

Vocal tract shape and acoustic adjustments of children during phonation into narrow flow-resistant tubes

Rita R. Patel, Steven M. Lulich, and Alessandra Verdi

Citation: *The Journal of the Acoustical Society of America* **146**, 352 (2019); doi: 10.1121/1.5116681

View online: <https://doi.org/10.1121/1.5116681>

View Table of Contents: <https://asa.scitation.org/toc/jas/146/1>

Published by the *Acoustical Society of America*

ARTICLES YOU MAY BE INTERESTED IN

[Unsupervised speaker adaptation for speaker independent acoustic to articulatory speech inversion](#)

The Journal of the Acoustical Society of America **146**, 316 (2019); <https://doi.org/10.1121/1.5116130>

[Vocal fold contact pressure in a three-dimensional body-cover phonation model](#)

The Journal of the Acoustical Society of America **146**, 256 (2019); <https://doi.org/10.1121/1.5116138>

[Influence of vocal fold cover layer thickness on its vibratory dynamics during voice production](#)

The Journal of the Acoustical Society of America **146**, 369 (2019); <https://doi.org/10.1121/1.5116567>

[How the peak glottal area affects linear predictive coding-based formant estimates of vowels](#)

The Journal of the Acoustical Society of America **146**, 223 (2019); <https://doi.org/10.1121/1.5116137>

[Automatic speech and singing classification in ambulatory recordings for normal and disordered voices](#)

The Journal of the Acoustical Society of America **146**, EL22 (2019); <https://doi.org/10.1121/1.5115804>

[Non-native consonant acquisition in noise: Effects of exposure/test similarity](#)

The Journal of the Acoustical Society of America **146**, 297 (2019); <https://doi.org/10.1121/1.5116575>



Vocal tract shape and acoustic adjustments of children during phonation into narrow flow-resistant tubes

Rita R. Patel,^{a)} Steven M. Lulich, and Alessandra Verdi

Department of Speech and Hearing Sciences, Indiana University, 200 South Jordan Avenue, Bloomington, Indiana 47405-7002, USA

(Received 29 November 2018; revised 21 June 2019; accepted 25 June 2019; published online 23 July 2019)

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. 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). All measurements were made on eight boys and ten girls (6–9 years) during sustained /o:/ production at typical pitch and loudness, with and without flow-resistant tube. Phonation with the flow-resistant tube resulted in a significant decrease in F_1 , F_2 , and $P_{vt}:P_{sg}$ and a significant increase in D2, D3, and F_{Sg1} . A statistically significant gender effect was observed for D1, with D1 higher in boys. These findings agree well with reported findings from adults, suggesting common acoustic and articulatory mechanisms for narrow flow-resistant tube phonation. Theoretical implications of the findings are discussed. © 2019 Acoustical Society of America. <https://doi.org/10.1121/1.5116681>

[ZZ]

Pages: 352–368

I. INTRODUCTION

Pediatric dysphonia (hoarseness) is a common condition with prevalence estimates ranging from 1.4% (Bhattacharyya, 2015) to 23.9% (Powell *et al.*, 1989). 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, but, to the best of our knowledge, only one investigation into the theoretical and physiological underpinnings of SOVT exercises and their efficacy has been carried out in children (Ramos and Gama, 2017). On the other hand, several studies of SOVT exercises have been conducted on adult subjects with normal and disordered voice, as discussed below.

SOVT exercises involve semi-occluding the supraglottal vocal tract anteriorly (at the lips) by phonating through (1) narrow “flow-resistant” tubes (Kapsner-Smith *et al.*, 2015; Titze *et al.*, 2002); (2) glass tubes inserted in water (resonant tube method) (Granqvist *et al.*, 2014; Laukkanen *et al.*, 1995; Simberg and Laine, 2007); (3) silicone tubes submerged in water (LaxVox technique) (Sihvo and Dienizoglu, 2018); (4) “hand-over mouth” semi-occlusions (Andrade *et al.*, 2014); (5) lip trills (Linklater, 1976; Nix, 1999); (6) tongue trills (Nix, 1999); or (7) sustained bilabial fricative /β:/ (Laukkanen, 1992; Laukkanen *et al.*, 1996) and vowel /u:/ (Laukkanen *et al.*, 2012a; Laukkanen *et al.*, 2008).

SOVT exercises with flow-resistant tubes are particularly advantageous for the pediatric population as these exercises are accessible and engaging, and the target productions can be achieved with ease.

The lengths and inner diameters of flow-resistant tubes are variable (Andrade *et al.*, 2016; da Silva *et al.*, 2018; Smith and Titze, 2017). Smith and Titze (2017) identify the range of flow-resistant tubes most commonly used for voice training and therapy as having inner diameters between 2 and 6 mm. High flow resistance provided by the small diameter tube (stirring straws) are reported to be ideal for voice training purposes (Smith and Titze, 2017; Titze, 2006). Tubes smaller than 5 mm in diameter (Andrade *et al.*, 2016; da Silva *et al.*, 2018; Smith and Titze, 2017) provide high intraoral pressure, suggesting high flow resistance with the small diameter tubes. The high intraoral pressure with the small diameter tube results in an increase in subglottal pressure and decrease in the transglottal pressure which pushes the top edges of the vocal folds slightly apart thereby reducing vocal fold collision and the vibratory amplitude even at high fundamental frequencies leading to improved vocal economy (Guzman *et al.*, 2013c; Horacek *et al.*, 2014; Radolf *et al.*, 2014; Smith and Titze, 2017; Titze, 2006; Titze *et al.*, 2002). SOVT exercises have been also shown to reduce the phonation threshold pressure (Titze, 2009) and result in balanced activation of the thyroarytenoid and the cricothyroid muscles (Laukkanen *et al.*, 2008). The ideal flow-resistant tube is reported to be a tube that has the same resistance as the glottal resistance, resulting in an oral pressure equal to half the subglottal pressure and providing maximum aerodynamic power transfer from the source to the

^{a)}Electronic mail: patelrir@indiana.edu

vocal tract (Smith and Titze, 2017). Smith and Titze (2017) specifically suggest that a tube with diameter 2.5 mm and length between 6 and 12 cm is appropriate for producing oral pressures near 1.0 kPa at flow rates near 0.1 L/s. This study therefore focuses on the use of narrow (2.53 mm inner diameter) flow-resistant tubes, with the goal of understanding how vocal tract shape and acoustics are affected *during* the SOVT exercise with the flow-resistant tubes in children with no voice complaints.

Theoretical studies have consistently indicated the importance of the vocal tract impedance for the therapeutic and/or voice training effects of flow-resistant tube exercises. Vocal tract impedance is the sum of reactance and resistance, the former capturing the oscillatory behavior of the air column above the glottis, and the latter capturing its dissipative characteristics. While vocal tract resistance is always positive, reactance can take on either negative (compliant) or positive (inertive) values. It is well established that positive inertive reactance (henceforth “inertance”) is a critical factor for self-sustained vocal fold vibrations (Titze, 1988, 2008; Zañartu *et al.*, 2007). Due to the non-linear coupling between the source and the filter (Titze, 2008; Titze *et al.*, 2008) theoretical investigations have posited the role of an increase in the inertance as the main regulatory factor responsible for the therapeutic benefit of the flow-resistant tube exercises. Increase in the inertance with flow-resistant tube exercises (Titze, 2006) can be achieved as a result of either the lengthening of the anterior vocal tract (Story *et al.*, 2000), the narrowing or semiocclusion of the lips for bilabial fricative /β:/ (Story *et al.*, 2000), the narrowing of the epilaryngeal tube (Titze, 2001, 2006; Titze and Story, 1997), or the enhancement of low frequency inertance through vibration of the vocal tract walls (Fric and Hruska, 2017; Horacek *et al.*, 2017). Vocal tract lengthening and semiocclusion at the lips increases the inertance primarily by driving vocal tract resonances to lower frequencies (Horacek *et al.*, 2017; Story *et al.*, 2000), whereas narrowing of the epilaryngeal tube and enhancing vibrations of the vocal tract walls increases the inertance without substantially altering the vocal tract resonances (Titze, 2006; Titze and Story, 1997). The therapeutic benefit of SOVT exercises is thought to come from learning compensatory strategies that mimic the effects of the semi-occlusion. Specifically, while the semi-occlusion raises vocal tract inertance by increasing the acoustic mass anteriorly (at the lips), it is suggested that patients learn to compensate for the subsequent removal of the semi-occlusion by narrowing the epilarynx tube (Titze, 2006).

Acoustic models of vocal tract inertance during flow-resistant tube exercises show a downward shift in the frequency of the first formant (F_1). Story *et al.* (2000) found that F_1 decreased from 475 Hz for the vowel /i/ to between 135 and 270 Hz for an 8 mm diameter tube of varying lengths. Horacek *et al.* (2017) reported that F_1 decreased from between 290 and 550 Hz to between 175 and 185 Hz for the vowel /u/ when using a tube of length 50 cm and an inner diameter of 6.8 mm. Using a soft silicon rubber model of the vocal folds and a plexiglass model of the vocal tract for the vowel /u:/, the F_1 was reported to decrease from 260 Hz to approximately 80 Hz during soft phonation

(Horacek *et al.*, 2014). Limited empirical studies on vocally healthy adult subjects with flow-resistant tubes of 2.5, 4, and 8 mm diameter report a decrease in F_1 (Andrade *et al.*, 2014); decrease in f_0 , F_1 , F_3 with increase in F_2 and F_4 (Radolf *et al.*, 2014) and a decrease in F_1, F_2 , F_3 , and F_4 , (Gaskill and Erickson, 2010), respectively. A slight increase in the fundamental frequency during flow-resistant phonation was reported in 20 adult male subjects with 8 mm diameter flow-resistant tube phonation (Gaskill and Quinney, 2012), although Gaskill and Erickson (2010) and Andrade *et al.* (2014) failed to find significant changes in f_0 . The acoustic theory of speech production (Fant, 1960; Stevens, 2000) predicts that all formant frequencies should decrease when the lip opening is semi-occluded, consistent with the limited empirical data on adults, and we therefore predict a similar decrease in all formant frequencies during flow-resistant tube phonation in children. Based on the mixed empirical data regarding f_0 , we predict no change in fundamental frequency during flow-resistant tube phonation in children.

Narrow flow-resistant tubes with diameters between 2.5 and 6 mm have high resistance (Andrade *et al.*, 2016; da Silva *et al.*, 2018; Smith and Titze, 2017) as well as large inertance (Titze, 2006). Varying the length of a small diameter flow-resistant tube does not substantially impact the back-pressure generated in the vocal tract. Several studies have theoretically (Horacek *et al.*, 2014; Titze, 2006; Titze *et al.*, 2002) and empirically (Guzman *et al.*, 2013c; Radolf *et al.*, 2014; Titze *et al.*, 2002) demonstrated an increase in subglottal pressure and a decrease in transglottal pressure due to large back-pressure generated from phonating into the narrow flow-resistant tubes. The increase in back pressure affects the mechanical stiffness of the subglottal airway walls, characterized by the Young’s Modulus, by stretching the airway smooth muscle and connective tissues (this phenomenon is called strain stiffening). This strain stiffening in turn influences the frequencies of the subglottal airway resonances (Gunst and Stropp, 1988; Lulich and Arsikere, 2015). To the best of our knowledge, investigations of the influence of flow-resistant tube exercises on subglottal airway resonances are nonexistent. We therefore investigate changes in the subglottal acoustics during flow-resistant tube phonation in children. Since increases in the subglottal pressure affects the Young’s Modulus of the tracheal wall soft tissues, causing the tissues to stiffen and raise the frequency of the wall vibration resonance (Lulich and Arsikere, 2015), we predict that subglottal resonance frequencies will increase during flow-resistant tube phonation.

Non-nasal vocal tract acoustic sound typically radiates primarily from the lip opening, but will also radiate from the vocal tract walls (Fant *et al.*, 1976; Mik *et al.*, 2018; Munger and Thomson, 2008; Švec *et al.*, 2005). During flow-resistant tube exercises, sound radiation from the lips is expected to decrease as the impedance of the tube increases. Subglottal acoustic sound radiates from the skin of the neck below the level of the glottis in the absence of flow-resistant tube phonation (Lulich and Arsikere, 2015; McKenna *et al.*, 2017; Zañartu *et al.*, 2013), and is not expected to be affected by the use of flow-resistant tubes. We therefore

predict that the ratio of sound energy radiating from the vocal tract to the sound energy radiating from the neck will decrease as the impedance of the tube increases. Conversely, an observed decrease in the ratio during flow-resistant tube phonation serves as empirical evidence that the impedance of the flow-resistant tube is large enough to divert sound radiation from the lip opening to the vocal tract walls.

Changes in vocal tract and tongue shape during flow-resistant tube exercises have been investigated in a number of imaging studies involving adult subjects (Guzman *et al.*, 2013a; Guzman *et al.*, 2017a; Guzman *et al.*, 2013c; Guzman *et al.*, 2017b; Laukkanen *et al.*, 2012b). Theory predicts that the therapeutic target of flow-resistant tube exercises is a lengthening of the vocal tract and narrowing of the epilaryngeal tube. In 4 out of 5 studies the changes in vocal tract shape resulting from flow-resistant tube exercises have been small and in the direction of *larger* volumes throughout most of the vocal tract, including in the region of the epilaryngeal tube during flow-resistant tube phonation. Guzman *et al.* (2013c), in a computerized tomography (CT) study of one classically trained male, reported increased elevation of the velum, lowering of the larynx, widening of the hypopharynx, widening of the outlet of the epilaryngeal tube, and increased tongue dorsum height during phonation into a narrow flow-resistant tube. Similar changes in oropharyngeal physiology were also reported in a cohort of 10 adults (6 females and 4 males) with vocal hyperfunction on CT study (Guzman *et al.*, 2017b) and one vocally normal female with extensive experience with voice exercises in an magnetic resonance imaging (MRI) study (Laukkanen *et al.*, 2012b) and CT study (Vampola, Laukkanen, Horacek, and Švec, 2011). Contrary to the CT and MRI studies, however, Guzman *et al.* (2013a) found anterior-posterior compression of the epilaryngeal tube during phonation into a narrow flow-resistant tube in 21 patients with vocal hyperfunction on videostroboscopic examination. The lack of clear support for narrowing of the epilaryngeal tube during flow-resistant tube exercises does not necessarily disprove the theoretical predictions, since it could be that expansion of the pharyngeal cavity counteracts the expansion of the epilaryngeal tube and yields the same effect as a narrowed epilarynx (Vampola *et al.*, 2011). Nevertheless, since narrowing of the epilaryngeal tube is posited by theory as an adaptation based on increased vocal tract inertance experienced during flow-resistant tube exercises, it is important to discover both how the vocal tract inertance increases, and the role that changes in vocal tract and tongue shape may play in the pediatric population.

The anatomy and physiology of the vocal folds are different in children than in adults, with children exhibiting an underdeveloped *vocalis* muscle (Boseley and Hartnick, 2006; Sato *et al.*, 2001), increased pliability of the lamina propria, increased subglottal pressures (Keilmann and Bader, 1995), and altered patterns of vibration (Patel *et al.*, 2015; Patel *et al.*, 2014). Structural differences within the vocal folds may affect nonlinear source-filter interaction (Story and Bunton, 2015) and impact the outcomes of SOVT exercises in children. Findings from adults cannot be applied to infer performance of children during flow-resistant tube

phonation, and empirical data describing changes in the supraglottal vocal tract and acoustic output during flow-resistant tube phonation in children are lacking. The current study aims to quantify the most salient acoustic and vocal tract characteristics during narrow flow-resistant tube phonation in children using microphone, accelerometer, and ultrasound recordings. Compared to CT and MRI, ultrasound imaging is non-invasive and poses minimal risk, and hence is particularly advantageous for investigating the changes in vocal tract and tongue shape in children during flow-resistant tube phonation.

II. METHOD

A. Participants

A total of 21 children (boys = 9, girls = 12) in the age range of 6–9 years were recruited at the Speech Production Laboratory, Department of Speech and Hearing Sciences at Indiana University after signing of IRB approved informed consent/assent forms. Data from one boy were excluded as the participant was currently receiving Speech Language Pathology services for Developmental Speech Delay. Recordings from two females were not included due to errors in the recording protocol and failure of the participants to adequately perform the experimental tasks. Hence data from a total of 18 participants, 8 boys ($M = 8.5$ years, $SD = \pm 11$ months) and 10 girls ($M = 8.4$ years, $SD = \pm 12$ months) were analyzed for the current study. All of the participants were typically developing, had no history of speech, language and hearing disorders, were non-singers, and were perceived to be native speakers of Midwest American English; we did not specifically inquire regarding speaker dialect.

B. Experimental protocol

Simultaneous recordings of 3D/4D ultrasound, acoustic properties of participants' speech, and subglottal acoustics were made with a Philips EPIQ 7G ultrasound system and xMatrix x6–1 digital 3D/4D transducer, a SHURE KSM32 microphone, and a KandK Sound HotSpot accelerometer, respectively, while the participants were seated comfortably in a double-walled sound treated booth.

The SHURE KSM32 cardioid condenser microphone was situated on a stand at a fixed distance of 60 cm in front of the child and positioned to the right in order to avoid distortions due to high oral airflow. A flat frequency response with a low-frequency roll off filter of 6 dB/octave below 115 Hz was used with the SHURE KSM32 microphone. The accelerometer was secured against the skin on the neck at the level of the cricoid cartilage below the thyroid notch to measure the subglottal resonances (Chi and Sonderegger, 2007; Lulich, 2010; Lulich *et al.*, 2012; Wade *et al.*, 2017). The placement of the accelerometer below the thyroid notch ensures that the vocal tract formants are not present in the accelerometer signal. The accelerometer was held in place by a small piece of medical tape (Wade *et al.*, 2017) rather than applying pressure with the fingers to hold the accelerometer (Lulich *et al.*, 2012) in order to capture reliable recordings for trial-to-trial comparisons. The sampling rate

for both the microphone and the accelerometer was 48 kHz and the quantization rate was 16 bits/sample.

The ultrasound transducer was manually held under the chin, between the anterior margin of the hyoid bone and the front of the mandible, to image tongue motion as the participants held their head as steady as possible for the duration of the recordings. Visual feedback of sagittal, coronal, and transverse image planes was also used to help ensure that the transducer was held in a stable orientation through each recording. The probe position was finely adjusted to obtain satisfactory midsagittal contours of the tongue. A layer of ultrasound gel was coated onto the transducer prior to placing the transducer under the chin, to ensure optimal image quality. In order to visualize the anatomical landmarks of interest in the raw ultrasound data, the recording frame rates and image sizes were adjusted based on the anatomical landmarks for each child. The median frame rate was 19.25 fps (Table I), and ultrasound intensity was recorded with a depth of 8 bits. Technical characteristics of the ultrasound system and similar studies using this system are described in Charles and Lulich (2018), Lulich *et al.* (2018), and Lulich *et al.* (2017).

Each participant produced the following phonatory tasks for simultaneous recordings from the ultrasound, microphone, and accelerometer, in the following sequence: (1) three sustained vowel /o:/ productions and (2) three sustained productions of a vowel (instructed to be the same as /o:/) while holding a narrow flow-resistant tube (a plastic stirring straw, 2.53 mm inner diameter and 124 mm length) tightly between the lips, with a target duration of 5 s each. The participants were instructed to keep the lips firmly sealed around the flow-resistant tube without puffing out their cheeks. A nose clip was used during the recordings with the flow-resistant tube, to ensure no leakage of air

TABLE I. Frame rates (frames per second, rounded to the nearest tenth of a second) and voxel sizes (mm, rounded to the nearest hundredth of a millimeter; in medial-lateral, anterior-posterior, and superior-inferior dimensions) for ultrasound recordings.

Gender	Medial-lateral (mm)	Anterior-posterior (mm)	Superior-inferior (mm)	Frame rate (FPS)
M	0.80	0.78	0.40	17.4
M	0.71	0.69	0.36	20.3
M	0.71	0.69	0.36	20.3
F	0.81	0.78	0.40	19.3
F	0.71	0.69	0.35	19.3
M	0.71	0.69	0.35	19.3
F	0.71	0.69	0.35	16.6
F	0.71	0.69	0.36	20.3
F	0.81	0.78	0.40	19.3
F	0.81	0.78	0.40	19.3
F	0.71	0.69	0.35	19.3
F	0.71	0.78	0.41	20.3
F	0.81	0.78	0.41	20.3
M	0.81	0.78	0.41	20.3
F	0.71	0.69	0.35	19.3
M	0.71	0.69	0.35	19.3
M	0.81	0.78	0.40	20.3
M	0.71	0.69	0.36	20.3

through the nose (Titze, 2009). Prior to recordings with the flow-resistant tube, practice trials were conducted for each participant with a larger drinking straw (5.18 mm inner diameter and 196.85 mm length) and subsequently with the flow-resistant tube (stirring straw). Each participant required approximately 1 trial with the drinking straw and 2–3 trials with the flow-resistant tube to achieve the target flow-resistant tube phonation for the experiment. The training with the flow-resistant tubes started with the examiner demonstrating the production emphasizing lip seal around the straw and air flow at the end of the straw. The participants were subsequently instructed to produce the target phonations with the flow-resistant tube. Both auditory and kinesthetic cues of airflow through the straw were provided if the target production was not achieved after 1 trail. Production of nasal sounds was monitored perceptually and through visual inspection of the acoustic recordings of the oral sounds. All training was provided by the first author who has over 20 years of clinical experience in the area of voice.

Three trials for each task were recorded. Additional trials of each production were conducted if the productions were judged to be unsatisfactory. Both tasks were performed at the participants’ self-selected vocal frequency and sound levels to minimize the effects of increased effort during targeted productions of vocal frequency and sound level (Hanson *et al.*, 1990). The total duration for data collection was approximately 30 min for each participant.

The experimental tasks were not randomized for this study in order to obtain a representative sample of the two groups (with and without flow-resistant tube), since flow-resistant tube phonation could have an immediate adaptive effect (Guzman *et al.*, 2013b; Laukkanen *et al.*, 2012b; Titze, 2009) and could therefore influence subsequent data collected without the flow-resistant tube. Pilot experimentation with different vowels (e.g., /o:/, /a:/, /i:/) revealed that it was easiest for children to produce the vowel /o:/ with the flow-resistant tube. Hence, the experimental tasks with and without flow-resistant tube were restricted to a single vowel /o:/.

C. Data analysis

Overall 108 acoustic recordings, 108 accelerometer recordings, and 108 ultrasound recordings were analyzed from 18 children in this study. Instantaneous synchronous measurements of the sustained /o:/ productions were made from the microphone, accelerometer, and ultrasound recordings. These instantaneous measurement points were generally located within the steady-state portion either after the beginning (discarding the onset) or the mid-point of the vowel and were selected based on visual inspection of the microphone and accelerometer waveforms and the relative strength of the formants and subglottal resonances on spectrographic displays of both the microphone and accelerometer signals. All study ratings were conducted by student research assistants and the first author. The students were given specific training to make the measurements reported in this study. The first and second authors verified all measurements. Additional training to the students was provided as needed so as to correct any measurement errors. Corrections

were also verified by the first and second authors. The study analyses began only after all corrections were made and verified.

1. Microphone and accelerometer spectral analysis

Microphone and accelerometer signals were analyzed using WAVESURFER (version 1.8.8p4) (Sjölander and Beskow, 2000). Both the signals were down-sampled to 8000 Hz for analysis. The measure of fundamental frequency was obtained from the discrete Fourier transform (DFT) line spectrum using a hamming window and 512 points from only the microphone (f_o) (Titze *et al.*, 2015). Manual measurements of formant frequencies F_1 , F_2 , and F_3 from the microphone [Fig. 1(A)] and the subglottal resonance frequencies F_{Sg1} , F_{Sg2} , and F_{Sg3} from the accelerometer signal [Fig. 1(B)] were guided by visual inspection of both the instantaneous DFT and linear predictive coding (LPC) spectra (Atal and Hanaver, 1971). The LPC order of 10 was generally used to estimate the formant frequencies and the subglottal resonances.

In 14 microphone signals and 9 accelerometer signals, minor adjustments to the analysis criteria were necessary in order to accurately represent the spectral envelope while still keeping the analyses comparable across trials. The use of long-term average spectra rather than instantaneous spectra was required for ten microphone signals ($range = 0.24\text{--}3.86$ s, $M = 1.16$ s) and 5 accelerometer trials ($range = 0.30\text{--}5.00$ s, $M = 1.43$). In 1 microphone trial and 3 accelerometer trials the LPC order was varied in order to obtain an accurate representation of the DFT spectral envelope (Buder, 1996; Vallabha and Tuller, 2002). An LPC order of 9 was used in 1 microphone trial, and LPC orders of 9, 11, and 15 were used for three accelerometer trials in instances where the 10-order LPC failed to adequately represent the DFT. In three trials from the microphone signal, adjusting the LPC order did not result in an adequate representation of the spectral envelope and hence the DFT line spectrum was used to estimate the F_1 and F_2 directly.

2. Microphone and accelerometer waveform analysis

Microphone and accelerometer waveform analysis. Microphone and accelerometer signals were analyzed using WAVESURFER (version 1.8.8p4) (Sjölander and Beskow, 2000). Both the signals were down-sampled to 8000 Hz for analysis. In order to determine the relative acoustic energy levels of the microphone and accelerometer signals, peak-to-peak (positive peak to negative peak) amplitudes (Baken and Orlikoff, 2007; Titze *et al.*, 2002) were measured at similar instantaneous time point as the DFT [Fig. 2(A)] and LPC spectra [Fig. 2(B)]. A ratio $P_{vt}:P_{sg}$ was obtained by dividing the peak-to-peak amplitude of the microphone waveform (P_{vt}) by the peak-to-peak amplitude of the accelerometer waveform (P_{sg}) to represent the ratio of the sound energy radiating from the vocal tract to the sound energy radiating from the neck. Since both P_{vt} and P_{sg} were measured directly from uncalibrated voltage signals, the $P_{vt}:P_{sg}$ represents an interval scale and is dimensionless. Related peak-to-peak amplitudes with intraoral pressure (P_o) and estimated subglottal pressure (P_{sub}) have been reported previously in similar studies by Radolf *et al.* (2014) and Horacek *et al.* (2014), but the ratio has not previously been used as a measure of the ratio of sound energy radiating from the vocal tract versus the neck. A close analog of this ratio has been used in several studies of the contribution of vocal tract resonances to the sound produced by clarinets and saxophones (Lulich *et al.*, 2017; Scavone *et al.*, 2008; Wilson, 1996). In the analog case, the ratio has been defined by the sound pressure level in the vocal tract and the sound pressure level in the instrument mouthpiece.

3. Ultrasound image analysis

Ultrasound image analysis. The ultrasound recordings were analyzed using a custom MATLAB toolbox (Mathworks Inc., Natick, MA) 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

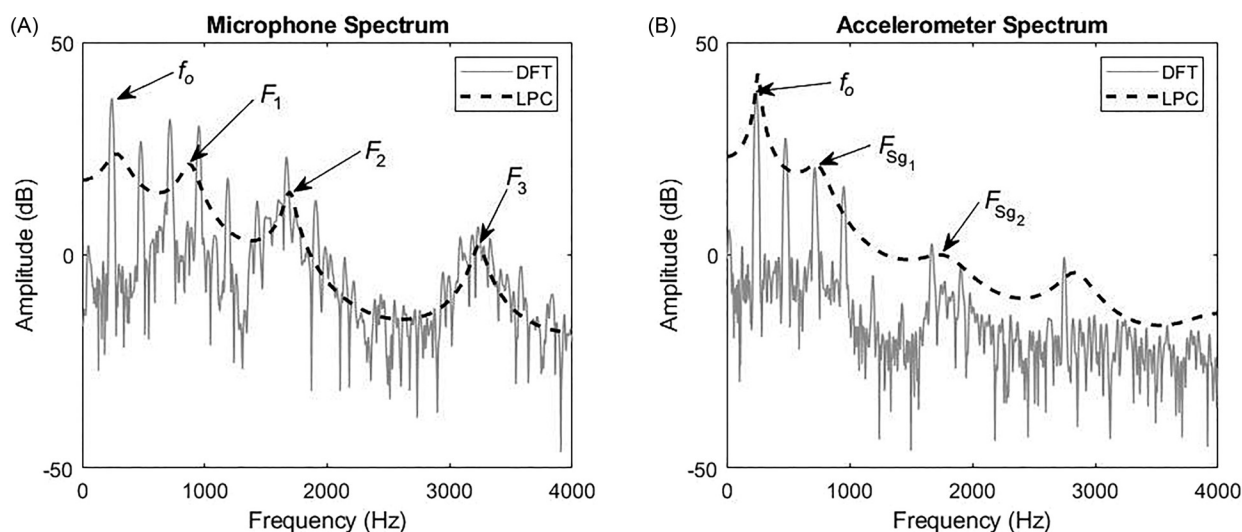


FIG. 1. DFT and LPC of a 9 year 1 month old boy for vowel /o:/ without the flow-resistant tube phonation with the (A) microphone signal depicting the fundamental frequency (f_o), first formant frequency (F_1), second formant frequency (F_2), and third formant frequency (F_3). (B) Accelerometer signal depicting f_o , first subglottal resonance (F_{Sg1}) and second subglottal resonance (F_{Sg2}).

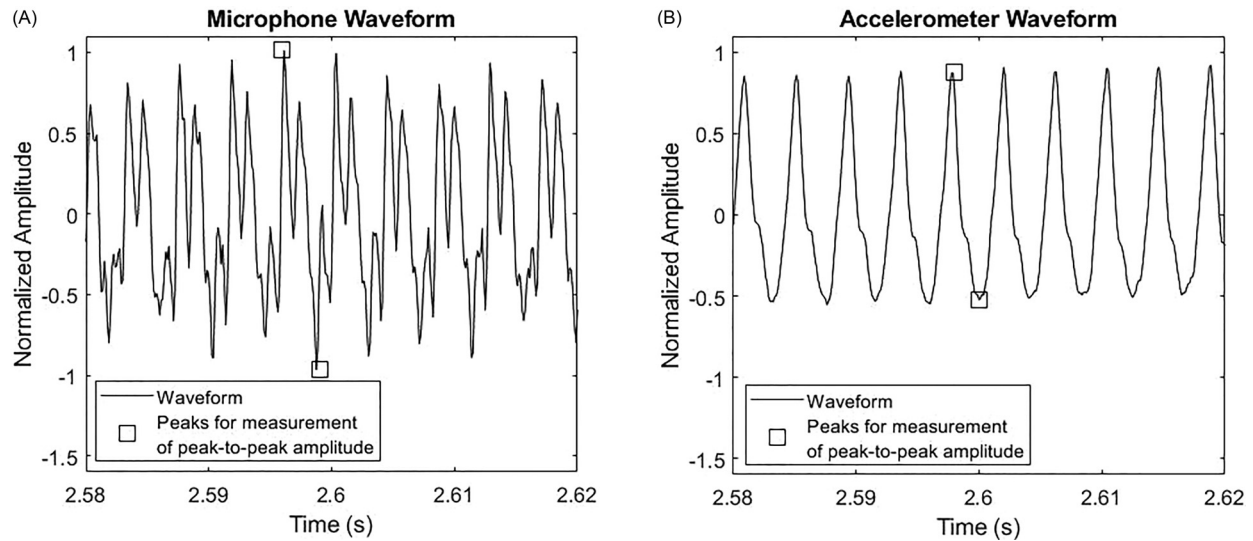


FIG. 2. Normalized peak-to-peak amplitude ($P_{vt}:P_{sg}$) measurement from the (A) acoustic waveform and the (B) accelerometer waveform from a 9 year 1 month old boy for vowel /o:/ without the flow-resistant tube phonation.

accelerometer waveforms. A single still image corresponding to the time point of the acoustic and accelerometer measurements was selected whenever possible for analysis. Since the tongue shape and position were steady for the sustained vowel production, a clear frame of the oral cavity representing the landmarks of interests could be selected in instances when the corresponding time point of the acoustic and accelerometer measurements did not yield clear landmarks (Lee *et al.*, 2015). One ultrasound recording did not contain any clear frames from which to make measurements, hence a total of 107 (of out 108 possible) ultrasound recordings were included in the analysis.

Various distances were manually measured in centimeters (cm) on the midsagittal ultrasound slice. The coronal and transverse ultrasound slices were used to guide the identification of landmarks of interest on the midsagittal ultrasound slice. These landmarks included: (1) the anterior insertion of the tendon of the genioglossus muscle into the mandibular symphysis (TnG) (Shawker *et al.*, 1984); (2) the anterior margin of the hyoid bone casting an acoustic shadow (HB); (3) the tongue root (TR) defined as the point where the surface of the posterior tongue surface intersects with the hyoid shadow; (4) the tongue dorsum (TD), which was defined as the highest part of the tongue body; (5) The tongue tip (TT) which was defined as the point where the surface of the anterior tongue intersects with the mandible shadow (Fig. 3).

In order to investigate the articulatory-acoustic relationship between tongue position and F_1 and F_2 , three vertical, two horizontal, and two additional Euclidian distances were measured. The three vertical distance measurements included: (1) Posterior tongue height (D1) measured as the vertical distance from TnG (the anterior insertion of the tendon of the genioglossus muscle into the mandibular symphysis) to TR; (2) tongue dorsum height (D2) measured as the vertical distance from TnG to TD; (3) tongue tip height (D3) measured as the vertical distance from TnG to TT. The tongue length (D4) was measured as the horizontal distance between TR and TT, and the width of the oral cavity (D5) was measured as the horizontal distance between TR and TnG.

Ultrasound does not allow direct measurement of epilar-yngeal or pharyngeal tube size because only the anterior wall of the pharynx/epilarynx is visible in ultrasound images of the tongue. In order to infer the enlargement or reduction of the epilaryngeal tube, measurements of hyoid elevation (D6) were obtained as the Euclidean distance between TnG and HB. An indirect measure of the anterior-posterior width of the pharynx (D7) was obtained as the Euclidian distance between HB and TR.

Two further observations were made as indirect indicators of pharyngeal expansion or reduction. First, the presence or absence of a midline groove in the tongue root was noted by examining coronal slices at the location of HB. The presence of a groove at this location was interpreted to suggest contraction of the posterior genioglossus muscle, with corresponding advancement of the tongue root and enlargement of the pharyngeal cavity. Second, the presence or absence of a midline groove in the tongue blade was noted by examining coronal slices at the location of TnG. The presence of a groove at this location was interpreted to suggest contraction of the anterior genioglossus muscle and displacement of tongue mass posteriorly, with corresponding reduction of the pharyngeal cavity.

D. Statistical analysis

Shapiro-Wilk tests were performed to test normality of the 12 dependent continuous variables of f_o , F_1 , F_2 , F_{Sg1} , $P_{vt}:P_{sg}$, D1, D2, D3, D4, D5, D6, D7. A linear mixed model analysis was used for investigating the normally distributed dependent variables to account for repeated measures. Kruskal-Wallis tests were used for analysis of the non-normally distributed dependent variables. Bonferroni correction was used to determine the significance level for all variables. The p values for each comparison were multiplied by the number of comparisons ($n=12$) to obtain the Bonferroni-adjusted p values, which is equivalent to using a critical value of $\alpha=0.05/12=0.0042$. A linear mixed effects logistic regression model was used to analyze the binary

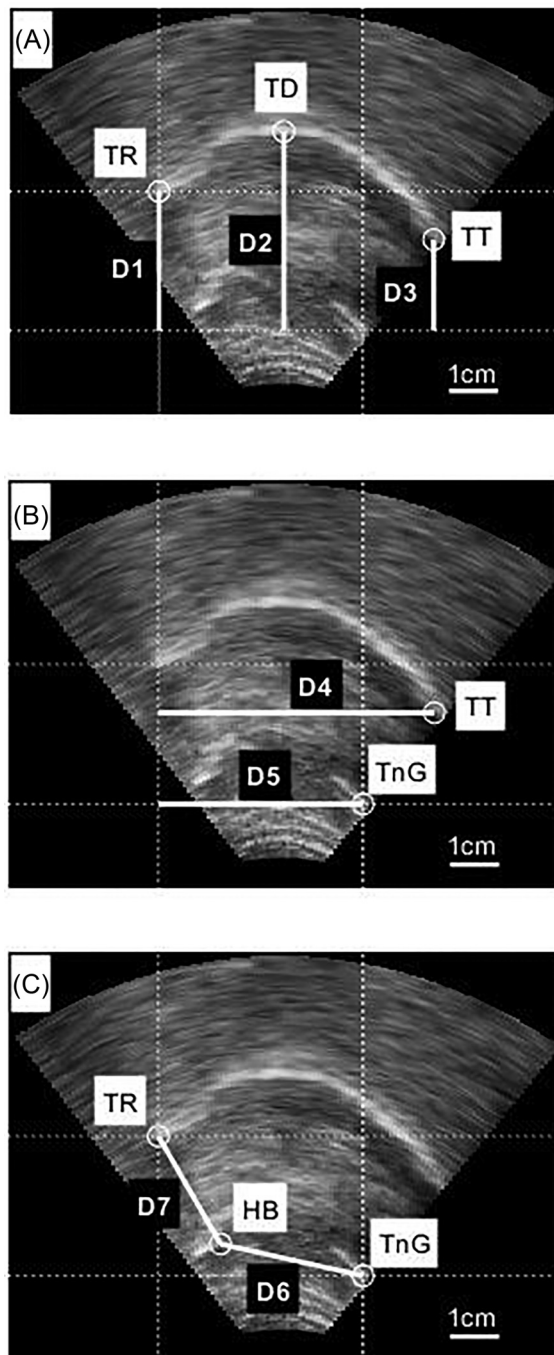


FIG. 3. Schematic representation of the (A) three vertical distances (D1-D3), (B) two horizontal distances (D4-D5), and (C) the two Euclidian distances measured (D6-D7); and the landmarks of TR, TD, TT, HB, and the anterior insertion of the tendon of the genioglossus muscle into the mandibular symphysis (TnG) from a 9 year 1 month old boy for vowel /o:/ with-out the flow-resistant tube phonation.

variables of presence or absence of posterior and anterior tongue groove, with alpha level of $p \leq 0.05$. Statistical analysis was performed using SAS statistical software version 9.4, for PCs (SAS Institute, n.d.) for all parametric variables, and using MATLAB software (Mathworks, Inc., Natick, MA) for all non-parametric variables.

Post hoc analyses were conducted on the statistically significant variables across condition and gender to determine whether variables were correlated with each other.

Cohen's d effect sizes were calculated for each of the statistically significant variables.

E. Reliability of measurements

A total of 11 trials were randomly selected for analysis of inter-rater and intra-rater reliability for the acoustic and accelerometer measurements. The inter-rater and intra-rater reliability for acoustic and accelerometer measurements were conducted between two trained undergraduate students majoring in Speech Language Pathology. The Pearson product-moment correlations for each measure were calculated between the initial and subsequent measurements for each variable. The Pearson product moment correlations for interrater reliability for the f_o , F_1 , F_2 , and $P_{vt}:P_{sg}$ measurements from the microphone signal exceeded 0.96 with a mean absolute difference of 6.64, 20.27, 48.82 Hz, and 47.46 (less than 2.86%) respectively. Similarly, the intrarater reliability for f_o , F_1 , F_2 , and $P_{vt}:P_{sg}$ measurements from the microphone signal exceeded 0.90 with a mean absolute difference of 1.82, 43.18, 96.73 Hz, and 58.73 (less than 5%). For the accelerometer signal the correlations were 0.38 and 0.98 for F_{Sg1} and $P_{vt}:P_{sg}$, respectively for interrater reliability, and 0.50 and 1, respectively, for intrarater reliability. The mean absolute difference was 275.36 Hz and 88.73 for F_{Sg1} and $P_{vt}:P_{sg}$, respectively for interrater reliability, and 0 Hz and 73.91 for intrarater reliability (less than 8%). One F_{Sg1} required average long-term average selection and changing the LPC order to 11. After removing this one trial the F_{Sg1} interrater reliability was 0.82 and intrarater reliability was 0.91 for F_{Sg1} .

For the ultrasound recordings a total of 12 trials were randomly selected for inter-rater and intra-rater reliability analysis. The intrarater reliability for the ultrasound recordings was conducted by a trained doctoral student and the interrater reliability was across two raters (doctoral student and the first author). The Pearson product-moment correlations for interrater reliability for the measurements of D1, D2, D3, D4, D5, D6, and D7 exceeded 0.80 with mean absolute differences of 0.26, 0.11, 0.18, 0.13, 0.16, 0.12, and 0.26 cm, respectively. The intrarater reliability for D1, D2, D3, D4, D5, D6, and D7 exceeded 0.80 with mean absolute differences of 0.33 cm (16.49%), 0.11 cm (3.16%), 0.22 cm (9.500%), 0.22 cm (4.69%), 0.16 cm (4.09%), 0.14 cm (5.08%), 0.25 cm (11.12%), respectively.

III. RESULTS

The third formant (F_3) was identifiable in only 11 (9 = without flow-resistant tube phonation, 2 = with flow-resistant tube phonation) of 108 recordings, and hence further analyses with F_3 and higher formants were not pursued. In this study both the second (F_{Sg2}) and third subglottal (F_{Sg3}) resonances could not be measured for either the flow-resistant tube phonation or the no flow-resistant tube phonation conditions for any recordings. Measurements of F_3 , F_{Sg2} , and F_{Sg3} could not be made reliably due to the high fundamental frequency of the children (Story and Bunton, 2015) combined with the low-pass filtering effects of the skin of the neck (Lulich et al., 2012) for sound radiated from

the vocal tract walls, and the added resistance of the flow-resistant tube.

Results from the Shapiro-Wilk test revealed that all variables except the width of the pharynx (D7), fundamental frequency (f_o), and $P_{vt}:P_{sg}$ were normally distributed (Table II). Descriptive statistics of the dependent variables are presented in Table III.

A. Main effect of flow resistant tube phonation across acoustic and ultrasound variables

The fundamental frequency was not significantly different between the two flow-resistant tube phonation conditions [$H(1)=0.52$, $p=0.47$] [Fig. 4(A)]. As expected, F_1 [$F(1, 88)=609.96$, $p<0.01$, Cohen's $d=4.03$] and F_2 [$F(1, 88)=108.32$, $p<0.01$, Cohen's $d=1.38$] frequencies were significantly lower during the flow-resistant tube phonation [Figs. 4(B) and 4(C)]. The F_1 during flow-resistant tube phonation had a mean frequency of 248.48 Hz (SD = 21.03 Hz), and a mean frequency of 621.15 Hz (SD = 128.96 Hz) without the flow-resistant tube phonation. The F_2 during the flow-resistant tube phonation had a mean frequency of 957.80 Hz (SD = 259.77 Hz) during the flow-resistant tube phonation and a mean frequency of 1284.81 Hz (SD = 212.48 Hz) without the flow-resistant tube phonation. The subglottal resonance (F_{Sg1}) was significantly higher for the flow-resistant tube phonation ($M=755.07$ Hz, SD = 95 Hz) compared to without the flow-resistant tube phonation ($M=693.54$ Hz, SD = 109.84 Hz) [$F(1, 88)=9.74$, $p<0.01$, Cohen's $d=0.60$] [Fig. 4(D)]. The ratio $P_{vt}:P_{sg}$ was significantly lower during the flow-resistant tube phonation ($MDN=0.81$) compared to without the flow-resistant tube phonation ($MDN=1.21$) [$H(1)=7.11$, $p<0.01$, Cohen's $d=0.40$] [Fig. 4(E)].

Phonation using the flow-resistant tube resulted in a significant increase in the tongue dorsum height (D2) ($M=3.55$ cm, SD = 0.37 cm) compared to the no flow-resistant tube phonation condition ($M=3.28$ cm, SD = 0.43 cm) [$F(1, 87)=35.58$, $p<0.01$, Cohen's $d=0.68$] [Fig. 5(A)]. Tongue tip height (D3) was also significantly higher ($M=2.20$ cm, SD = 0.66 cm) in the flow-resistant tube phonation condition compared to the no flow-resistant tube phonation condition ($M=2.01$ cm, SD

= 0.51 cm) [$F(1, 87)=8.96$, $p<0.01$, Cohen's $d=0.33$] [Fig. 5(B)]. There was no significant difference in the posterior tongue height (D1) [$F(1, 87)=1.65$, $p=0.20$]; tongue length (D4) [$F(1, 87)=0.08$, $p=0.78$]; width of the oral cavity (D5) [$F(1, 87)=0.08$, $p=0.78$]; hyoid elevation (D6) [$F(1, 87)=0.83$, $p=0.36$], and width of the pharynx (D7) [$H(1)=0.96$, $p=0.33$]. The presence of an anterior groove was 1.9 times more common in the condition without the flow-resistant tube phonation than in the flow-resistant tube phonation condition [$t(1, 87)=1.30$, $p=0.20$], however, the group difference in the presence versus absence of an anterior groove was not statistically significant. The linear mixed effects logistic regression model did not converge for the posterior groove binary variable, as the posterior groove was predominantly present in both conditions.

B. Main effect of gender for flow resistant tube phonation across acoustic and ultrasound variables

For all the acoustic variables, there was no significant main effect of gender for F_1 [$F(1, 88)=0.96$, $p=0.33$]; F_2 [$F(1, 88)=0.39$, $p=0.53$]; F_{Sg1} [$F(1, 88)=0.87$, $p=0.35$]; f_o [$H(1)=0.22$, $p=0.64$]; or $P_{vt}:P_{sg}$ ratio [$H(1)=3.44$, $p=0.06$].

Posterior tongue height (D1) was significantly higher in boys ($M=1.84$ cm, SD = 0.60 cm) compared to girls ($M=1.35$ cm, SD = 0.65 cm) [$F(1, 87)=4.02$, $p=0.05$, Cohen's $d=0.78$] [Fig. 5(C)]. There was no significant main effect of gender for the tongue dorsum height (D2) [$F(1, 87)=1.32$, $p=0.25$]; tongue tip height (D3) [$F(1, 87)=0.06$, $p=0.81$]; tongue length (D4) [$F(1, 87)=0.01$, $p=0.94$]; width of the oral cavity (D5) [$F(1, 87)=0.01$, $p=0.92$]; hyoid elevation (D6) [$F(1, 87)=0.31$, $p=0.58$]; and width of the pharynx (D7) [$H(1)=3.36$, $p=0.07$]. The presence of an anterior tongue groove was 5.6 times more likely in girls than in boys [$t(1, 87)=1.87$, $p=0.06$], however, this finding was not statistically significant.

Interaction between condition and gender for flow resistant tube phonation across ultrasound and acoustic variables were not statistically significant.

C. Correlations between articulatory and acoustic variables

The tongue dorsum height (D2) was significantly correlated with both the first formant (F_1) [$\rho=-0.30$; $p<0.01$] [Fig. 6(A)] and the second formant (F_2) [$\rho=-0.36$; $p<0.01$] after Bonferroni correction [Fig. 6(B)]. Correlations between F_1 (or F_2) and tongue tip height (or posterior tongue height) were not significant, nor was the correlation between F_{Sg1} and $P_{vt}:P_{sg}$ ratio [$-23 \leq \rho \leq 0.05$; $p \geq 0.01$; Bonferroni corrected alpha = $0.05/7 = 0.007$].

IV. DISCUSSION

This study examines vocal tract acoustic, subglottal acoustic, and vocal tract shape adjustments during narrow flow-resistant tube phonation in typically developing pre-pubertal children. Eighteen speakers (boys = 8, girls = 10) participated in the experiment using a narrow flow-resistant

TABLE II. Shapiro-Wilk test of normality for the ultrasound and acoustic dependent variables.

Dependent variables	Shapiro-Wilk statistic (s-w)	df	p
Posterior tongue height (D1)	0.977	18	0.066
Tongue dorsum height (D2)	0.988	18	0.472
Tongue tip height (D3)	0.980	18	0.107
Tongue length (D4)	0.988	18	0.423
Width of the oral cavity (D5)	0.992	18	0.779
Hyoid elevation (D6)	0.978	18	0.073
Width of pharynx (D7)	0.973	18	0.028 ^a
Fundamental frequency (f_o)	0.923	18	<0.001 ^a
First formant frequency (F_1)	0.978	18	0.065
Second formant frequency (F_2)	0.981	18	0.125
First subglottal resonance frequency (F_{Sg1})	0.993	18	0.889
Peak-to-peak amplitude ratio ($P_{vt}:P_{sg}$)	0.863	18	<0.001 ^a

^aIndicates that the variables were non-normally distributed.

TABLE III. Mean and standard deviations of the dependent variables with and without flow-resistant tube phonation: posterior tongue height (D1), tongue dorsum height (D2), tongue tip height (D3), tongue length (D4), width of the oral cavity (D5), hyoid elevation (D5), width of the pharynx (D7), fundamental frequency (f_o), first formant frequency (F_1), second formant frequency (F_2), first subglottal resonance (F_{Sg1}), and peak-to-peak amplitude ratio ($P_{vt}:P_{sg}$).

Condition	Gender	Variable	N	Mean	Standard deviation	Minimum	Maximum	Range
Without flow-resistant tube	F	D1	29	1.23	0.78	-0.19	2.45	2.64
		D2	29	3.22	0.32	2.72	3.84	1.12
		D3	29	2.01	0.46	1.40	2.89	1.49
		D4	29	4.79	0.45	4.04	5.46	1.42
		D5	29	3.83	0.28	3.30	4.30	1.00
		D6	29	2.78	0.36	2.09	3.72	1.63
		D7	29	2.07	0.74	0.66	3.49	2.83
		f_o	30	250.77	25.93	212.00	313.00	101.00
		F_1	30	599.00	100.72	425.00	840.00	415.00
		F_2	30	1298.80	191.58	911.00	1630.00	719.00
		F_{Sg1}	30	668.13	103.24	455.00	881.00	426.00
		$P_{vt}:P_{sg}$	30	1.35	0.93	0.31	4.34	4.04
	M	D1	24	1.86	0.73	0.24	3.11	2.87
		D2	24	3.35	0.53	2.55	4.14	1.59
		D3	24	2.01	0.59	1.19	3.47	2.28
		D4	24	4.72	0.57	3.48	5.64	2.16
		D5	24	3.84	0.44	2.71	4.36	1.65
		D6	24	2.61	0.29	2.08	3.20	1.12
		D7	24	2.44	0.63	0.67	3.38	2.71
		f_o	24	261.96	51.12	192.00	394.00	202.00
		F_1	24	648.83	155.21	364.00	941.00	577.00
		F_2	24	1267.33	239.13	749.00	1670.00	921.00
		F_{Sg1}	24	725.29	111.67	516.00	982.00	466.00
		$P_{vt}:P_{sg}$	24	1.56	0.86	0.57	4.35	3.79
With flow-resistant tube	F	D1	30	1.47	0.47	0.58	2.37	1.79
		D2	30	3.44	0.29	2.67	3.97	1.30
		D3	30	2.15	0.56	1.03	3.48	2.45
		D4	30	4.73	0.47	3.49	5.42	1.93
		D5	30	3.84	0.27	3.12	4.31	1.19
		D6	30	2.73	0.38	2.04	3.62	1.58
		D7	30	2.11	0.66	0.80	3.47	2.67
		f_o	30	243.67	18.45	222.00	273.00	51.00
		F_1	30	246.37	19.22	222.00	283.00	61.00
		F_2	30	994.30	283.89	749.00	1721.00	972.00
		F_{Sg1}	30	759.70	89.91	627.00	992.00	365.00
		$P_{vt}:P_{sg}$	30	0.96	0.81	0.11	3.40	3.29
	M	D1	24	1.81	0.45	0.93	2.52	1.59
		D2	24	3.68	0.42	3.21	4.59	1.38
		D3	24	2.27	0.77	1.20	3.71	2.51
		D4	24	4.81	0.52	3.43	5.74	2.31
		D5	24	3.86	0.37	3.03	4.60	1.57
		D6	24	2.75	0.28	2.42	3.36	0.94
		D7	24	2.21	0.53	1.20	3.25	2.05
		f_o	24	248.58	25.35	212.00	303.00	91.00
		F_1	24	251.13	23.24	212.00	303.00	91.00
		F_2	24	912.17	223.54	526.00	1326.00	800.00
		F_{Sg1}	24	749.29	102.67	506.00	911.00	405.00
		$P_{vt}:P_{sg}$	24	1.26	0.92	0.28	3.70	3.42

tube with a 2.53 mm inner diameter and 124 mm length, with simultaneous recordings of microphone, accelerometer, and 3D/4D ultrasound, resulting in a total of 108 acoustic recordings, 108 accelerometer recordings, and 107 ultrasound recordings. The results of the present study indicate that the magnitude of F_1 , and F_2 adjustments in children is similar to findings from adults during flow-resistant tube phonation. Relative tongue dorsum height adjustments made by children

during flow resistant-tube phonation are also similar to findings from adults. Additional new findings of increased first subglottal resonance and $P_{vt}:P_{sg}$ ratio in children are also reported.

A. Acoustic variables

The average f_o , F_1 , and F_2 values for the sustained /o:/ in the no-flow-resistant tube phonation condition were

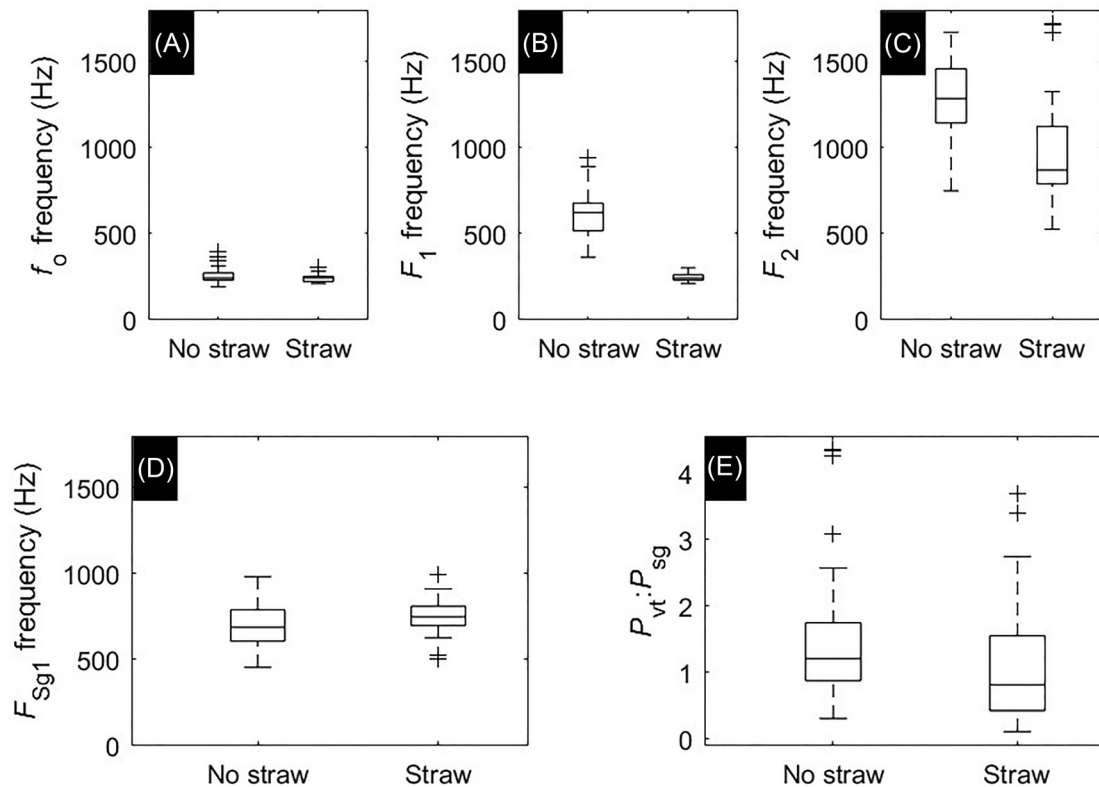


FIG. 4. Box and whisker plots showing in fundamental frequency (f_o) (A), first formant frequency (F_1), second formant frequency (F_2), first subglottal resonance (F_{Sg1}), and peak-to-peak amplitude ($P_{vt}:P_{sg}$) in the two phonation conditions of with- and without- narrow flow resistant tube [(B), (C), (D), (E), respectively].

similar to, but slightly higher than, measurements reported for 10- to 12-year-old children by Hillenbrand *et al.* (1995), who found average $f_o = 236$ Hz, $F_1 = 597$ Hz, and $F_2 = 1137$ Hz. The measured values in the current study were slightly higher in frequency ($f_o = 243$ Hz, $F_1 = 621$ Hz, and $F_2 = 1285$ Hz) because our participants were in a younger age bracket (6- to 9-year-old children). Lee *et al.* (1999) carried out a more detailed acoustic study of children's vowels at each age from 5 to 17 years, but did not include the vowel /o/ because it is typically diphthongized in speech. The participants in Hillenbrand *et al.* (1995) were from Michigan, Illinois, Wisconsin, Minnesota, Ohio, and Indiana, and therefore offer an excellent match for our participants in terms of dialect. More recently, Ramos and

Gama (2017) measured f_o in the vowels /a/ and /e/ from 27 children aged 5 to 10 years, and reported average values between 255 and 267 Hz following the intrasubject control condition of vocal rest, although their participants were children with vocal nodules or cysts.

The f_o was not affected during the flow-resistant tube phonation, which is in agreement with two previous reports for adults (Andrade *et al.*, 2014; Gaskill and Erickson, 2010). One study reported a slight decrease in f_o from 164 Hz during vowel production to 158 Hz during flow-resistant tube phonation for one female subject (Radolf *et al.*, 2014). A similar decrease in f_o is observed in the current study in both boys ($n = 8$) and girls ($n = 10$), however, the group differences were not statistically significant.

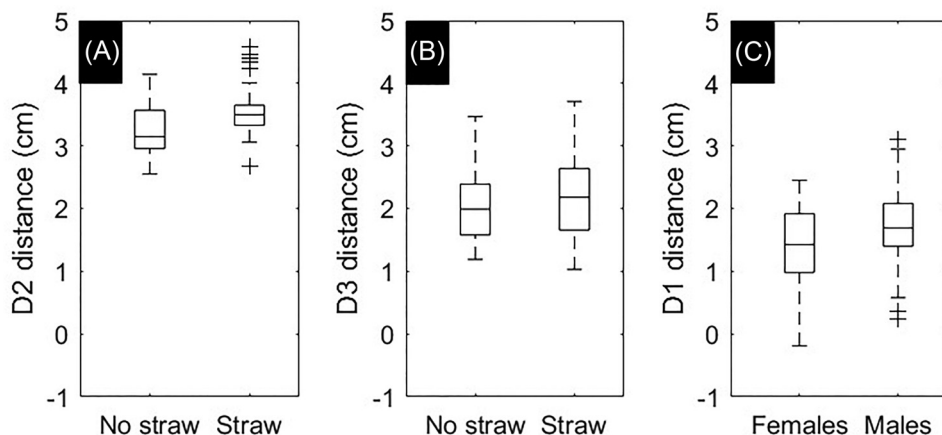


FIG. 5. Box and whisker plots showing statistically significant differences in tongue dorsum height (D2) and tongue tip height (D3) in the two phonation conditions of with- and without- narrow flow resistant tube [(A) and (B)], and statistically significant difference across gender in posterior tongue height (D1) (C).

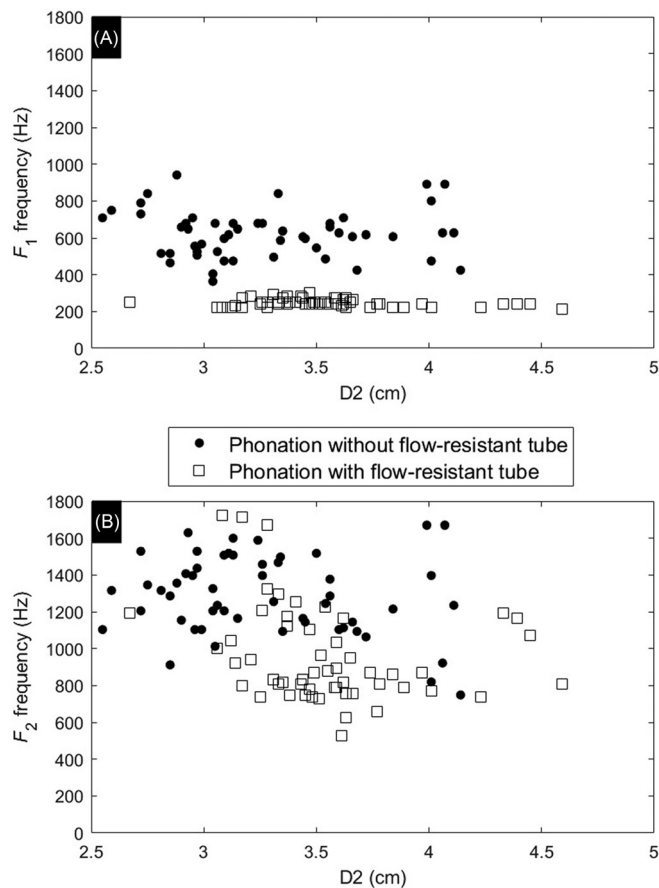


FIG. 6. Scatter plots showing the relationships between first formant frequency (F_1) and tongue dorsum height (D2) (A), and between second formant frequency (F_2) and tongue dorsum height (D2) (B).

Ramos and Gama (2017) reported no change in f_o following flow-resistant tube phonation in children with vocal nodules or cysts; however, did not report the effect of flow-resistant tube phonation on f_o during the exercise.

The frequency of F_1 was lowered substantially during flow-resistant tube-phonation to a mean value of 248 Hz. The lowering of F_1 during flow-resistant tube phonation agrees well with findings from adults in which the mean F_1 value during phonation with a larger diameter tube (8 mm inner diameter) was between 146 and 333 Hz (Andrade *et al.*, 2014; Gaskill and Erickson, 2010; Gaskill and Quinney, 2012), with mean values around 250 Hz, and with findings from Radolf *et al.* (2014) where an F_1 of 160 Hz was observed in a female subject during phonation with a 2.5 mm inner diameter flow-resistant tube. The frequency of F_2 was lowered during flow-resistant tube phonation, from 1285 to 958 Hz. Our finding of lowering of F_2 is in agreement with Gaskill and Erickson (2010) findings on adults who reported a mean F_2 of 979 Hz during tube phonation by adults, but contrasts with Radolf *et al.* (2014), who reported an increase from 720 Hz for normal phonation of the vowel /u:/ to 1220 Hz during phonation into a narrow flow-resistant tube. Overall, the close agreement of our findings from children with those from studies with adults suggest that the physiological mechanisms responsible for influencing change with the narrow flow-resistant tube exercise in children are similar to those of adults.

Subglottal resonances have been investigated in both adults (Chi and Sonderegger, 2007; Cranen and Boves, 1987; Csapo *et al.*, 2009; Dogil *et al.*, 2010; Hanna *et al.*, 2018; Ishizaka *et al.*, 1976; Lulich, 2010; Lulich *et al.*, 2012; van den Berg, 1960) and in children (Lulich, 2010; Lulich *et al.*, 2011b; Yeung *et al.*, 2018). Among children in the age range 6- to 9-years-old, the mean frequency of F_{Sg1} was 780 Hz in 3 boys, 2 girls (Lulich *et al.*, 2011b) and 744 Hz in 12 boys and 3 girls (Yeung *et al.*, 2018). The mean first subglottal frequency of 694 Hz found in the present study in the no-flow resistant tube condition is somewhat lower compared to previous findings in children (Lulich *et al.*, 2011b; Yeung *et al.*, 2018), but each are within one standard deviation of our mean value. The frequency of F_{Sg1} increased to 755 Hz during flow-resistant tube phonation, as expected based on Lulich and Arsikere (2015). Although this was a statistically significant increase, it represents an increase of less than one standard deviation, and is still very similar to the mean values reported for 6- to 9-year-old children by Lulich *et al.* (2011b) and for 6- to 18-year-old children in the study by Yeung *et al.* (2018). Furthermore, we could not rule out the possibility that the change observed in F_{Sg1} was due to other causes, such as larynx lowering, instead of, or in addition to, wall stiffening. Irrespective of the possible mechanism/s of change, the increase in F_{Sg1} frequency during flow-resistant tube phonation changes the input impedance of the subglottal airways, and alters the overall load impedance experienced by the vibrating vocal folds (Lulich and Arsikere, 2015; Titze, 2008).

To the best of our knowledge, $P_{vt}:P_{sg}$ measurements have not been reported during flow-resistant tube phonation, although Radolf *et al.* (2014) reported peak-to-peak fluctuations in intraoral pressure (related to our P_{vt}) and Horacek *et al.* (2014) reported both intraoral (P_o) and subglottal (P_{sub}) peak-to-peak pressure fluctuations separately. The statistically significant decrease in this ratio during flow-resistant tube phonation indicates that the signal obtained from the accelerometer was louder relative to the microphone signal during flow-resistant tube phonation. It is possible that the ratio of the peak-to-peak amplitudes in the microphone and accelerometer signals could decrease simply as a result of decreasing the mouth opening, however, we were careful to elicit a close /o:/ vowel in the trials without the straw, thus minimizing the change in mouth opening between conditions.

B. Ultrasound variables

Most articulatory measures of the tongue and pharynx were not found to statistically differ between the flow-resistant tube and no flow-resistant tube conditions. Raising the tongue dorsum and the tongue tip were the only statistically significant differences in children during flow-resistant tube phonation. The tongue dorsum was found to rise by an average distance of 0.27 cm. This is similar to, but somewhat smaller than, the amount of tongue dorsum raising of 0.35 cm (for their distance 3) observed by Laukkanen *et al.* (2012a) for one adult female, and an average difference of 1.2 cm for ten adults observed by Guzman *et al.* (2017b).

The greater magnitude of raising of the tongue dorsum in adults (Guzman *et al.*, 2017b; Laukkanen *et al.*, 2012b) compared with the present observations in children may be due to the larger size of these adults compared with the children enrolled in the present study. A further study, by Vampola *et al.* (2011), did not report measures directly comparable to those reported here, but did find for one adult that the vocal tract area function was contracted in the region of the tongue dorsum, consistent with the finding of tongue dorsum rising in the present study. In contrast to the finding of rising of tongue dorsum in the present study in children, Guzman *et al.* (2013c) reported tongue dorsum lowering during straw phonation in one adult male, with an average difference of 0.8 cm.

The tongue tip was found to be higher by an average distance of 0.19 cm in children during flow-resistant tube phonation. Laukkanen *et al.* (2012a) observed a lowering of the tongue tip (their distance 1) in one adult female by an average distance of 1.72 cm, which is the opposite of the effect reported here, but this could be due both to the smaller size of the children's vocal tracts in the present study, as well as to differences between the anatomical definitions of their distance 1 and our distance D3. Vampola *et al.* (2011) found that the vocal tract area function was expanded in the region of the tongue tip, which is also inconsistent with tongue tip elevation experienced by children in the current study.

Although investigations of vocal tract imaging in adults have consistently found expansion of the pharyngeal and epilaryngeal cavities (Guzman *et al.*, 2013c; Laukkanen *et al.*, 2012a; Vampola *et al.*, 2011) except for one endoscopic study (Guzman *et al.*, 2013a) during flow-resistant tube phonation, the ultrasound findings from our study do not statistically support expansion of the pharyngeal and epilaryngeal cavities in children. Even though statistically significant group differences were not observed for variables D6 and D7, overall, there was a trend to indicate expansion of the pharyngeal and epilaryngeal cavities during flow-resistant tube phonation in children. This trend of pharyngeal and epilaryngeal expansion is further strengthened by the fact that the tongue is a muscular hydrostat (Chi-Fishman and Stone, 1996; Kier and Smith, 1985; Lundberg and Stone, 1999; Stone *et al.*, 2001), and expansion of the pharynx and epilarynx displaces tongue mass upward, as observed in the significantly raised tongue dorsum (D2) during flow-resistant tube phonation. The lack of statistical significance could be due to the sample size, effect size, or that our ultrasound distance measurements of hyoid elevation (D6) and width of pharynx (D7) were inherently inadequate for making such observations. For instance, investigations of vocal tract imaging in adults have used MRI (Laukkanen *et al.*, 2012a), CT (Guzman *et al.*, 2013c; Vampola *et al.*, 2011), and endoscopic (Guzman *et al.*, 2013a) imaging modalities, which are capable of imaging the epilaryngeal and pharyngeal cavities. The 3D/4D ultrasound used in the current study is generally restricted to imaging only the suprahoid parts of the vocal tract. It is also possible that the children in this study lowered the larynx or raised the soft palate during flow-resistant tube phonation. Both of these

gestures have been observed before in adults (Guzman *et al.*, 2013c; Guzman *et al.*, 2017b; Vampola *et al.*, 2011), and would have the effect of expanding the vocal tract volume.

C. Acoustic theory of narrow flow-resistant tube phonation

The study findings suggests that ultrasound measurements from children reveal similar (though smaller) changes in tongue shape (raised tongue dorsum) to adults in response to flow-resistant tube phonation. During flow-resistant tube phonation children's F_1 and F_2 frequencies are remarkably similar to adults' F_1 and F_2 frequencies. While the differences between adult and child tongue shape data are explicable on the basis of the difference in vocal tract and tongue size, the nearly identical acoustic measures are contrary to what we should expect based on size differences alone. We here present a theoretical account of SOVT exercises with narrow flow-resistant tubes that accounts for the present acoustic findings from children, while remaining consistent with previous results from adults. While the following theoretical account is suggested as it successfully accounts for the empirical findings, it is better to consider it *proposed* rather than *proven*. Additional research is needed to test and evaluate the theory.

Previous theoretical acoustic models have predominantly been concerned with the tube length and diameter effects of SOVT exercises (Gaskill and Erickson, 2010; Horacek *et al.*, 2017; Mills *et al.*, 2017; Story *et al.*, 2000; Titze, 2006; Titze and Laukkanen, 2007). However, the low F_1 finding during narrow flow-resistant tube phonation from the current study cannot be fully explained by the length and diameter of the tube alone. We demonstrate this by showing that the predicted F_1 accounting for length and diameter of the tube is much lower than the observed values, as follows. Children in the age range from 6 to 9 years have relatively slow-growing vocal tracts with vocal tract lengths of approximately 12 cm (Vorperian *et al.*, 2009), and mean vocal tract cross-sectional areas around 3 cm² (Bunton *et al.*, 2013). The flow-resistant tubes used in the present study had a length of 12.4 cm and cross-sectional area of $A_{ftr} = \pi r^2 = 0.0502$ cm². To a first degree of approximation, the vocal tract and flow-resistant tube can be modeled simply as two uniform tubes with lossless input impedances of $Z_{vt} = -\rho c / A_{vt} \cot(kL_{vt})$ and $Z_{ftr} = \rho c / A_{ftr} \tan(kL_{ftr})$, respectively, where ρ is the density of air, c is the speed of sound in the vocal tract, k is the wave number ($k = 2\pi f / c$, where f is frequency), A_{vt} and A_{ftr} are the cross-sectional areas of the vocal tract and tube, respectively, and L_{vt} and L_{ftr} are the lengths of the vocal tract and tube, respectively. The natural frequencies (which are approximations of the formant frequencies) of this system are those for which the two impedances are equal (Stevens, 2000, pp. 136–148). For the vocal tract and tube dimensions given above, the frequency F_1 is 59 Hz, 75% lower than the observed values. The failure of this model to predict F_1 in children indicates that the low F_1 during narrow flow-resistant tube phonation must have another explanation beyond the effects of the length and diameter of the tube.

The low F_1 in children during narrow flow-resistant tube phonation can alternatively be explained by taking into account the vibration of the vocal tract walls. In this approach, the resonance F_1 is not directly affected by the length or diameter of the flow-resistant tube itself, but primarily by the volume of the vocal tract and the mass of the vocal tract walls, including the cheeks, neck, and lips. Damping losses within the air column and walls will affect the formant bandwidth, but will not have a substantial impact on the frequency (Stevens, 2000, pp. 152–162). The similar mechanism is responsible for the “voice bar” observed in voiced stop consonants (Badin and Fant, 1984; Fant, 1972, 1960; Fant *et al.*, 1976; Stevens, 2000), and the “breathing mode” observed in vocal tract acoustics during clarinet performance (Chen *et al.*, 2009; Fritz and Wolfe, 2005; Lulich *et al.*, 2017).

Before examining the predictions of this alternative voice bar account of the observed F_1 frequencies during flow-resistant tube phonation, it is necessary to determine the basic validity of the approach. In order to be valid, it must be shown that the impedance of the narrow flow-resistant tube is larger than the impedance of the vocal tract walls, for frequencies near f_o and higher. A large tube impedance will behave like an open-circuit, driving acoustic sound propagation to flow through the lower impedance of the vocal tract walls. Note that this acoustic flow is vibratory in nature, unlike direct airflow. The direct airflow must still escape from the mouth through the flow-resistant tube (Stevens, 2000, pp. 94–97). Magnitude spectra for the vocal tract wall impedance and the narrow flow-resistant tube impedance are shown in Fig. 7. At very low frequencies (below approximately 75 Hz), the impedance of the flow-resistant tube is smaller than the impedance of the vocal tract walls (Švec *et al.*, 2000), and thus low frequency sound and direct airflow escapes through the tube. At all other frequencies, including all frequencies near or above f_o , the narrow flow-resistant tube impedance is larger than the vocal tract

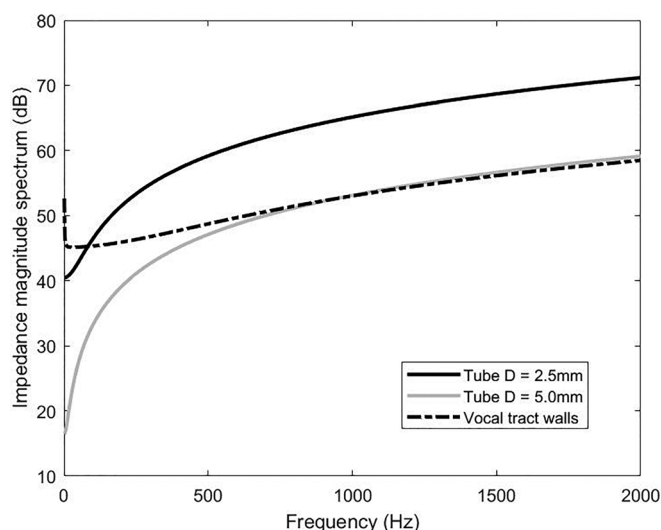


FIG. 7. Impedance magnitude spectra for (1) flow-resistant tubes with inner diameters of 2.5 and 5.0 mm, and (2) the vocal tract walls over the range of frequencies from 0 to 2000 Hz. The calculation of these impedances is presented in the Appendix.

wall impedance. Therefore, during narrow flow-resistant tube phonation, it is reasonable and valid to pursue the alternative account that F_1 during flow-resistant tube phonation is equivalent to the voice bar or breathing mode, and dependent primarily on the vocal tract volume and the mass of the vocal tract walls, rather than the length and diameter of the tube.

The frequency of F_1 according to this alternative voice bar approach accounting for the vocal tract volume and mass has been measured and modeled in a number of studies, and is found to fall between 150 and 300 Hz, depending on the speaker (Badin and Fant, 1984; Fant, 1972, 1960; Fant and Lindquist, 1968; Fant *et al.*, 1976; Fant and Sonesson, 1964; Fujimura and Lindqvist, 1971; Stevens, 2000; Story *et al.*, 2000; Titze, 2001). The mean F_1 frequency of 248 Hz during narrow flow-resistant tube phonation observed in the present study falls within this expected range of 150 and 300 Hz, and is thereby well accounted for. Thus, we propose that the predominant effect on F_1 of phonation through a narrow flow-resistant tube is likely to be attributable to the wall vibration rather than to the added length and narrow diameter of the tube itself.

It is important to emphasize that only narrow flow-resistant tube effects can be explained predominantly by vocal tract wall vibration. The degree of the effective occlusion is dependent on the inner diameter of the tube. Since the tube's inertance is inversely related to its cross-sectional area, i.e., $L_a = \rho c/A$, where $\rho = 1.14 \times 10^{-3}$ g/cm³ is the density of air, $c = 35900$ cm/s is the speed of sound in humid air at body temperature, $A = \pi(D/2)^2$ is the cross-sectional area, and D is the inner diameter (Fant, 1960; Stevens, 2000; Titze, 2006), every doubling of the inner diameter (e.g., from 2.5 to 5.0 mm) results in a decrease in tube impedance magnitude by approximately 12 dB. According to Fig. 7, the impedance of a tube with inner diameter 5.05 mm is just below the impedance of the vocal tract walls, and it is no longer valid to analyze F_1 as an analog of the voice bar. SOVT exercises using flow-resistant tubes with inner diameter smaller than 5 mm are therefore fundamentally different from SOVT exercises using larger diameter tubes. This fundamental difference should not be confused with the frequently made distinction that the predominant effective mechanism for large-diameter tubes is increased inertance in the f_o range, while the predominant effective mechanism for narrow-diameter tubes is increased resistance (Smith and Titze, 2017; Story *et al.*, 2000). The narrow-diameter tube also has an increased inertance as well as increased resistance (Story *et al.*, 2000), and both of these combined lead to the narrow-diameter tube behaving like a full occlusion (open circuit) in the f_o range, and the consequent recruitment of inertive vibrations of the vocal tract walls (Stevens, 2000, pp. 94–97).

Previous empirical and modeling studies that found decreases in F_1 to frequencies as low as 80 Hz made use of large diameter tubes (Horacek *et al.*, 2014; Horacek *et al.*, 2017; Story *et al.*, 2000). The study by Story *et al.* (2000) is particularly noteworthy in this respect because they explicitly included vocal tract wall vibrations in their model with a parameter $F_T \approx 200$ Hz, described as “the lowest resonance frequency of the tract when closed at both the glottal and lip ends” and explained as “a lower bound on the first formant frequency when the shunting effect of the yielding wall is

included.” They accordingly found a low F_1 of 190 Hz for a complete bilabial occlusion. Significantly, this means that the lowest F_1 they observed with a 100 cm extension tube (≈ 135 Hz) is more a resonance of the tube than a resonance of the vocal tract. This resonance will affect the load impedance seen by the vibrating vocal folds, but it will do so in a way that is very different from the mechanism we are proposing for narrow diameter tubes, namely by enhancing resonant vibrations within the extension tube itself rather than by enhancing vibrations of the vocal tract walls close to the phonation source.

Given that the predominant effect of the narrow flow-resistant tube is to occlude the vocal tract for all frequencies at or greater than f_o , the second formant (F_2) during narrow flow-resistant phonation must be considered a resonance of the vocal tract closed at both the glottal and lip ends. In a uniform, rigid-walled tube closed at both ends, a pressure node exists at the middle of the tube. Decreasing the cross-sectional area at this point (e.g., by raising the tongue dorsum) lowers all resonance frequencies of the tube except the voice bar resonance F_1 (Chiba and Kajiyama, 1941; Fant, 1960; Stevens, 2000). Similarly, expanding the cross-sectional area at the posterior end of the tube (e.g., by expanding the pharynx or epilarynx) lowers all resonance frequencies. Thus, the combined effects of tongue dorsum raising (which was observed in the present study) and pharyngeal/epilaryngeal expansion (observed in the present study, but without statistical significance) are to lower all resonance frequencies of the vocal tract except the voice bar resonance F_1 . Such a lowering of F_2 has been observed for vocal tract acoustics during clarinet performance, where adjusting the tongue shape for the vowel /o/ results in an F_2 value of approximately 900 Hz for adults (Lulich *et al.*, 2017), which is similar to the mean F_2 frequency of 958 Hz observed for children during narrow flow-resistant tube phonation in the present study.

The large impedance of the narrow flow-resistant tube also has the effect of blocking sound radiation through the tube. During narrow flow-resistant tube phonation, sound radiates preferentially from the vibrating walls of the vocal tract. The vocal tract walls act as a low-pass filter, attenuating high frequencies (Lulich *et al.*, 2011a; Meltzner *et al.*, 2003; Wu *et al.*, 2014), and it is this low-pass filtered signal that reaches the microphone during data collection. The attenuation of sound energy by the vocal tract walls accounts for the significant decrease we observed in $P_{vt} \cdot P_{sg}$ and provides empirical evidence of increased vocal tract inertance during flow-resistant tube phonation. It may also account for some of the difficulty we experienced in identifying and measuring formant frequencies and subglottal resonance frequencies above F_2 and F_{Sg2} , respectively.

V. CONCLUSIONS

The following conclusions could be drawn from the results of this study.

Children exhibit similar lowering of F_1 , F_2 and raising of tongue dorsum compared 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.

We propose that 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 ($P_{vt} \cdot P_{sg}$), together with increased first subglottal resonance frequency (F_{Sg1}), provides empirical support for increased vocal tract inertance near and above f_o predominantly due to vocal tract occlusion and enhancement of vocal tract wall vibrations rather than vocal tract lengthening for phonation through narrow flow-resistant tubes.

The flow-resistant tubes with inner diameter greater than 5 mm are fundamentally different from SOVT exercises using narrow diameter tubes as these larger tubes do not increase the vocal tract inertance predominantly due to vocal tract occlusion since the impedance of tubes with inner diameter of 5.05 mm is just below the impedance of the vocal tract walls.

The lack of statistically significant differences in the indirect measures of pharyngeal and epilaryngeal tube widths points to the need for additional future studies directly investigating changes in these cavities during SOVT exercises in children using more direct visualization methods, such as MRI and CT imaging.

The conclusion that F_1 during narrow flow-resistant tube phonation in both children and adults may be analogous to the voice bar, and hence due to inertive vibration of the vocal tract walls, warrants further investigation, especially using methods appropriate to directly measure vocal tract wall vibration, such as accelerometry (Meltzner *et al.*, 2003; Munger and Thomson, 2008).

The finding that F_{Sg1} changes in response to flow-resistant tube phonation was not conclusively linked to an increase in subglottal wall stiffness rather than larynx lowering or other factors. It is nevertheless noteworthy that a change was observed, suggesting a corresponding change in the load impedance experienced by the vibrating vocal folds (Lulich and Arsikere, 2015; Titze, 2008). It may therefore be important to consider the contribution of the subglottal impedance to voice production during SOVT exercises, and the mechanisms by which the F_{Sg1} changes.

The present study provides quantitative acoustic and ultrasound targets during narrow-flow resistant tube phonation in children which could potentially serve as quantitative predictive targets for training the narrow flow-resistant tube phonation voice therapy technique in children with voice disorders.

ACKNOWLEDGMENTS

This work was supported by the Indiana University, Department of Speech and Hearing Sciences Undergraduate Research Grant. The authors thank Brandon Merritt, Jennifer Philp, and Abigail Matthews for their assistance with data analysis, and three anonymous reviewers for helpful feedback.

APPENDIX

Flow-resistant tube impedances were calculated using Eq. (A1) (Stevens, 2000):

$$Z = \frac{128\mu}{\pi D^4} + k \frac{\rho U_g}{2A^2} + j\omega \frac{\rho l}{A}. \quad (A1)$$

In Eq. (A1), the first term refers to viscous losses from laminar airflow within the tube; the second term refers to losses due to turbulence at the tube outlet; and the third term refers to the acoustic mass of the air column within the tube. The diameter of the tube is D and its cross-sectional area is A . The length of the tube is 124 mm. The volume velocity flow was assumed to be $U_g = 500 \text{ cm}^3/\text{s}$ the density of air was $\rho = 1.14 \times 10^{-3} \text{ g/cm}^3$ the coefficient of viscosity of air was $\mu = 1.86 \times 10^{-4} \text{ dyn s/cm}^2$ The dimensionless constant k representing the pressure loss at the outlet of the tube had a value of 0.875.

The impedance of the vocal tract walls was calculated using Eq. (A2) and the mechanical constants reported by Ishizaka *et al.* (1975) for the cheeks,

$$Z = \frac{B + j(\omega M + K/\omega)}{S}. \quad (A2)$$

In Eq. (A2), the mass of the vocal tract walls is M , the stiffness is K , and the tissue viscosity is B . Ishizaka *et al.* (1975) reported values of M between 1.5 and 2.1 g for relaxed versus tensed cheeks, respectively; values of K between 33 300 and 84 500 dyn/cm; and values of B between 800 and 1060 g/s. The exact values used in the present study were $M = 2 \text{ g}$, $K = 75\,000 \text{ dyn/cm}$, and $B = 900 \text{ g/s}$ The circumference of the vocal tract was, $S = 6.14 \text{ cm}$ corresponding to a cross-sectional area equal to $A = 3 \text{ cm}^2$.

- Andrade, A. P., Wistbacka, G., Larsson, H., Sodersten, M., Hammarberg, B., Simberg, S., Svec, J. G., and Granqvist, S. (2016). "The flow and pressure relationships in different tubes commonly used for semi-occluded vocal tract exercises," *J. Voice* **30**(1), 36–41.
- Andrade, A. P., Wood, G., Ratcliffe, P., Epstein, R., Pijper, A., and Svec, J. G. (2014). "Electroglottographic study of seven semi-occluded exercises: LaxVox, straw, lip-trill, tongue-trill, humming, hand-over-mouth, and tongue-trill combined with hand-over-mouth," *J. Voice* **28**(5), 589–595.
- Atal, B. S., and Hanaver, S. L. (1971). "Speech analysis and synthesis by linear prediction of the speech wave," *J. Acoust. Soc. Am.* **50**(2), 637–655.
- Badin, P., and Fant, G. (1984). "Notes on vocal tract computation," *STL QPSR* **25**(2–3), 53–108.
- Baken, R. J., and Orlikoff, R. F. (2007). *Clinical Measurement of Speech and Voice*, 2nd ed. (Delmar Cengage Learning, Clifton Park, NY).
- Bhattacharyya, N. (2015). "The prevalence of pediatric voice and swallowing problems in the United States," *Laryngoscope* **125**(3), 746–750.
- Boseley, M. E., and Hartnick, C. J. (2006). "Development of the human true vocal fold: Depth of cell layers and quantifying cell types within the lamina propria," *Ann. Otol. Rhinol. Laryngol.* **115**(10), 784–788.
- Buder, E. H. (1996). "Experimental phonology with acoustic phonetic models: Formant measures from child speech," in *Proceedings of the UBC International Conference on Phonological Acquisition*, pp. 254–265.
- Bunton, K., Story, B., and Titze, I. (2013). "Estimation of vocal tract area functions in children based on measurement of lip termination area and inverse acoustic mapping," *Proc. Mtgs. Acoust.* **19**, 060054.
- Charles, S., and Lulich, S. M. (2018). "Case study of Brazilian Portuguese laterals using a novel articulatory-acoustic methodology with 3D/4D ultrasound," *Speech Commun.* **103**, 37–48.
- Chen, J. M., Smith, J., and Wolfe, J. (2009). "Pitch bending and glissandi on the clarinet: Roles of the vocal tract and partial tone hole closure," *J. Acoust. Soc. Am.* **126**(3), 1511–1520.
- Chi, X., and Sonderegger, M. (2007). "Subglottal coupling and its influence on vowel formants," *J. Acoust. Soc. Am.* **122**(3), 1735–1745.
- Chiba, T., and Kajiyama, M. (1941). *The Vowel: Its Nature and Structure* (Kaiseikan, Tokyo).
- Chi-Fishman, G., and Stone, M. (1996). "A new application for electropalatography: Swallowing," *Dysphagia* **11**(4), 239–247.
- Cranen, B., and Boves, L. (1987). "On subglottal formant analysis," *J. Acoust. Soc. Am.* **81**(3), 734–746.
- Csapo, T. G., Barkanyi, Z., Graczi, T. E., Bohm, T., and Lulich, S. M. (2009). "Relation of formants and subglottal resonances in Hungarian vowels," in *Proceedings of Interspeech*, pp. 484–487.
- da Silva, A. R., Ghirardi, A. C., Reiser, M. R., and Paul, S. (2018). "An exact analytical model for the relationship between flow resistance and geometric properties of tubes used in semi-occluded vocal tract exercises," *J. Voice* (published online).
- Dogil, G., Lulich, S. M., Madsack, A., and Wokurek, W. (2010). "Crossing the quantal boundaries of features: Subglottal resonances and Swabian diphthongs," in *Tones and Features: Phonetic and Phonological Perspectives*, edited by J. A. Goldsmith, E. Hume, and W. Wetzels (Walter de Gruyter, Berlin), pp. 137–148.
- Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, the Hague, the Netherlands).
- Fant, G. (1972). "Vocal tract wall effects, losses, and resonance bandwidths," *Speech Trans. Lab. Quart. Prog. Stat. Rep.* **13**(2–3), 28–52.
- Fant, G., and Lindqvist, J. (1968). "Pressure and gas mixture effects on diver's speech," in *Quarterly Progress and Status Report STL-QPSR*, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, Vol. 9(1), 7–17.
- Fant, G., Nord, L., and Branderud, P. (1976). "A note on the vocal tract wall impedance," *KTH Quart. Prog. Stat. Rep.* **17**(4), 13–20.
- Fant, G., and Sonesson, B. (1964). "Speech at high ambient air-pressure," in *Quarterly Progress and Status Report, Speech Transmission Laboratory*, Royal Institute of Technology (KTH), Stockholm, Sweden, Vol. 5(2), 9–21.
- Fric, M., and Hruska, V. (2017). "The effect of resonance tubes on facial and laryngeal vibration—A case study," *Biomed. Sign. Process. Control* **37**, 50–60.
- Fritz, C., and Wolfe, J. (2005). "How do clarinet players adjust the resonances of their vocal tracts for different playing effects?," *J. Acoust. Soc. Am.* **118**(5), 3306–3315.
- Fujimura, O., and Lindqvist, J. (1971). "Sweep-tone measurements of vocal tract characteristics," *J. Acoust. Soc. Am.* **49**(2), 541–558.
- Gaskill, C. S., and Erickson, M. L. (2010). "The effect of an artificially lengthened vocal tract on estimated glottal contact quotient in untrained male voices," *J. Voice* **24**(1), 57–71.
- Gaskill, C. S., and Quinney, D. M. (2012). "The effect of resonance tubes on glottal contact quotient with and without task instruction: A comparison of trained and untrained voices," *J. Voice* **26**(3), e79–e93.
- Granqvist, S., Simberg, S., Hertegard, S., Holmqvist, S., Larsson, H., Lindstad, P. A., Sodersten, M., and Hammarberg, B. (2014). "Resonance tube phonation in water: High-speed imaging, electroglottographic and oral pressure observations of vocal fold vibrations—A pilot study," *Logoped Phoniatr. Vocol.* **40**, 113–121.
- Gunst, S. J., and Stropp, J. Q. (1988). "Pressure-volume and length-stress relationships in canine bronchi in vitro," *J. Appl. Physiol.* **64**(6), 2522–2531.
- Guzman, M., Castro, C., Testart, A., Munoz, D., and Gerhard, J. (2013a). "Laryngeal and pharyngeal activity during semioccluded vocal tract postures in subjects diagnosed with hyperfunctional dysphonia," *J. Voice* **27**(6), 709–716.
- Guzman, M., Higuera, D., Fincheira, C., Munoz, D., Guajardo, C., and Dowdall, J. (2013b). "Immediate acoustic effects of straw phonation exercises in subjects with dysphonic voices," *Logoped Phoniatr. Vocol.* **38**(1), 35–45.
- Guzman, M., Jara, R., Olavarria, C., Caceres, P., Escuti, G., Medina, F., Medina, L., Madrid, S., Munoz, D., and Laukkanen, A. M. (2017a). "Efficacy of water resistance therapy in subjects diagnosed with behavioral dysphonia: A randomized controlled trial," *J. Voice* **31**(3), e381–e385.
- Guzman, M., Laukkanen, A. M., Krupa, P., Horacek, J., Svec, J. G., and Geneid, A. (2013c). "Vocal tract and glottal function during and after vocal exercising with resonance tube and straw," *J. Voice* **27**(4), e519–e534.
- Guzman, M., Miranda, G., Olavarria, C., Madrid, S., Munoz, D., Leiva, M., Lopex, L., and Bortnem, C. (2017b). "Computerized tomography measures during and after artificial lengthening of the vocal tract in subjects with voice disorders," *J. Voice* **31**(1), e121–e124.
- Hanna, N., Smith, J., and Wolfe, J. (2018). "How the acoustic resonances of the subglottal tract affect the impedance spectrum measured through the lips," *J. Acoust. Soc. Am.* **143**(5), 2639–2650.

- Hanson, D. G., Gerratt, B. R., and Berke, G. S. (1990). "Frequency, intensity, and target matching effects on photoglottographic measures of open quotient and speed quotient," *J. Speech Hear. Res.* **33**(1), 45–50.
- Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am.* **97**(5), 3099–3111.
- Horacek, J., Radolf, V., Bula, V., and Laukkanen, A. M. (2014). "Air-pressure, vocal folds vibration and acoustic characteristics of phonation during vocal exercising. Part 2. Measurement on a physical model," *Eng. Mech.* **21**(3), 193–200.
- Horacek, J., Radolf, V., and Laukkanen, A. M. (2017). "Low frequency mechanical resonance of the vocal tract in vocal exercises that apply tubes," *Biomed. Sign. Process. Control* **37**, 39–49.
- Ishizaka, K., French, J., and Flanagan, J. (1975). "Direct determination of vocal tract wall impedance," *IEEE Trans. Acoust. Speech Sign. Process.* **23**(4), 370–373.
- Ishizaka, K., Matsudaira, M., and Kaneko, T. (1976). "Input acoustic-impedance measurement of the subglottal system," *J. Acoust. Soc. Am.* **60**(1), 190–197.
- Kapsner-Smith, M. R., Hunter, E. J., Kirkham, K., Cox, K., and Titze, I. R. (2015). "A randomized controlled trial of two semi-occluded vocal tract voice therapy protocols," *J. Speech Lang. Hear. Res.* **58**, 535–549.
- Keilmann, A., and Bader, C. A. (1995). "Development of aerodynamic aspects in children's voice," *Int. J. Pediatr. Otorhinolaryngol.* **31**(2-3), 183–190.
- Kier, W., and Smith, K. (1985). "Tongues, tentacles and trunks: The biomechanics of movement in muscular-hydrostats," *Zoolog. J. Linn. Soc.* **83**(4), 307–324.
- Laukkanen, A. M. (1992). "Voiced bilabial fricative [β:] as a vocal exercise: An electroglottographic and acoustic investigation," *Scand. J. Logoped. Phoniater.* **17**(3-4), 181–189.
- Laukkanen, A. M., Horacek, J., and Havlik, R. (2012a). "Case-study magnetic resonance imaging and acoustic investigation of the effects of vocal warm-up on two voice professionals," *Logoped Phoniater. Vocol.* **37**(2), 75–82.
- Laukkanen, A. M., Horáček, J., Krupa, P., and Švec, J. (2012b). "The effect of phonation into a straw on the vocal tract adjustments and formant frequencies. A preliminary MRI study on a single subject completed with acoustic results," *Biomed. Sign. Process. Control* **7**, 50–57.
- Laukkanen, A. M., Lindholm, P., and Vilkman, E. (1995). "On the effects of various vocal training methods on glottal resistance and efficiency: A preliminary report," *Folia Phoniater. Logop.* **47**, 324–330.
- Laukkanen, A. M., Lindholm, P., Vilkman, E., Haataja, K., and Alku, P. (1996). "A physiological and acoustic study on voiced bilabial fricative [β:] as a vocal exercise," *J. Voice* **10**(1), 67–77.
- Laukkanen, A. M., Titze, I. R., Hoffman, H., and Finnegan, E. (2008). "Effects of a semioccluded vocal tract on laryngeal muscle activity and glottal adduction in a single female subject," *Folia Phoniater. Logop.* **60**(6), 298–311.
- Lee, S., Potamianos, A., and Narayanan, S. (1999). "Acoustics of children's speech: Developmental changes of temporal and spectral parameters," *J. Acoust. Soc. Am.* **105**(3), 1455–1468.
- Lee, S. H., Yu, J. F., Hsieh, Y. H., and Lee, G. S. (2015). "Relationships between formant frequencies of sustained vowels and tongue contours measured by ultrasonography," *Am. J. Speech-Lang. Pathol.* **24**(4), 739–749.
- Linklater, K. (1976). *Freeing the Natural Voice* (Drama Book, New York).
- Lulich, S. M. (2010). "Subglottal resonances and distinctive features," *J. Phon.* **38**(1), 20–32.
- Lulich, S. M., Alwan, A., Arsikere, H., Morton, J. R., and Sommers, M. S. (2011a). "Resonances and wave propagation velocity in the subglottal airways," *J. Acoust. Soc. Am.* **130**(4), 2108–2115.
- Lulich, S. M., and Arsikere, H. (2015). "Tracheo-bronchial soft tissue and cartilage resonances in the subglottal acoustic input impedance," *J. Acoust. Soc. Am.* **137**(6), 3436–3446.
- Lulich, S. M., Arsikere, H., Morton, J. R., Leung, G. K. F., Alwan, A., and Sommers, M. S. (2011b). "Analysis and automatic estimation of children's subglottal resonances," in *12th Annual Conference of the International Speech Communication Association 2011 (Interspeech 2011)*, Vols. 1–5, pp. 2817–2820.
- Lulich, S. M., Berkson, K., and de Jong, J. M. (2018). "Acquiring and visualization 3D/4D ultrasound recordings of tongue motion," *J. Phon.* **71**, 410–424.
- Lulich, S. M., Charles, S., and Lulich, B. (2017). "The relation between tongue shape and pitch in clarinet playing using ultrasound measurements," *J. Acoust. Soc. Am.* **141**, 1759–1768.
- Lulich, S. M., Morton, J. R., Arsikere, H., Sommers, M. S., Leung, G. K., and Alwan, A. (2012). "Subglottal resonances of adult male and female native speakers of American English," *J. Acoust. Soc. Am.* **132**(4), 2592–2602.
- Lundberg, A. J., and Stone, M. (1999). "Three-dimensional tongue surface reconstruction: Practical considerations for ultrasound data," *J. Acoust. Soc. Am.* **106**(5), 2858–2867.
- McKenna, V. S., Llico, A. F., Mehta, D. D., Perkell, J. S., and Stepp, C. E. (2017). "Magnitude of neck-surface vibration as an estimate of subglottal pressure during modulations of vocal effort and intensity in healthy speakers," *J. Speech Lang. Hear. Res.* **60**(12), 3404–3416.
- Meltzner, G. S., Kobler, J. B., and Hillman, R. E. (2003). "Measuring the neck frequency response function of laryngectomy patients: Implications for the design of electrolarynx devices," *J. Acoust. Soc. Am.* **114**(2), 1035–1047.
- Mik, L., Krol, D., Lorenc, A., and Wielgat, R. (2018). "Fusing the electromagnetic articulograph, high-speed video cameras and a 16-channel microphone array for speech analysis," *Bull. Polish Acad. Sci.: Tech. Sci.* **66**(3), 257–266.
- Mills, R. D., Rivedal, S., DeMorett, C., Maples, G., and Jiang, J. J. (2017). "Effects of straw phonation through tubes of varied lengths on sustained vowels in normal-voiced participants," *J. Voice* **32**, e21–e29.
- Munger, J. B., and Thomson, S. L. (2008). "Frequency response of the skin of the head and neck during production of selected speech sounds," *J. Acoust. Soc. Am.* **124**(6), 4001–4012.
- Nix, J. (1999). "Lip trills and raspberries: 'High-spit factor' alternatives to the nasal continuant consonants," *J. Sing.* **55**, 15–19.
- Patel, R., Donohue, K. D., Unnikrishnan, H., and Kryscio, R. J. (2015). "Kinematic measurements of the vocal-fold displacement waveform in typical children and adult populations: Quantification of high-speed endoscopic videos," *J. Speech Lang. Hear. Res.* **58**(2), 227–240.
- Patel, R., Dubrovskiy, D., and Döllinger, M. (2014). "Measurement of glottal cycle characteristics between children and adults: Physiological variations," *J. Voice* **28**(4), 476–486.
- Powell, M., Filter, M. D., and Williams, B. (1989). "A longitudinal study of the prevalence of voice disorders in children from a rural school division," *J. Commun. Disord.* **22**(5), 375–382.
- Radolf, V., Laukkanen, A. M., Horacek, J., and Liu, D. (2014). "Air-pressure, vocal fold vibration and acoustic characteristics of phonation during vocal exercising. Part 1: Measurement in vivo," *Eng. Mech.* **21**(1), 53–59.
- Ramos, L. A., and Gama, A. C. C. (2017). "Effect of performance time of the semi-occluded vocal tract exercises in dysphonic children," *J. Voice* **31**(3), 329–335.
- Sato, K., Hirano, M., and Nakashima, T. (2001). "Fine structure of the human newborn and infant vocal fold mucosae," *Ann. Otol. Rhinol. Laryngol.* **110**(5), 417–424.
- Scavone, G. P., Lefebvre, A., and da Silva, A. R. (2008). "Measurement of vocal-tract influence during saxophone performance," *J. Acoust. Soc. Am.* **123**(4), 2391–2400.
- Shawker, T. H., Sonies, B. C., and Stone, M. (1984). "Soft tissue anatomy of the tongue and floor of the mouth: An ultrasound demonstration," *Brain Lang.* **21**(2), 335–350.
- Sihvo, M., and Dienizoglu, I. (2018). "Lax vox voice therapy technique," www.laxvox.com (Last viewed July 26, 2018).
- Simberg, S., and Laine, A. (2007). "The resonance tube method in voice therapy: Description and practical implementations," *Logoped Phoniater. Vocol.* **32**(4), 165–170.
- Sjölander, K., and Beskow, J. (2000). "Wavesurfer—an open source speech tool," paper presented at the *Proceedings of International Conference on Spoken Language Processing*.
- Smith, S. L., and Titze, I. R. (2017). "Characterization of flow-resistant tubes used for semi-occluded vocal tract voice training and therapy," *J. Voice* **31**(1), e111–e118.
- Stevens, K. N. (2000). *Acoustic Phonetics* (The MIT Press, Cambridge, MA).
- Stone, M., Davis, E. P., Douglas, A. S., NessAiver, M., Gullapalli, R., Levine, W. S., and Lundberg, A. (2001). "Modeling the motion of the internal tongue from tagged cine-MRI images," *J. Acoust. Soc. Am.* **109**(6), 2974–2982.
- Story, B. H., and Bunton, K. (2015). "Formant measurement in children's speech based on spectral filtering," *Speech Commun.* **76**, 93–111.

- Story, B. H., Laukkanen, A. M., and Titze, I. R. (2000). "Acoustic impedance of an artificially lengthened and constricted vocal tract," *J. Voice* **14**(4), 455–469.
- Švec, J. G., Horacek, J., Sram, F., and Vesely, J. (2000). "Resonance properties of the vocal folds: *In vivo* laryngoscopic investigation of the externally excited laryngeal vibrations," *J. Acoust. Soc. Am.* **108**(4), 1397–1407.
- Švec, J. G., Titze, I. R., and Popolo, P. S. (2005). "Estimation of sound pressure levels of voiced speech from skin vibration of the neck," *J. Acoust. Soc. Am.* **117**(3), 1386–1394.
- Titze, I. R. (1988). "The physics of small-amplitude oscillation of the vocal folds," *J. Acoust. Soc. Am.* **83**(4), 1536–1552.
- Titze, I. R. (2001). "Acoustic interpretation of resonant voice," *J. Voice* **15**(4), 519–528.
- Titze, I. R. (2006). "Voice training and therapy with a semi-occluded vocal tract: Rationale and scientific underpinnings," *J. Speech Lang. Hear. Res.* **49**(2), 448–459.
- Titze, I. R. (2008). "Nonlinear source-filter coupling in phonation: Theory," *J. Acoust. Soc. Am.* **123**(5), 2733–2749.
- Titze, I. R. (2009). "Phonation threshold pressure measurement with a semi-occluded vocal tract," *J. Speech Lang. Hear. Res.* **52**(4), 1062–1072.
- Titze, I. R., Baken, R. J., Bozeman, K., Granqvist, S., Henrich, N., Herbst, C. T., Howard, D. M., Hunter, E. J., Kaelin, D., Kent, R. D., Kreiman, J., Kob, M., Löfqvist, A., McCoy, S., Miller, D. G., Noé, H., Scherer, R. C., Smith, J. R., Story, B. H., Švec, J. G., Ternström, S., and Wolfe, J. (2015). "Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization," *J. Acoust. Soc. Am.* **137**(5), 3005–3007.
- Titze, I., and Laukkanen, A. M. (2007). "Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study," *Logoped Phoniater. Vocol.* **32**, 147–156.
- Titze, I. R., Laukkanen, A. M., Finnegan, E. M., and Jaiswal, S. (2002). "Raising lung pressure and pitch in vocal warm-ups: The use of flow-resistant straws," *J. Sing.* **58**(4), 329–338.
- Titze, I. R., Riede, T., and Popolo, P. (2008). "Nonlinear source-filter coupling in phonation: Vocal exercises," *J. Acoust. Soc. Am.* **123**(4), 1902–1915.
- Titze, I. R., and Story, B. H. (1997). "Acoustic interactions of the voice source with the lower vocal tract," *J. Acoust. Soc. Am.* **101**(4), 2234–2243.
- Vallabha, G. K., and Tuller, B. (2002). "Systematic errors in the formant analysis of steady-state vowels," *Speech Commun.* **38**, 141–160.
- Vampola, T., Laukkanen, A. M., Horacek, J., and Švec, J. G. (2011). "Vocal tract changes caused by phonation into a tube: A case study using computer tomography and finite-element modeling," *J. Acoust. Soc. Am.* **129**(1), 310–315.
- van den Berg, J. P. (1960). "An electrical analogue of the trachea, lungs and tissue," *Acta Physiol. Pharmacol. Neerl* **9**, 361–385.
- Vorperian, H. K., Wang, S., Chung, M. K., Schimek, E. M., Durtschi, R. B., Kent, R. D., Ziegert, A. J., and Gentry, L. R. (2009). "Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study," *J. Acoust. Soc. Am.* **125**(3), 1666–1678.
- Wade, L., Hanna, N., Smith, J., and Wolfe, J. (2017). "The role of vocal tract and subglottal resonances in producing vocal instabilities," *J. Acoust. Soc. Am.* **141**(3), 1546.
- Wilson, T. D. (1996). "The measured upstream impedance for clarinet performance and its role in sound production," doctoral dissertation, University of Washington, Seattle.
- Wu, L., Xiao, K., Dong, J., Wang, S., and Wan, M. (2014). "Measurement of the sound transmission characteristics of normal neck tissue using a reflectionless uniform tube," *J. Acoust. Soc. Am.* **136**(1), 350–356.
- Yeung, G., Lulich, S. M., Guo, J., Sommers, M. S., and Alwan, A. (2018). "Subglottal resonances of American English speaking children," *J. Acoust. Soc. Am.* **144**(6), 3437–3449.
- Zañartu, M., Ho, J. C., Mehta, D. D., Hillman, R. E., and Wodicka, G. R. (2013). "Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration," *IEEE Trans. Audio Speech Lang. Process.* **21**(9), 1929–1939.
- Zañartu, M., Mongeau, L., and Wodicka, G. R. (2007). "Influence of acoustic loading on an effective single mass model of the vocal folds," *J. Acoust. Soc. Am.* **121**(2), 1119–1129.