Published in final edited form as:

Laryngoscope. 2012 December ; 122(12): 2773–2780. doi:10.1002/lary.23549.

Classification of glottic insufficiency and tension asymmetry using a multilayer perceptron

Matthew R. Hoffman, B.S., Ketan Surender, M.S., Erin E. Devine, B.S., and Jack J. Jiang, M.D., Ph.D.

University of Wisconsin-Madison School of Medicine and Public Health, Department of Surgery, Division of Otolaryngology – Head and Neck Surgery, Madison, WI, 53792

Abstract

Objective—Laryngeal function can be evaluated from multiple perspectives, including aerodynamic input, acoustic output, and mucosal wave vibratory characteristics. To determine the classifying power of each of these, we used a multilayer perceptron artificial neural network (ANN) to classify data as normal, glottic insufficiency, or tension asymmetry.

Study design—Case series analyzing data obtained from excised larynges simulating different conditions.

Methods—Aerodynamic, acoustic, and videokymographic data were collected from excised canine larynges simulating normal, glottic insufficiency, and tension asymmetry. Classification of samples was performed using a multilayer perceptron ANN.

Results—A classification accuracy of 84% was achieved when including all parameters. Classification accuracy dropped below 75% when using only aerodynamic or acoustic parameters and below 65% when using only videokymographic parameters.

Conclusions—Samples were classified with the greatest accuracy when using a wide range of parameters. Decreased classification accuracies for individual groups of parameters demonstrate the importance of a comprehensive voice assessment when evaluating dysphonia.

Level of evidence—Not applicable – study was performed on excised canine larynges.

Keywords

voice analysis; artificial neural network; multiparameter assessment; recurrent laryngeal nerve paralysis; superior laryngeal nerve paralysis

INTRODUCTION

Voice production is a complex physiological process requiring integration of the nervous system, respiratory tract, and larynx. Accordingly, dysphonia is multidimensional in nature, with different pathologic processes affecting different aspects of the voice.¹⁻² Despite this complexity, assessment is primarily perceptual and often consists of subjective interpretation of vocal quality and stroboscopic exams. Perceptual assessment is regarded as the gold

Corresponding author: Jack Jiang, M.D., Ph.D., 1300 University Avenue, 2725 Medical Sciences Center, Madison, WI 53706, Telephone: 608-265-9854, Fax: 608-265-2139, jjjiang@wisc.edu.

This paper was accepted for oral presentation at the 2012 American Laryngological Association's Spring Meeting, April 18-22, 2012, San Diego, CA.

Conflicts of interest: None.

standard;³ however, it is imprecise and unreliable when analyzing the results of a therapeutic intervention or when comparing patient groups.⁴ A widely used metric to evaluate vocal health, the GRBAS (grade, roughness, breathiness, asthenia, strain) scale, lacks detail and sensitivity⁵ and displays only low to moderate interrater reliability.^{4,6-7} The low temporal resolution of videostroboscopy is adequate when evaluating periodic vocal fold vibration, but cannot reliably evaluate the aperiodic vocal fold vibration which is characteristic of dysphonia.⁸⁻⁹ A systematic approach implementing a series of objective, quantitative methods is warranted.

Quantitative acoustic measurements are a common assessment method used to evaluate vocal pathology.¹⁰ Acoustic measures are valuable and provide different information about laryngeal function than self-assessment.¹¹ However, acoustic parameters do not provide a global assessment and cannot predict severity of dysphonia.² Also, single acoustic parameters display poorer correlation than a combination of several objective parameters with perceptual analysis.⁴

Aerodynamic assessment provides important information on the inputs to normal and disordered phonation¹² and the effort required to produce voice. Aboras et al. measured acoustic and aerodynamic parameters and found that subglottal pressure was the only measure which was predictive of a patient's self-perception of dysphonia.¹ Though valuable, aerodynamic measurements alone cannot describe vocal fold vibratory characteristics or resultant sound quality.

To provide a complete picture of vocal health and laryngeal function, a range of parameters must be considered simultaneously. While such an assessment would certainly provide valuable information, the number of parameters would also make analysis laborious. An algorithm for efficient, automated data interpretation would be valuable and could potentially facilitate widespread clinical application. Classification models, including artificial neural networks (ANNs), are powerful mathematical models which can classify data according to nonlinear statistical analysis.¹³ Further, ANNs can handle extremely large data sets. Of particular interest to this study, they can be used to determine the classifying power of individual parameters and groups of parameters.

The multilayer perceptron (MLP) is one type of ANN and is the most commonly used in medical applications.¹⁴ It consists of an input layer that receives data, at least one hidden layer, and an output layer which provides the classification result. Data are presented to the input layer, computations are performed in the hidden layers, and an output value is obtained at each node of the output layer.¹⁵ These values determine the class into which the data set is classified. Before an ANN can be used to classify an unlabeled data set, it must be trained. Back propagation is one of the most common methods of training for an MLP¹⁶ and minimizes the mean-square error of the output for the training set.¹³ During the learning process, weights associated with connections between nodes are varied with the objective of decreasing the mean-square error of the output.^{13,17} The more input parameters and examples in the training set that are included, the better the ANN typically performs. The ability to synthesize a large amount of information and provide a simple output is a key benefit relevant to medical decision-making and specifically, multi-parameter voice assessment.

We extracted feature vectors containing aerodynamic, acoustic, and videokymographic parameters from excised larynx experiments simulating normal and various severities of glottic insufficiency, simulating recurrent laryngeal nerve paralysis (RLNP), and tension asymmetry, simulating superior laryngeal nerve paralysis (SLNP). A machine learning algorithm was used to train the multilayer perceptron neural network, which was then used

to classify the data. The number of hidden nodes was modified to achieve a higher correct classification rate, and the components of the feature vector were examined to consider their individual contribution to classification. We hypothesized that classification accuracy would be higher with a larger number of parameters.

MATERIALS AND METHODS

Larynges

Thirty-two larynges were excised postmortem from canines sacrificed for non-research purposes according to the protocol described by Jiang and Titze.¹⁸ Canine larynges are much more widely available at our institution than human larynges and have been used extensively to study laryngeal physiology.¹⁹⁻²⁰ There are several anatomical differences between the human and canine larynx. The thyroid and cricoid cartilages are more angulated and not as tall in the canine larynx, and there is no well-defined vocal ligament.¹⁹ These differences did not negatively impact the study. Larynges were examined for evidence of trauma or disorders and any larynges exhibiting trauma or disorders were excluded. Following visual inspection, larynges were frozen in 0.9% saline solution.

Classification ability generally improves as more data are included in the analysis.²¹ To increase the amount of data, we used both previously collected²²⁻²⁵ and newly collected data. In total, 389 trials from 32 larynges were included. Of the 389 trials, 179 simulated normal, 100 simulated tension asymmetry (representative of superior laryngeal nerve paralysis), and 110 trials simulated glottic insufficiency (representative of recurrent laryngeal nerve paralysis).

Apparatus

Prior to the experiment, the supraglottic tissues were removed to expose the true vocal folds. The superior cornu and posterosuperior part of the thyroid cartilage were also removed to facilitate insertion of a lateral 3-pronged micrometer into the arytenoid cartilage. The larynx was mounted on the apparatus (figure 1) specified by Jiang and Titze.¹⁸ A metal hose clamp stabilized the trachea to a tube connected to a constant pressure source, or pseudolung. The pseudolung was designed to simulate the human respiratory system. Pressurized airflow was passed through two Concha Therm III humidifiers (Fisher & Paykel Healthcare Inc., Laguna Hills, California) in series to humidify and warm the air. The potential for dehydration was further decreased by application of 0.9% saline between trials. Airflow was controlled manually and measured using an Omega airflow meter (model FMA-1601A, Omega Engineering Inc., Stamford, Connecticut). Pressure measurements were recorded immediately before the air passed into the larynx using a Heise digital pressure meter (901 series, Ashcroft Inc., Stratford, Connecticut).

Acoustic data were collected using a dbx microphone (model RTA-M, dbx Professional Products, Sandy, Utah) placed at a 45° angle to the vocal folds. The microphone was placed approximately 10 cm from the glottis to minimize acoustic noise produced by turbulent airflow. Acoustic signals were subsequently amplified by a Symetrix preamplifier (model 302, Symetrix Inc., Mountlake Terrace, Washington). A National Instruments data acquisition board (model AT-MIO-16; National Instruments Corp, Austin, Texas) and customized LabVIEW 8.5 software were used to record aerodynamic and acoustic signals. Aerodynamic data were recorded at a rate of 100 Hz and acoustic data at 40,000 Hz. Experiments were conducted in a triple-walled, sound-attenuated room to reduce background noise and stabilize humidity and temperature.

The vocal fold mucosal wave was recorded for approximately 200 milliseconds per trial using a high-speed camera (model Fastcam-ultima APX; Photron, San Diego, CA). Videos were recorded with a resolution of 512×256 pixels at a rate of 4000 frames/second.

Experimental Methods

Trials were conducted as a sequence of 5 second periods of phonation followed by 5 seconds of rest. Five trials were performed for each condition. To simulate normal, both arytenoids were adducted with lateral prongs and the vocal folds were elongated via a suture placed at the midline of the thyroid cartilage, just superior to the vocal folds. This approach was also used for the larynges simulating glottic insufficiency. To simulate glottic insufficiency, only the left arytenoid was adducted to the midline; the right was left unadducted (figure 2a).²⁶ The size of the glottal gap was varied across trials and larynges to simulate paralysis of differing severity.

The method described in Devine et al. was used to simulate tension asymmetry.²⁵ Asymmetry was created using weights which simulated cricothyroid muscle function. The cricothyroid muscle bellies were dissected away to facilitate suture placement. Insertion points for the muscles were noted as fibers were removed. Placement of the sutures and resulting force vectors can be seen in figure 2b. A suture fixing the cricoid cartilage to the trachea was first placed along the midline; this prevented displacement of the cricoid relative to the trachea due to the force of the weights. Sutures simulating the oblique and vertical bellies were inserted at the center of insertion of each belly on the thyroid cartilage and extended along the line of action of the muscle. After unilateral suture placement, the distance and angle between the suture line and the midline of the larynx were measured. The angle of the vertical belly was zero since these sutures were set parallel to the midline. These measurements were then translated to the other side of the larynx to maintain symmetry. Suture angle and the thyroid insertion point were equivalent between sides since the motion of the thyroid cartilage (relative to the cricoid cartilage) was the targeted output of suture forces. Suture placement is shown in figure 2. During tension asymmetry trials, weights were only placed on the sutures corresponding to the left cricothyroid muscle bellies. The mass of the weights was varied across trials and larynges to simulate paralysis of differing severity.

Data analysis

Airflow and pressure at the phonation onset were recorded as the phonation threshold flow (PTF) and phonation threshold pressure (PTP), respectively. Phonation threshold power (PTW) was calculated as the product of these values. PTF, PTP, and PTW were determined manually using customized LabVIEW 8.5 software.

Measured acoustic parameters included fundamental frequency (F_0), signal-to-noise ratio (SNR), percent jitter, and percent shimmer. Acoustic signals were trimmed to produce three 1-second segments per trial using GoldWave 5.1.2600.0 software (GoldWave Inc., St. John's, Canada) and these segments were analyzed using TF32 software (Madison, WI).

High speed video recordings of the mucosal wave were analyzed using a customized MATLAB program (The MathWorks, Natick, MA). Vibratory properties of the four vocal fold lips (right upper, right lower, left upper, left lower) were quantified via digital videokymography. Threshold-based edge detection, manual wave segment extraction, and non-linear least squares curve fitting using the Fourier Series equation were applied to determine the most closely fitting sinusoidal curve. This curve was used to derive the amplitude and phase difference of the mucosal wave for each vocal fold lip. Mucosal wave amplitude was calculated as the average of the amplitudes of the upper and lower vocal fold

lips. While only relative rather than absolute values could be obtained due to current technological limitations, this was sufficient for comparisons across conditions.

Data processing

MATLAB software including the Neural Network Toolbox (The MathWorks) was used for all data processing. In total, 389 trials were analyzed and the derived feature sets were used as a basis for determining models of normal, glottic insufficiency, and tension asymmetry voice production. By attaching the known status of a trial to its feature vector, machine learning techniques can be applied with the goal of modeling the relationship between the input features and the classification of a given trial. The data were randomly split 70/15/15 into training, validation, and test sets, respectively. This division is recommended by the software and is also a common split used in the field. Using 99% of the data in the training set may train the model well, but the resulting classification system would likely be poorly generalizable. Conversely, using a small percentage of the data during training may ensure the system is generalizable, but it would likely perform poorly.

The ANN is presented with the known data, goes through a training and validation stage, and finally is presented with new data during a test stage. The training data and testing data are kept separate in order to evaluate the generalizing ability of the classification.

Data were normalized with each variable in the data set ranging from -1 to 1 , with a mean of 0 and a standard deviation of 1 . Normalizing data improves the efficiency and accuracy of the classification algorithm.²⁷ As random influences may occur during the partitioning process, a more stable performance measurement was obtained by repeating each classification task ten times and averaging over the individual results. Classification rates were calculated based on evaluations occurring during all stages of the machine learning process. A standard multi-layer perceptron (figure 3) was created using sigmoidal activation functions in one hidden layer, and the number of nodes in the hidden layer was varied in increments of 20 from $N=20$ to $N=200$. An upper limit of 200 was selected because using a number of hidden nodes significantly higher than one half of the number of data points can adversely affect generalization. A scaled conjugate backpropagation learning algorithm was used. The goal of the learning algorithm in this model is to modify the weights associated with the connections between the nodes (represented by lines in figure 3) such that an input vector will produce the specified desired output vector.

Separate from the variation of models, the feature set was selectively reduced in an attempt to discover the classification ability of individual parameters and subgroups of parameters. This included the categorical elimination of aerodynamic, acoustic, and videokymographic parameters. In addition to their inclusion in these subsets, all parameters were used on their own as a singular input. The number of hidden nodes in these analyses was determined based on the number attaining the highest classification accuracy when considering all parameters.

Receiver operating characteristic analysis

To determine the ability of the ANN to correctly analyze normal, glottic insufficiency, and tension asymmetry trials, receiver operating characteristic (ROC) analysis was performed and area under the curve (AUC) was determined.

RESULTS

Summary data are provided in table 1. Overall classification accuracy was $84.02 \pm 1.90\%$, including $83.58 \pm 3.95\%$ for normal trials, $70.56 \pm 5.78\%$ for tension asymmetry trials, and $98.11 \pm 0.28\%$ for glottic insufficiency trials (table 2). These classification rates

corresponded to the use of 180 hidden nodes. Total classification rates varied from a minimum of approximately 82% with 20 and 200 hidden nodes to the maximum of 84% when using 180 hidden nodes. Classification accuracy was $74.33 \pm 2.05\%$ when using only aerodynamic parameters, $73.25 \pm 2.55\%$ when using only acoustic parameters, and $64.22 \pm 2.56\%$ when using only videokymographic parameters (table 2). Phonation threshold flow (PTF) demonstrated the greatest individual classification accuracy at $73.91 \pm 2.02\%$.

ROC analysis yielded curves with AUC of 0.8795 for normal (figure 4a), 0.6639 for SLNP (figure 4b), and 0.9878 for glottic insufficiency (figure 4c).

DISCUSSION

Classification accuracy was highest when including all parameters and decreased when considering single groups of parameters or pairs of groups. As expected, classification rates were lower for individual parameters, ranging from 52% (interfold mucosal wave phase difference) to nearly 74% (phonation threshold flow). Interestingly, the classification rate of phonation threshold flow approached that of the aerodynamic parameters as a group, indicating airflow was the most distinguishing parameter in this set of data. As airflow is more sensitive than pressure to changes in glottal abduction,²⁸ it could be expected to classify glottic insufficiency effectively. It did not, however, differentiate well between normal and tension asymmetry, though symmetric elongation-dependent changes in phonation threshold flow have been reported.²⁹ Signal-to-noise ratio displayed a similar classification pattern, classifying normal and glottic insufficiency effectively while displaying very limited ability to detect tension asymmetry. Signal-to-noise ratio is a good parameter to describe glottic insufficiency, as the presence of a wide glottal gap decreases the signal (voice) and increases the noise (turbulent airflow). This parameter is not as useful in distinguishing between subtle differences caused by asymmetric vocal fold elongation.

Consistently high classification rates were predictably observed for glottic insufficiency. As the normal and tension asymmetry conditions are more similar to each other than either is to glottic insufficiency, a high classification rate is expected. Aerodynamic and acoustic parameters displayed high group classification rates. Although the classifying ability of videokymographic parameters was lower, classification of both aerodynamic and acoustic parameters improved when coupled with information on the vibratory characteristics of the mucosal wave. This finding also illustrates an important point about artificial neural network analysis: including more parameters generally increases classifying power. While an individual parameter such as mucosal wave amplitude may not distinguish among groups in isolation, evaluating it in relation to other parameters can improve the ability to identify a given condition. Additionally, classification rates for mucosal wave parameters (vibratory amplitude, phase difference) may have been relatively low because they are not the optimal parameters to describe superior and recurrent laryngeal nerve paralyses. While the parameters used in this study were selected because they could be applied clinically with minimal difficulty, complex parameters such as global entropy and correlation length provided by spatiotemporal analysis³⁰ may better describe irregular vocal fold vibration. Pursuing methods which can expedite the extraction of these parameters to facilitate clinical application is warranted.

Two main limitations will be the subject of future studies. First, it may be interesting to evaluate more complex voice parameters such as those provided by spatiotemporal analysis or nonlinear dynamic acoustic analysis. These measures may prove more valuable than some of the parameters included in this study such as percent jitter and shimmer, which, though capable of distinguishing between normal and glottic insufficiency here as well as between normal and vocal fold polyps in a previous study,³¹ cannot identify more subtle voice

disorders such as nodules or tension asymmetry. Patient-based measures such as the voice handicap index could also be included, if it was found that inclusion of this subjective augmented classification accuracy. Including these parameters in future analyses may improve the ability to distinguish among normal and various voice disorders. Second and most importantly, data from excised larynx experiments rather than human patients were used. This was done to allow us to examine a wide range of parameters that are not typically collected in the clinical setting. Phonation threshold power, for example, has not yet been measured clinically. Using data from excised larynx experiments also provided more inputs than could be usually included if using data from human subjects; however, the excised larynx model can only approximate the dynamic conditions which occur in living patients. Specifically, we could not simulate the effects of thyroarytenoid contraction, vocal fold asymmetry due to paralysis-induced muscle atrophy, or compensatory phenomena. While the models of glottic insufficiency and tension asymmetry used in this study have been applied in previous investigations,^{19,25,26} they do not encapsulate the subtleties of the clinical entities which they represent.

While perceptual analysis and patient self-reporting are frequently used to evaluate dysphonia in the clinic, they are subjective and can introduce bias into treatment decisions.³² The quantitative parameters in this study which best parallel perceptual analysis are likely fundamental frequency and perturbation measurements. Though valuable as part of a more comprehensive voice evaluation, these parameters exhibited individual classification rates between 59 and 68%. This is much lower than the overall classification rate of 84% when considering all parameters. The importance of considering multiple parameters is particularly evident when evaluating tension asymmetry. The aforementioned acoustic parameters displayed classification rates of 13-25%, compared to nearly 71% for the entire feature set. Diagnosis of superior laryngeal nerve paralysis is difficult³³ and recurrent laryngeal nerve paralysis is likely underdiagnosed.³⁴ Developing a standardized comprehensive, multiparameter assessment may aid in the evaluation of these disorders.

There is a wide spectrum of vocal dysfunction³⁵ and no single parameter can adequately characterize vocal quality or dysphonia severity.³ An admitted limitation of quantitative multiparameter assessment is the time required to record the range of measurements; however, noninvasive devices such as the airflow interrupter³⁶ or KayPENTAX Phonatory Aerodynamic System can record seven of the eleven parameters used in this study in less than one minute. Videokymographic analysis of high-speed video images could be performed in an additional few minutes, and even less as improved automated analysis techniques are developed.³⁷ Employing artificial neural network analysis eliminates the most time-consuming aspect of the process – data interpretation. If this method is applied on a larger scale, databases could be generated which could then serve as inputs, thus increasing the classifying power of the algorithm and allowing for a more complex classification scheme with more disorders included.

CONCLUSION

Superior classification rates obtained with a multiparameter assessment compared to subgroup or individual parameter results demonstrate the value of a comprehensive voice assessment. Individual parameter classification rates, particularly for superior laryngeal nerve paralysis, were rather low. Considering a wide range of parameters as well as the relationships among those parameters allows for an evaluation of laryngeal function from multiple perspectives. Additional work developing new parameters able to improve current classification rates as well as automated extraction of these parameters would be beneficial.

Acknowledgments

The authors thank Jason Mielens for providing consultation on the artificial neural network analysis. This study was funded by NIH grant numbers R01 DC008153, R01 DC05522, R01 DC008850, and T32 DC009401 from the National Institute on Deafness and other Communicative Disorders.

REFERENCES

1. Aboras Y, El-Banna M, El-Magraby R, Ibrahim A. The relationship between subjective self-rating and objective voice assessment measures. *Logoped Phoniatr Vocol*. Apr; 2010 35(1):34–8. [PubMed: 20350074]
2. Yu P, Ouaknine M, Revis J, Giovanni A. Objective voice analysis for dysphonic patients: A multiparametric protocol including acoustic and aerodynamic measurements. *J Voice*. 2001; 15(4): 529–542. [PubMed: 11792029]
3. Ma EP, Yiu EM. Multiparametric evaluation of dysphonic severity. *J Voice*. 2006; 20:380–90. [PubMed: 16185841]
4. Hakkesteegt MM, Brocaar MP, Wieringa MH, Feenstra L. The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. *J Voice*. 2008; 22(2): 138–45. [PubMed: 17570633]
5. Hartl DA, Hans S, Vaissiere J, Brasnu DA. Objective acoustic and aerodynamic measures of breathiness in paralytic dysphonia. *Eur Arch Otorhinolaryngol*. 2003; 260:175–82. [PubMed: 12709799]
6. De Bodt MS, Wuyts FL, Van de Heyning PH, Croux C. Test-retest study of the GRBAS scale: influence of experience and professional background on perceptual rating of voice quality. *J Voice*. 1997; 11:74–80. [PubMed: 9075179]
7. Dejonckere PH, Remacle M, Fresnel-Elbaz E, Woisard V, Crevier L, Millet B. Reliability and clinical relevance of perceptual evaluation of pathological voices. *Rev Laryngol Otol Rhinol (Bord)*. 1998; 119:247–8. [PubMed: 9865100]
8. Krausert CR, Olszewski AE, Taylor LN, McMurray JS, Dailey SH, Jiang JJ. Mucosal wave measurement and visualization techniques. *J Voice*. 2011; 25(4):395–405. [PubMed: 20471798]
9. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*. 2008; 117:413–24. [PubMed: 18646437]
10. Titze, IR. Workshop on Acoustic Voice Analysis: Summary Statement; Iowa City: National Center for Voice and Speech, USA. 1995;
11. Hanschmann H, Lohmann A, Berger R. Comparison of subjective assessment of voice disorders and objective voice measurement. *Folia Phoniatr Logop*. 2011; 63:83–7. [PubMed: 20938187]
12. Baken, RJ.; Orlikoff, RF. Clinical Measurement of Speech and Voice. Singular Publishing Group; San Diego, CA: 2000.
13. Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *The Lancet*. Oct 21; 1995 346(8982):1075–9.
14. Bishop, CM. Neural networks for pattern recognition. Oxford University Press; Oxford: 1995.
15. Yan H, Jiang Y, Zheng J, Peng C, Li Q. A multilayer perceptron-based medical decision support system for heart disease diagnosis. *Expert Syst Appl*. 2006; 30:272–81.
16. Rumelhart D, Hinton G, Williams R. Learning representations by back-propagating errors. *Nature*. 1986; 323:533–6.
17. Ruck DW, Rogers SK, Kabrisky M. Feature selection using a multilayer perceptron. *J Neural Network Comp*. 1990; 2(2):40–8.
18. Jiang JJ, Titze IR. A methodological study of hemilaryngeal phonation. *Laryngoscope*. Aug; 1993 103(8):872–82. [PubMed: 8361290]
19. Noordzij JP, Perrault DF, Woo P. Biomechanics of arytenoid adduction surgery in an ex vivo canine model. *Ann Otol Rhinol Laryngol*. 1998; 107:454–61. [PubMed: 9635454]

20. Alipour F, Jaiswal S, Finnegan E. Aerodynamic and acoustic effects of false vocal folds and epiglottis in excised larynx models. *Ann Otol Rhinol Laryngol*. 2007; 116(2):135–44. [PubMed: 17388238]
21. Daelemans, W.; Hoste, V. Evaluation of machine learning methods for natural language processing tasks; Proceedings of LREC-2002, the 3rd International Language Resources and Evaluation Conference; 2002; p. 755–60.
22. Hoffman MR, Witt RE, Chapin WJ, McCulloch TM, Jiang JJ. Multiparameter comparison of injection laryngoplasty, medialization laryngoplasty, and arytenoid adduction in an excised larynx model. *Laryngoscope*. Apr; 2010 120(4):769–76. [PubMed: 20213797]
23. Hoffman MR, Surender K, Chapin WJ, Witt RE, McCulloch TM, Jiang JJ. Optimal arytenoid adduction based on quantitative real-time voice analysis. *Laryngoscope*. Feb; 2011 121(2):339–45. [PubMed: 21271585]
24. Hoffman MR, Witt RE, McCulloch TM, Jiang JJ. Preliminary investigation of adjustable balloon implant for type I thyroplasty. *Laryngoscope*. Apr; 2011 121(4):793–800. [PubMed: 21305554]
25. Devine EE, Bulleit EE, Hoffman MR, McCulloch TM, Jiang JJ. Aerodynamic and nonlinear dynamic acoustic analysis of tension asymmetry in excised canine larynges. *J Speech Lang Hear Res*. 2012 In press.
26. Czerwinka L, Ford CN, Machi AT, Levenson GE, Jiang JJ. A-P positioning of medialization thyroplasty in an excised larynx model. *Laryngoscope*. Mar; 2009 119(3):591–6. [PubMed: 19235760]
27. Saarinen S, Bramley R, Cybenko G. Ill-conditioning in neural network training problems. *SIAM J Sci Comp*. 1993; 14:693–714.
28. Hottinger DG, Tao C, Jiang JJ. Comparing phonation threshold flow and pressure by abducting excised larynges. *Laryngoscope*. Sep; 2007 117(9):1695–9. [PubMed: 17762794]
29. Jiang JJ, Regner MF, Tao C, Pauls S. Phonation threshold flow in elongated excised larynges. *Ann Otol Rhinol Laryngol*. Jul; 2008 117(7):548–53. [PubMed: 18700432]
30. Zhang Y, Jiang JJ, Tao C, Bieging E, MacCallum JK. Quantifying the complexity of excised larynx vibrations from high-speed imaging using spatiotemporal and nonlinear dynamic analyses. *Chaos*. Dec.2007 117(4):043114. [PubMed: 18163778]
31. Jiang JJ, Zhang Y, MacCallum J, Sprecher A, Zhou L. Objective acoustic analysis of pathological voices from patients with vocal nodules and polyps. *Folia Phoniatri Logop*. 2009; 61(6):342–9. [PubMed: 19864916]
32. Carding PN, Wilson JA, MacKenzie K, Deary IJ. Measuring voice outcomes: state of the science review. *J Laryngol Otol*. Aug; 2009 123(8):823–9. [PubMed: 19454129]
33. Dursun G, Sataloff RT, Spiegel JR, Mandel S, Heuer RJ, Rosen DC. Superior laryngeal nerve paralysis and paresis. *J Voice*. 1996; 10(2):206–11. [PubMed: 8734396]
34. Myssiorek D. Recurrent laryngeal nerve paralysis: anatomy and etiology. *Otolaryngol Clin N Am*. 2004; 37:25–44.
35. Painter C. The incidence of voice disorders. *Eur Arch Otorhinolaryngol*. 1990; 247:197–8. [PubMed: 2350513]
36. Jiang J, O'Mara T, Conley D, Hanson D. Phonation threshold pressure measurements during phonation by airflow interruption. *Laryngoscope*. Mar; 1999 109(3):425–32. [PubMed: 10089970]
37. Jiang JJ, Zhang Y, Kelly MP, Bieging ET, Hoffman MR. An automatic method to quantify mucosal waves via videokymography. *Laryngoscope*. Aug; 2008 118(8):1504–10. [PubMed: 18545215]

Figure 1.
Schematic of the excised larynx bench apparatus.

Figure 2.

A) Schematic demonstrating simulation of glottic insufficiency. One arytenoid is adducted to the midline while the contralateral arytenoid is not manipulated. B) Schematic demonstrating simulation of tension asymmetry. Sutures attached to weights are used to simulate cricothyroid muscle function. One suture is used for the oblique belly and one for the vertical belly. To simulate tension asymmetry, sutures are only placed on one side of the larynx. Dotted lines indicate the sutures are falling into the page (vertical) and solid lines

indicate the sutures are in the plane of the page (horizontal). Also depicted are the micrometer prongs used to medialize the arytenoids.

\$watermark-text

\$watermark-text

\$watermark-text

Figure 3.

Schematic of a multilayer perceptron artificial neural network. Each parameter of interest in the input vector has a corresponding node in the input layer. The hidden layer contains the nodes, the number of which was varied during the experiment from 20 to 200. The output vectors are the possible classifications of data, which were normal, superior laryngeal nerve paralysis, and recurrent laryngeal nerve paralysis in this study.

Figure 4.

A) Receiver operating characteristic (ROC) curve for classification of normal (area under the curve (AUC) = 0.8795). B) ROC curve for classification of tension asymmetry (AUC = 0.6639). C) ROC curve for classification of glottic insufficiency (AUC = 0.9878).

Table 1

Summary data from the three groups. Values are presented as mean $\pm$ standard deviation. SLNP = superior laryngeal nerve paralysis; RLNP = recurrent laryngeal nerve paralysis; PTP = phonation threshold pressure; PTF = phonation threshold flow; PTW = phonation threshold power; F_0 = fundamental frequency; SNR = signal-to-noise ratio; VKG = videokymography.

<i>Parameter</i>	<i>Normal</i>	<i>Tension asymmetry</i>	<i>Glottic insufficiency</i>
<i>Aerodynamic parameters</i>			
PTP (cmH ₂ O)	14.58 $\pm$ 6.84	20.61 $\pm$ 13.37	19.36 $\pm$ 8.40
PTF (L/min)	27 $\pm$ 15	25 $\pm$ 15	122 $\pm$ 40
PTW (cmH ₂ O*L/min)	407 $\pm$ 372	642 $\pm$ 712	2610 $\pm$ 1914
<i>Acoustic parameters</i>			
F_0 (Hz)	392 $\pm$ 142	389 $\pm$ 136	194 $\pm$ 81
% Jitter	0.83 $\pm$ 0.83	1.10 $\pm$ 1.31	5.39 $\pm$ 2.76
% Shimmer	6.24 $\pm$ 7.04	6.30 $\pm$ 5.94	31.08 $\pm$ 15.49
SNR	18.14 $\pm$ 6.11	16.80 $\pm$ 7.04	4.27 $\pm$ 2.67
<i>VKG parameters</i>			
Ipsilateral amplitude (pixels)	4.48 $\pm$ 2.67	4.59 $\pm$ 2.75	4.97 $\pm$ 2.51
Contralateral amplitude (pixels)	5.70 $\pm$ 3.57	6.37 $\pm$ 3.45	3.88 $\pm$ 2.03
Intrafold phase difference	0.17 $\pm$ 2.24	-0.30 $\pm$ 0.24	-1.22 $\pm$ 2.95
Interfold phase difference	0.52 $\pm$ 2.89	0.20 $\pm$ 2.95	-0.54 $\pm$ 2.84

Table 2

Total classification accuracies (%) at each number of hidden nodes evaluated. Values are presented as mean $\pm$ standard deviation.

<i>Nodes</i>	<i>Normal</i>	<i>Tension asymmetry</i>	<i>Glottic insufficiency</i>	<i>Total</i>
20	84.02 $\pm$ 6.62	61.84 $\pm$ 15.22	98.2 $\pm$ 0.01	81.88 $\pm$ 3.26
40	85.20 $\pm$ 1.95	67.61 $\pm$ 11.38	98.11 $\pm$ 0.28	83.93 $\pm$ 3.02
60	84.87 $\pm$ 3.29	67.61 $\pm$ 6.33	97.93 $\pm$ 0.61	83.76 $\pm$ 2.94
80	83.07 $\pm$ 3.01	64.59 $\pm$ 9.34	97.47 $\pm$ 1.49	82.00 $\pm$ 2.41
100	84.48 $\pm$ 3.61	61.92 $\pm$ 13.60	97.75 $\pm$ 0.87	81.97 $\pm$ 3.37
120	80.56 $\pm$ 3.80	70.28 $\pm$ 8.28	97.56 $\pm$ 1.16	82.44 $\pm$ 2.91
140	83.24 $\pm$ 3.22	66.88 $\pm$ 10.13	97.75 $\pm$ 0.76	82.75 $\pm$ 2.70
160	80.50 $\pm$ 3.69	68.36 $\pm$ 8.06	97.84 $\pm$ 0.63	81.96 $\pm$ 2.40
180	83.58 $\pm$ 3.95	70.56 $\pm$ 5.78	98.11 $\pm$ 0.28	84.02 $\pm$ 1.90
200	81.85 $\pm$ 2.77	65.24 $\pm$ 11.91	98.11 $\pm$ 0.28	81.80 $\pm$ 3.53

Table 3

Summary classification accuracies for each category and group of parameters. Values are presented as mean $\pm$ standard deviation. Values are presented as mean $\pm$ standard deviation. VKG = videokymography.

<i>Parameter set</i>	<i>Normal</i>	<i>Tension asymmetry</i>	<i>Glottic insufficiency</i>	<i>Total</i>
All parameters	83.58 $\pm$ 3.95	70.56 $\pm$ 5.78	98.11 $\pm$ 0.28	84.02 $\pm$ 1.90
Aerodynamic	92.19 $\pm$ 4.57	19.31 $\pm$ 7.35	94.78 $\pm$ 0.89	74.33 $\pm$ 2.05
Acoustic	85.64 $\pm$ 5.35	28.38 $\pm$ 16.81	93.45 $\pm$ 2.02	73.25 $\pm$ 2.55
Videokymographic	85.04 $\pm$ 3.25	26.98 $\pm$ 9.11	63.91 $\pm$ 4.52	64.22 $\pm$ 2.26
Aero + Acoustic	89.89 $\pm$ 4.99	26.36 $\pm$ 12.13	98.92 $\pm$ 0.38	76.73 $\pm$ 2.08
Aero + VKG	92.01 $\pm$ 5.23	24.14 $\pm$ 17.66	96.92 $\pm$ 1.79	76.10 $\pm$ 3.47
Acoustic + VKG	86.81 $\pm$ 3.65	48.40 $\pm$ 7.83	93.98 $\pm$ 1.52	79.05 $\pm$ 7.49

Table 4

Classification rates for individual parameters. Values are presented as mean $\pm$ standard deviation. PTP = phonation threshold pressure; PTF = phonation threshold flow; PTW = phonation threshold power; F_0 = fundamental frequency; SNR = signal-to-noise ratio; VKG = videokymography.

<i>Parameter</i>	<i>Normal</i>	<i>Tension asymmetry</i>	<i>Glottic insufficiency</i>	<i>Total</i>
<i>Aerodynamic parameters</i>				
PTP	73.18 $\pm$ 5.32	30.01 $\pm$ 13.18	47.65 $\pm$ 12.59	54.93 $\pm$ 4.39
PTF	91.22 $\pm$ 5.51	19.82 $\pm$ 16.16	94.41 $\pm$ 0.28	73.91 $\pm$ 2.02
PTW	86.55 $\pm$ 4.89	16.25 $\pm$ 7.31	78.27 $\pm$ 3.08	66.27 $\pm$ 1.56
<i>Acoustic parameters</i>				
F_0	68.37 $\pm$ 22.98	18.80 $\pm$ 11.27	81.53 $\pm$ 11.56	59.47 $\pm$ 9.73
% Jitter	88.65 $\pm$ 2.81	13.62 $\pm$ 5.89	83.54 $\pm$ 5.26	68.06 $\pm$ 1.40
% Shimmer	78.32 $\pm$ 5.71	24.96 $\pm$ 8.92	84.82 $\pm$ 5.48	66.53 $\pm$ 1.20
SNR	88.39 $\pm$ 5.04	16.97 $\pm$ 13.09	93.66 $\pm$ 3.01	71.65 $\pm$ 1.35
<i>VKG parameters</i>				
Ipsilateral amplitude	65.36 $\pm$ 23.35	22.33 $\pm$ 23.99	27.56 $\pm$ 14.74	43.66 $\pm$ 8.05
Contralateral amplitude	75.03 $\pm$ 5.17	20.80 $\pm$ 4.87	32.91 $\pm$ 9.69	49.26 $\pm$ 1.74
Intrafold phase difference	72.45 $\pm$ 12.49	22.01 $\pm$ 18.41	38.47 $\pm$ 9.91	49.94 $\pm$ 3.85
Interfold phase difference	80.91 $\pm$ 13.20	17.27 $\pm$ 8.56	38.00 $\pm$ 10.67	52.50 $\pm$ 4.66