation of properties of the collagen distribution following methods reported in a past study².

Conclusions

The results suggest that a combination of MRI, contrast-enhanced CT, and NLSM can be used to characterize vocal fold injury over time and at different spatial scales. The label-free visualization mechanism of MRI motivates its use for injury localization in live-animal imaging, and each technique serves as a platform for qualitative and quantitative image analysis.

Acknowledgments: This work was supported by the National Institutes of Health (Grant R01 DC005788), the Natural Sciences and Engineering Research Council, and the Canadian Foundation for Innovation.

References

1. Kazarine, A.; et al. Multimodal virtual histology of rabbit vocal folds by nonlinear microscopy and nano computed tomography. *Biomed. Opt. Express* **2019**, *10*, 1151–1164.
2. Miri, A.K.; et al. Nonlinear laser scanning microscopy of human vocal folds. *Laryngoscope* **2012**, *122*, 356–363

6. Session 3

6.1. The Causes and Laryngeal Electromyography Characteristics of Unilateral Vocal Fold Paralysis

Please add author names here.

Department of Otorhinolaryngology–Head and Neck Surgery, Beijing Tongren Hospital, Capital Medical University, Beijing, China; hurong_ronghu@126.com; Tel.: +86-186-1198-7192

* Correspondence: xuwendoc@126.com

Keywords: vocal fold paralysis; vocal fold immobility; recurrent laryngeal nerve; laryngeal electromyography

Objectives

To investigate the causes and laryngeal electromyography (LEMG) characteristics of unilateral vocal fold paralysis.

Methods

We retrospectively analyzed the history and laryngeal electromyography of the patients with unilateral vocal fold immobility in the department of otolaryngology head and neck surgery in our hospital from September, 2009 to December, 2015. A total of 337 patients with unilateral vocal fold paralysis were involved. The etiology was reviewed and the characteristics of LEMG (including spontaneous potential, recruitment potential, evoked potential, synkinesia, and et al.) was analyzed.

Results

Among the 337 patients, 180 were female and 157 were male. Their age ranged from 13 to 80 years, with an average age of 45.1 ± 15.1 years. There were 232 cases of left vocal fold paralysis and 105 cases of right vocal fold paralysis. The causes were injury (177 cases, 52.5%), idiopathic (72 cases, 21.4%), infectious factors (61 cases, 18.1%), neoplasm and oppression factors (27 cases, 8.0%). In the injury group, 161 cases (47.8%) were surgically injured, including 123 cases of neck surgery (111 cases of thyroid surgery, 5 cases of cervical neoplasms resection, 4 cases of anterior cervical surgery, 2 cases of esophageal cancer surgery, and 1 case of debridement and suture in neck trauma), 38 cases of chest surgery. Other 16 cases were trauma injured, including 15 cases of neck trauma and 1 case of chest trauma. The neoplasm and oppression factors group included 27 cases of neck neoplasms (16 cases of benign thyroid disease, 3 cases of thyroid carcinoma, 1 case of parathyroid adenoma), 4 cases of pulmonary hypertension, and 3 cases of mediastinal neoplasms. LEMG showed that the proportion of complete recurrent laryngeal nerve injury was 72.9% in the injured group, 66.7% in the mass and compression group, 49.2% in the infectious factor group, and 44.4% in the idiopathic group. The denervated potential (fibrillations and positive sharp waves) or the regenerative potential can be detected. The recruitment potentials were rest or simple phase, and the evoked potential disappeared. Among the 337 patients, 136 cases (40.4%) were found synkinesia in the posterior cricoarytenoid (PCA) muscles, while 2 cases were found synkinesia in the thyroarytenoid (TA) muscles. The proportion of complete recurrent laryngeal nerve injury was higher in those with synkinesia than those without synkinesia.

Conclusions

The major cause of vocal fold paralysis is neck surgery, and the thyroid surgery is the most common causes. Different causes led to different degrees of recurrent laryngeal nerve injury. LEMG showed the highest proportion of complete nerve damage causing by surgery or trauma. Patients with severe recurrent laryngeal nerve injury are prone to found synkinesia. The PCA muscles were more likely to detected synkinesia than the TA muscles.

Acknowledgments: This study was supported by Natural Science Foundation of Beijing (7172051) and Beijing Municipal Administration of Hospitals' Youth Programme (QML20170201).

6.2. Arytenoid Adduction and Type 1 Thyroplasty for Unilateral Vocal Fold Paralysis: Measurements from Six Excised Canine Larynges

Alexandra Maddox ¹, Charles Farbos de Luzan ², Liran Oren ², Sid M. Khosla ² and Ephraim Gutmark ¹

¹ Department of Aerospace Engineering and Engineering Mechanics, Univ. of Cincinnati medical Center, Cincinnati, OH, USA

² Department of Otolaryngology-Head and Neck Surgery, Univ. of Cincinnati medical Center, Cincinnati, OH, USA

Keywords: voice; unilateral vocal fold paralysis; arytenoid adduction; thyroplasty type 1

Objectives

Using excised canine larynx models, characterize the effects of an arytenoid adduction (AA) on Thyroplasty Type 1 (TT1) procedures. Specifically, on the efficiency of the glottis during phonation and the quality of the acoustic signal.

Introduction

Patients with unilateral vocal fold paralysis (UVFP) complain of a soft breathy voice that is hard to understand in noisy environments. TT1, the most common surgical intervention for UVFP¹, uses an implant to push over the membranous fold^{2,3}. Surgically there remains controversy on how to restore optimal vocal function. For example, regarding TT1 procedures, questions include the optimal shape and size for the implant and whether an AA should be added. Many laryngologists prefer a TT1 with an AA while others prefer TT1 alone. The literature does not clearly answer the question of when and if an AA should be added.

Methods

Six excised canine larynges were tested with a TT1 alone and a TT1 with AA at various subglottal pressures (Psg). In each case measurements were taken of the acoustics, supplied flow rate (Q) and Psg. From these measurements, the glottal efficiency (Eg) was calculated using Schutte's equation, where $E_g = I/P_{sg} \cdot Q$, I being the acoustic intensity. Additionally, the acoustic signal was analyzed using the Analysis of Dysphonia in Speech and Voice (ADSV) program (Model 5109, KayPENTAX) to determine the quality of the sound wave produced during phonation.

Results

On average cases with TT1 and AA had higher vocal efficiency than cases with TT1 alone. This means that for the same aerodynamic power ($P_{sg} \cdot Q$) cases with an AA produced a stronger acoustic wave. Additionally, cases with TT1 and AA had higher cepstral peak prominence (CPP). CPP is higher in acoustic signals with a well-defined harmonic structure and has a strong negative correlated with breathiness, roughness, and harshness. Thus, a higher CPP in cases with TT1 and AA is indicative of a higher quality sound source than cases with TT1 alone.

Conclusions

The addition of an AA to TT1 procedures results in more efficient phonation of the vocal folds and a higher quality sound source.

Acknowledgments: This project is supported by NIH Grant no. R01 DC009435 from the National Institute of Deafness and Other Communication Disorders.

References

1. Shen, T.; Damrose, E.J.; Morzaria, S. A meta-analysis of voice outcome comparing calcium hydroxylapatite injection laryngoplasty to silicone thyroplasty. *Otolaryngol. Head Neck Surg.* **2013**, *148*, 197–208.
2. Isshiki, N.; Okamura, H.; Ishikawa, I. Thyroplasty type 1 (Lateral compression) for dysphonia due to paralysis or atrophy. *Acta Oto-Laryngol.* **1975**, *80*, 465–473.
3. Isshiki, N.; Morita, H.; Okamura, H.; et al. Thyroplasty as a new phonosurgical technique. *Acta Otolaryngol.* **1974**, *78*, 451–457.

6.3. Increased Calcium Channel in the Lamina Propria of Aging Rat

Byung-Joo Lee^{1,3,*}, Ji Min Kim¹, Hyoung-Sam Heo², Sung Chan Shin³, Han-Seul Na³, Jin-Choon Lee⁴ and Eui-Suk Sung⁴

¹ Pusan National University Medical Research Institute, Pusan National University School of Medicine, Pusan National University, Busan, Democratic People's Republic of Korea

² Division of Bio-Medical Informatics, Center for Genome Science, Korea National Institute of Health, Korea Centers for Disease Control and Prevention, Cheongju-si, ChoongchungBuk-do 28159, Democratic People's Republic of Korea

³ Department of Otorhinolaryngology, Head and Neck Surgery, Pusan National University Hospital, Busan, Democratic People's Republic of Korea

⁴ Department of Otorhinolaryngology, Head and Neck Surgery, Pusan National University Yangsan Hospital, Yangsan, Gyeongnam, Democratic People's Republic of Korea

Keywords: aging; vocal folds; laminal propreaia; extracellular matrix; voltage-gated calcium channels

Objectives

The aim of this study was to investigate the differences in gene expression of lamina propria using next generation sequencing (NGS) in young and aging rats and to identify genes that affect aging-related ECM changes for developing novel therapeutic target molecule.

Introduction

The alterations of the extracellular matrix (ECM) in lamina propria of the vocal folds are important changes that are associated with decreased vibrations and increased stiffness in aging vocal fold. The mechanism of aging-related ECM changes has not yet been fully understood. The studies for the molecular mechanism of the alterations of aging-related ECM in lamina propria may be better understood for the basic mechanism of presbylarynx and may be applied to develop the novel therapeutic target molecules for aging voice disorder.

Methods

To investigate the difference of genes expression of age-related lamina propria in rat vocal fold during aging process, we used 6 and 22 months old male Sprague-Dawley rats (n = 8, each group) for NGS. Immunohistochemical staining of the calcium channel genes suggested to increase in NGS

analysis was performed in the lamina propria of vocal fold. The expression of calcium channel genes were observed in human vocal fold fibroblast cell lines (hVFFs). The changes of ECM in hVFFs after the treatment with calcium channel blocker were examined by PCR and western blotting.

Results

Among the 40 genes suggested in the NGS analysis, voltage-gated calcium channels (VGCC) subunit alpha1 S (CACNA1S), VGCC auxiliary subunit beta 1 (CACNB1), and VGCC auxiliary subunit gamma 1 (CACNG1) were increased in the lamina propria of the old rats compared to the young rats. The synthesis of collagen I and III in hVFFs decreased after si-CACNA1S and verapamil treatment. However, there was no effect on the synthesis of HA or elastin. The expression and activity of matrix metalloproteinases (MMP)-1 and -8 were increased in hVFFs after the treatment of verapamil. However, there was no change in the expression of MMP-2 and -9.

Conclusions

These results suggest that some calcium channels may be related with the alteration of aging-related ECM in vocal folds. Calcium channel has promising potential as a novel therapeutic target for the remodeling ECM of aging lamina propria.

Acknowledgments: This work was supported by NRF-2016R1D1A3B01015539.

Reference

1. Sachs; et al. Treatment effectiveness for aging changes in the larynx. *Laryngoscope* **2017**, *127*, 2572–2577.

6.4. Localization of the Tight Junction Proteins Claudin Family in the Laryngeal Glands: A Rat Study

Ryo Suzuki ¹, Yo Kishimoto ¹, Nao Hiwatashi ¹, Masanobu Mizuta ^{1,2}, Atsushi Suehiro ¹, Ichiro Tateya ¹ and Koichi Omori ¹

¹ Department of Otolaryngology-Head and Neck Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan

² Department of Otolaryngology-Head and Neck Surgery, Kurashiki Central Hospital

Keywords: tight junction; claudin; laryngeal glands

Objectives

The mucus secreted by the laryngeal glands not only serves as a lubricant for the vocal fold vibration, but also protects the vocal folds from inhaled irritants or acid reflux. Therefore, when the laryngeal glands become dysfunctional, secretions are reduced, and voice disorders due to drying, inflammation and infection may occur. Secretory fluids from acinar and mucous cells are considered to be subject to the reabsorption of water and electrolytes when passing through the ducts, and then become the air-way mucus. The modification of mucus is considered to be regulated by tissue-specific tight junctions (TJs) with permselective barrier to the diffusion of solutes and ions through the paracellular pathway. Claudins (cldns), the essential integral membrane proteins of TJs, constitute the large gene family of more than 20 members, and the permselectivity of the epithelium is defined by the combination and ratio of the cldn family members within the individual TJs. Therefore, to understand the physiological function of the laryngeal glands, this study aimed to clarify the localization of cldn family proteins in the laryngeal glands.

Methods

Five 13-week-old male Sprague-Dawley rats were used. Larynges were harvested for histological studies (n = 1 for hematoxylin and eosin, and PAS staining; n = 1 for scanning electron microscope) and immunohistochemistry (n = 3). The antibodies against occludin, ZO-1 and cldn1, 3, 4, 5, 6, 7, 8, 10, 11, 12, 17 were used to clarify the localization of each TJ-related protein in the laryngeal glands. Anti-cldn antibodies were selected among those whose gene expression was confirmed by previously performed RT-PCR analysis with rat larynx.

Results

Histological studies indicated that many laryngeal glands are distributed at supra and subglottis of rat larynx. They were mostly PAS positive in the supraglottis, whereas PAS positive glands were sparse in the subglottis. Furthermore, many duct cells were distributed in the subglottis, forming large conduits which opened to the luminal surface. Sharp signals of cldn3, 5, 8, 10 and 11 were observed to the cell-cell junctions at gland cells, and cldn3, 5, 7, 8, 10 and 11 were observed at duct cells.

Conclusions

Many mucous glands are distributed in the supraglottis, whereas serous and mucous glands are mixedly distributed in the subglottis of rat larynx. The proportion of ducts in the subglottis was very large, suggesting that large amounts of mucus may be constantly accumulated in the subglottic tissues. The localization of cldn family proteins in the rat laryngeal glands was elucidated in this study. Cldn3, 5 and 11 were reported to be functionally classified as barrier-forming TJ proteins and cldn10 as channel-forming TJ protein, whereas the precise function of cldn7 and 8 were unknown. These cldns expressed in the cell-cell junctions of the gland and duct cells may fine-tune the solutions secreted from the gland cells, and finally they become the air-way mucus. To understand these cldn-based pathophysiology of the laryngeal glands, future study is necessary to perform the morphological and functional analysis using specific cldn deficient, aged or irradiated animals.

References

1. Nassar, V.H.; Bridger, G.P. Topography of the laryngeal mucous glands. *Arch Otolaryngol.* **1971**, *94*, 490–498.
2. Sato, K.; Hirano, M. Age-related changes in the human laryngeal glands. *Ann. Otol. Rhinol. Laryngol.* **1998**, *107*, 525–529.
3. Tsukita, S.; Furuse, M.; Itoh, M. Multifunctional strands in tight junctions. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 285–293.
4. Günzel, D.; Yu, A.S. Claudins and the modulation of tight junction permeability. *Physiol. Rev.* **2013**, *93*, 525–569.

6.5. Macrophages in the Vocal Fold

Yo Kishimoto ¹, Shinji Kaba ¹ and Nathan V. Welham ²

¹ Department of Otolaryngology Head and Neck Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan

² Div. of Otolaryngology, Department of Surgery, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA

Keywords: vocal fold; macrophage; wound healing; scarring; fibrosis; loss-of-function

Objectives

To investigate the distributions, phenotypes and roles of macrophages (MQs) during the wound healing process of the murine vocal fold.

Introduction

The inflammatory phase is a key phase in initiating and controlling the wound healing cascades, and MQs play significant roles in this phase. In general, MQs are essential for either efficient or impaired wound healing as they promote debridement, cell proliferation, angiogenesis, collagen deposition, and matrix remodeling [1]. However, previous studies on vocal fold wound healing have focused on the behavior of vocal fold fibroblasts because they primarily synthesize extracellular matrices [2], and the role of MQs has been rarely described and the information about their distributions, phenotypic changes and functions is still limited. MQs are known to be a regulator of wound healing in other organs, and if we can control wound healing process via controlling the function of MQs in the injured vocal fold, it would contribute for an establishment of novel therapeutic strategy for scarred vocal fold.

Methods

Unilateral vocal fold stripping was performed on C57BL/6 mice [3], and larynges were harvested. Immunohistochemical analysis of the vocal fold lamina propria was performed to clarify the distributions of MQs and their polarization phenotypes.

Further, based on a loss-of-function paradigm, the effects of targeted depletion of monocyte (MO) lineage cells on VF wound healing were investigated by immunohistochemistry and qRT-PCR using transgenic mice.

Results

In the naïve vocal fold, the majority of F4/80+ macrophages reside at the superficial layer of the lamina propria and they were positive for CD206. In the injured vocal fold, F4/80+ cells were recruited to the wound. During the wound healing, F4/80+ iNOS+ cells increased at post-operative day 3, and then gradually decreased. As for F4/80+ CD206+ cells, they didn't exhibit any significant changes during the process. It was difficult to distinguish the polarization phenotypes in all the MQs clearly and coexistence of M1 and M2 markers in the same macrophages was observed. Selective depletion of monocyte lineage cells in the early stage of wound healing resulted in reduced scar formation of the vocal fold.

Conclusions

Macrophage phenotypes are regulated by complex tissue-derived signals and exhibit dynamic changes during wound healing. Selective depletion of monocyte lineage cells attenuates scar formation in injured vocal fold, and MOs/MQs therefore seem to be a promising therapeutic target for the prevention or restoration of vocal fold scarring.

References

1. Novak, M.L.; Koh, T.J. Phenotypic transitions of macrophages orchestrate tissue repair. *Am. J. Pathol.* **2013**, *183*, 1352–1363, doi:10.1016/j.ajpath.2013.06.034.
2. Kishimoto, Y.; Kishimoto, A.O.; Ye, S.; Kendzierski, C.; Welham, N.V. Modeling fibrosis using fibroblasts isolated from scarred rat vocal folds. *Lab. Investig.* **2016**, doi:10.1038/labinvest.2016.43.
3. Yamashita, M.; Bless, D.M.; Welham, N.V. Surgical method to create vocal fold injuries in mice. *Ann. Otol. Rhinol. Laryngol.* **2009**, *118*, 131–138.

6.6. Vocal Fold-Mimetic Environment for the Modulation of Stem Cell Functions

Aidan B. Zerdoum ¹, Alexander J. Stuffer ², Zhixiang Tong ¹ and Xinqiao Jia ^{1,2,3,*}

¹ Department of Biomedical Engineering, University of Delaware, Newark, DE, USA

² Department of Biological Sciences, University of Delaware, Newark, DE, USA

³ Department of Materials Science and Engineering, University of Delaware, Newark, DE, USA

Keywords: vocal fold; wound healing; vibration; stem cells

Objectives

The goal of this study is to evaluate how phonation relevant vibratory stimulations affect the phenotype and function of human bone marrow-derived mesenchymal stem cells (MSCs).

Introduction

Vocal fold scarring is a significant clinical issue and remains a therapeutic challenge in the field of laryngology. Stem cell-based regenerative therapy has been implicated in mediating vocal fold wound healing and fibrosis. Mechanistically, it is not clear how injected stem cells interact locally with the extracellular matrix (ECM) of the lamina propria (LP) and how such interactions affect stem cell behaviors to improve function. Molecular and cellular analyses of MSCs cultured in a LP-mimetic scaffold in the presence of phonation-mimicking mechanical stimulations will provide insight on the applicability of MSCs for vocal fold repair and regeneration.

Methods

A dynamic culture system capable of generating vibratory stimulations at human phonation frequencies was constructed and characterized.¹ The bioreactor was composed of two metal bars, each housing four parallel vibration modules. The individual module containing a sandwiched silicone elastomer was directly mounted on top of a speaker controlled by a speaker selector. A watertight vibration chamber was created by sandwiching an elastomeric silicone disk between a pair of hollow acrylic blocks. The vibration signals were translated to the chamber aerodynamically by the oscillating air pressure underneath. The vibration units and speaker selector were enclosed in an anti-humidity acrylic chamber. MSC-laden electrospun poly(ϵ -caprolactone) (PCL) mats were incorporated into the bioreactor and were subjected to high frequency vibrations at 200 Hz for a total of 7 days. A continuous (CT) or a 1h-on-1h-off (OF) vibration mode was applied for a total of 12 h daily. Cellular responses were analyzed by qPCR and immunofluorescence.

Results

Vibratory stimulations did not cause any physiological trauma to the cells. The reinforcement of actin filaments and the enhancement of $\alpha 5 \beta 1$ integrin expression were observed under selected dynamic conditions. A 7-day OF regime significantly up-regulated the gene expression (in terms of fold increase relative to the static controls) of collagen III $\alpha 1$ (Col3A1), fibronectin (FN), hyaluronan synthase 1 (HAS1), elastin (ELN), tenascin C (TNC), cyclooxygenase-2 (COX2), integrin $\alpha 5$ subunit (ITGA5) and matrix metalloproteinase 1 (MMP1). Analyses of genes encoding alternative MSC differentiation markers substantiated MSC's adaptation of fibroblastic phenotype. Cellular production of essential ECM components, such as elastin, HA, and matrix MMP1 was enhanced by the vibrations, with the OF mode being more conducive.² To further improve the regenerative outcome, the high frequency vibration and growth factor (connective tissue growth factor, CTGF) stimuli were sequentially introduced to the MSC/fiber construct.³ Our results show that the initial vibratory culture rendered MSCs more sensitive to the subsequent CTGF treatment, particularly with respect to the expression of MMP1, HA, TNC, decorin and collagen III. The vibrations and soluble growth factors cooperatively mediated MSCs functions, leading to an accelerated ECM synthesis and balanced ECM remodeling; and the classical Erk1/2 pathway was critically engaged in the mechano-biochemical cooperation.

Conclusions

Our study underscores the significance of reproducing a physiologically relevant microenvironment to modulate stem cell behaviors for successful functional vocal fold assembly *in vitro*.

Acknowledgments: This work is supported by NIH/NIDCD (R01 DC008965, R01 DC011377, R01 DC014461).

References

1. Zerdoum; et al. *J. Vis. Exp.* **2014**, *90*, e51594.
2. Tong; et al. *Tissue Eng. Part A* **2013**, *19*, 1862.
3. Tong; et al. *Tissue Eng. Part A* **2014**, *20*, 1922.

6.7. Bioprinting Highly Porous Chitosan-Based Scaffolds With Tunable Stiffness and Viscoelasticity for Vocal Fold Repair

Guangyu Bao¹, Tao Jiang¹, Hossein Ravanbakhsh¹, Huijie Wang¹, Joseph Kinsella², Jianyu Li¹ and Luc Mongeau¹

¹ Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

² Department of Bioengineering, McGill University, Montreal, QC, Canada

Keywords: tissue engineering; bioprinting; porous scaffolds; tunable mechanical properties

Objectives

The goal is to fabricate chitosan-based hierarchical porous structure using bioprinting with orthogonal control over stiffness and viscoelasticity. The fabricated scaffolds can be used as wound filling materials or grafts for vocal fold repair.

Introduction

Hydrogel scaffolds have been widely used as implant materials for vocal fold tissue engineering [1,2]. Ideally, the implanted scaffolds should match the physical architecture and mechanical properties of vocal folds to provide familiar environment for native cells. However, the intrinsic pore sizes for most hydrogel networks are smaller than 100 nm. Such pore sizes are 100–10,000 fold less than the desirable sizes, which limits the ability of scaffolds to recruit native cells nor allow the migration of loaded cells to remodel the matrix. Challenges still remain even if porous scaffolds can be fabricated. The high porosity usually impedes the structural stability and reduce the stiffness of hydrogels. Besides, simple and robust approach for tuning viscoelasticity has yet been revealed.

Methods

The macro pores were created through precise position control of robotic extrusion dispenser (GeSiM). Chitosan solution (95% DDA, medium/high molecular weight) were printed inside sacrificial Bingham fluid made of gelatin (Type A, Sigma), which supported low viscous hydrogel printing and triggered chitosan phase separation to form micro pores. Nano pores were from the intrinsic crosslinked meshes. Stiffness was regulated using sodium bicarbonate (Fisher Scientific) embedded inside sacrificial gelatin slurry with various pH. Orthogonally controlled viscoelasticity was realized through introducing free poly(ethylene glycol) (PEG) (4000 kDa, Sigma), which altered crosslinking density and strengthened hydrogen bonds between chitosan polymeric chains to maintain desirable stiffness. Pore sizes was characterized by confocal and scanning electron microscopy (SEM). For confocal imaging, Rhodamine B isothiocyanate was conjugated to chitosan's primary amine groups to

gain fluorescence signal. For SEM imaging, CO₂ super critical point dryer (Leica) was used to prepare the hydrogel samples without disrupt porous structure. Biocompatibility of the hydrogel system was assessed using Live/Dead viability kit. All rheological tests were performed with a torsional rheometer with parallel plates (TA Instruments).

Results

The high porosity was verified by both confocal and SEM imaging. The pore size of scaffolds was around 15 μm, which was significantly greater than most available hydrogels, such as alginate, gelatin, and covalently crosslinked chitosan. Human vocal fold fibroblasts were observed to migrate and proliferate within the 3D matrix. In terms of mechanical properties, the hydrogel system exhibited a strong pH-controlled stiffness. The storage modulus was ~700 Pa at pH = 6.5, ~5.5 kPa at pH = 6.8, and ~25 kPa at pH = 7.1. The stiffness range here covers most needs for soft tissue engineering. Meanwhile, at the same pH condition, the addition of 4% free PEG decreased the hydrogel relaxation time by almost 30-fold without affecting the original hydrogel stiffness. The proposed hydrogel system also presented great 3D printability and can be used to fabricate complicated structures, such as vocal fold M5 model.

Conclusions

The presented hydrogel system showed great potential for customized vocal fold repair. The highly porous structure and orthogonally tunable stiffness and viscoelasticity can also be applied to other tissue engineering fields.

Acknowledgments: This research was supported by grant number R01DC014461 (Xinqiao Jia, PI) and R01DC005788 (Luc Mongeau, PI) from the National Institute on Deafness and other Communication Disorders.

References

1. Heris, H.K.; Latifi, N.; Vali, H.; Li, N.; Mongeau, L. *Procedia Eng.* **2015**, *110*, 143.
2. Latifi, N.; Asgari, M.; Vali, H.; Mongeau, L. *Sci. Rep.* **2018**, *8*, doi:10.1038/s41598-017-18523-3.

6.8. The Effects of Laryngeal Massage and Nebulized Saline on High-Voice Users

Matti Groll^{1,2}, Daniel Buckley^{1,3}, Kimberly Dahl¹ and Cara Stepp^{1,2,3}

¹ Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA

² Department of Biomedical Engineering, Boston University, Boston, MA, USA

³ Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, Boston, MA, USA

Keywords: voice therapy; laryngeal massage; vocal effort

Objectives

This study investigates the effects of two common voice therapy techniques on vocal function, in a population of high-voice users with voice complaints.

Introduction

Laryngeal massage (LM) to reduce muscular tension and nebulized saline (NS) to promote vocal fold hydration are both common therapy techniques that have been shown to result in vocal improvements in individuals with voice disorders (1, 2). High-voice users often report voice complaints consistent with voice disorders, such as dysphonia, increased levels of effort, and soreness of the throat (3). However, it is unclear whether LM or NS can improve vocal function in these individuals with

less severe symptoms. Here we examine the short-term effects of both interventions on individuals' self-perception of vocal tract discomfort (VTD) and vocal effort (VE), as well as auditory-perceptual, acoustic, and videoendoscopic measures of their resulting vocalization.

Methods

Participants were selected based on survey responses, which indicated that they had high voice use and vocal symptoms. Currently, eleven participants (4 male, $M = 22.5$, $SD = 3.31$) have been enrolled. Each participant experiences the two therapy techniques over two separate days, on average one week apart, with the order randomized across participants. Subjective measures of VTD and VE, speech acoustics, and high-speed videoendoscopy (HSV) are recorded before and after each technique is administered, resulting in four time-points of data for each participant. Both LM and NS are administered by a certified speech language pathologist for a total of fifteen minutes. Five additional participants are expected to complete the study, for a total of sixteen participants. Outcome measures include: VTD, VE, auditory-perceptual measures, acoustic measures (relative fundamental frequency and cepstral peak prominence), and quantitative measures of laryngeal function based on HSV. Following the conclusion of data collection, a two factor repeated measures analysis of variance (ANOVA) will be completed for each outcome measure with pre-/post-intervention and the type of technique as factors.

Results

Preliminary investigation of the trends from the eleven completed participants suggests that VTD and VE decrease following both forms of therapy. Auditory-perceptual and acoustic outcome changes are inconsistent based on this small sample size. HSV data processing is ongoing.

Conclusions

Based on preliminary results, the subjective symptoms (VTD and VE) in this population appear to improve following both therapies. However, currently the results from auditory-perceptual and acoustic measures are inconclusive. Furthermore, based on these data, there are no substantive differences between the effects of LM and NS.

Acknowledgments: This work was supported by the National Institutes of Health grant R01DC015570 (CES) from the National Institute on Deafness and Other Communication Disorders.

References

1. Roy, N.; Ford, C.; Bless, D. Muscle tension dysphonia and spasmodic dysphonia: the role of manual laryngeal tension reduction in diagnosis and management. *Ann. Otol. Rhinol. Laryngol.* **1996**, *105*, 851–856.
2. Tanner, K.; Nissen, S.L.; Merrill, R.M.; Miner, A.; Channell, R.W.; Miller, K.L.; Elstad, M.; Kendall, K.A.; Roy, N. Nebulized isotonic saline improves voice production in Sjogren's syndrome. *Laryngoscope* **2015**, *125*, 2333–2340.
3. Martins, R.; Pereira, E.; Hidalgo, C.; Tavares, E. Voice disorders in teachers. A review. *J. Voice* **2014**, *28*, 716–724.

6.9. Investigating the Pathobiology of Vocal Fold Dehydration and Rehydration

Abigail Durkes¹ and Preeti M. Sivasankar²

¹ Comparative Pathobiology, Purdue University, West Lafayette, IN, USA

² Speech, Language and Hearing Sciences, Purdue University, West Lafayette, IN, USA

Keywords: voice; physiology; animal models

Objectives

We will present data addressing the complex relation between altered hydration state of the body and the effects of altered hydration state on vocal fold tissue. The objectives of this study include (1) identifying the optimal animal model for investigations of systemic dehydration and systemic rehydration, (2) developing a reliable and physiologically relevant methodology of systemic dehydration, (3) developing markers of systemic dehydration and (4) demonstrating whether systemic dehydration induces vocal fold dehydration through a combination of methodologies.

Introduction

A central tenet in voice physiology is that optimal hydration is necessary for maintaining healthy vocal folds. Data from excised vocal folds support this assertion however in vivo evidence is inconclusive. The variable findings between ex vivo and in vivo models are likely explained by the presence of homeostatic mechanisms that regulate water balance in live subjects that are not available in excised tissue. Our programmatic research goal is to investigate the complex relationship between altered hydration state and vocal function in vivo, and to evaluate whether optimal hydration promotes healthy vocal folds.

Specific to this presentation, we will focus on showing data on establishing markers of systemic dehydration and that systemic dehydration can induce vocal fold dehydration. Data will be presented from two animal models: rat and rabbit. Methodologies for two different systemic dehydration protocols will also be shared: water withholding and diuretic injection.

Methods

Sprague Dawley (SD) rats (males and females) and New Zealand White rabbits (males) were used for systemic dehydration and rehydration studies. Methodologies include parameters of hemoconcentration, proton-density weight MRI at variable dehydration status, genomic analysis, and histopathology.

Results

Proton-density weighted MR imaging can be used to demonstrate that systemic dehydration (defined by body weight loss) reliably induces vocal fold dehydration as detected by signal intensity changes. These changes however, are only detected at high body weight loss levels (>6% body weight loss). Acute episodes of systemic dehydration do not produce reliable, adverse pathological changes to the vocal folds as assessed by histopathology and gene expression studies. Access to water does not induce rehydration if defined by body weight.

Conclusions

There are challenges in inducing physiologically-relevant systemic dehydration and developing a robust and reliable animal model to study the pathobiology of vocal fold dehydration. A combination of techniques are necessary to confirm that dehydration, in an otherwise healthy animal, is occurring; and that dehydration of the body is dehydrating the vocal folds. The sequelae of chronic versus acute dehydration, and systemic dehydration versus surface dehydration also need to be parsed out.

Acknowledgments: We acknowledge the contributions of Professor Sarah Calve and Professor Zhongming Liu, both at the Weldon School of Biomedical Engineering, Purdue University. This work is funded by a R01 grant from the NIDCD to Durkes and Sivasankar.

6.10. Increased Laryngeal Mucosal Cellular Proliferation in Mice Exposed Short-Term to Cigarette Smoke

Elizabeth Erickson-DiRenzo, Meena Easwaran and Joshua Martinez

Department of Otolaryngology—Head & Neck Surgery, Stanford University School of Medicine, Palo Alto, CA, USA

Keywords: laryngeal mucosa; cellular proliferation; cigarette smoke; mucus production

Introduction/Objectives

The larynx is a vital organ situated at the divergence of the upper and lower respiratory tract that helps coordinate important human functions including swallowing, breathing, coughing and voice production.¹ Despite the high frequency of benign and malignant tobacco product-induced laryngeal diseases, the mechanisms by which cigarette smoke (CS) affects the health of the larynx have been largely unexplored. In order to identify key cellular events that result in smoke-induced laryngeal disease, we examined the early responses of the laryngeal mucosa to cigarette smoke-induced injury. Specifically, we investigated cell proliferation and mucus production in the murine larynx in order to determine the histologic regions of greatest mucosal injury after short-term exposure to mainstream cigarette smoke.

Methods

Adult C57BL/6J male mice were assigned to a cigarette smoke exposure (CSE), reversibility (REV), or air exposed control group. CSE mice were exposed to CS for 2 h/day for 1, 5 or 10 days using the SCIREQ inExpose inhalation system (Montreal, Canada). REV mice were exposed to CS for 5 days then air-exposed for an additional 5 days. All mice were administered 5-Bromo-2'-deoxyuridine (BrdU), 2 h before euthanasia. Laryngeal tissues were harvested and stained with Alcian blue/Periodic acid schiff (AB/PAS) to evaluate mucus production and BrdU to label proliferative cells.

Results

The AB/PAS-stained mucus-producing glands of the laryngeal subglottis had a phenotypic area expansion but were not significantly increased in area in the CSE or REV groups. Area of AB-positive acidic and PAS-positive neutral mucus levels were unaltered in CSE group at days 1 and 5. However, acidic mucus significantly decreased in the CSE and REV groups ($p \leq 0.0001$) and neutral mucus increased in CSE group ($p \leq 0.05$) on day 10. Cellular proliferation expressed as a measure of BrdU labeling was significantly increased in the vocal folds and subglottic region. Specifically, the number of BrdU-positive cells was increased in the vocal folds of the CSE group at 1 day ($p < 0.001$) and 5 days ($p < 0.01$) as compared to controls. BrdU-positive cells were increased in the subglottic region of the CSE group at 5 days ($p < 0.0001$) and 10 days ($p < 0.01$) as compared to controls. Proliferative index in the subglottic region of the REV group was lower than the CSE group ($p < 0.01$) and comparable to the controls. Simple hyperplasia, or thickening, of the vocal fold or subglottic epithelium was not observed.

Conclusions

Overall, short-term mainstream CS exposure promotes cellular proliferation in the laryngeal mucosa and alters mucus production. Rates of proliferation were differentially affected depending on larynx site with increased BrdU-positive cells identified earlier in the vocal folds as compared to the subglottic region. These findings suggest short-term exposure to cigarette smoke may induce cytotoxicity and subsequent regeneration within the laryngeal mucosa.² Findings further help identify the location and types of cells at risk for injury following long-term exposures and provide insights into the cellular mechanisms underlying tobacco-product induced laryngeal diseases.

Acknowledgments: We acknowledge HISTOWIZ (Brooklyn, NY) for their assistance with histological processing of laryngeal samples.

References

1. Thibeaut, S.; Rees, L.; Pazmany, L.; Birchall, M.A. At the crossroads: Mucosal immunology of the larynx. *Mucosal. Immunol.* **2009**, *2*, 122–128.
2. Dodmane, P.; Arnold, L.L.; Pennington, K.L.; Cohen, S.M. Orally administered nicotine induces urothelial hyperplasia in rats and mice. *Toxicology* **2012**, *315*, 49–54.

6.11. Effects of Voice Changes under Testosterone Therapy on Listener Perception of Gender: A Transgender Case Study

Kimberly L. Dahl¹, Gabriel J. Cler², Victoria S. McKenna¹ and Cara E. Stepp^{1,2,3,4}

¹ Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA

² Graduate Program in Neuroscience, Boston University, Boston, MA, USA

³ Department of Biomedical Engineering, Boston University, Boston, MA, USA

⁴ Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, Boston, MA, USA

Keywords: transgender; transmasculine; acoustics; perception

Objective: This study aimed to measure the effects of changes in fundamental frequency (f_0) and estimated vocal tract length on listener perceptions of a transmasculine speaker's gender over a one-year course of testosterone therapy.

Introduction

The scarcity of research on transmasculine voice prevents clinicians from providing transmasculine speakers with accurate prognoses of voice changes under testosterone. Few studies (e.g., [1]) have tracked these voice changes longitudinally or assessed how listeners perceive the gender of transmasculine speakers [2,3]. None has measured how listener perceptions may change as transmasculine speakers undergo testosterone therapy. A fuller understanding of expected voice changes and their potential impact on the perceptions of others is necessary to best serve transmasculine clients with an evidence-based approach.

Methods

A 30-year-old transmasculine individual underwent voice assessments three times prior to starting testosterone therapy and every two weeks thereafter for a year. The speaker's f_0 during reading and formants during sustained vowels were collected at each timepoint. The fourth vowel formant (F4) was used to estimate vocal tract length. In a single-session perceptual experiment, 8 inexperienced listeners (4 cisgender female, 4 cisgender male; $M = 20.9$ years, $SD = 2.8$ years) provided ratings of the participant's gender for each timepoint. Listeners based their ratings on excerpts of the participant reading the rainbow passage and marked them on a 100-mm visual analog scale ranging from "definitely male" to "definitely female." Spearman's rank-order correlation coefficients were calculated to measure the relationships between listener perception of gender and f_0 and estimated vocal tract length.

Results

The participant's f_0 during reading decreased from 183 Hz at baseline to 134 Hz during the final three sessions. The fourth formant had a mean value of 3693 Hz at baseline and 3554 Hz during the final three sessions. These correspond to an estimated vocal tract length (VTL) of 16.3 cm at baseline and 16.9 cm in the final three sessions. The mean gender perceptual rating of the participant's voice was 96.2 ($SD = 6.3$; 0 = definitely male, 100 = definitely female) during the baseline sessions and

30.0 ($SD = 26.0$) during the final sessions. All changes were significant ($p < 0.05$). The participant was reliably identified as female (≥ 65) through the first 15 weeks of testosterone therapy and reliably identified as male (≤ 35) after 37 weeks. Gender perceptual ratings correlated strongly with f_o ($r = 0.908$, $p < 0.05$) and moderately with estimated vocal tract length ($r = -0.647$, $p < 0.05$).

Conclusions

Over one year on testosterone, a transmasculine speaker's f_o during reading dropped to a typical male range, and his vocal tract lengthened by 0.6 cm, per estimates derived from F4 values. The latter suggests that the participant's larynx may have dropped or tilted as during cisgender male puberty. Finally, listeners consistently attributed a male gender to the participant after 37 weeks on testosterone. This shift in gender perception may be influenced by changes in f_o and estimated vocal tract length.

Acknowledgments: This work was supported by grants DC015570, DC013017, and DC014872 from the National Institute on Deafness and Other Communication Disorders.

References

1. Nygren, U.; Nordenskjöld, A.; Arver, S.; Södersten, M. Effects of voice fundamental frequency and satisfaction with voice in trans men during testosterone treatment: A longitudinal study. *J. Voice* **2016**, *30*, 766.e23–766.e34.
2. Scheidt, D.; Kob, M.; Willmes, K.; Neuschaefer-Rube, C. Do we need voice therapy for female-to-male transgenders? In *2004 IALP Congress Proceedings*; Brisbane, Australia, 2004. Available online: <https://www.researchgate.net/publication/240743853> (accessed on).
3. Van Borsel, J.; de Pot, K.; De Cuypere, G. Voice and physical appearance in female-to-male transsexuals. *J. Voice* **2009**, *23*, 494–497.

7. Poster Session 3

7.1. The Use of Nasalance for Voice Stabilisation during the Tenors' Passaggio

Matthias Echternach¹, Michael Döllinger², Catalina Högerle¹, Marie-Anne Kainz¹, Marie Köberlein³ and Bernhard Richter³

¹ Division of Phoniatics and Pediatric Audiology, Department of Otorhinolaryngology, LMU, Munich, Germany

² Division of Phoniatics and Pediatric Audiology, Department of Otorhinolaryngology, Erlangen University, Germany

³ Institute of Musicians' Medicine, Freiburg University, Germany

Keywords: voice; singing; passaggio; high speed videoendoscopy

Introduction

The passaggio of untrained voices is characterized by sudden pitch jumps due to nonlinear properties of the phonatory system. Professional tenors are able to stabilize vocal fold oscillations during the passaggio. However, the mechanisms for stabilization are not clarified, yet. In a historical study by Trendelenburg in 1937 and also in a recent study by Sundberg et al., it was shown that increased nasalance could stabilize the voice within the passaggio. The present study aims to analyze if nasalance is frequently used for stabilization of vocal fold oscillations during the passaggio.

Methods

In this prospective study eight vocally healthy professional tenors were asked to perform pitch glides from A3 (f_o 220 Hz) to A4 (f_o 440 Hz) on the vowel /i/ (1) with a register shift from modal

to falsetto and (2) from modal with continuation to the stage voice above the passaggio (SVaP). During the experiment transnasal highspeed videoendoscopy (HSV, 20.000 fps) with simultaneous electroglottography (EGG), audio, accelerometer, oral and nasal flow signals were simultaneously recorded. As in previous studies, detection of EGG derived sample entropy was used to verify greatest instability during the transition.

Results and Conclusions

For almost all voices the transition to SVaP showed greater vocal fold oscillatory stability than the transition to falsetto. However, only a minority increased nasalance during the passaggio for the modal to SVaP task. In some other subjects there was a greater supraglottic compression visible during the passaggio. As a consequence, it seems that nasalance is only used in some of the subjects in order to stabilize vocal function during the passaggio.

7.2. Numerical Analysis of the Airflow Downstream from a Tracheoesophageal Voice Prosthesis

Fernando H. T. Santos, André M. C. Tourinho and Andrey R. da Silva

Department of Mechanical Engineering, Federal University of Santa Catarina, Florianópolis, Santa Catarina, Brazil

Keywords: voice prosthesis; tracheoesophageal speech; voice aerodynamics

Objectives

Investigate the influence of position and angle of a typical voice prosthesis on the voice production of tracheoesophageal speakers by changing the characteristics of the flow in the pharyngoesophageal segment.

Introduction

The tracheoesophageal prosthesis is probably the most appealing alternative of voice recovery for patients who have undergone a total laryngectomy, when considering voice quality and voice control. The vibration of the pharyngoesophageal segment—the main voice source—will be highly influenced by the flow within the prosthesis and by the flow structures downstream from the prosthesis outlet. Previous works have investigated the pressure drop across different prosthesis designs with both in-vitro and in-vivo experiments. However, these studies do not provide much information on the correlation between pressure drop and flow behavior downstream from the prosthesis. Furthermore, the aerodynamic aspects of the flow in the esophageal region have been only investigated for an idealized geometry representing the tracheoesophageal system. In the present study, the pressure drop between the trachea and the esophagus, as well as the distribution of aerodynamic forces on the walls of the pharyngoesophageal segment are investigated as a function of the prosthesis position and angle.

Methods

A numerical model based on a finite volume scheme is used to analyze the airflow through the tracheoesophageal system. The solution is obtained with a RANS-based solver using a $k-\epsilon$ turbulence model. The approach is validated with the experimental results provided by an idealized model proposed by Erath and Hemsing (2016). After validating the approach, a new computational model of the tracheoesophageal system was created based on tomographic images of patients who have undergone a total laryngectomy. In this model, the representation of the voice prosthesis was based on the Provox 2 prosthesis, developed by Athos Medical. After validating the numerical scheme, the pressure drop and the distribution of aerodynamic forces on the pharyngoesophageal segment were assessed for different valve positions and angles.

Results

The results obtained for the pressure drop and the aerodynamic force distribution in the pharyngoesophageal segment indicate that variations on the valve position and angle produce considerable variations on these two parameters. Moreover, it was observed that the position of the valve gate cannot be overlooked.

Conclusions

The results have shown that the position and angle of the voice prosthesis change the pressure drop and the distribution of aerodynamic forces in the pharyngoesophageal segment. This suggests that controlling the position of the valve may significantly affect the voice production mechanism of tracheoesophageal speakers. Nevertheless, further investigations involving different prosthesis types and a dynamic model of the pharyngoesophageal system are necessary before establishing a general protocol for the positioning voice prostheses.

Acknowledgments: The authors would like to acknowledge the financial support provided by CAPES, CNPq and FINEP, and Byron D. Erath for providing the experimental data obtained in his studies.

Reference

1. Erath, B.D.; Hemsing, F.S. Esophageal aerodynamics in an idealized experimental model of tracheoesophageal speech. *Exp. Fluids* **2016**, *57*, 34.

7.3. Beneficial Effects of Choral Singing on Speech and Voice in Normal Aging

Valérie Brisson ^{1,2}, Maxime Perron ^{1,2}, Émilie Belley ^{1,2}, Lisa-Marie Deschênes ^{1,2}, Julie Poulin ^{1,2}, Johanna-Pascale Roy ³, Josée Vaillancourt ⁴, Philip Jackson ^{2,5} and Pascale Tremblay ^{1,2}

¹ Département de Réadaptation, Université Laval, Quebec City, Canada

² CERVO Brain Research Centre, Quebec City, Canada

³ Département de langues, linguistique et traduction, Université Laval, Quebec City, Canada

⁴ Faculté de Musique, Université Laval, Quebec City, Canada

⁵ École de psychologie, Université Laval, Quebec City, Canada

Keywords: choral singing; communication; aging

Objectives

The main objective of this study is to clarify the protective effects of choral singing on communication, including voice, articulation and prosody, during normal aging.

Introduction

Aging is associated with multiple changes that affect voice, speech, language and hearing. It is well established that, with age, the voice becomes less stable and less intense, and that pitch undergoes important changes [1]. Articulation also declines: the speech of elderly adults is less accurate, slower, and more variable [2]. There is also some evidence to suggest that older adults are less expressive. Moreover, evidence from our team suggests that voice-related changes are perceived and can have an impact on social interactions [3]. Yet, very few studies have investigated potential prevention or compensation strategies against these age-related declines. One such avenue is choral singing, a universal and agreeable social activity that has been shown to have beneficial impacts on voice, and, to a lesser extent, speech as well [4]. However, the nature and extent of the protective effects of singing on communication remain to be clarified.

Methods

142 healthy adults aged 20 to 98 years ($M = 53.0 \pm 12.16$) with no speech or voice disorders and no gastric reflux were recruited. Voice, speech, hearing and prosody were evaluated through several tasks: (1) reading aloud a standardized passage (*La bise et le soleil*), (2) producing vowels at different intensities and pitches, (3) narrating three personal stories with different valences (positive, negative, neutral), and (4) repeating two-syllables non-words with various levels of complexity.

To evaluate voice and prosody, acoustical parameters (amplitude, jitter, shimmer, pitch) were extracted through *Praat* scripts from the vowel, stories and passage tasks. To evaluate speech, non-words were transcribed to the phonetic alphabet, and the number of errors and reaction time were calculated. Linear mixed model and moderation analyses were conducted to examine the impact of age and singing on voice, articulation and prosody with each parameter (e.g., maximal intensity) as dependant variable and Age and Choral singing as independent variable.

Preliminary Results

Analyses are still underway but preliminary results indicate that, with age, there is a reduction in the distinctiveness of the acoustical signatures of the emotional voice, especially the happy voice, in terms of maximal pitch ($p \leq 0.05$), and maximal intensity ($p \leq 0.05$), with potentially important effects on social communication. Moreover, pitch ($p \leq 0.05$) and intensity ($p \leq 0.05$) of the neutral voice becomes more variable with age, but only in non-singers. The vowel tasks also revealed decline in pitch range and maximal pitch. However, pitch range, minimum and maximum pitch were better in singers compared to non-singers, regardless of age. Additional analyses are still underway.

Conclusions

Choral singing is a widely accessible potential activity to prevent age-related communication difficulties. Our preliminary analyses suggest benefits to voice functions. Analyses of speech data is underway, but it is expected that singing will be associated with increased rate and better articulation.

Acknowledgments: This project was funded by grants from the Drummond Foundation, SSHRC, and NSERC.

References

1. Lortie, C.L.; Rivard, J.; Thibeault, M.; Tremblay, P. *J. Voice* **2016**, *31*, 112.e1–112.e12.
2. Bilodeau-Mercure, M.; Kirouac, V.; Langlois, N.; Ouellet, C.; Gasse, I.; Tremblay, P. Movement sequencing in normal aging: speech, oro-facial, and finger movements. *Age* **2015**, *37*, 9813.
3. Lortie, C.L.; Thibeault, M.; Guitton, M.J.; Tremblay, P. *JSLHR* **2018**, *61*, 227–24.
4. Fogg-Rogers, L.; Buetow, S.; Talmage, A.; McCann, C.M.; Leao, S.H.; Tippett, L.; Purdy, S.C. *Disabil. Rehabil.* **2015**.

7.4. Esophageal Wall Compliance and Its Influence on the Driving Pressures of Tracheoesophageal Speech

Byron D. Erath¹ and Sean D. Peterson²

¹ Department of Mechanical and Aeronautical Engineering, Clarkson University, Potsdam, NY, USA

² Department Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON, Canada

Keywords: tracheoesophageal speech; voice remediation; lumped element model; laryngectomy

Objectives

The objective of this work is to determine how compliance of the esophageal tract during tracheoesophageal speech (TES) impacts the unsteady pressure within the esophagus, which ultimately drives the flow that produces sound.

Introduction

TES, the gold standard for voice remediation following laryngectomy, produces sound by redirecting airflow from the trachea into the esophagus via a tracheoesophageal prosthesis (TEP), which then passes through the pharyngoesophageal segment (PES) producing self-sustained oscillations; this modulates the airflow and produces sound, which is formed into intelligible speech via posturing of the lips, tongue, soft palate, etc.¹ Unfortunately, success rates of TES are surprisingly low and, when successful, the sound that is produced is of very poor quality.² Prior work has indicated that the esophageal pressure field influences successful TES sound production.³ Unfortunately, the interaction between the esophageal pressure field and the compliance of the walls of the esophagus during TES and its impact on sound product remain unclear and uninvestigated.

Methods

An electrical circuit analogy is adopted to model flow from the trachea, through the TEP into the esophagus, and through the PES. Flow resistance and inertance, and esophageal wall resistance and compliance are modeled using 1-D circuit elements. Flow resistance through the PES is determined by coupling the circuit model to a self-oscillating, lumped-element model of PES vibration and solving both systems simultaneously. The esophageal diameter (volume) and pressure are then computed as a function of wall compliance over a range of values representative of human physiology, and for various representations of PES muscle tension (e.g., hyper- versus isotonicity).

Results

The second-order response of the system produces a highly-unsteady esophageal pressure that is strongly-dependent upon the wall compliance of the esophagus, with resonance conditions capable of generating esophageal pressures that are much greater than the magnitude of the driving lung pressure. Interestingly, over the range of normal compliance values, largely periodic, ordered oscillations of the PES are observed. However, for abnormal values of esophageal compliance, irregular and chaotic vibratory behavior of the PES develops, indicating that pathological conditions (e.g., esophageal hardening due to radiation treatments) may have adverse impacts on the ability to successfully produce TES. Modeling hypertonicity of the PES is also shown to have a significant impact on the coupling between the esophageal pressure and the resultant PES oscillation.

Conclusions

The results present, for the first time, a description of how the esophageal pressure is highly unsteady, and is significantly impacted by compliance of the esophageal wall. Esophageal pressures are found to reach magnitudes greater than the driving lung pressure due to resonance of the system. Future research directions include incorporating sound production to quantify acoustical impacts.

References

1. Singer, M.; Blom, E.D. An endoscopic technique for restoration of voice after total laryngectomy. *Ann. Otol. Rhinol. Laryngol.* **1980**, *89*, 529–532.
2. Blom, E.D.; Singer, M.; Hamaker, R. Pharyngeal plexus neurectomy for alaryngeal speech. *Laryngoscope* **1986**, *96*, 50–54.
3. Erath, B.D.; Hemsing, F.S. Esophageal aerodynamics in an idealized experimental model of tracheoesophageal speech. *Exp. Fluids* **2016**, *57*, 34.

7.5. On the Role of Simultaneous Observations for a Bayesian Estimation of Subglottal Pressure and Laryngeal Muscle Activation

Gabriel A. Alzamendi ¹, Sean D. Peterson ², Byron D. Erath ³ and Matías Zañartu ¹

¹ Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

² Department of Mechanical and Mechatronics Engineering, University of Waterloo, ON, Canada

³ Department of Mechanical & Aeronautical Engineering, Clarkson University, Potsdam, NY, USA

Keywords: Bayesian estimation; lumped mass model; subject-specific modeling; lung pressure; muscle activation

Objective

To assess the relevance of incorporating simultaneous signal observations in the framework of a Bayesian estimation of subglottal pressure and muscle activation via a state space representation of a reduced-order computational model of voice production.

Introduction

Computational models of voice production can provide access to parameters and measures that are difficult, if not impossible, to obtain with current clinical technologies. As such, recent efforts have been devoted to estimate subglottal pressure from kinematic observations using high speed video via optimization methods [1]. The Bayesian estimation framework offers a probabilistic view of the estimation problem and allows for “virtual sensing” of additional signals through the subject-specific model construction [2]. In this study, we explore Bayesian inference using a state space method of the complete voice production system to estimate lung pressure and laryngeal muscle activation using both glottal area and airflow signals as system observations, to assess the effect of adding glottal airflow signal in the estimation process.

Methods

Bayesian estimation is performed through an extended Kalman filter (EKF) [3]. A three-mass body-cover model is applied for describing vocal fold dynamics, and physiologically-inspired rules are considered to simulate the effects of cricothyroid (CT) and thyroarytenoid (TA) muscles [4]. The three-way coupling between vocal fold oscillations, airflow, and acoustic pressures is also accounted for in the proposed method. An EKF is applied for estimating subglottal pressure (P_L) and activation levels for CT and TA from noisy simulated signals generated using the body-cover model.

Results

Results using synthetic data illustrates that combining different observations helps to improve the estimates, especially for lung pressure. Using solely glottal area yields biased estimates and broad confidence intervals that are improved when glottal airflow is accounted for. Percent RMS estimation errors are computed, where median (IRQ) measures are reported: 4.24% (2.85%) for P_L , 3.09% (2.26%) for CT, and 6.23% (6.61%) for TA.

Conclusions

Current results indicate that combining different biomedical signals leads to more accurate and precise estimates of clinical data, which is a feature of Bayesian estimation. That said, not all the desired estimates benefit in the same way. Additional efforts applying the proposed method to clinical data are currently underway, as well as the analysis of signal degradation on the estimation procedure.

7.6. Comparing Accelerometer and Oral Airflow Based Aerodynamic Measures in Patients with Vocal Hyperfunction

Víctor M. Espinoza ^{1,2}, Daryush D. Mehta ^{3,4,5}, Jarrad H. Van Stan ^{3,4,5}, Robert E. Hillman ^{3,4,5} and Matías Zañartu ¹

¹ Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaiso, Chile

² Department of Sound, Universidad de Chile, Santiago, Chile

³ Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, MA, USA

⁴ MGH Institute of Health Professions, Massachusetts General Hospital, Boston, MA, USA

⁵ Harvard Medical School, Harvard University, Boston, MA, USA

Keywords: voice assessment; aerodynamic measures; vocal hyperfunction; neck skin acceleration

Objective

To determine if aerodynamic measures derived from Impedance-based Inverse Filtering (IBIF) of the neck skin acceleration are comparable to those obtained from intraoral pressure, radiated sound pressure, and oral airflow for discriminating between normal subjects and patients with vocal hyperfunction.

Introduction

Laboratory estimates of glottal aerodynamic parameters obtained from measurements of oral volume velocity (OVV) and intraoral air pressure (IOP) have been shown to discriminate between matched controls and two types of vocal hyperfunction, phonotraumatic (PVH) and nonphonotraumatic (NPVH), thus providing important insights into the pathophysiology of these disorders [1]. Previous work has also shown that it is possible to obtain estimates of glottal aerodynamic measures from the type of neck-placed accelerometer (ACC) that is used for voice monitoring [2,3,4] which could greatly enhance the clinical utility of ambulatory monitoring. An important step in the clinical development of this approach is to determine whether ACC-based estimates of glottal aerodynamic measures are comparable to those from OVV and IOP in terms of discriminating between healthy controls and patients with PVH or NPVH.

Methods

Participants were two groups of adult females with PVH (vocal fold nodules or polyps, $n = 16$) or NPVH (primary muscle tension dysphonia, $n = 14$) and two groups of matched (normal) controls. Each subject produced strings of five consecutive /pae/ syllables tokens using comfortable and loud (approximately 6 dB increase) voice while synchronized recordings of oral airflow, radiated sound pressure, intraoral pressure and neck skin acceleration were obtained [1]. The IBIF algorithm was calibrated for every token to yield estimates of peak-to-peak glottal airflow (ACFL), maximum flow declination rate (MFDR) and open quotient (OQ). Subglottal pressure (SGP) and Sound Pressure Level (SPL) were estimated from the ACC signal using subject-specific linear regression models [4]. These estimates were compared with SPL-Normalized measures derived from OVV and IOP signals.

Results

Statistically significant differences were found between PVH and control subjects in the comfortable and loud voice conditions for both the ACC and OVV/IOP-based measures of ACFL, OQ and SGP—significant differences in OVV-based MFDR were not found using IBIF analysis of the ACC signal. Statistically significant differences were found between NPVH and control subjects in the comfortable voice condition for both the ACC and OVV/IOP-based measures of OQ and SGP—this

was also true for OQ in the loud voice condition, but not for SGP which was only significantly different based on IBIF analysis of the ACC signal.

Conclusions

SPL-Normalized aerodynamic measures [1] derived from the ACC are mostly comparable to those obtained from OVV/IOP-based in terms of discriminating between patients with vocal hyperfunction and matched healthy controls. These findings have the potential to enhance the clinical utility of ambulatory monitoring by providing better insights into the pathophysiologic mechanisms associated with disorders in which daily voice use is assumed to play a role.

Acknowledgments: This research was supported by the NIDC (grants R21 DC015877 and P50 DC015446), CONICYT (grants FONDECYT 1151077 and BASAL FB0008), UTFSM (FSM1204), and UChile (PEEI and VID Grants). The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

References

1. Espinoza, V.M.; et al. *J. Speech Lang. Hear. Res.* **2017**, *60*, 2159–2169.
2. Zañartu, M.; et al. *IEEE Trans. Audio Speech Lang. Proc.* **2013**, *21*, 1929–1939.
3. Cortes, J.P.; et al. *PLoS ONE* **2018**, *13*, e0209017.
4. Fryd, A.S.; et al. *J. Speech Lang. Hear. Res.* **2016**, *59*, 2159–2169.

7.7. Development of a Vocal Warm Up Protocol for Vocal Fatigue Prevention

Elaine Kwong, Suen Yue Sarah Poon, Cheuk Yiu Tse and Yuk Wun Natalie Yeung

Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong, China

Keywords: voice; vocal fatigue prevention; vocal warm up; semi-occluded vocal tract

Objectives

The present study aimed to develop a vocal warm up protocol that is practical, feasible and evidence-based for the prevention of vocal fatigue.

Introduction

Vocal fatigue is frequently experienced by occupational voice users who have consistently high demand on their voice. It may be presented as distorted voice quality, dynamic range and pitch range; reduced respiratory support; increased muscular and structural tension or discomfort; reduced vocal mechanism control; and/or increased vocal effort (Welham & Maclagan, 2003). Regular vocal fatigue is believed to be associated with functional and/or organic voice disorders. One of the theories regarding vocal fatigue suggests that prolonged voice use would result in increase in viscosity of the vocal folds, which will in turn increase the friction and heat dissipation in the vocal folds during vibration. Besides that of the vocal folds, viscosity of the intrinsic laryngeal muscles may also be increased after prolonged voicing (Titze, 2006). From the literature sport sciences, increase in muscle viscosity may be avoided through warm up exercises. Vocal warm up exercises had been practiced in the field of singing for decades. Judiciously warming up the laryngeal musculature is believed to be able to prevent vocal fatigue. Straw phonation, one of the semi-occluded vocal tract (SOVT) variations, was employed as a kind of physiological vocal warm up exercise in the present study.

Methods

Twenty adult subjects of both genders who are vocally healthy participated in the study. Each subject attended seven data collection sessions. In each session, subjects practiced straw phonation for various durations (i.e., 1 min, 3 min, and 5 min) and the straw might end in air or water. Vocal fatigue was induced by requiring participants to read continuously with a loudness of 10dB above the participants' habitual loudness for 90 min. Preventive effect of the straw phonation exercises was compared with voice rest and outcome measures included phonatory threshold pressure (PTP), perceived phonatory effort (PPE), acoustic analysis and auditory-perceptual analysis of voice quality. Measurements were taken in the time points: (1) prior to voice rest/straw phonation, (2) after voice rest/straw phonation but prior to the fatigue-inducing task, and (3) after the fatigue-inducing task.

Results

Preliminary results shows that PTP and PPE were maintained, despite the fatigue-inducing task, if the subject had practiced straw phonation that ends in water for 3 min and 5 min. The changes from time points (1) to (3) in PTP and PPE is different in the above straw phonation conditions as compared to the voice rest condition ($p \leq 0.016$).

Conclusions

Practicing straw phonation that ends in water for 3 and 5 min have better preventive effects for vocal fatigue than voice rest. This study lays the groundwork for future studies to confirm the effectiveness of straw phonation as a vocal warm up exercise in occupational voice users.

Acknowledgments: Nil.

References

1. Titze, I.R. Voice training and therapy with a semi-occluded vocal tract: rationale and scientific underpinnings. *J. Speech Lang. Hear. Res.* **2006**, *49*, 448–459.
2. Welham, N.V.; Maclagan, M.A. Vocal fatigue: current knowledge and future directions. *J. Voice* **2003**, *17*, 21–30.

7.8. Evaluation of Anti-Fibrotic Activity of Wound Healing Macrophages in a 3D In Vitro Model for Vocal Fold Scar Treatment

Sepideh Mohammadi and Luc Mongeau

Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

Keywords: vocal fold scar; macrophages; fibroblasts; 3D cell culture; hydrogel

Objectives

The aim of the current study is to investigate how encapsulated vocal fold fibroblasts and different types of wound healing macrophages interact within hydrogel scaffolds to promote or prevent scar formation.

Introduction

There is plentiful evidence that the immune system regulates the healing response of various tissues following injury, including the degree of scarring, and the restoration of the organ structure and function [1]. Monocytes and macrophages are key players within the innate immune system. Following tissue damage, macrophages direct the restoration of tissue homeostasis the recruitment of other immune cells to the site of infection, the clearance of pathogens and dead cells by phagocytosis, and the synthesis of multiple cytokines and growth factors. Macrophages are also involved in

angiogenesis, organ regeneration, and tumor growth. They are also known for modulating fibrosis and scarring during different stages of wound healing [1,2]. In the context of scar treatment, many studies have confirmed the anti-fibrotic contribution of certain types of macrophages and their stimulators in preventing, limiting or even reversing scars in various organs [3]. However, the beneficial or detrimental effects of macrophages in the scarring of injured vocal fold, are not yet clear.

Methods

Human monocytic THP-1 cells were cultured in Roswell Park Memorial Institute medium (RPMI 1640) culture medium supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1% penicillin-streptomycin. THP-1 monocytes were differentiated into inactivated macrophages (M0) by 48 h culturing at 0.5×10^6 cells/mL of complete media with 200 Nm phorbol 12-myristate 13-acetate (PMA) followed by 24 h incubation in RPMI medium. Macrophages were polarized over the next two days by treatment with 20 ng/mL IFN-gamma and 100 ng/mL lipopolysaccharide (LPS) for M1, 20 ng/mL IL 4 and 20 ng/mL IL 13 for M2a or 20 ng/mL IL 10 for M2C. The expression of the cell surface markers CD68, CD14, CCR7, CD206, CD23 and CD 163 were used to determine the macrophage subtypes using a LSRFortessa flow cytometer (BD Biosciences). In future experiments, fibrogenesis will be induced by treating vocal fold fibroblasts with complete Dulbecco's modified Eagle's medium (DMEM) containing 5 ng/mL TGF beta-1 and then encapsulating them together with activated macrophages in glycol-chitosan hydrogel discs. A custom multimodal nonlinear laser scanning microscope will be then used to scan the deposition and organization of collagen fibers in 4% formaldehyde-fixed hydrogel samples [4].

Results

Cells became adherent after 48 h incubation with PMA and the expression of recognized macrophage markers, CD14 and CD68 was clearly upregulated. CCR7 was more highly expressed in M1 macrophages, whereas CD206 was stained more strongly in M2a and M2c macrophages. Furthermore, M2c showed increased expression of CD23 whereas CD163 was more highly expressed in the M2c subtype. These results confirm the successful differentiation of the THP-1 cells into different subtypes. It is also hypothesized that the incorporation of M1 and M2a macrophages in the 3D co-culture system will result in increased and disorganized deposition of collagen bundles, while the encapsulation of M2c subtype will result in organized collagen fibers and reduced fibrosis.

Conclusions

Our finding may potentially help to devise strategies for designing scaffolds that can effectively prevent and reduce fibrosis.

Acknowledgments: The financial support of National Institutes of Health (Grant #R01 DC-005788) and the Canadian National Science and Engineering Research Council is gratefully acknowledged.

References

1. Julier, Z.; et al. Promoting tissue regeneration by modulating the immune system. *Acta Biomater.* **2017**, *53*, 13–28.
2. Sridharan, R.; et al. Biomaterial based modulation of macrophage polarization: a review and suggested design principles. *Mater. Today* **2015**, *18*, 313–325.
3. Vannella, K.M.; Wynn, T.A. Mechanisms of organ injury and repair by macrophages. *Ann. Rev. Physiol.* **2017**, *79*, 593–617.
4. Miri, A.K.; et al. Nonlinear laser scanning microscopy of human vocal folds. *Laryngoscope* **2012**, *122*, 356–363.

7.9. Characterizing Injury Recovery in Rabbit Vocal Folds with Multimodal Imaging

Ksenia Kolosova ¹, Marius Tuznik ², Qiman Gao ³, Sarah Bouhabel ⁴, Huijie Wang ⁵, Luc Mongeau ⁵ and Paul W. Wiseman ^{1,6}

¹ Department of Physics, McGill University, Montreal, QC, Canada

² McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, QC, Canada

³ Department of Dentistry, McGill University, Montreal, QC, Canada

⁴ Department of Otolaryngology, Head and Neck Surgery, McGill University, Montreal, QC, Canada

⁵ Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

⁶ Department of Chemistry, McGill University, Montreal, QC, Canada

Keywords: vocal fold injury; nonlinear microscopy; microcomputed tomography; magnetic resonance imaging

Objectives

We visualized vocal fold injury recovery in a rabbit model using three imaging modalities. These techniques can be applied to characterize the injury recovery process and test treatments for vocal fold scarring.

Introduction

Rabbits are frequently used for studying vocal fold injury and testing injectable biomaterial treatments for vocal fold scarring. In these studies, imaging-based evaluation is most often conducted by tissue slicing and histological staining. To obtain three-dimensional information without physical slicing, we recently used nonlinear laser-scanning microscopy and nanoscale computed tomography to visualize a dissected rabbit vocal fold specimen¹. As they require a small specimen size, these techniques oblige precise injury localization prior to dissection. We sequentially applied magnetic resonance imaging (MRI), microscale computed tomography (CT), and nonlinear laser-scanning microscopy (NLSM) to visualize injury with and without labelling.

Methods

A unilateral injury was created using microcup forceps in the left vocal fold of three New Zealand White rabbits. Animals were sacrificed at 3, 10, and 39 days post injury, and the larynx was excised and dissected. Three imaging methods were sequentially applied to each specimen. MRI was performed using the 7-T Bruker Pharmascan at the McConnell Brain Imaging Center. CT was performed using the Bruker SkyScan 1172 at the McGill Institute for Advanced Materials. NLSM was performed with optical clearing sample preparation following methods described in our previous study¹. Images were analyzed using Dragonfly (Object Research Systems) and a Python script.

Results

The MRI modality allowed visualization of the injury location label-free with 100 µm resolution. The CT modality achieved finer resolution down to the micrometer scale. However, the intrinsic density contrast of CT was insufficient to visualize features of the injury beyond the contour, necessitating heavy metal staining for contrast enhancement. The NLSM modality provided simultaneous specific visualization of second harmonic generation from fibrillar collagen and two-photon autofluorescence of elastin with near diffraction-limited spatial resolution, allowing clear resolution of collagen fibers in the vocal fold lamina propria, muscle, and surrounding cartilages at submicrometer scales. It allowed quantitative evaluation of properties of the collagen distribution following methods reported in a past study².

Conclusions

The results suggest that a combination of MRI, contrast-enhanced CT, and NLSM can be used to characterize vocal fold injury over time and at different spatial scales. The label-free visualization mechanism of MRI motivates its use for injury localization in live-animal imaging, and each technique serves as a platform for qualitative and quantitative image analysis.

Acknowledgments: This work was supported by the National Institutes of Health (Grant R01 DC005788), the Natural Sciences and Engineering Research Council, and the Canadian Foundation for Innovation.

References

1. Kazarine, A.; et al. Multimodal virtual histology of rabbit vocal folds by nonlinear microscopy and nano computed tomography. *Biomed. Opt. Express* **2019**, *10*, 1151–1164.
2. Miri, A.K.; et al. Nonlinear laser scanning microscopy of human vocal folds. *Laryngoscope* **2012**, *122*, 356–363

7.10. Stress Relaxation in Carbon Nanotube Composite Hydrogels for Vocal Fold Tissue Regeneration

Hossein Ravanbakhsh, Guangyu Bao and Luc Mongeau

Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

Keywords: injectable hydrogels; carbon nanotubes; stress relaxation; tissue engineering

Objectives

The aim of this study is to investigate the effect of carboxylic functionalized carbon nanotubes (CNTs) on the stress relaxation characteristics of composite hydrogels. The hypothesis is that functionalized CNTs play a role in ionic and covalent crosslinking of glycol chitosan hydrogels, which alters their viscoelasticity.

Introduction

Stress relaxation is a key characteristic of hydrogels for cell culture. It indicates whether the substrate is suitable for three-dimensional cultured cells to grow, proliferate, adhere, and migrate through the hydrogel network. Previous researchers have shown that a lower relaxation time can improve cell adhesion [1]. Adding carboxylic functionalized CNTs to the covalently cross-linkable hydrogels generate ionic bonds. Different concentrations of ionic bonds will modulate the stress relaxation time, which may alter cell adhesion within the hydrogel.

Methods

Carboxylic CNTs were dispersed in 1% Triton X-100. The suspension was then sonicated for 5 min to ensure homogenous dispersion. The composite hydrogels were prepared using 2% glycol chitosan solution as the precursor, 0.005% glyoxal as the principal crosslinker, and different concentrations of CNTs (0, 250, 500, and 750 µg/mL) [2]. A DHR-2 torsional rheometer (TA Instruments) with parallel plates was employed to measure the stress relaxation. A volume of 350 µL of hydrogel was prepared and directly added on the bottom plate of the rheometer. The top plate was then lowered until the hydrogel filled the gap, and the temperature was increased to 37 °C. The sample was isolated with mineral oil and a solvent trap to minimize dehydration during experiments. After curing for 14 h, a constant torsional strain of 10% was applied to the hydrogel and the stress was measured for 3 h.

Results

The relaxation timescale (τ^*) is defined as the time required for the stress level to be relaxed to half of its initial value. The average τ^* for samples with CNT concentration of 0, 250, 500, and 750 $\mu\text{g/mL}$ was measured to be 6375, 6308, 5721, and 5452 s, respectively. Adding more carboxylic CNTs resulted in a higher concentration of COOH groups in the hydrogel solution, which further increased the ionic crosslinking. Samples with higher CNT concentrations generally yielded a lower relaxation time. This trend is in good agreement with that reported in previous studies for ionically crosslinked hydrogels, in comparison with covalently crosslinked hydrogels.

Conclusions

Stress relaxation time was measured for hydrogels with various CNT concentrations. The relaxation time of the hydrogels was modulated by applying different concentrations of CNTs. High concentration of CNT caused an increase in hydrogel relaxation time by 15%, which may provide a better substrate for human vocal fold fibroblasts' adhesion. In conclusion, we found that CNTs enhance the mechanical properties of the hydrogels while supporting cell viability and migration.

Acknowledgments: This research was supported by NIH (NIDCD) grant R01-DC005788 (Mongeau, PI).

References

1. Chaudhuri, O.; et al. Hydrogels with tunable stress relaxation regulate stem cell fate and activity. *Nature Mater.* **2015**, *15*, 326.
2. Ravanbakhsh, H.; Bao, G.; Latifi, N.; Mongeau, L. The rheological properties of carbon nanotube-based composite hydrogels as an injectable biomaterial for vocal fold tissue engineering. In Proceedings of the 8th World Congress of Biomechanics (WCB 2018), Dublin, Ireland, July 2018

7.11. *Three-Dimensional Vocal Fold Deformation under Simulated Lateral Cricoarytenoid Muscle Activation in an Excised Human Larynx*

Liang Wu ^{1,2}, Dinesh Chhetri ¹ and Zhaoyan Zhang ¹

¹ Department of Head and Neck Surgery, UCLA School of Medicine, Los Angeles, CA, USA

² Department of Biomedical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, China

Keywords: MRI imaging; vocal fold posturing; lateral cricoarytenoid muscle activation; arytenoid motion

Objectives/Introduction

Voice production control is achieved through laryngeal muscle activation, which postures the vocal folds into desired geometry. Because of limited access to the larynx in live humans, direct quantification of vocal fold deformation due to laryngeal muscle activation in human is difficult. The current study presents an experimental approach to directly measure the three-dimensional motion of the laryngeal cartilages and vocal folds under mechanically simulated activation of the lateral cricoarytenoid (LCA) muscle using magnetic resonance imaging (MRI). Characterizing the geometric changes of the vocal folds and the glottis from laryngeal muscle activation would provide insights into vocal control in human communication.

Methods

The Bruker Biospec 7 Tesla MRI (Bruker Biospin GmbH, Rheinstetten, Germany) was used to scan two excised hemi-larynges (left half and right half from one cadaver male larynx) with a spatial

resolution of $0.1 \times 0.1 \times 0.1 \text{ mm}^3$. The left half larynx was surgically manipulated to simulate the activation of the LCA muscle with tightening sutures along the LCA muscle fibers. The right half served as a control group representing a neutral shape at resting position. The MRI data was processed to reconstruct the three-dimensional structures of the cartilages, intrinsic laryngeal muscles, and cover layer. Finally, the two half larynges were matched to compare the laryngeal movement and geometric changes due to the simulated LCA muscle activation.

Results

The results showed that the LCA muscle activation caused the arytenoid cartilage to rotate medially and downward about the long axis of the cricoid cartilage. This motion was effective in closing the posterior glottal gap but left a gap in the anterior glottis, indicating that LCA activation alone may be insufficient in complete closing the membranous glottis. The arytenoid motion also caused a medial rotation of the vocal folds and medial bulging of the superior medial surface, which is expected to facilitate onset of voice production.

Conclusions

The MRI experimental approach is feasible to capture the details of vocal fold deformation due to the simulated LCA muscle activation, and it could be directly applied to investigate the three-dimensional posturing changes due to other intrinsic laryngeal muscles.

Acknowledgments: This study was supported by NIH/NIDCD and the National Natural Science Foundation of China No. 11874049.

7.12. High throughput Drug & Kinase Inhibitor Screening for Idiopathic Subglottic Stenosis

Jordan Malenke ¹ and Alexander Gelbard ¹

Department of Otolaryngology—Head & Neck Surgery, Vanderbilt University Medical Center, Nashville, TN, USA

Keywords: airway; immunology

Objectives

Recent work has shown that several inflammatory pathways may contribute to the development of the mucosal scar seen in idiopathic subglottic stenosis (iSGS) including upregulation of the IL-17 axis. Here we sought to develop an instrument to help assess iSGS fibroblast cell line responses to 1018 unique FDA-approved compounds to further clarify potential drugs of interests or target signaling pathways for further investigation.

Introduction

Idiopathic subglottic stenosis is a rare condition in which local fibroinflammatory scar formation leads to progressive upper airway obstruction. Recent investigation has demonstrated extensive fibrosis without cartilage remodeling. A number of associations have been investigated regarding the pathophysiology and studies reveal consistent findings. iSGS mucosal scar show abundant fibroblasts, a disordered deposition of extracellular matrix (ECM), activation of the inflammatory cascade. There is insufficient evidence for effective medical therapies to prevent the progression of scar in iSGS, prompting an interest in finding new medical options. We have established fibroblast cell lines from 5 patients with iSGS. Here we investigate application of multiple drug libraries on our established cell lines in an effort to study fibroblast response to assess for potential new targets and new treatments.

Methods

Fibroblast cell lines (n = 5) from established iSGS scar biopsies derived from five patients demonstrated a proliferative response to IL-17. Fibroblasts were induced with IL-17A and tested against the commercially available drug library of 1018 FDA-approved compounds (Selleck), 80 kinase inhibitors, and 51 epigenetic modulators in a high throughput drug screen. Drugs were presented in a standard concentration of 10 micromolar. Cell response were measured using established high-content imaging live/dead assay using fluorescence staining.

Results

1145 established FDA-approved compounds were successfully tested against five iSGS fibroblast cell lines. 191 compounds showed a reduction of living cells by at least 25%, while 11 compounds showed an increase in living cells by at least 25%. Anti-neoplastic, kinase inhibitors, and anti-helminth drugs showed the strongest inhibition, while some estrogens, anti-fungals, and calcium channel blockers increased proliferation.

Conclusions

High throughput drug screening using well-known compounds against IL-17A induced fibroblasts showed reliable results and may lead to new drug discovery or new signaling pathways to study in the treatment of iSGS. This can be successfully performed using fibroblast lines derived from iSGS scar. This produced several classes of drugs of interest for further study including anti-bacterials, steroids, anti-malarials, anti-helminth, and beta-blockers among others.

Acknowledgments: We would like to acknowledge Joshua Bauer, and the Vanderbilt Institute of Chemical Engineering High Throughput Screening Core.

7.13. Clinical and Surgical Implications of Intraoperative Optical Coherence Tomography Imaging for Benign Pediatric Vocal Fold Lesions

Fouzi Benboujja and Christopher Hartnick

Department of Otolaryngology, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, MA, USA

Keywords: voice disorders; oct imaging; vocal fold; benign laryngeal lesions

Objectives

To investigate the potential of intraoperative optical coherence tomography (OCT) to delineate pediatric benign laryngeal lesions.

Introduction

Vocal lesions in children include an extensive list of debilitating levels of dysphonia. An accurate functional assessment of pathological voices in children remains challenging. As many laryngeal pathologies have subepithelial roots, some remain invisible or hard to distinguish with the current imaging techniques, such as a rigid or transnasal endoscopy or videolaryngostroboscopy.

Methods

Optical coherence tomography (OCT) was explored to delineate pediatric benign laryngeal lesions. Under general anesthesia, direct laryngoscopy and three-dimensional OCT imaging were performed on 25 dysphonic pediatric patients (n = 12 male, n = 13 female) ranging from 1 year to 16 years of age.

Results

An assessment of the optical contrast between healthy and abnormal tissue revealed distinct and specific morphological differences among vocal fold lesions, such as nodules, cysts, Reinke's edema, vocalis sulcus, and papilloma. The underlying tissue optical properties (scattering and absorption) suggest a remodeling of the vocal fold mucosa. Furthermore, OCT enables clear depth margins of exophytic fibrovascular lesions indicating the severity of the papilloma invasion.

Conclusions

Optical features of benign pediatric laryngeal lesions using intraoperative OCT may help towards a more qualitative and quantitative approach to current standard of care enabling more personalized therapeutic treatments, especially when diagnosis remains unclear. The ability to assess margins and depth of invasion of papilloma lesions raise the possibility of combining OCT with angiolytic lasers for patient-tailored treatments.

8. Session 4

8.1. *The Relationship between Speech Rate, Voice Quality and Listeners' Purchase Intentions*

May M.W. Poon^{1,2}, Karen M.K. Chan¹ and Edwin M.L. Yiu¹

¹ Division of Speech and Hearing Sciences, The University of Hong Kong, Hong Kong, China

² ENT Laser Hearing & Speech Therapy Centre, Hong Kong, China

Keywords: speech rate; voice quality severity level; speakers' perceived personalities; listeners' purchase intentions

Objectives

This study aims at investigating how do speech rate and voice quality influence speakers' perceived personalities and listeners' purchase intentions.

Introduction

The human voice is a tool for conveying messages. Other than the message content itself, the characteristics in voice contribute to the non-content meaning of particular message to be conveyed [1,2]. From the literature, speaking in a faster rate helped to convey a more competent and attractive image of the salespeople, which could further promote the purchase intentions of the potential customers [3]. However, not many studies edited the speech rate by modifying the pause duration only and keeping the syllable length unchanged in order to best preserve the overall speech intelligibility. Besides, no comparative studies have been conducted to investigate the effects of healthy and disordered voices in direct selling, given that occupational voice disorders such as fatigue and dysphonia are frequently observed in vocally demanding professions.

Methods

This study consisted of two phrases. Phase One aimed at identifying the just noticeable difference (JND) in speech rate (SR) and the best speech samples representing different levels of voice quality severity (VQ) for use as stimuli in the main study in Phase Two. Two speakers recorded a selling script at normal speech rate with normal, mildly and severely dysphonic voice. The speech rate of the recorded stimuli was manipulated by adjusting the pause duration. Forty listeners participated in a listening task to identify the JND in SR and VQ of the stimuli. Phase Two aimed at investigating the effects of SR and VQ on the speakers' perceived personalities and the listeners' purchase intents. Another forty listeners rated the speakers' personalities (attractiveness and competence) and how

likely they would purchase from the speaker as rated with a 5-point Likert Scale after listening to SR stimuli set, VQ stimuli set and mixed stimuli set.

Results

The JND in SR was found to be $\pm 10\%$ of the normal SR. In the SR stimuli set, all ratings on the four parameters representing speakers' personalities and listeners' purchase intentions were highest when the SR ranged from -10% to normal. However, the correlation was not statistically significant. On the other hand, all the four parameters were found to be significantly correlated ($r_s = 0.82$ to 0.92 , $p < 0.05$) among themselves. In the VQ stimuli set, VQ was found to be significantly correlated with all four parameters ($r_s = 0.53$ to 0.64 ; $p < 0.05$). In the mixed stimuli set, all the possible effects of SR on the four parameters were masked by poor voice quality.

Conclusions

Overall voice quality severity level is negatively related to both speakers' perceived personalities and listeners' purchase intentions. Even mild voice problem affects listeners' perception and intentions. Keeping syllable length unchanged in modifying SR may minimize its impact on both speakers' perceived personalities and listeners' purchase intentions. When the speaker's voice quality is poor, speech rate does not affect listeners' perception and purchase intentions at all. Trainers in the industry of direct selling can consider providing voice enhancement training to their salesforce to promote the business outcome.

Acknowledgments: We would like to thank Nicole Li and Richard Wong for their advices and support to our study.

References

1. Borkowska, B.; Pawlowski, B. Female voice frequency in the context of dominance and attractiveness perception. *Anim. Behav.* **2011**, *82*, 55–59.
2. Peterson, R.A.; Cannito, M.P.; Brown, S.P. An exploratory investigation of voice characteristics and selling effectiveness. *J. Pers. Sell. Sales Manag.* **1995**, *15*, 1–15.
3. G elinas-Chebat, C.; Chebat, J.-C.; Boivin, R. Voice and Information Processing. Presented at the 15th International Congress of Phonetic Sciences, 2003.

8.2. Predicting Emphatic Speech: Classification of Non-Literal Utterances

Richard Yanaky

School of Information Studies, McGill University, Montreal, QC, Canada

Keywords: speech categorization; prosody; supervised learning; phonetics

Objectives

This study aims to (1) identify the acoustic features of emphatic speech, and (2) determine the extent to which phonetic information can predict emphatic speech. Supervised learning via k-nearest neighbors (kNN) and random forests are used for speaker-independent classification of emphasis.

Introduction

Previous studies show that by placing emphasis on words, we can change an utterance's meaning; it's not just what you say, but how you say it. Identification of emphatic words through their increased pitch, duration and intensity can act as a trigger to re-evaluate sentential meaning (Wagner & Watson, 2010). This is one key to machine understanding of non-literal utterances. Previous complex unsupervised models have taken this into account (Cernak et al., 2016); however, I take a simpler

supervised approach which provides a more thorough phonetic breakdown of the acoustic cues while providing an equally predictive system.

Methods

Twenty monolingual speakers of Canadian English (16 female, age 18–30) were recorded producing 19 different sentence pairs using the frame, “Mary [verbed] the ball?”, where the verb was emphasized or not. Different monosyllabic verbs were chosen to avoid lexical stress interference. All verbs ended in a coronal consonant for consistent cue measurements. Each pair was recorded twice for a total of 38 sentences per participant, producing a total of 760 of each emphasized and unemphasized utterances for contrast. Recordings were performed in a sound-treated booth. Participants were trained to read the words with 3 s between each reading, but were never demonstrated how to say them; they were asked to say the sentence first regularly, and second as though they were surprised at the action, to elicit emphasis.

Recordings were hand segmented in Praat by a trained phonetician. The pitch, duration, and intensity of the verb’s phonemes were extracted individually, the surrounding words were analyzed at the word level. Data was cleaned (7 sentences with erroneous pitch-tracking removed) and normalized where applicable. Data was split into a 70/30 training & testing set to categorize sentential emphasis. Hyperparameters of the kNN and random forest algorithms were optimized via GridSearchCV (via Scikit-learn) with 5-fold cross-validation to prevent overfitting.

Results

(1) Descriptive: When emphasized, the verb increased in duration (mean 43 ms, $p < 0.001$), increased in pitch (mean 43 Hz, $p < 0.001$), and intensity (mean 2.6dB, $p < 0.001$). Pitch and intensity were measured as the difference from the mean pitch of the preceding word, to the peak of the verb. Surrounding words decreased in duration (mean –15 ms, $p < 0.001$) and short pauses often followed the emphasized word (mean 37 ms, $p < 0.001$). The preceding word ‘Mary’ had small but significant decreases in intensity and pitch vs unemphasized utterances as well (–0.5 dB, $p = 0.008$ and –6 Hz, $p < 0.001$).

(2) Predictive: The best model achieved a 96% correct classification rate when including all phoneme variables (3rd degree polynomial random forest). Duration variables alone classified 85% sentences correct (kNN), pitch alone 86% (2nd degree polynomial kNN), and intensity alone 75% (kNN).

Conclusions

This study indicates that a very high classification rate can be achieved by using detailed prosodic information from both the phonemic level of the emphasized word, as well as from its surrounding context.

Acknowledgments: Data collection took place at the Alberta Phonetics Laboratory under the supervision of Benjamin Tucker.

References

1. Cernak, M.; Asaei, A.; Honnet, P.; Garner, P.N.; Boulard, H. Sound Pattern Matching for Automatic Prosodic Event Detection. In Proceedings of the Interspeech 2016, San Francisco, CA, USA, 8–12 September 2016; pp. 170–174.
2. Wagner, M.; Watson, D.G. Experimental and theoretical advances in prosody: A review. *Lang. Cogn. Process.* **2010**, *25*, 905–945.

8.3. Cortical Mechanisms Controlling the Speech Production during Lombard Effect: An EEG Study

Pavel Prado ¹, Christian Castro ², Alejandro Weinstein ^{1,3}, Lucía Zepeda ¹, Juan Mucarquer ¹ and Matías Zañartu ^{1,4}

- ¹ Advanced Center for Electric and Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile
- ² Department of Speech, Language and Hearing Sciences, Universidad de Valparaíso, Valparaíso, Chile
- ³ Biomedical Engineering School, Universidad de Valparaíso, Valparaíso, Chile
- ⁴ Department of Electronic, Universidad Técnica Federico Santa María, Valparaíso, Chile

Keywords: electroencephalogram; lombard effect; loreta; speech control

Objectives

Describing the cortical mechanisms controlling the voice production during the Lombard Effect (LE) in healthy volunteers.

Introduction

Lombard Effect (LE) is defined as the automatic and involuntary tendency of speaking louder in noisy environments. Studies in animal suggest that LE primary results from sensorimotor integration processes taking place in subcortical structures, i.e., in the brainstem. Although a regulatory role of the cerebral cortex has been established in the LE, the cortical mechanisms underlying this compensatory behavior have not been completely described. Recording the electroencephalogram (EEG), we analyzed the event related potential (ERP) elicited in response to noise masking during vocal production. Furthermore, cortical areas involved in auditory-motor integration processes during the LE were estimated.

Methods

Healthy volunteers (N = 20) were asked to utter series of syllables by reading a controlled sequence of texts presented in a screen. Vocalizations were produced in three experimental conditions: quiet, Lombard (elicited by speech noise at 80 dB HL) and recovery (quiet, 5 min after the end of LE conditions). The electroencephalogram (EEG) was recorded from 64 scalp-electrodes (10/20 system). The event related potentials (ERP) elicited by the auditory feedback of one's own voices were computed to analyze the amplitude and the neural generators of the N1-P2 complex. Cortical activations were estimated using standardized low-resolution brain electromagnetic tomography (sLORETA).

Results

Increased N1-P2 amplitudes were obtained in the Lombard condition in comparison with that obtained in quiet. The amplitude of the ERP decreased in the recovery condition but did not return to its basal level. The auditory feedback of one's own voice induced the activation of left temporal and frontal areas, including the Broca's and Wernicke's areas, primary auditory cortex, primary motor cortex and temporal language areas. The cortical activation increased during the LE. Five minutes later (in the recovery condition), the cortical activity elicited by one's own voices were still significantly higher than that obtained in quiet (prior to the masking onset). Furthermore, visual associative and parietal language areas typically silent in the baseline condition where, however, significantly active in the recovery condition.

Conclusions

The auditory motor integration processes controlling the compensatory increase in vocal intensity during the LE are mediated by the activation of auditory, motor, visual, and language cortical regions, which varies as a function of the signal to noise ratio of the auditory feedback of the one's own vocal production. The remaining greater activation obtained five minutes after the masking offset might

reflect priming cortical mechanisms which increase the communicational efficiency in a potential new noise acoustic environment.

Acknowledgments: This work was supported by CONICYT grants BASAL FB0008 and FONDECYT 1151077, as well as the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under award number P50DC015446. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health

8.4. Auditory Acuity to Fundamental Frequency in Children with and without Vocal Fold Nodules

Elizabeth Heller Murray^{1,2}, Anne Hseu², Roger Nuss², GERALYN HARVEY WOODNORTH² and Cara Stepp^{1,3}

¹ Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA

² Department of Otolaryngology and Communication Enhancement, Boston Children's Hospital, Boston, MA, USA

³ Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, Boston, MA, USA

Keywords: voice; vocal motor control; pediatric voice; auditory acuity

Objective

To investigate auditory acuity to fundamental frequency (f_0) in children with and without vocal fold nodules (VFNs).

Introduction

Based on the historically held belief that children with VFNs are not aware of their own vocal deviations, a non-specific focus on general awareness to vocal changes is often incorporated into therapeutic interventions (e.g., [1]). However, no study to date has directly examined auditory acuity in children with VFNs as compared to vocally healthy children. Understanding auditory acuity in children with VFN is necessary to build more efficient and targeted therapeutic interventions.

Methods

Twenty-one children with VFNs (16 males, Mean (M) = 9.8 years, standard deviation (stdev) = 1.8 years) and thirty-five control speakers (19 males, M = 8.4 years, stdev = 1.6 years) completed a two-alternative forced choice (TAFC, [2,3]) listening task. During each trial, participants heard two sustained tokens of / α / produced by the same speaker and were asked to judge if the tokens were the 'same' or 'different.' One token (base) always had a f_0 of 216 Hz, whereas the other token (test) was experimentally manipulated to have an f_0 that was either the same, higher, or lower. The f_0 of the test token was adaptively modified via the TAFC procedure based on the participant's judgements: correct judgements moved the f_0 of the test token closer to the base token in a subsequent trial, whereas incorrect judgements moved the f_0 further away. Using this staircase procedure, a just-noticeable-difference (JND) was acquired for each participant, as a measure of their auditory acuity to f_0 . A linear regression was used to examine whether JND was significantly predicted by age, group (VFNs, controls), or an interaction between age and group.

Results

Neither group nor a group \times age interaction were significant predictors of JND. Age was a significant predictor of JND ($\beta = -1.3$, $t(55) = -4.0$, $p < 0.001$), explaining a significant portion of the variance in JND (R^2 (adj) = 21.7, $F(1, 55) = 16.3$, $p < 0.001$).

Conclusions

There was no group difference in JNDs between children with and without VFNs, suggesting that extensively targeting awareness may not be an efficient therapeutic intervention strategy. Additionally, as age explained a significant portion of the variance in JND scores, it is important to consider this developmental trend of reduced auditory acuity when designing targeted therapies for younger children.

Acknowledgments: This work was supported by grants DC016197 and DC015446 from the National Institute on Deafness and Other Communication Disorders.

References

1. Andrews, M.L. *Voice Therapy for Children: The Elementary School Years*; Longman Publishing Group: Harlow, UK, 1986.
2. Levitt, H. Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* **1971**, *49*, 467–477.
3. Macmillan, N.A.; Creelman, C.D. Adaptive methods for estimating empirical thresholds. In *Detection Theory: A User's Guide*; Psychology Press: London, UK, 2004; pp. 269–296.

8.5. Phonation Type and Amplitude of Voice Source Fundamental

Johan Sundberg

Department of speech Music Hearing, School of Electrical Engineering and Computer Science, KTH, Stockholm, Sweden

Keywords: hyperfunctional phonation; subglottal pressure; glottal adduction; flow glottogram

Objectives

Identification of characteristics of phonation type in flow glottogram and spectrum.

Introduction

The voice source is of prime relevance to clinical as well as artistic aspects of voice. It can be visualized in terms of flow glottograms, showing glottal airflow versus time. Previous studies have shown that most flow glottogram parameters are strongly correlated with the negative peak of the flow derivative, which, in turn, is strongly correlated with subglottal pressure [1]. Also glottal adduction, regulating degree of hyperfunction, affects this correlation; in hyperfunctional phonation closed phase is long and AC flow pulse amplitude is low [2]. Attempts to identify acoustical correlates of hyperfunction have not been completely successful [3–5].

Methods

Measurements were made on trained voices so as to minimize random variation of data. Five males produced samples of hyperfunctional, neutral, and hypofunctional phonation at varied degrees of vocal loudness. The audio pressure signal, picked up by an omnidirectional microphone a few cm from the corner of the mouth, was inverse filtered by means of the Sopran software. For tuning the filters a ripple-free closed phase of the flow glottogram and a source spectrum envelope as free as possible of local dips and peaks near formant frequencies were used as criteria. Subglottal pressure was captured as the oral pressure during /p/-occlusion. The relationship between the AC amplitude of the flow pulse and the amplitude of the voice source fundamental was examined, as well as the dependence of this relationship on subglottal pressure.

Results

For a given subglottal pressure, the amplitude of the voice source fundamental could be approximated by a linear function of the AC amplitude of the flow glottogram.

Conclusions

Weak glottal adduction produces high flow glottogram AC amplitude, strong voice source fundamental and large airflow. Such airflow is a sign of flow phonation if associated with complete glottal closure. It should be possible to derive information on glottal adduction from the level of the fundamental in the radiated spectrum, if the contributions from vocal tract transfer function are taken into account. In long-term-average analysis of speech and singing the level in the frequency range of the fundamental should be a parameter related to phonation type.

References

1. Sundberg, J. Flow glottogram and subglottal pressure relationship in singers and untrained voices. *J. Voice* **2018**, *32*, 23–31
2. Gauffin, J.; Sundberg, J. Spectral correlates of glottal voice source waveform characteristics. *J. Speech Hear. Res.* **1989**, *32*, 556–565.
3. Alku, P.; Bäckström, T.; Vilkmán, E. Normalized amplitude quotient for parameterization of the glottal flow. *J. Acoust. Soc. Am.* **2002**, *112*, 701–710
4. Sundberg, J.; Thalén, M.; Alku, P.; Vilkmán, E. Estimating perceived phonatory pressedness in singing from flow glottograms. *J. Voice* **2004**, *18*, 56–62
5. Millgård, M.; Sundberg, J.; Fors, T. Flow glottogram characteristics and perceived degree of phonatory pressedness. *J. Voice* **2016**, *30*, 287–292

8.6. Comparison of Voice Onset Measures with Glottal Pulse Identification in Acoustic Signals: Preliminary Analyses

Catherine Madill and Duy Duong Nguyen

Voice Research Laboratory, Dr Liang Voice Program, The University of Sydney, Sydney, Australia

Keywords: voice onset; vocal rise time; vocal attack time; voice onset coordination

Objectives

This study aimed to examine the relationship between the temporal position of the first peak of the acoustic derivative waveform (ADW1) and Vocal Rise Time (VRT), Vocal Attack Time (VAT), and Voice Onset Coordination (VOC). It was hypothesized that ADW1 would correlate with these voice onset measures.

Introduction

Voice onset of vowel phonation provides useful information regarding vocal function in both normal and dysphonic voices. Previous research has attempted to quantify vowel onset using a number of measures including VRT, VAT, and VOC. However, their dependence upon time-consuming and sophisticated analyzing protocols has limited their wide application in clinical settings. Comparing these measures with the derivative of the acoustic signal may be useful in developing a reliable and easily utilized acoustic measure of voice onset.

Methods

Thirty female vocally healthy speakers read three vowels /a/, /i/, and /ou/ using their habitual pitch and loudness for simultaneous recording of acoustic and electroglottographic (EGG) signals and

phonatory airflow. The ADW1 was measured in LabChart as the latency between the onset of acoustic signal deviation and the first peak of the derivative waveform. The VRT was calculated from acoustic signals using a Praat script. The VAT was measured using an algorithm implemented in MATLAB. The VOC was also measured in LabChart as the time interval between the onset of airflow derivative and the first peak of EGG derivative. Correlation was calculated using Spearman's rho (r_s).

Results

The correlation between ADW1 and the voice onset measures was not consistent across measures and vowel types. This measure had a low correlation with VRT ($r_s = 0.449$, $p = 0.013$) which was observed for /a/ only. There was a moderate correlation between ADW1 of /ou/ and VAT of /i/ ($r_s = 0.502$, $p = 0.005$). The ADW1 of /a/ had a high correlation with VOC of /a/ ($r_s = 0.761$, $p < 0.001$) and low correlation with VOC of /ou/ ($r_s = 0.462$, $p = 0.01$). There was a moderate correlation between ADW1 and VOC for /ou/ ($r_s = 0.505$, $p = 0.004$).

Conclusions

The ADW1 showed more correlations with VOC than with VRT and VAT. These correlations appeared to depend on vowel types. Further studies are needed to clarify the role of this measure in the accurate identification of voice onset using acoustic analyses.

8.7. Differences in Ambulatory Vocal Behavior between Patients with Phonotraumatic Lesions and Matched Healthy Controls

Jarrad H. Van Stan ^{1,2,3,*}, Mark Vangel ^{1,3}, Daryush D. Mehta ^{1,2,3}, Andrew J. Ortiz ¹, James A. Burns ^{1,3}, Laura E. Toles ^{1,2}, Katherine L. Marks ^{1,2} and Robert E. Hillman ^{1,2,3}

¹ Massachusetts General Hospital, Boston, MA, USA

² MGH Institute of Health Professions, Charlestown, MA, USA

³ Harvard Medical School, Boston, MA, USA

Keywords: ambulatory voice monitoring; vocal fold nodules; vocal fold polyps; voice disorders

Objectives

This study used ambulatory voice monitoring to identify differences in vocal behavior during daily life in a large cohort of patients with phonotraumatic lesions and matched healthy controls.

Introduction

It is assumed that phonotraumatic vocal fold lesions (nodules, polyps) are caused by, or associated with, voice misuse and/or overuse in daily life. However, previous work using ambulatory voice recordings has not shown large differences in average vocal behavior (sound pressure level/SPL, fundamental frequency/ f_0 , cepstral peak prominence/CPP, or vocal doses) between patients with phonotrauma and matched-controls. This study replicated previous null results and expanded the analysis approach through investigations into distributional characteristics of SPL, f_0 , CPP, and vocal doses, as well as adding a spectral measure of physiological significance (H1-H2).

Methods

Subjects were 100 adult females: 50 with vocal fold nodules or polyps and 50 age-, sex-, and occupation-matched vocally normal individuals. Weeklong summary statistics of voice use were computed from anterior neck-surface acceleration recorded using a smartphone-based ambulatory voice monitor. Once significant differences between patients and controls were identified in SPL, f_0 , CPP, H1-H2, and vocal doses, a step-wise logistic regression was used to produce a final set of

minimal features which maximized discrimination between patients and controls while minimizing feature redundancy.

Results

Paired t-tests resulted in significant differences between patients and matched controls with SPL (skew), f_0 (standard deviation/ SD, 95th percentile, middle 90% range, kurtosis), CPP (skew), H1-H2 (mean, SD, interquartile range, 95th percentile, middle 90% range, kurtosis), and phonatory/silent segments (phonatory segments between 0.1–0.316 s and silent segments between 1000–3160 s). A step-wise logistic regression reduced these significant features to 4 final statistics: f_0 middle 90% range, SPL skew, the cumulative duration of silent segments between 1000–3160 s, and H1-H2 kurtosis. The overall classification from this final group was 79%, and the area under the ROC curve was 0.88 (95% confidence interval = 0.82–0.95).

Conclusions

Compared to controls, voice production in patients with phonotraumatic lesions is associated with [i] less f_0 variability, [ii] higher SPL values more often than lower SPL values, [iii] H1-H2 values that tend to cluster within a lower and narrower range, and [iv] fewer occurrences of longer non-phonatory (silent) periods (especially within the 1000–3160 s intervals). Taken together, these results support prevailing clinical assumptions that patients with phonotraumatic lesions tend to speak louder using a more restricted pitch range, with more forceful/abrupt glottal closure, and with fewer prolonged periods of silence (periods of vocal recovery) than normal speakers. Further work will examine post-treatment data in patients after laryngeal surgery and voice therapy to determine if these differences in vocal function are related solely to the presence of pathology and/or play a role in the etiology of phonotraumatic lesions.

Acknowledgments: Funding provided by the Voice Health Institute and the National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders (Grants R33 DC011588 and P50 DC015446). The paper's contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

8.8. Automatic Voice Signal Typing Using Classic and Nonlinear Dynamics Features

J. M. Miramont¹, J. F. Restrepo¹, J. Codino², G. Schlotthauer¹ and C. Jackson—Menaldi^{2,3}

¹ Instituto de Investigación y Desarrollo en Bioingeniería y Bioinformática, UNER-CONICET, Oro Verde, Entre Ríos, Argentina

² Lakeshore Professional Voice Center, Lakeshore Ear, Nose and Throat Center, St. Clair Shores, MI, USA

³ Department of Otolaryngology, School of Medicine, Wayne State University, Detroit, MI, USA

Keywords: voice typing; support vector machines; nonlinear dynamics

Objectives

The aim of this research is to evaluate the use of classic and nonlinear dynamics features as objective measures for automatic voice classification in three types as proposed by Titze [1], where type 1 voices are nearly periodic, type 2 voices have strong modulating and subharmonic frequencies, and type 3 three voices lack of an apparent periodic structure.

Introduction

Perturbation measures are ubiquitous in voice clinical evaluation, but they fail to assess signals that suffer from heavy fluctuations. To determine the suitability of voice signals to perturbation

analysis, a classification scheme was proposed by Titze [1]. Nevertheless, distinguishing among voice types is still rather subjective. As a solution, we propose an automatic algorithm for signal typing, based on quantitative descriptors.

Methods

Correlation dimension, correlation entropy and noise level, were estimated using the recently proposed U-Correlation Integral [2]. In contrast to previous works [3], this method for estimating attractor's invariants is automatic and user-independent. Additionally, were included Shimmer, Jitter, Harmonic-to-Noise Ratio (HNR), First Harmonic (R1), and a novel feature called Principal Component Normalized Variance (PCNV) which measures the variance explained by the principal component of the set of the signal's periods.

Pathological voices from the MEEI [4] database were labeled by experts as type 1, 2 or 3 (207, 313 and 137 voices, respectively). Firstly, a linear Support Vector Machine (SVM) was trained to separate types 1 and 2 voices from type 3. Secondly, another SVM was trained to separate type 1 from type 2 voices. The rationale behind this is that some descriptors cannot be reliably measured for type 3 voices. The 80% of the data were used to train and validate the model, while it was tested with the remaining 20%. Validation measures were estimated by 10-fold cross-validation. A subset of the extracted features was selected by forward feature selection.

Results

For types 1 and 2 vs. type 3 classification R1, HNR and noise level were used. The accuracy obtained was $93.18 \pm 1.6\%$ (mean and standard deviation), where $91.06 \pm 2.0\%$ of types 1 and 2 voices, and $93.82 \pm 1.73\%$ of type 3 voices, were correctly classified. The accuracy for the test set was 90.25%. For type 1 vs. type 2 classification, PCNV, R1, noise level and correlation entropy were selected. The accuracy obtained was $83.64 \pm 1.8\%$, where $85.46 \pm 2.0\%$ of types 1 voices and $81.27 \pm 2.02\%$ of type 2 voices were correctly classified. The accuracy for the test set was 82.69%.

Conclusions

The nonlinear dynamics features used were estimated with a user-independent method, which is a further step towards a fully automatic tool for objective voice type classification. Our results showed that the proposed features can be used as objective measures to distinguish between voice types. Further research will include a statistical evaluation of inter-rater agreement to assess the generalizability of the proposed approach.

References

1. Titze, I.R. *Workshop on acoustic voice analysis: Summary statement*. National Center for Voice and Speech, 1995.
2. Restrepo, J.F.; et al. Invariant Measures Based on the U-Correlation Integral: An Application to the Study of Human Voice. *Complexity* **2018**, *2018*, 2173640.
3. Lin, L.; et al. An objective parameter for quantifying the turbulent noise portion of voice signals. *J. Voice* **2016**, *30*, 664–669.
4. Massachusetts Eye and Ear Infirmary. *Voice Disorders Database*; Kay Elemetrics Corp.: Lincoln Park, NJ, USA, 1994.

8.9. *Vocal Tract Shape and Acoustic Adjustments of Children during Phonation into Narrow Flow-Resistant Tubes*

Rita Patel and Steven Lulich

Department of Speech & Hearing Sciences, Indiana University, Bloomington, IN, USA

Keywords: semi-occluded vocal tract exercises; narrow-Flow-resistant tubes; pediatric voice; ultrasound tongue imaging

Objectives

The goal of the study is to quantify the salient vocal tract acoustic, subglottal acoustic, and vocal tract physiological characteristics during phonation into a narrow flow-resistant tube with 2.53 mm inner diameter and 124 mm length in typically developing vocally healthy children using simultaneous microphone, accelerometer, and 3D/4D ultrasound recordings.

Introduction

Pediatric dysphonia (hoarseness) is a common condition with prevalence estimates ranging from 1.4%¹ to 23.9%.² Depending on the cause of the dysphonia, management options for pediatric dysphonia typically involve medication, surgery, and/or speech therapy. Anecdotally, voice exercises called 'semi-occluded vocal tract' (SOVT) exercises are widely used in speech therapy for rehabilitation of injured voice and for training vocal performers in the pediatric population, nonetheless to our knowledge only one investigation³ into the theoretical and physiological underpinnings of SOVT exercises and their efficacy has been carried out in children. On the other hand, several studies of SOVT exercises have been conducted on adult subjects with normal and disordered voice.

Methods

Acoustic measurements included fundamental frequency (f_0), first formant frequency (F_1), second formant frequency (F_2), first subglottal resonance (F_{Sg1}), and peak-to-peak amplitude ratio ($P_{vt}:P_{sg}$). Physiological measurements included posterior tongue height (D1), tongue dorsum height (D2), tongue tip height (D3), tongue length (D4), oral cavity width (D5), hyoid elevation (D6), pharynx width (D7). The ultrasound recordings were analyzed using a custom MATLAB toolbox (Mathworks Inc., Natick, Massachusetts, USA) called 'WASL,' which provided synchronous display and full-speed playback of the sagittal, coronal, and transverse views from the ultrasound along with the acoustic and accelerometer waveforms. All measurements were made on 9 boys and 12 girls (6–9 years) during sustained /o:/ production at typical pitch and loudness, with and without flow-resistant tube. Linear mixed model analysis was used for normally distributed variables and Kruskal-Wallis test for non-normally distributed variables. Bonferroni correction was used to determine the significance levels.

Results

Phonation with the flow-resistant tube resulted in a significant decrease in F_1 , F_2 , $P_{vt}:P_{sg}$ and D3, and a significant increase in D2 and F_{Sg1} . A statistically significant gender effect was observed for D1, with D1 higher in boys. The flow-resistant tubes with inner diameter greater than approximately 5mm are fundamentally different from SOVT exercises using narrow diameter tubes as these larger tubes primarily lengthen the vocal tract while the narrow tubes increase vocal tract inertance by occluding the vocal tract and enhancing vocal tract wall vibrations.

Conclusions

Children exhibit lowering of F_1 , F_2 , and tongue tip height, and raising of the tongue dorsum are similar to adults during narrow flow-resistant tube phonation, suggesting that the physiological mechanisms responsible for influencing change with the narrow flow-resistant tube exercise in children are similar to those of adults. The lowering of F_1 , F_2 , and the new findings of the decrease in the ratio of the sound energy radiated from the vocal tract to the sound energy radiated from the neck ($P_{vt}:P_{sg}$), and increased first subglottal resonance frequency (F_{Sg1}) provides empirical evidence of increased vocal tract inertance predominantly due to vocal tract occlusion and enhancement of vocal tract wall vibrations for phonation through narrow flow-resistant tubes.

Acknowledgments: Department of Speech & Hearing Sciences Undergraduate Research Grant. Brandon Merritt, Jennifer Philp, and Abigail Matthews for their assistance with data analysis. Alessandra Verdi for her assistance with data collection.

References

1. Bhattacharyya, N. The prevalence of pediatric voice and swallowing problems in the United States. *Laryngoscope* **2015**, *125*, 746–750.
2. Powell, M.; Filter, M.D.; Williams, B. A longitudinal study of the prevalence of voice disorders in children from a rural school division. *J. Commun. Disord.* **1989**, *22*, 375–382.
3. Ramos, L.A.; Gama, A.C.C. Effect of Performance Time of the Semi-Occluded Vocal Tract Exercises in Dysphonic Children. *J. Voice* **2017**, *31*, 329–335.

9. Poster Session 4

9.1. Estimating Patient-Specific Contact Pressures Using a Finite Element Model

Paul J. Hadwin¹, Mohsen Motie-Shirazi², Byron D. Erath², Matías Zañartu³ and Sean D. Peterson^{1,*}

¹ Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON, Canada

² Mechanical and Aeronautical Engineering, Clarkson University, Potsdam, NY, USA

³ Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

* Correspondence: peterson@uwaterloo.ca

Keywords: Patient-Specific Modeling; Silicon Vocal Fold Models; Contact Stresses; Bayesian Inverse Analysis

Objective

This work examines whether a patient-specific finite element (FE) model of vocal fold kinematics can provide accurate estimates of the contact pressures experienced by self-oscillating silicon vocal folds, as a step towards similar estimations in vivo.

Introduction

Excessive contact pressure on the vocal folds (VFs) during phonation is thought to be a progenitor of organic pathologies (nodules and polyps) [1]. Attempts to directly measure contact pressures in vivo have, thus far, had little success. This work proposes “virtual” measurements wherein contact pressures are extracted from a patient-specific numerical VF model. The patient-specific model is formed by estimating model parameters from more standard clinical measures, e.g., high speed video, flow rate, etc. This is tested with a silicone VF experimental facility in a hemi-larynx configuration. A 2D finite element (FE) model of the silicone VFs is developed using inverse analysis to estimate the model parameters. The model is then used to generate estimates of the collision pressures, which are compared with measured values from the experiment for validation.

Methods

Contact pressures of the silicone VFs were simulated using a 2D FE model whose material properties were estimated by fitting the glottal area waveform (GAW) from the model to that extracted from high speed video of the silicone VFs. Estimates of the elastic moduli, material densities, and subglottal pressure were computed using the Bayesian technique of importance sampling. This approach allocates probabilities to potential parameter values based on how well a simulated signal fits the corresponding measured signal; each parameter is estimated by choosing the value with the highest probability. In this study, the VF geometry is assumed known and is based upon the silicone VF layout.

However, all material properties estimated via inverse analysis; as such, the FE model is considered “patient-specific” since it is established for a particular set of silicone VFs. Using the patient-specific 2D FE model, virtual contact pressure measurements were generated based on the stress of the VFs during contact and compared with the contact pressures measured at the midline of the silicone VFs using a static pressure tap. Additionally, the Bayesian framework admits an approximation of the uncertainty in parameter estimates. This uncertainty was propagated through the model producing an uncertainty estimate for the virtual contact pressures [3].

Results

The estimates of the material properties have reasonable accuracy; each value was within 6% of the nominal material properties of the silicon VFs, with at most an average uncertainty of 8% of the estimated value. Simulated contact pressures were qualitatively on the same order as the hemi-larynx measurements. Difficulty in differentiating contact pressure from aerodynamic pressure in the experiments has made direct quantitative comparison challenging.

Conclusions

This study suggests that using a FE model of VF dynamics is a promising approach for the development of patient-specific models capable of accurately estimating values that are difficult or impossible to measure in the clinic, when computational cost is not an impediment.

Acknowledgments: This research was supported by the NIDCD of the NIH under award P50DC015446 and the Ontario Ministry of Research and Innovation through the Early Researcher Award program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Hoffman, H.T.; Overholt, E.; Karnell, M.; McCulloch, T.M. Vocal process granuloma. *Head Neck* **2001**, *23*, 1061–1074.
2. Alipour, F.; Berry, D.A.; Titze, I.R. A finite-element model of vocal-fold vibration. *J. Acoust. Soc. Am.* **2000**, *108*, 3003–3012.
3. Hadwin, P.J.; Galindo, G.E.; Daun, K.J.; Zañartu, M.; Erath, B.D.; Cataldo, E.; Peterson, S.D. Non-stationary Bayesian estimation of parameters from a body cover model of the vocal folds. *J. Acoust. Soc. Am.* **2016**, *139*, 2683–2696.

9.2. Methodological Barriers in Building an Audiovideo Database for Automatic Identification of Fatigue Levels through Speech and Facial Expressions in People with a Neurological Condition

Madeleine Borgeat^{1,2}, Imane Hocine^{1,2}, Patrick Cardinal^{3,4}, Eric Granger^{3,4}, François Michaud^{4,5}, Claire Croteau^{1,2}, Claudine Auger^{1,3,6} and Ingrid Verduyck^{1,2,3}

¹ Centre de recherche interdisciplinaire en réadaptation du montréal métropolitain

² École d’orthophonie et d’audiologie de l’université de montréal

³ Regroupement ingénierie de technologies interactives en réadaptation

⁴ École de technologie supérieure

⁵ Université de Sherbrooke

⁶ École de réadaptation de l’Université de Montréal

Keywords: voice; speech; face; fatigue; automatic recognition

Objectives

The objective of this study is to identify the barriers encountered in developing an audio-video database of speech samples of people with a neurological condition for the development of automatic identification of fatigue.

Introduction

Technologies using voice commands are more and more present in our daily lives, enabling us to perform different tasks via voice control. These technologies could greatly assist the clientele with a neurological disability impacting mobility, including improving the quality of life and promoting adherence to rehabilitation therapies (Tapus, Mataric, and Scassellati, 2007). The results of research conducted with people with neurological condition reveals a need that is to account for fatigue during interactions with technologies using voice commands (Boisvert et al., in press; Jobin et al., In press). The ability to recognize fatigue in the voice is a factor desired by rehabilitation clients in order to improve their interaction experience with voice interaction technologies. To meet these needs, a project entitled “Integrated speech processing system for human-robot interaction in rehabilitation” is currently underway. One of the objectives of this project is to design a voice interaction system able to detect fatigue in the voice so that the technologies with voice commands can adjust to the level of fatigue of people living with a neurological condition. Five major stages are planned for this project:

1. Build a database of speech and face samples documenting experience experienced with fatigue in people with a neurological condition;
2. Experimentally validate the developed system;
3. Study fatigue in people with neurological impairment and its impact on the functioning in everyday life;
4. Study communicative strategies used by people with neurological disabilities in different situations of conversations and exchanges.

This presentation focusses on the barriers encountered in the building of the audio-video database.

Methods

We recruited adult participants with a neurological disability whose main communication channel is voice. Participants were recruited through several rehabilitation centers in Montreal. Participants were invited to take part in two interviews of 60 min each with a research assistant through the Reacts[®] videoconferencing system. During this interview, speech and video data were collected continuously and fatigue measurements on a visual analogue scale were made every 10 min by the participants. Data regarding the general fatigue state of the participants were also collected. One hundred speech and face samples are expected at the end of the project.

Results

In 3 months, 47 persons meeting the inclusion criteria were identified, 38 could be contacted, 17 accepted to participate and 8 were able to participate in both video interviews.

Conclusions

The main barriers identified to participating in the study regarded participants' access to a computer or a tablet permitting them to use the video conference system. The second barrier concerned the instability of the video conference system that yielded disturbances in the quality of both the audio and the video signal at times.

Acknowledgments: The research is funded by the Fonds de Recherche du Québec—Nature et Technologie (FRQNT).

9.3. Simultaneous Measurements of Glottal Velocities and Vocal Folds Geometry in a Canine Larynx Model

Charles Farbos de Luzan ¹, Alexandra Maddox ², Liran Oren ¹, Ephraim Gutmark ² and Sid Khosla ¹

¹ Department of Otolaryngology, University of Cincinnati, Cincinnati, OH, USA

² Department of Aerospace Engineering, University of Cincinnati, Cincinnati, OH, USA

Keywords: particle image velocimetry; digital image correlation; flow structure interaction

Objectives

Demonstrate the feasibility of synchronizing digital image correlation (DIC) with particle image velocimetry (PIV), in order to measure simultaneously instantaneous velocity fields, as well as the 3D shape of the vocal folds' geometry in the intraglottal region of a canine larynx model. The relationship between volume velocity (flow rate) and the geometry of the glottis (especially during the closing phase) is of importance because of its effect on the quality of the sound produced by the vocal source.

Introduction

During phonation, the flow-structure interaction (FSI) is the driving mechanism for the vibrations of the vocal folds. In our previous studies, we managed to measure the volume glottal flow velocity at the glottal exit and showed how it correlated with the vocal efficiency. Our studies have also showed that intraglottal vortices occur near the superior aspect of the folds during the closing phase. Yet the role of these vortices in the vibration mechanism is still an ongoing debate in the community. Miri (2014) reviewed the studies that characterized the biomechanical properties of the vocal folds during vibrations, and showed that none measured time-resolved FSI of a vibrating larynx. Computational FSI models have predicted formation of intraglottal vortices, but also suggested that their contribution to the fold vibration mechanism was negligible. These computational models, however, lack experimental data for validation because true FSI measurements in a larynx model did not exist. The goal of the current study is to obtain time-resolved measurements of the volume flow simultaneously with the 3-D displacement measurements of the glottis.

Methods

Our experiments employed canine larynx models after removing all the tissue above the level of the vocal folds. Flow measurements were conducted using particle image velocimetry (PIV) and simultaneous glottal geometry was obtained using digital image correlation (DIC). Acoustic data were also collected. Each larynx was tested at different subglottal pressures.

Results

Glottal geometry and corresponding flow velocity fields are shown.

Conclusions

The current FSI data is valuable to further our understanding of the phonation process and can also aid in validating existing models for vocal folds vibrations. Comparison of current data with previous studies will be discussed.

Acknowledgments: This project is supported by NIH Grant no. R01 DC009435 from the National Institute of Deafness and Other Communication Disorders.

References

1. Miri, A.K. Mechanical Characterization of Vocal Fold Tissue: A Review Study. *J. Voice* **2014**, doi:10.1016/j.jvoice.2014.03.001.

- Oren, L.; Sid, K.; Ephraim, G. Intraglottal Geometry and Velocity Measurements in Canine Larynges. *J. Acoust. Soc. Am.* **2014**, *135*, 380–388, doi:10.1121/1.4837222.
- Xue, Q.; Zheng, X.; Mittal, R.; Bielałowicz, S. Subject-Specific Computational Modeling of Human Phonation. *J. Acoust. Soc. Am.* **2014**, *135*, 1445–1456.

9.4. Application of a Promotion of Vocal Health Program (Virtual + Face to Face) for College Professors

Ángela Patricia Atar-Piraquive and Lady Catherine Cantor-Cutiva

Department of Collective Health. Universidad Nacional de Colombia, Bogot, Colombia

Keywords: vocal health; vocal training; tele practice

Objectives

To identify the effect of a promotion of vocal health program on the voice quality of college professors.

Introduction

College professors use their voice as primary tool of work. Therefore, it is important they keep good vocal performance. Vocal health programs have been widely used to promote voice care and prevent the appearance of vocal problems due to overuse or misuse. In general, vocal health programs have included face to face sessions and practice at home to reinforce training. Currently, the use of technological tools can be used to support vocal health programs in the application of healthy behaviors.

Methods

Ten professors will participate in the promotion of vocal health program. The program consists of 4 sessions. The first and last session will be face to face and the second and third online. Acoustic measurements will be obtained pre- and post-intervention. In addition, the teachers will give a self-report of the vocal functioning and benefits of the program in each of its modalities (face-to-face + virtual).

Expected Results

Changes on acoustic voice parameters will be associated with a self-reported perception of better voice quality.

Conclusions

Currently, the use of technologies has allowed strengthening the daily practice of vocal exercises to improve vocal health. The interaction with virtual tools promotes the adoption of healthy practices in university teachers. In addition, the application of virtual sessions can be used as a strategy for teachers who have difficulty attending face-to-face sessions. However, it is important to highlight the work of the speech and language pathologist in the accompaniment and feedback of the process.

References

- Titze, I.R.; Lemke, J.; Montequin, D. Populations in the U.S. workforce who rely on voice as a primary tool of trade: a preliminary report. *J. Voice* **1997**, *11*, 254–259, doi:10.1016/S0892-1997(97)80002-1.
- Roy, N.; Merrill, R.M.; Thibeault, S.; Parsa, R.A.; Gray, S.D.; Smith, E.M. Prevalence of Voice Disorders in Teachers and the General Population. *J. Speech Lang. Hear. Res.* **2004**, *47*, 281, doi:10.1044/1092-4388(2004/023).

3. Yiu, E.M.-L. Impact and Prevention of Voice Problems in the Teaching Profession: Embracing the Consumers' View. *J. Voice* **2002**, *16*, 14.
4. Chen, S.H.; Chiang, S.-C.; Chung, Y.-M.; Hsiao, L.-C.; Hsiao, T.-Y. Risk Factors and Effects of Voice Problems for Teachers. *J. Voice* **2010**, *24*, 183–192, doi:10.1016/j.jvoice.2008.07.008.
5. Rinsky-Halivni, L.; Klebanov, M.; Lerman, Y.; Paltiel, O. Adherence to Voice Therapy Recommendations Is Associated with Preserved Employment Fitness Among Teachers with Work-Related Dysphonia. *J. Voice* **2017**, *31*, 386.e19–386.e26, doi:10.1016/j.jvoice.2016.09.011.

9.5. Investigation of Vocal Folds Poroelastic Behaviour under Mechanical Loading in Different Bath Concentrations

Pooya Tavakoli-Saberi ¹ and Luc Mongeau ^{1,2}

¹ Department of Mechanical Engineering, McGill University, Montreal, QC, Canada

² Canada Research Chair, Tier 1, Voice Biomechanics and Mechanobiology

Keywords: inverse poroelasticity; electro-chemical potentials; vocal folds; computer modelling

Objectives

The aim of the current study was to investigate the effects of electro-chemical potentials on vocal folds mechanical properties. Experimental measurements were performed to study the osmotic pressures and interstitial fluid mobility within the tissue. The collected data was used to further calibrate and validate the simulation model based on inverse poroelastic theory [1]. The simulations were used to evaluate the interstitial fluid mobility and damping characteristics of the tissue under static and dynamic loading.

Introduction

Biological tissues can contain large amounts of fluid, mainly water, accounting for between 60-80% of their mass [2,3]. The interstitial fluid mobility is governed by hydrostatic and osmotic pressures within the tissue. Negatively charged proteoglycans trapped in the tissue result in a greater cation concentration inside and lower anion concentration outside the tissue. This creates an osmotic pressure difference known as the Donnan osmotic pressure. In absence of mechanical load, the tissue tends to swell or dehydrate in hypotonic and hypertonic environments, respectively. Under mechanical loading, the pressure gradient governs the interstitial fluid mobility as described by Darcy's law. A change in volume fraction of the fluid within the tissue due to swelling or dehydration causes changes in the ionic concentrations and consequently osmotic pressure. This complex electro-chemical-mechanical coupling modulates the viscoelastic properties of the tissue. Characterizing this behavior in vocal folds tissue maybe useful to better understand the effects of mechanical loading on tissue response.

Methods

Parallel plate rheological measurements were performed using the Hybrid Rheometer DHR-2 (TA instruments, DE, USA). Relaxation, compression and strain sweep tests were carried on fresh porcine vocal folds and rabbit muscle tissues. The tests were carried on in a closed chamber where the tissue was fully submerged in isotonic (0.9% saline), hypertonic (30.0% saline) and hypotonic (distilled water) solutions. The theory of inverse poroelasticity was developed in the COMSOL commercial software (COMSOL® Multiphysics Version 5.3a). Osmotic pressures and Darcy's law were implemented to simulate the interstitial fluid effects within the solid pores. A 3rd order Ogden hyperelastic model was used to describe the matrix of the porous medium. Data collected from the experiments and previously published data was used to calibrate and validate the computational model.

Results

In simulations, the damping ratio of the tissue under mechanical loading varies between 0.1–0.9 depending on bath concentration. Under dynamic loading, up to 60% of water could be expelled from dehydrated tissue. The tissue had a 20% volume increase when submerged in hypotonic solutions. The damping parameters of tissue depend on the fluid volume fraction. When hydrated, tissue had a damping ratio nearly sevenfold greater than that of dehydrated tissue.

Conclusions

The bath environment regulates the fluid content, interstitial fluid mobility and osmotic pressures. The electro-chemical potentials play an important role in regulating the energy dissipation within the tissue under mechanical loading. The viscoelastic response of the tissue depends on the interstitial fluid transport and matrix deformation.

Acknowledgments: The financial support of National Institutes of Health (Grant #R01 DC-005788) and the National Science and Engineering Research Council is gratefully acknowledged.

References

1. Wilson, W. A Comparison Between Mechano-Electrochemical and Biphasic Swelling Theories for Soft Hydrated Tissues. *J. Biomech. Eng.* **2005**, *127*, 158. ISSN 0148-0731.
2. Miri, A.K.; Barthelat, F.; Mongeau, L. Effects of dehydration on the viscoelastic properties of vocal folds in large deformations. *J. Voice* **2012**, *26*, 688–697. ISSN 08921997.
3. Ehret, A.E.; et al. Inverse poroelasticity as a fundamental mechanism in biomechanics and mechanobiology. *Nat. Commun.* **2017**, *8*, 1002. ISSN 2041-1723.

9.6. In Vitro Analysis of Polymeric Microspheres Containing Human Vocal Fold Fibroblasts for Vocal Fold Lamina Propria Regeneration

Alicia Reyes ¹, Guangyu Bao ², Qiman Gao ³, Nikita Lomis ⁴, Satya Prakash ^{1,4} and Luc Mongeau ^{1,2}

¹ Biological and Biomedical Engineering, McGill University, Montréal, QC, Canada

² Mechanical Engineering, McGill University, Montréal, QC, Canada

³ Faculty of Dentistry, McGill University, Montréal, QC, Canada

⁴ Experimental Medicine, McGill University, Montréal, QC, Canada

Keywords: alginate; cell encapsulation; cell therapy; lamina propria; microspheres; scarring; vocal folds

Objectives

The aim of the present study was to evaluate the feasibility of using microspheres (Ms) containing human vocal fold (VF) fibroblasts to induce lamina propria (LP) regeneration.

Introduction

Material injection is a frequently used strategy to treat injured or dysfunctional VF-LP. The regeneration function of this strategy is compromised because injectable materials may be rapidly cleared by the mononuclear phagocyte system after injection. The morphology of Ms capsules helps to reduce the immune response, allowing the therapeutics it contains to remain in the body for a longer time period. Ms with a diameter around 500 μm are considered ideal substrates for cell delivery due to their optimal diffusion properties.

Methods

Electrospraying and layer-by-layer assembly of polyelectrolytes were used to fabricate Alginate-Poly-L-Lysine-Alginate (APA) and Alginate-Chitosan (Al-Cs) Ms. The optimal loaded

cell concentration of the Ms was determined based on the morphology and integrity of the Ms. The mechanical toughness of the Ms was determined using mechanical stability and osmotic pressure tests. The stiffness of the materials was determined using a torsional rheometer and an atomic force microscope. Swelling properties were evaluated by measuring the diameter changes of the Ms when stored in PBS for 24 h. Live/Dead and MTT assays were used to monitor the viability of the cells after incubation with Ms for 48 h. To evaluate the immunoprotection of the microspheres, the concentration of IL-1 β was measured using ELISA kits when Ms were co-cultured with monocytes derived from mice.

Results

A homogeneous size distribution, and a spherical morphology with a diameter of $552.74 \pm 7.72 \mu\text{m}$ were obtained. The optimal cell concentration for viability was found to be 4×10^5 cells/mL. Mechanical stability and osmotic pressure tests showed that alginate Ms were the toughest. The Young's modulus of pure alginate hydrogel was 3 kPa. The added layers of PLL and alginate decreased the stiffness to 1.92 kPa. The added Cs layer increased the local stiffness to 12.23 kPa. Alginate, APA, and Al-Cs Ms had swelling percentages of 33.67%, 2.32%, and 52.76%, respectively. None of the three configurations of Ms compromised the viability of fibroblasts. The levels of IL-1 β after 24 h of incubation with empty microspheres were 7.08 pg/mL, and 16.77 pg/mL with encapsulated hVFF. The concentration of IL-1 β in the control group was 3.74 pg/mL, and in the free hVFF was 40.88 pg/mL.

Conclusions

Microspheres for the encapsulation of hVFF cells and other therapeutics were fabricated and evaluated for possible application to promote VF regeneration. Alginate microspheres were found to have significant potential as cell delivery tool. They were resistant to mechanical challenges, with a Young's modulus similar to that of the VF-LP. Their swelling rate did not cause bursting. They were found to be friendly to hVFFs. Alginate Ms reduced the expression levels of IL-1 β , and thus they may provide cell immunoprotection. These properties make alginate Ms adequate biomaterials to be tested in future animal studies. Ms are extensively used for other organs, but this is the first time that they are used and characterized to induce the regeneration of the VF-LP, to the author's knowledge.

Acknowledgments: The National Institutes of Health is acknowledged for providing funding for this research through grant DC-005788 (Mongeau, PI).

Reference

1. Chang, T.M.S. *Artificial Cells: Biotechnology, Nanomedicine, Regenerative Medicine, Blood Substitutes, Bioencapsulation, Cell/Stem Cell Therapy, Regenerative Medicine, Artificial Cells and Nanomedicine*; World Scientific: Hackensack, NJ, USA, 2007.

9.7. *Laser-Projection System and Method for 3D Calibrated Laryngeal Measurements Using Transnasal Flexible High-Speed Videendoscopy*

Dimitar D. Deliyski ¹, Hamzeh Ghasemzadeh ^{1,2}, David S. Ford ¹, Daryush D. Mehta ^{3,4}, Milen Shishkov ⁵, Brett E. Bouma ^{4,5}, James B. Kobler ^{3,4}, Matias Zanartu ⁶, Alessandro de Alarcon ⁷ and Robert E. Hillman ^{3,4}

¹ Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, MI, USA

² Department of Computational Mathematics Science and Engineering, Michigan State University, East Lansing, MI, USA

³ Department of Surgery, Massachusetts General Hospital, Boston, MA, USA

⁴ Harvard Medical School, Boston, MA, USA

⁵ Wellman Center for Photomedicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

- ⁶ Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile
- ⁷ Division of Pediatric Otolaryngology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

Keywords: instrumental voice assessment; high-speed videoendoscopy; laser calibration; flexible endoscopy

Objectives

A laser-projection transnasal flexible endoscope coupled to a high-speed videoendoscopy (HSV) system is described. The system is designed to provide three-dimensional (3D) calibrated measurements of the laryngeal structures during in-vivo recordings. The protocol and procedure for calibration of the system are also presented. The calibration method accounts for variations due to different makes of the system, as well as for variations due to different recording conditions, including rotation of the camera, position of field of view (FOV), and variations in the distance between the target recording surface and the endoscopic tip.

Methods

A transnasal flexible endoscope with surgical channel is used. A diffraction-based laser system projecting a green laser pattern at the wavelength of 520 nm was incorporated through the surgical channel. The laser pattern is a 7×7 grid of laser dots creating a square of 16×16 mm at a working distance of 20 mm. The imaging channel of the endoscope is coupled to a color HSV system, allowing the recording of the projected laser pattern superimposed on the superior view of the larynx. The angle difference between the projection and recording axes encodes the axial dimension (target surface to the endoscopic tip), whereas the known distance between different laser points encodes the horizontal dimension. The fiberoptic light-delivery system of the endoscope is coupled with a 300-W xenon light source. The calibration protocol consists of multiple benchtop recordings at controlled working distances. The automatic calibration procedure starts by detecting the center and radius of the FOV and the fiducial marker using statistical image processing. The variations in recording conditions and image quality are compensated at this step using the statistical information and by mapping all images into a pre-defined template. Next, a gradient-based approach is implemented for automatic detection of laser points. The information from this step is used to construct the trajectory of all laser points corresponding to different working distances. Finally, this information is used for decoding the distance from each laser point to the endoscopic tip.

Results

Visual inspection of the in-vivo recordings demonstrated clear visibility of the laser points even at maximum intensity of light-source illumination and frame rates up to 6000 fps, confirming the applicability of the instrument. The results from the calibration experiments demonstrated that the proposed approach could handle effectively the variability in the recordings. Furthermore, the distance to the tip of the endoscope was estimated with acceptable accuracy.

Conclusions

The initial tests demonstrated a successfully developed laser-calibrated transnasal flexible HSV system capable of 3D-calibrated measures of field distance and distance to the endoscopic tip. The proposed method for calibration effectively compensated for the variability in recording conditions and for different makes/brands of the system.

Acknowledgments: Funding provided by the Voice Health Institute and the National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders (Grant P50 DC015446). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest: The authors declare no conflict of interest.



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).