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, *Madison, Wisconsin*

Summary: 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 the 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 10 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; and +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.

Key Words: Vocal tract–Supra glottal structure–Source-filter coupling–Phonation threshold pressure–Excised larynx.

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[1] 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. This linear phonation model has been shown to serve as a useful approximation of adult male speech[2] and 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[11,12] in a theoretical and experimental companion-article series. 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 the (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 article, 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 setup with canine larynges to confirm that narrowing the epilarynx area helps facilitate phonation by decreasing PTP through source-tract impedance matching. To date, however, no research has observed the effects of source-tract coupling in a complete excised larynx because of 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,[14] it would be beneficial to introduce a vocal tract model to a full-size excised larynx to provide a more accurate approximation of *in vivo* phonation, which is influenced by the supraglottis.

Accepted for publication December 21, 2012.
This study was supported by National Institutes of Health grant number R01 DC05522 from the National Institute on Deafness and other Communicative Disorders.
Conflict of interest: None.
From the Division of Otolaryngology – Head and Neck Surgery, Department of Surgery, University of Wisconsin-Madison School of Medicine and Public Health, Madison, Wisconsin.
Address correspondence and reprint requests to Jack J. Jiang, 1300 University Avenue, 2745 Medical Sciences Center, Madison, WI 53706. E-mail: jjjiang@wisc.edu
Journal of Voice, Vol. 27, No. 3, pp. 261-266
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2012.12.012

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 killed for the 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.[15,16] 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. After 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[14] and further adapted by Döllinger.[13] Döllinger, 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 and rapid-prototyped on a three-dimensional printer from an exported Solidworks file (EDU 2010-11; SolidWorks, 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 three-pronged micrometer used to adduct the arytenoids, the arytenoid cartilages were exposed by removing the superior cornu and the posterosuperior 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[15] and shown in Figure 2.

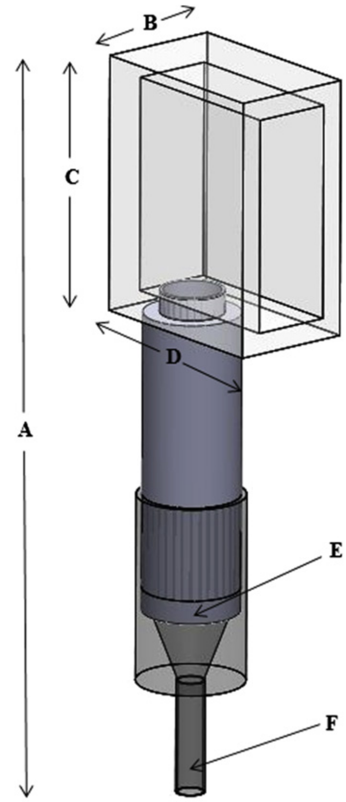

FIGURE 1. Schematic of the vocal tract. A = 207 mm, B = 20.6 mm, C = 57.4 mm, D = 48 mm, E = 25 mm (diameter), and F = 6 mm (diameter). Lengths B–D refer to the inner dimensions of the piece simulating the oral cavity.

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 three-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 two 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).