Nonlinear source-filter coupling due to the addition of a simplified vocal tract model for excised larynx experiments

Benjamin L. Smith, **Steven P. Nemcek**, **Krzysztof A. Swinarski**, 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

Objectives—Traditional excised larynx dissection and setup calls for the removal of all supraglottal structures, eliminating any source-filter interactions that measurably affect the acoustic properties of phonation. We introduce a simplified vocal tract model that can be used in excised larynx experiments and tested the nonlinear source-filter interactions that are present with the addition of highly-coupled, supraglottal structures.

Methods—Aerodynamic and acoustic data were measured at phonation threshold pressure (PTP) and +25% PTP in ten excised canine larynges using a modified dissection technique. PTP and phonation threshold flow (PTF) were defined as the pressure and flow at the phonation onset; phonation threshold power (PTW) is the product of these values. Data were recorded for four experimental conditions: PTP without vocal tract; +25% PTP without vocal tract; PTP with vocal tract; +25% PTP with vocal tract. Differences in PTP, PTF, and PTW were evaluated. For trials conducted at +25%PTP, differences in airflow were evaluated.

Results—PTP (p = 0.009) and PTW (p = 0.002) were significantly reduced with the addition of the novel vocal tract. A reduction in PTF was also present with the vocal tract (p = 0.021) but airflow was not significantly reduced in +25% PTP trials (p = 0.196).

Conclusion—The proposed vocal tract can be used with complete larynges when conducting excised larynx experiments. The effects of nonlinear source-filter interaction were observed during trials with the vocal tract, as evidenced by changes in threshold aerodynamic parameters.

Introduction

Our understanding of the acoustic interaction between the source and filter during voice and speech production has been advanced since its inception by Fant in 1960. Fant described the relationship between the source of speech (pulsatile airflow through the larynx) and the filter (supraglottal structures defined as the vocal tract) as a linear interaction in which the filter does not affect the source; thus, the individual acoustic output of the source and filter can be superimposed to yield the overall production of sound expelled from the oral cavity (1). This

linear phonation model has been shown to serve as a useful approximation of adult male speech (2) as well as provide a mathematical basis, by the superposition principle, for computational speech analysis and production (3, 4).

The validity of the linear source-filter theory, however, has been questioned when applied to more complex types of phonation. An early experiment using circuit-element modeling (5) suggested a more intricate, nonlinear interaction between the source and filter by observing the effect of the vocal tract on glottal wave motion. Glottal flow was found to change with different vocal tract configurations, suggesting that the two are coupled in a nonlinear fashion. Additionally, one (6) and multiple-mass (7–10) computer modeling simulations have shown similar nonlinear interactions between the source and the filter. The most extensive analysis of nonlinear source-filter coupling was conducted by Titze in a theoretical and experimental companion-paper series (11, 12). Titze confirmed that the source-filter interactions were heightened with a sufficiently narrow epilarynx tube and classified the interactions into two levels of interactions based on 1) subglottal and supraglottal pressures and 2) vocal tract reactance (11). It was found that phonation onset is influenced by this interaction. Other phenomena induced by source-tract coupling include subharmonics and frequency jumps, or bifurcations, seen when the fundamental frequency (F_0) crossed the first formant (F_1) during a frequency glide (11). These bifurcations were observed in most subjects in the vocal exercise companion paper, suggesting that humans have some flexibility in controlling the type of source-filter interaction (linear or nonlinear coupling) during certain forms of singing and speech (12).

The effect of a vocal tract on phonation threshold pressure (PTP) was further studied using excised hemilarynx experiments. Döllinger et al (13) used a hemilarynx set-up with canine larynges to confirm that narrowing the epilarynx area helps facilitate phonation by decreasing phonation threshold pressure through source-tract impedance matching. To date, however, no research has observed the effects of source-tract coupling in a complete excised larynx due to the difficulties associated with securing an airtight seal between the vocal folds and the vocal tract while keeping the natural structure of the larynx. As suggested by Montequin in 2003 (15), it would be beneficial to introduce a vocal tract model to a full-size excised larynx in order to provide a more accurate approximation of *in vivo* phonation, which is influenced by the supraglottis.

We describe a novel, simplified vocal tract model that can be used during full excised larynx experiments. To test the vocal tract, we compared the phonatory properties of an excised larynx with and without the vocal tract. We hypothesized that adding the vocal tract would create nonlinear interactions between the vocal folds and the supraglottis, resulting in a decrease in PTP.

Method

Larynges

Ten canine larynges were harvested from dogs sacrificed for purposes unrelated to this study. Canine larynges are used frequently in the study of laryngeal physiology and have been shown to be an appropriate model for the human larynx (14,16). Larynges were dissected according to the protocol described by Jiang and Titze (15). Each larynx was visually inspected for any signs of obvious trauma or disorders before dissection that may have been introduced before or during the primary excision. Signs of discoloration, edema, lesions, nodules or bowing would have merited rejection (17); however, no such lesions were identified in our sample and thus no larynges were excluded. Following preparation, specimens were rinsed, placed in 0.9% saline solution, and frozen at −20 ° C before use.

Vocal tract model

A schematic of the vocal tract model can be seen in Figure 1. All dimensions of the vocal tract were adapted from the rectangular prism hemilarynx vocal tract model designed by Montequin (15) and further adapted by Dollinger (13). Dollinger, comparing several rectangular epilaryngeal areas, found that source-tract coupling was greatest with a cross sectional area of 28.4 mm^2. For this reason, the epilaryngeal area of our vocal tract model was 28.4 mm^2.

The vocal tract model was constructed as three separate parts. The epilarynx was made of acrylonitrile butadiene styrene (ABS) and rapid-prototyped on a 3D printer from an exported Solidworks file (SolidWorks EDU 2010–11, Waltham, MA) The epilarynx was fitted into a straight pharynx tube. The epilarynx and pharynx tubes were mounted to a hollow rectangular prism made of clear acrylic, acting as the oral cavity. Clear acrylic cement was used to join the individual faces of the oral cavity. All connections and joints were checked for air leakage using Sherlock Gas and Air Leak Detection (Winston Products Co., Charlotte, NC).

Experimental apparatus

All supraglottal structures were left intact including the epiglottis and ventricular folds. To facilitate insertion of a 3-pronged micrometer used to adduct the arytenoids, the arytenoid cartilages were exposed by removing the superior cornu and the postero superior part of the thyroid cartilage. A 3–0 nylon suture passed through the thyroid cartilage, superior to the anterior commissure, was used to control vocal fold elongation. Approximately 3 cm of trachea were preserved to allow the larynx to be mounted on the excised bench apparatus as specified by Jiang and Titze (10) and shown in Figure 2.

The larynx was mounted on a barbed hose fitting and clamped using a metal hose clamp. Arytenoid adduction and manipulation of the larynx was accomplished by laterally inserting a 3-pronged device—controlled by micrometers—into each arytenoid cartilage (Figure 3). A third micrometer was attached to the elongation suture and used to control vocal fold elongation. Constant, pressurized airflow was passed through 2 humidifiers (MR-410; Fisher & Paykel Healthcare Inc, Laguna Hills, CA) to ensure that the air passing through the vocal tract was adequately humidified to avoid dehydration of the vocal folds during trials. The humidified air was then passed through a pseudolung designed to mimic the human respiratory anatomy with a total subglottal length of 20 cm before passing through the excised larynx and vocal tract. Pressure and airflow measurements were taken immediately inferior to the barbed hose connection using a Heise Model HPO pressure transducer (Ashcroft Inc, Stanford, CT) and Omega airflow meter (model FMA-1601A; Omega Engineering Inc, Stamford, CT), respectively. Both pressure and airflow data were recorded at a sampling rate of 100 Hz. A flat response dbx microphone (model RTA-M; dbx Professional Products, Sandy, UT) was placed approximately 10 cm from the larynx or vocal tract outlet to reduce the noisy effects of airflow traveling past the microphone. Acoustic data were collected at a sampling rate of 40,000 Hz. The acoustic signal was amplified using a Symetrix preamplifier (model 302, Symetrix Inc, Mountlake Terrace, WA). Airflow, pressure, and acoustic signals were digitized and recorded using customized LabVIEW 8.5 software (National Instruments Corp, Austin, TX) coupled with a National Instruments data acquisition board (model USB6229, National Instruments Corp). To reduce background noise, all experiments were conducted in a triple-walled, sound-attenuated room.

Protocol

The larynx was mounted on the excised bench apparatus and phonation was elicited. Airflow was controlled using a needle valve and was slowly increased until phonation onset. Phonation was stabilized for 2–4 seconds and five trials were conducted for each of the following conditions: no vocal tract at PTP; no vocal tract at +25% PTP; vocal tract at PTP; and vocal tract at +25%PTP. Each larynx served as its own internal control and all larynges were subjected to the four experimental conditions, resulting in a total 20 trials per larynx. The +25% PTP values were obtained by recording the pressure at phonation onset and multiplying that value by 1.25. These trials were conducted to determine if there were any changes in airflow at higher levels of subglottal pressure input. Thus, for trials performed at +25%PTP, independent variables were subglottal pressure and the presence/absence of the vocal tract, and the dependent variable was airflow.

The vocal tract was added to the excised larynx by lowering the distal end of the epilaryngeal tube to the level of the ventricle. The vocal tract mounting mechanism used the supraglottal structures as attachment points onto which the model epilarynx could be secured with glue. Several attempts (latex connection, sutures) were made to achieve the airtight seal before glue was chosen due to the fast bonding time and reports of successful biomedical applications, as cyanoacrylate-based adhesives are commonly used in tissue closure and wound repair (18,19). Loctite 401 Instant Adhesive (Loctite Corp, Tocky Hill, CT), a cyanoacrylate derivative, was applied to the epiglottis, superior aspect of the ventricular folds, and epilaryngeal tube to create and airtight seal. The primary tissue creating the seal was the epiglottis, which was wrapped around the epilarynx tube and secured with adhesive. To ensure the seal was airtight, air was passed through the larynx prior to beginning experimental trials and the specimen was evaluated for areas of leakage. To account for potential dehydration, 0.9% saline solution was applied to the vocal folds between trials.

Data and statistical analysis

A spectrogram (Figure 4) was used to visualize phonation onset and the associated time was recorded to the nearest hundredth of a second (0.01). PTP and phonation threshold flow (PTF) values were found by recording the pressure and flow, respectively, at the previously determined time of phonation onset. Phonation threshold power (PTW) was calculated as the product of PTP and PTF. Finally, for the +25% PTP trials, airflow was determined as the average airflow over the length of the entire trial.

Values of PTP, PTF, and PTW obtained with and without the vocal tract were compared using a paired t-test. Airflow was compared between the +25% no vocal tract and +25% vocal tract configurations using a paired t-test. A Wilcoxon Signed Rank Test was used if the data did not meet assumptions for parametric testing and equal variance. A significance level of $\alpha = 0.05$ was used for all analyses. All statistical analysis was performed using SigmaPlot 11.0 software (Systat Software, Inc, San Jose, CA).

Results

The average values for PTP without and with the vocal tract were 8.16 ± 2.71 and 7.00 ± 1.70 cm H_2O respectively. PTF was 7.28 ± 3.90 L/min without the vocal tract and 4.32 ± 4.01 L/min with the vocal tract. PTW was 67.36 ± 47.612 cm H_2O*L/min without the vocal tract and 29.568 ± 25.315 cm H_2O*L/min with the vocal tract. The addition of the vocal tract to the excised larynx resulted in a significant decrease in PTP and PTW (n = 10, p = 0.009 & p= 0.002, respectively) and PTF (n=10, p = 0.021). Airflow decreased in the +25%

PTP trials; however, there was not a significant decrease with the addition in the vocal tract (p = 0.196). Summary data are presented in Table 1 and Figure 5.

Discussion

The nonlinear effects of a vocal tract model with a complete excised larynx were examined. Pressure measurements show that PTP was significantly lower for trials with compared to trials without the vocal tract. This lowering of pressure necessary to initiate phonation is consistent with previous findings in hemilarynx experiments (9) and computer model simulations (7–10). The introduction of the vocal tract model in this experiment had a noticeable effect on the source, requiring less driving force to initiate and sustain phonation. This facilitation of phonation is an effect of the nonlinear source-filter relationship, most likely due to the introduction of vocal tract inertance. Vocal tract dimensions were chosen to introduce a high degree of source-filter coupling, heightening the nonlinear effect on the source; specifically, a very narrow cross-sectional area (A) for the epilarynx increased the inertance (I) introduced by the vocal tract and described by the following equation:

$$I = \frac{\rho L}{A} \quad (1)$$

where ρ is air density and L is the length of the air column. The inertive air column feeds energy back to the system by creating favorable pressure gradients during the opening and closing of the glottis (18). Experimentally, Dollinger found that the narrower the epilarynx insert, the more drastic effect their vocal tract had on the hemilarynx (13). Thus, the epilarynx has been deemed the most important part of the vocal tract (9), as it serves as an impedance coupler between the source and the filter. The dimensions of both the pharynx and the oral cavity were adopted from previously published vocal tract models (13, 15), which were focused on replicating the cross-sectional area of the human vocal tract, the most important variable contributing to impedance, not the overall physical shape of the human vocal tract. We adopted the vocal tract designs of these previous hemilarynx studies, but added a novel full-size larynx mounting mechanism to introduce the vocal tract experiment to the excised larynx. Future studies may include the use of imaging data to not only replicate the epilarynx, but also the shape of the vocal tract.

A limitation of the proposed model is the inability to visualize the vocal folds. Creating a completely clear vocal tract model could address this issue; however, an acrylic or plastic supraglottal structure may distort images recorded using standard high-speed imaging in excised larynx experiments. Also, vocal fold dynamics with the addition of a vocal tract have been thoroughly investigated by Dollinger et al using high-speed video analysis of 30 surgical microsutures placed on the vocal folds in hemilarynx experiments with a vocal tract(13). The inability to visualize the vocal folds did not prevent application of saline to prevent dehydration. It was possible to hydrate the vocal folds using 0.9% saline solution; the solution was simply poured down the model epilaryngeal tube with care to avoid accumulation of liquid. Potential effects of repositioning the epiglottis when adding the vocal tract were not quantified in this study. Due to proximity of the epiglottis to the true and false vocal folds, minor alterations in elongation or adduction may have occurred when securing the vocal tract. Repositioning the epiglottis to an upright position similar to its position with the vocal tract has been found to alter the phonatory properties of excised canine larynges, causing a decrease in PTP and airflow (21). Elevation of the epiglottis caused the epilarynx tube to assume a more natural tubular shape and subsequently reduced laryngeal resistance.. Finally, airflow in the +25% PTP trials did not show a significant decrease with the vocal tract because of a high degree of standard deviation in the airflow

values; however, there appears to be a slight decrease in airflow from 7.86 ± 3.91 L/min without compared to 6.02 ± 4.75 L/min with the vocal tract.

The full-size excised larynx vocal tract model presented in this study could be useful to approximate *in vivo* phonation more accurately in future excised larynx studies. Further improvements to the vocal tract model are also possible, such as creating a design based on human imaging. This could potentially be used to simulate different vocal tract configurations due to speech token or presence of pathology. With further development, it may be possible to quantitatively examine how structural abnormalities affect voice production in a controlled environment, and how interventions for those disorders may restore normal phonation. It may also be possible to study the phonatory effects of changes in vocal tract configurations, as used in therapeutic techniques and vocal exercises.

Conclusion

This study presents a novel, simplified vocal tract model to be used for full excised canine larynx experiments and examines the phonatory effects induced by the addition of the vocal tract. Pressure and airflow at onset were found to decrease significantly with the addition of the vocal tract, suggesting a nonlinear interaction between the source and filter. With further development, the addition of a vocal tract to excised larynx experiments may better approximate *in vivo* phonation and could be used to study a range of pathologies and treatments that involve supraglottal structures.

Acknowledgments

References

1. Fant, G. The Acoustic Theory of Speech Production. Moulton; The Hague: 1960.

2. Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. J Acoust Soc Am. 1990; 116:2234–2439.

3. Markel, JD.; Gray, AJH. Linear Prediction of Speech. Springer; New York: 1976.

4. Atal BS, Scherer RC. Linear prediction analysis of speech based on a pole-zero representation. J Acoust Soc Am. 1978; 64:1310–1318. [PubMed: 744832]

5. Flanagan JL. Source-system interaction in the vocal tract. Annals of the New York Academ of Sciences. 1968; 155(1):9–17.

6. Flanagan JL, Landgraf LL. Self-oscillating source for vocal tract synthesizers. IEEE Trans Audio Electroacoust. 1968; AU-16(1):57–64.

7. Hatzikirou H, Fitch WT, Herzel H. Voice instabilities due to source-tract interactions. Acta Acusticaunited with Acustica. 2006; 92(3):468–475.

8. Titze IR. The physics of small-amplitude oscillation of the vocal folds. J Acoust Soc Am. 1988; 83(4):1536–1552. [PubMed: 3372869]

9. Rothenburg, M. Acoustic interaction between the glottal source and the vocal tract. In: Stevens; Hirano, editors. Vocal Fold Physiology. University of Tokyo Press; 1980. p. 305-328.

10. Tokuda IT, Zemke M, Kob M, Herzel H. Biomechanical modeling of register transitions and the role of vocal tract resonators. J Acoust Soc Am. 2010; 127(3):1528–36. [PubMed: 20329853]

11. Titze IR. Nonlinear source-filter coupling in phonation: theory. J Acoust Soc Am. 2008; 123(5): 2733–49. [PubMed: 18529191]

12. Titze IR, Riede T, Popolo P. Nonlinear source-filter coupling in phonation: Vocal exercises. J Acoust Soc Am. 2008; 123(4):1902–1915. [PubMed: 18396999]

13. Dollinger M, Berry DA, Luegmair G, Huttner B, Bohr C. The influence of epilarynx area on vocal fold dynamics. Otolarygol Head Neck Surg. 2006; 135(5):724–729.

14. Jiang JJ, Titze IR. A methodological study of hemilaryngeal phonation. Laryngoscope. 1993; 103:872–882. [PubMed: 8361290]

15. Montequin, DA. PhD thesis. Iowa City, IA: University of Iowa; 2003. Developing a Methodology to Study the Effect of the Epilarynx Tube on Phonation Theshold Pressure and Driving Pressure.

16. Noordizij JP, Perrault DF Jr, Woo P. Biomechanics of arytenoid adduction surgery in an ex vivo canine model. Ann Otol Rhinol Laryngol. 1998; 107:454–461. [PubMed: 9635454]

17. Titze, IR. Myoelastic Aerodynamic Theory of Phonation. National Center for Voice and Speech; Iowa City, IA: 2006. p. 7-16.

18. Titze, IR. Principles of voice production. National Center for Voice and Speech; Iowa City, IA: 2000. p. 99-118.

19. Quinn JVV, Drewiecki A, Li MM, Stiell IG, Sutcliffe T, Elmslie TJ, et al. A randomized, controlled trial comparing a tissue adhesive with suturing in the repair of pediatric facial lacerations. Ann Emerg Med. 1993; 22:1130–5. [PubMed: 8517562]

20. Bruns TB, Simon HK, McLario DJ, Sullivan KM, Wood RJ, Anand KJ. Laceration repair using a tissue adhesive in a children's emergency department. Pediactrics. 1997; 98:673–5.

21. Finnegan EM, Alipour F. Phonatory effects of supraglottic structures in excised canine larynges. J Voice. 2009; 23(1):51–61. [PubMed: 17400425]
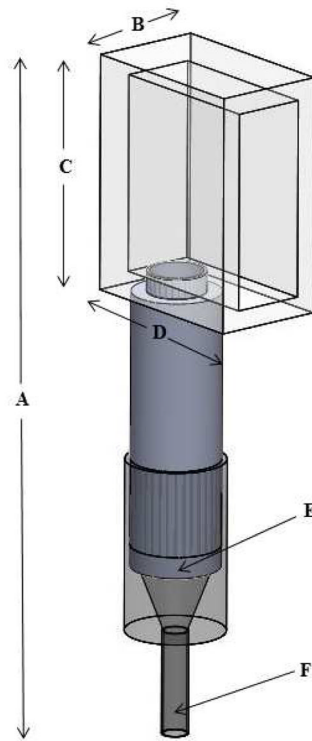
Figure 1.
Schematic of the vocal tract. A = 207 mm, B = 20.6 mm, C = 57.4 mm, D = 48 mm, E = 25 mm (diameter), F = 6 mm (diameter). Lengths B–D refer to the inner dimensions of the piece simulating the oral cavity.
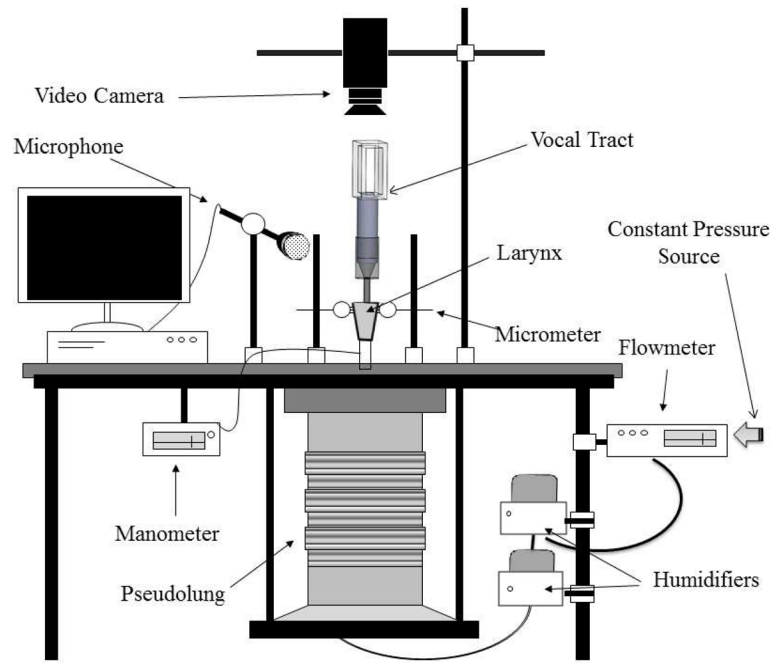
Figure 2.
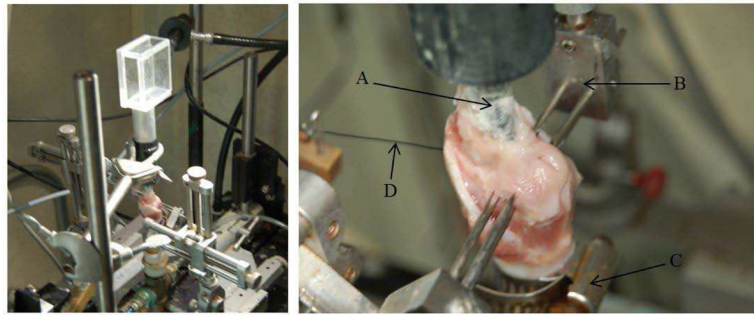Schematic of the experimental apparatus used to mount the excised larynx.

Figure 3.
Left: Digital image of the experimental setup. *Right:* An enlarged view of the excised larynx setup highlighting A) epilaryngeal insert, B) three-pronged adduction device, C) tracheal hose clamp and D) the elongation suture.
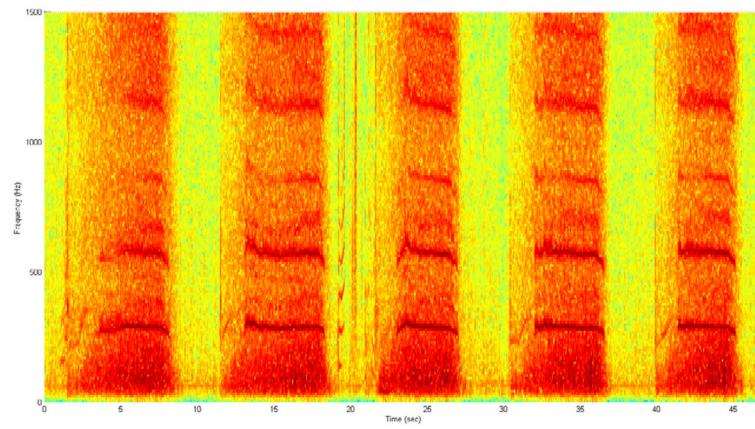
Figure 4.
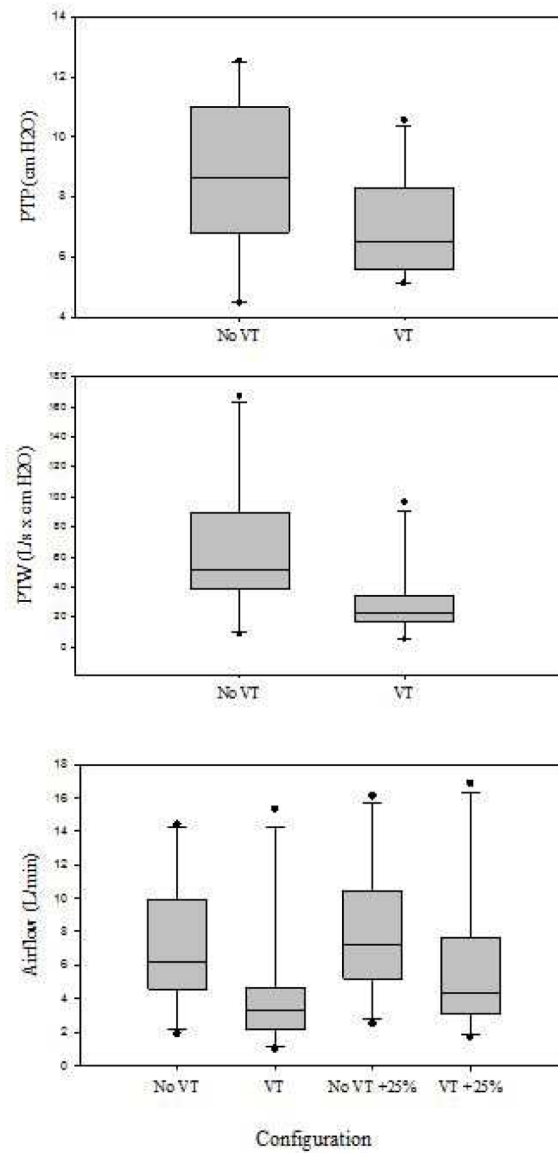Sample spectrogram showing five consecutive phonation tokens.

Figure 5.
A boxplot for PTP (cm H_2O, top), PTW (L/s x cm H_2O, middle) data gathered for the no vocal tract and vocal tract configurations. Boxplot for airflow (L/s, bottom) data gathered for the no vocal tract, vocal tract, no vocal tract +25%, and vocal tract +25% configurations. VT = vocal tract.

Table 1

Summary aerodynamic data for n = 10 larynges.

Parameter	Configuration		
	No VT	*VT*	*P-value*
PTP (cm H$_2$O)	8.62 ± 2.71	7.00 ± 1.70	0.009
PTF (L/min)	7.28 ± 3.90	4.32 ± 4.01	0.021
PTW (cm H$_2$O * L/min)	67.36 ± 47.62	29.57 ± 25.32	0.002
	PTP +25% no VT	*PTP +25% VT*	
Airflow (L/min)	7.86 ± 3.91	6.02 ± 4.75	0.196

VT = vocal tract; PTP = phonation threshold pressure; PTF = phonation threshold flow; PTW = phonation threshold power.