On the acoustic effects of the supraglottic structures in excised larynges

Fariborz Alipour^{a)} and Eileen Finnegan

Department of Communication Sciences and Disorders, The University of Iowa, 334E WJSHC, Iowa City, Iowa 52242-1012

(Received 12 June 2012; revised 27 February 2013; accepted 28 February 2013)

The acoustic effects of the supraglottic laryngeal structures (SGSs), including the false vocal folds (FVFs) laryngeal ventricle, and the epiglottis were investigated in an excised canine larynx model with and without these anatomical structures. The purpose of this study was to better understand the acoustic contributions of these structures to phonation. Canine larynges were prepared and mounted over a 3/4 in. tube, which supplied pressurized, heated, and humidified air. Glottal adduction was accomplished by rotating the arytenoids with a suture passed behind the vocal folds to simulate the lateral cricoarytenoid muscle action. The SGSs were kept intact in the first part of the experiment and were removed in the second part. Results indicated that when the FVFs vibrated, a low frequency component was observed in the spectral data. The excised larynx with a SGS had a limited range of frequency with subglottal pressure, while the larynx without a SGS had a larger frequency range. The excised canine larynx with a SGS oscillated with a higher phonation threshold pressure and significantly louder. © 2013 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4796109]

PACS number(s): 43.70.Gr, 43.70.Bk, 43.70.Aj, 43.70.Dn [BHS]

Pages: 2984–2992

I. INTRODUCTION

The laryngeal supraglottic structures (SGSs) including the laryngeal ventricles, the ventricular folds, and epiglottis are multifunctional organs that are involved in phonation, respiration, and swallowing. The ventricular folds [also known as the false vocal folds (FVFs) or vestibular folds] are thick folds lying above the true vocal folds (TVFs) and separated from the TVFs by the laryngeal ventricle. The ventricular folds contain glands which provide lubrication to the TVFs (Kutta *et al.*, 2002) and they smooth out the vortices around the glottal jet (Kucinski *et al.*, 2006). They have historically been linked to closure of the laryngeal lumen during deglutition and other primitive reflexes such as coughing, gagging, etc. (Pinho *et al.*, 1999; Olthoff *et al.*, 2007). They are commonly referred to as the FVFs as they have been thought to not be directly involved in the production of a normal voice. However, FVF activity is more often associated with production of voice with some disorders (Stager *et al.*, 2001; Maryn *et al.*, 2003). With respect to normal speech production, Stager *et al.* (2000, 2003) observed movements of SGSs during continuous speech, and categorized them as either static or dynamic. The static component represented a supraglottic configuration assumed on initiation of a speech task and maintained until its completion. The static component might be considered the typical or dominant laryngeal configuration for voice production regardless of speech task and be expected to be present within and across most speech tasks when identified within a particular individual. In subjects with normal laryngeal structure and function observed by Stager *et al.* (2003), 26%

of the males demonstrated static anterior to posterior compression of the FVF, compared to 9% of the females. The dynamic component involved short duration adductory FVF gestures. In the subjects with normal laryngeal structure and function observed by Stager *et al.* (2003), 94% demonstrated dynamic FVF movement associated with the production of glottal stops, suggesting the FVFs play a role in normal speech production.

Vocal-ventricular phonation, or throat singing, was investigated by Fuks *et al.* (1998) with aerodynamic and acoustic analysis, including laryngoscopy in one trained singer. While the subject was phonating, a sustained low-pitched tone of the vowel /a/, electroglottograph (EGG), audio, and flow signals were recorded and later a high-speed video laryngoscopy was performed on the same phonation type. They found that for this type of phonation, the ventricular folds closed every second cycle of TVF vibration. Their EGG results indicated an increase of the tissues' conductance between electrodes due to ventricular fold closure. They concluded that the ventricular and TVFs had closings of opposite phase.

Use of the ventricular folds during throat singing was also studied by Lindestad *et al.* (2001) who observed a male bass type singer alternating between modal and throat singing voices. Lindestad *et al.* (2004) also studied ventricular vibration. In both studies, vocal fold vibration was examined using high speed video and kymography. In a sonorous, slightly pressed, modal voice (140 Hz), TVF vibration was normal with lower amplitude and normal mucosal wave. It coexisted with low amplitude ventricular fold oscillations, with incomplete closures at the same frequency and phase as TVF vibration. The throat voice was perceived as low-pitched, almost one octave lower than the modal voice (70 Hz), slightly pressed, with a restricted F_0 range and high

^{a)}Author to whom correspondence should be addressed. Electronic mail: alipour@iowa.uiowa.edu

intensity sonority. The ventricular folds were found to vibrate at the same frequency as the TVFs; however, closure of the ventricular folds did not coincide with the vocal fold closure but preceded it. [Bailly et al. \(2010\)](#) investigated the effects of the FVFs on the phonation of a singer during throat singing using direct laryngoscopy with high-speed imaging. Using kymographic images and estimates of the opening areas of the TVFs and FVFs, they reported that the FVFs closed for every two closures of the TVFs with a phase difference.

The laryngeal ventricle has also been considered to be an acoustical filter by numerous researchers. [Pepinsky \(1942\)](#) described it as a Helmholtz resonator of small volume and large conductivity. This filtering effect was also considered by [van den Berg \(1955\)](#), who thought it resembled a condenser due to its small dimensions. He suggested that the dimension of the ventricle affects the frequency spectrum of the glottal source through low-pass filtering. [Kitzing and Sonesson \(1967\)](#) reported on the shape and size of ventricle changes during phonation. Through a radiographic study of 41 young normal male subjects, they observed that the laryngeal ventricle was elongated up to 50% during high pitched phonations and its height could expand about 60% during medium pitched phonations. This ability to control the configuration of the FVFs and the laryngeal ventricle has been suggested to play a role in the “singer’s formant” ([Sundberg, 1974](#); [Titze and Story, 1997](#); [Nemetz et al., 2005](#)).

[Drechsel and Thomson \(2008\)](#) studied the effects of the SGSs on the glottal jet. Using a self-oscillating synthetic model with a rigid acrylic FVF model and idealized straight tube vocal tract, they measured the glottal jet with the particle image velocimetry and high-speed imaging. They concluded that the presence of the FVF in the vocal tract obstructed the start of the downstream vortex. Also, they suggested that the FVF decreased the root-mean-square values of the glottal jet and acted as a stabilizer in the vocal tract.

The aerodynamic interaction between the FVFs and the TVFs was studied by [Bailly et al. \(2008\)](#) using an experimental set-up with static and dynamic TVFs and fixed FVFs separated by a large and long ventricle. They also analyzed the aerodynamic effects of the FVFs using the Bernoulli equation. They concluded that the presence of the FVFs may have positive effects on the oscillation of the TVFs.

The effects of the FVF and ventricle (e.g., including their size) on the sound source has been studied by some investigators using numerical models, but due to limitations of these models, such as lack of FVF vibration, they did not reveal any important acoustic contribution from the FVF and the ventricle. For example, [Zhang et al. \(2002\)](#) assessed the influence of the FVFs on sound generation using computational fluid dynamics and concluded that the FVFs may (in a small way) impact voice production by reducing glottal flow resistance. In another theoretical study of the effects of FVF on the voice source, [McGowan and Howe \(2010\)](#) calculated the unsteady drag of vortices in a forced-oscillation model of the vocal folds and concluded that the FVFs do not have a noticeable effect on the voice source.

Another laryngeal simulation model including the TVFs and FVFs that employ an immerse boundary for the air flow

modeling and a finite-element method for the tissue motion was presented by [Zheng et al. \(2009\)](#). Their simulation included ten cycles of oscillations with and without the FVFs. They concluded that the presence of the FVFs decreases translaryngeal flow impedance and increases flow rate through the glottis. These computation models were all limited by the lack of pliable moving FVFs.

[Alipour et al. \(2007\)](#) and [Finnegan and Alipour \(2009\)](#), using excised canine larynges, determined that medial compression of the FVF increased glottal flow resistance, as well as sound intensity. Also, in a study of the phonatory effects of SGSs using excised canine larynges, [Finnegan and Alipour \(2009\)](#) demonstrated that positioning of the epiglottis from a vertical rest condition to a horizontal position had major effects on the aerodynamics and acoustics of the larynx by increasing glottal resistance, subglottal pressure, and consequently fundamental frequency. They also reported that vibration of the FVFs induced some irregularity into the acoustic output of the larynx and manipulation of the SGSs resulted in major changes in the harmonic structure of the acoustic output. This work is an extension of that work with more emphasis on the spectral information and use of high-speed imaging for vocal fold dynamics interpretation. The purpose of this study was to better understand the acoustic contributions of the SGSs and correlate them to glottal dynamics. It was hypothesized that the canine SGSs significantly contribute to its acoustic energy due to their vibration.

The canine larynx was chosen due to its dimensional similarity to human larynges, relatively large ventricles, and availability in the fresh form.

II. METHODS

A. Data acquisition

Eight excised canine larynges were obtained following cardiovascular research experiments at the University of Iowa Hospitals and Clinics. The larynges were removed from the animal after death, quick frozen in liquid nitrogen, and stored in a -82°F freezer. The frozen larynges were later transferred to our lab freezer and stored until the day of experiment. They were thawed in saline solution overnight in the refrigerator before preparation for the experiments. The gender and weight of the larynges are provided in Table I. Excised larynges were mounted and operated according to previous work ([Alipour et al., 2007](#)).

Adduction of the TVFs was controlled by a pair of sutures pulling on the muscular process of each arytenoid cartilage to simulate lateral cricoarytenoid and (lateral) thyroarytenoid (TA) muscle action, as in an arytenoid adduction [see Fig. 1(a)]. The TVF adduction levels were adjusted using weights (50 to 150 g) that pulled the sutures attached to the muscular process of the arytenoid cartilages. FVF adduction was not controlled independently in this study. During the first half of the experiment the FVFs and epiglottis remained intact as data were obtained from the larynx at three adduction levels, then the SGSs, including the epiglottis and the FVFs, were removed with a small curved scissor such that the TVFs became visible in the superior view [Fig. 1(b)] and the experiment was repeated. The removal of

TABLE I. Canine larynges with their oscillating conditions.

Larynx	Larynx ID	Gender	Weight kg	Ps Ranges cm H ₂ O	Flow Ranges mL/s	F0 Ranges Hz
1	CL73	M	17	6.0–33.0	200–1400	82–119
2	CL81	F	15	5.0–30.0	475–1350	79–283
3	CL82	M	22	6.8–28.0	390–1400	68–195
4	CL83	F	18	6.9–24.8	400–1400	86–198
5	CL84	F	16	5.0–18.2	190–1100	79–249
6	CL85	F	18	3.2–25.0	80–1400	72–215
7	CL86	F	13	7.4–29.2	110–1150	66–213
8	CL87	F	16	6.2–25.4	120–1400	88–243

the FVFs exposed part of the lateral TA muscle as seen in Fig. 1(b), but this surgical procedure did not affect the tensioning or posturing of the TVFs. Due to anatomical considerations, the excision of a part of the SGSs in a canine larynx could leave a damaged boundary condition; the whole SGSs were removed.

Mean values of subglottal pressure and flow rate (controlled with a fine rotary valve) were read from a wall manometer and an in-line rotameter (model J197, Gilmont Instruments, Barrington, IL) (see Fig. 2 for the experimental setup). The subglottal pressure signal (which may be referred to as pressure for brevity) was recorded using a pressure transducer (Microswitch 136PC01G1) mounted perpendicular to the flow in the tracheal tube 10 cm below the vocal folds with the end of the transducer near the tracheal wall. The flow rate signal was recorded with a pneumatic flow meter (Rudolph 4700) and low-range pressure transducer (Validyne DP103) upstream of the humidifier (ConchaTherm® unit, RCI Laboratories). The tubing length from the TVFs to the humidifier was about 48 in., suggesting a subglottal resonance of approximately 140 Hz.

The EGG signal was obtained by placing electrode plates from a Synchrovoice EGG on the thyroid laminae (anterior side of the larynx) during phonation. The electrodes were positioned in place while the larynx was oscillating and the EGG signal was monitored on the oscilloscope to assure a good signal, and then it was secured in place with a narrow strip of duct tape.

The audio signal was obtained with a microphone (Sony ECM-MS907) at a distance of 15 to 20 cm from the larynx and recorded on a digital audio tape recorder (Sony PCM-M1). The sound pressure level (SPL) was measured with a type 2 sound level meter (Extec model 407738) with “C” weighting and “fast” averaging, placed about 15 cm from the larynx.

For each of the 8 excised larynx experiments, the manipulated variables were level of adduction (i.e., low, medium, and high adduction corresponding to 50, 100, and 150 g weights applied to both adduction sutures) and the presence or absence of the SGSs. Once the degree of TVF adduction was established, we performed two pressure-flow sweeps, in which flow was gradually increased or decreased with a rotary control valve (consequently altering pressure as well) to determine the range of aerodynamic conditions during which the vocal folds would vibrate (Table I). Once the range of conditions was established, we manipulated flow level in a stepwise fashion from the bottom of the oscillation

range to the top of the range and recorded a series of 6 to 14 sustained oscillations. Some larynges produced vibration in a narrow range of pressure/flow conditions and at only two levels of adduction, while others produced phonation over a wide range of pressure/flow condition and at all three levels of adduction. Each adduction level of an experiment including 2 pressure-flow sweeps and 8 to 10 sustained oscillations took about 8 to 10 min. Recordings of oscillation of the vocal folds in slow motion were visualized with a strobe light and with high-speed imaging and were obtained for selected conditions at each adduction level.

High-speed imaging was performed with a monochrome camera (Photron, model 100K Fastcam- 1024-PCI). For selected cases (usually, the last 3 to 4 oscillatory conditions), the larynx was illuminated with a Lowell Pro light at a distance of about 50 cm and video images were acquired during oscillations at a rate of 5000 to 10 000 frames per second (fps) for a short duration (1 to 2 s) due to the memory limitation and to avoid damage to the tissues by the intense light. Only 500 frames of the video image were downloaded from memory to disk for later analysis. This process lasted about 10 s, approximately the interval between oscillation experiments. This number of frames provided images of a few oscillation cycles that were adequate for comparing TVF and FVF oscillations.

B. Data analysis

Analog signals from the EGG, microphone, and pressure and flow transducers were recorded simultaneously onto a Sony SIR1000 digital tape recorder at a sampling rate of 40 kHz per channel. These recorded signals were later digitized using a 14-bit A/D converter (DATAQ Instruments). The signals were then converted to calibrated physical quantities in a MATLAB routine and used for the aerodynamic and acoustic analyses.

A spectrographic analysis of the acoustic signals produced by the excised larynx with and without SGSs was obtained for each larynx. Spectral analyses of the signals were obtained with a fast Fourier transform (FFT) of the microphone signal (with a Hamming window) within the MATLAB computing environment. Each FFT was calculated with at least 4096 data points for adequate resolution. To calculate the fundamental frequency, the EGG signal was low-pass filtered at 150% of its estimated *F0* value seen from the spectrogram or an oscilloscope. The fundamental frequency

(a)

(b)

FIG. 1. (Color online) (a) Mounted excised larynx with SGSs (epiglottis and FVFs) and control sutures, and (b) mounted larynx without SGSs with grids laid on the left vocal fold before the experiment for calibration.

was then calculated with a zero crossing method. First, the signal dc offset was removed, and then the periods of every cycle in the selected segment were calculated from consecutive zero crossings (Deem *et al.*, 1989) and averaged.

The frames recorded with the high-speed camera were analyzed and converted to kymographic images or kymograms by scanning a line from each image (Svec and Schutte, 1996) and augmenting these image lines across all frames in MATLAB using custom made software. To obtain

FIG. 2. (Color online) Excised larynx experimental setup including mounting and control fixtures, lights, cameras, and recording sensors.

these images, a typical frame of the high-speed image was displayed on the computer screen and a line was placed across the glottis (perpendicular to midsagittal plane) at approximately 30%, 50%, and 70% length locations from the posterior end. Then the pixel intensities of this line were obtained for every frame of the recorded video. By mapping the pixel intensities of each frame in the y-axis and the time of that frame in the x-axis, the kymogram for that location was generated.

Once the kymograms were generated, they were used to estimate the oscillation frequencies of the TVFs and FVFs. Since the images in this experiment were recorded at 10 000 fps, the processed kymograms of 500 frames generated images with 50 ms time windows (X-axis). Using a custom made algorithm in MATLAB, the visible peaks of the true folds were selected with the cursor and, by measuring their distances in time, the average period was estimated for the oscillation of the TVFs. The same procedure was repeated for the peaks of the FVFs to obtain its oscillation period. Also, within this process, the phase difference between the peaks as a percentage of the period was calculated.

A statistical analysis was performed using Statistica 10 (StatSoft, Inc., Tulsa, OK) to determine if there were significant differences across phonatory conditions with and without SGSs. Unpaired *t*-tests were performed to determine if there were significant differences between individual oscillatory conditions. Tests were two-tailed and the significance level was $\alpha = 0.05$. The *t*-test was used to compare mean values of PTP, SPL, and oscillation frequencies of the TVF and FVF.

III. RESULTS

The aerodynamic conditions, during which all larynges (with and without SGSs) produced vibration, are shown in Table I with individual gender and weight. The fourth column of Table I shows the subglottal pressure ranges, and the fifth column shows the flow rate ranges, suggesting that most of these larynges operate within the same range of aerodynamic conditions. Finally, the last column shows the frequency ranges of their oscillations. For example, larynx 1 produced phonation at airflow values ranging from 200 to 1400 mL/s and, as airflow was increased, pressure values varied between 6 and 33 cm-H₂O, and the *F*₀ of the phonation produced over this range of pressure/flow conditions varied from 82 to 119 Hz.

As seen in Fig. 3, the TVFs showed the expected increase in *F*₀ with rising subglottal pressure (Titze, 1989; Alipour and Scherer, 2007) and flow in larynges without ventricular folds, but when the ventricular folds were present, the TVFs' ability to increase frequency of vibration as a function of subglottal pressure was limited. Figure 3 compares alterations in frequency of vibration for excised larynx #1 with and without the FVF. The solid symbols (A, B) are for the larynx oscillated with the FVF at adduction levels of 100 and 150 g corresponding to medium and high adduction levels. The hollow symbols (C, D) represent oscillation of the same larynx at similar adduction levels when the FVFs were removed. This comparison suggests that the activity of the FVF prevents the TVFs' frequency to ascend with increasing pressure much less than the value of 2.9 Hz/cm H₂O that was reported in the previous studies (e.g., Alipour and Scherer, 2007).

The limitation in ability to increase frequency of vibration appears related to the onset of FVF vibration rather than to the presence (and positioning) of the FVFs themselves.

FIG. 3. (Color online) Frequency-pressure relations in excised larynx #1. The symbols are: A and B with SGSs at medium and high adductions; C and D without SGSs at medium and high adduction levels.

Figure 4 shows an example of pressure-flow sweep data from larynx #4 (CL83) at a medium adduction level with intact SGSs. This includes, from the top to bottom, (a) mean subglottal pressure, (b) mean flow rate, (c) fundamental frequency, and (d) pressure amplitude as a function of time. The entire oscillating portion of the sweep was divided into about 100 segments of 10 to 20 cycles each. Then, mean pressure and flow were calculated for each 10 to 20 cycles of the corresponding signal. The fundamental frequency was also calculated from the same number of cycles. The pressure amplitude was computed as half of the peak-to-peak value of subglottal pressure during those cycles. The pressure and flow started from a phonatory threshold pressure (PTP) of 11 cm H₂O and increased monotonically, but the fundamental frequency shows minor increases in the first 2 s of oscillations and remained almost steady until the FVFs started oscillation, at which time it dropped about one octave and remained almost constant. During the pressure-flow sweep, the superior surface of the larynx was inflated and puffed up and the FVFs oscillated after about 10 s when subglottal pressure was about 20 cm H₂O (1 cm H₂O = 98.1 Pa). The extra peaks of the subglottal pressure are probably due to the subglottal resonance. It is interesting to note that during this process, the pressure amplitude retained its high values, suggesting a large amplitude tissue movement that was probably caused by FVF vibration that was also observed from the stroboscopic video. The first sharp peak in the pressure amplitude signal corresponds to an increase of EGG signal (not shown here) and the second sharp peak corresponds to the flow rate drop and changes in the glottal amplitudes during the transition. During the first half of the sweep, the

FIG. 4. (Color online) Pressure-flow sweep in excised larynx #4 with the SGS, showing a change in *F*₀ (fundamental frequency) and *P*_a (pressure amplitude) with a gradual increase in *P*_s (subglottal pressure) and flow.

fundamental frequency increases to 199.4 Hz and then suddenly drops to 102.5 Hz. The drop in frequency at around 12 s was due to the initiation of FVF oscillations, which resulted in a decrease in the rate of vibration of the TVFs.

A change in the pattern of TVF vibration was observed with the onset of FVF vibration. The glottal waveforms of the excised canine larynx (CL83) with SGSs before and after the activity of the FVF are shown in Figs. 5 and 6, respectively. The three signals in each figure, from top to bottom, are EGG, subglottal pressure (in cm H₂O), and flow rate (in L/s). The oscillations before the FVF activity appear to be regular with very sharp EGG peaks indicating a brief contact of the TVFs. The fundamental frequency at this selected region of sweep was 187.5 Hz with a mean subglottal pressure of 17.8 cm H₂O and a mean flow rate of 0.71 L/s. The pressure signal appears to be sinusoidal with peaks corresponding to the glottal contact, if it was corrected for the distance from the larynx. A drastic change in these waveforms appeared after the onset of FVF activity (Fig. 6). The second pulse seen on the EGG was caused by the FVF contact. At this region of the sweep, the fundamental frequency dropped to 104.7 Hz and the mean subglottal pressure was 29 cm H₂O with a mean flow rate of 0.95 L/s. There are second peaks in the pressure waveform that may correspond to FVF closures. The flow signal (that was recorded much further upstream from the larynx, before humidification) has less prominent changes due to the filtering effect.

High-speed video images of the superior aspect of these larynges reveal that the true folds and false folds vibrate at similar frequency but out of phase. Figure 7(a) shows six kymograms made from the high-speed video of six of these excised larynges with SGSs, including larynges #1 to #5

FIG. 5. (Color online) Glottal waveforms of the excised larynx #4 with SGSs before FVF activity.

FIG. 6. (Color online) Glottal waveforms of the excised larynx #4 with SGSs after FVF activity.

and #7, from top to bottom, respectively. The larynges oscillated at 90, 162, 88, 142, 161, and 85 Hz, respectively. In each kymogram, the darker gray teeth represent the TVFs and the lighter gray teeth represent the FVFs. The bottom teeth are from the right side and the top teeth are from the left side. In some images, one of the TVFs is not visible due to the location of the camera. During removal of the FVF, a little TA muscle tissue moved over the vocal fold (which appears with a different shade) and generated extra teeth in the kymogram that look like oscillation irregularity in some larynges. The measurement of the distance between the peaks of these teeth provided the oscillation periods of the

FIG. 7. (a) Kymographic images of six different canine excised larynges with SGSs at various oscillating conditions. The darker gray in each kymogram represents the TVFs and the lighter gray represents the FVFs. (b) Kymographic images of six excised larynges without SGSs acquired at 1/3 vocal fold length from the posterior. The horizontal axis is time with a range of 0 to 50 ms.

TABLE II. Oscillating conditions of canine larynges with and without FVF.

Larynx	PTP with FVF cm H ₂ O	PTP w/o FVF cm H ₂ O	SPL with FVF dB	SPL w/o FVF dB
1	5.3 ± 2.3	4.5 ± 0.7	85.3 ± 6.8	85.1 ± 5.0
2	10.3 ± 0.6	6.7 ± 1.5	84.4 ± 3.9	80.0 ± 6.0
3	8.3 ± 0.6	5.3 ± 0.6	92.5 ± 5.5	83.4 ± 4.5
4	8.3 ± 3.0	8.0 ± 0.7	85.7 ± 7.6	79.5 ± 5.1
5	4.3 ± 0.5	3.7 ± 1.2	77.7 ± 4.3	75.1 ± 3.2
6	7.7 ± 1.5	4.3 ± 1.5	85.4 ± 5.3	78.0 ± 4.4
7	8.0 ± 2.0	5.3 ± 0.6	87.1 ± 4.3	82.2 ± 3.9
8	7.3 ± 1.2	5.7 ± 0.6	83.3 ± 4.7	81.5 ± 5.4
Mean	7.44	5.44	85.2	80.6

TVFs and FVFs. Similarly, Fig. 7(b) shows six kymograms from canine larynges #3 to #8 without SGSs for comparison. They oscillated at frequencies of 107, 129, 272, 83, 216, and 215 Hz, respectively. The top kymogram shows the extra teeth corresponding to the overlap of TA muscle on TVFs.

Table II compares phonatory and acoustic conditions of all larynges with and without SGSs. The second and third columns represent the mean and standard deviation of PTP during these two conditions, which were labeled “with FVF” and “w/o FVF” for brevity. The fourth and fifth columns represent the mean and standard deviation of SPL during these two conditions.

Larynges with intact SGSs produced louder phonation (mean = 85.2 dB) than those with SGSs removed (mean = 80.6 dB, $p = 0.0266$, paired two-tail t -test). The acoustic contribution of the FVFs and laryngeal ventricle was partly observed from the frequency variations of the excised larynx during pressure-flow sweeps and also based on the PTP and SPL data depicted in Table II. Figure 8 shows the SPL for

larynx #3 (CL82) as a function of subglottal pressure. The filled symbols (cases A–C) represent oscillations with SGSs and the hollow symbols (cases E–G) represent oscillations without it. The cases A–C and E–G contain data from three different adduction levels of low to high. While SPL increases with pressure in all cases, the SGSs appear to add more energy to the acoustic output due to their higher SPL values.

Another aspect of this contribution can be noticed graphically through the examination of SPL and spectral information from the microphone signal. Spectral comparisons of phonation with and without SGSs are shown in Fig. 9. Here the microphone signals, which were recorded separately into audio files at the 44.1 kHz sampling rate, were converted to FFT spectra with frequency in hertz on the X-axis and amplitude in dB on the Y-axis. The top figure was obtained from an audio signal recorded from larynx #3 with SGSs intact at a mean subglottal pressure of 20 cm H₂O, a

FIG. 8. (Color online) Sound pressure level of excised larynx #3 measure at 15 cm. The symbols are: A, B, and C for larynges with SGSs at low, medium, and high adductions; E, F, and G for larynges without SGSs at low, medium, and high adductions.

FIG. 9. (Color online) FFT spectrum of excised larynx #3 with SGSs (top graph) and without SGS (bottom graph).

mean flow rate of 1.3 L/s, and SPL of 76.1 dB. Besides the main high energy region around the fundamental, there are two areas of high energy around 2000 and 4000 Hz. The bottom figure was obtained from signals recorded from the same larynx without SGSs at a mean subglottal pressure of 17 cm H₂O, mean flow rate of 1.4 L/s, and SPL of 70 dB. There is only one region of high energy around fundamental frequency and another region around 1400 Hz, but at a much lower level. Thus it appears that SGSs contributed some acoustic energy in the high frequency ranges.

Phonation threshold pressure was also higher for larynges with intact SGSs in comparison to those without ($p = 0.0304$, paired two-tail t -test). During the pressure flow sweeps, oscillation started at an average PTP of 7.4 ± 1.9 cm H₂O for cases with SGSs and 5.4 ± 1.4 cm H₂O for cases without (averaged for 8 larynges and 24 observations).

IV. DISCUSSION

This study examined the acoustic effects of the FVFs and laryngeal ventricle using canine excised larynges. Although there are some histological and structural differences between canine and human larynges (Kurita *et al.*, 1983), the dimensional similarity and the existence of the FVFs and laryngeal ventricle make this larynx an ideal model for this study in the absence of excised human larynges. However, the degree of supraglottic activity that is induced aerodynamically or manually may be exaggerated in comparison to the human larynx due to the smaller epiglottis opening and FVF gap. Yet this may be an advantage to simulate the extreme conditions that occur in ventricular related voice disorders.

The fact that, during pressure-flow sweeps of the excised larynx with SGSs, the frequency rise of the larynx was modified to a reduced range was similar to what was reported by Lindestad *et al.* (2001) for throat singing, when the FVFs were active. Yet it was possible for the larynx to have regular sustained oscillations with these structures when there was enough FVF opening to allow airflow without FVF oscillations. Whenever these structures oscillated, there was a significant increase in acoustic energy registered in the sound pressure level (Fig. 8). This additional energy may have been attributed to the large mucosal wave in the FVFs and other SGSs that was observed in the stroboscopic videos but not measured.

Generally we are interested in determining the role of the FVFs during normal phonation. But more importantly we need to know the detrimental effects of FVFs on the acoustics and aerodynamics of phonation. We know that they are not simply passively pulled along by the TVFs. They play an active role during the production of glottal stops during speech (Stager *et al.*, 2000). However, during sustained phonation, the false folds generally maintain their distance from the midline, allowing the TVFs to function as the primary sound source. Given that the FVFs are excited by low frequency (<100 Hz) vibration (Svec *et al.*, 2000), it seems possible that the FVFs may also play a role during phonation. Alternatively, the added energy may result entirely from alternating pressures in the ventricular space.

Mucosal vibration is not limited to the medial and superior aspects of the vocal fold that we are able to easily visualize during stroboscopy. Mucosal vibration begins in the inferior aspect (underside) of the vocal fold and propagates superiorly (Boessenecker *et al.*, 2007). The amplitude and frequency of the mucosal wave seen on stroboscopy is the result of the interaction of aerodynamic forces and the viscoelasticity of the vocal fold tissue, but also by the amplitude and frequency of this wave that is propagating from below. Similarly, large pressure changes in the ventricular space may drive the FVFs from below. Small amplitude out of phase vibration of the FVFs may add to skewing of the air-flow signal.

The results from this study indicate that when the FVFs do vibrate during ventricular phonation (i.e., when the false folds are relatively closely adducted), most often they vibrate at a similar frequency as the TVFs but sometimes out of phase (Fig. 7). This phenomenon has been observed with kymographic image analysis in human subjects with hyperfunctional voice (Lindestad *et al.*, 2004). This out of phase vibration has also been reported by Fuks *et al.* (1998). This out of phase vibration may help to skew the glottal airflow pulses, further driving both TVF and FVF vibration.

Another finding of the study was that larynges with intact SGSs produced phonation of greater intensity, given the same subglottic pressure. This could be due to oscillation of the FVFs (in addition to the TVFs), providing a second sound source and adding to sound intensity in that way. This is consistent with findings of Fuks *et al.* (1998) for vocal ventricular phonation and Lindestad *et al.* (2001) on the voice source of Mongolian singing. Alternatively, it could be that the presence of the false folds, the ventricular space, and laryngeal vestibule help to reinforce higher partials of the glottal source signal, enhancing overall intensity. Spectrographic analysis of the excised larynx with and without SGSs indicated a difference in the distribution of acoustic energy produced by the larynges. Without the SGSs, the low frequency components of the signal were strong, while the high frequency components were relatively low in amplitude. With the SGSs intact, the energy distribution changed, with greater reinforcement of higher frequency (2000 to 4000 Hz) partials. This is similar to what has been reported by Lindestad *et al.* (2001) for modal voice and throat singing.

Also, the examination of Table II suggests that the canine larynx without SGSs has significantly lower PTP ($p = 0.0303$ for 23 observations). In other words it oscillated easier, suggesting that canine SGSs do not help its phonation. This is consistent with the findings of Bailly *et al.* (2008).

The activity of the FVFs in this study, which was a factor in the control of fundamental frequency (Figs. 3–6), has direct relevance to ventricular dysphonia, a voice disorder that is caused by either hyperatrophy or hyper adduction of the ventricular folds (Arnold and Pinto, 1960). Examination of glottal waveforms before and after FVF activity suggested a major change in the EGG waveform that can be used as a diagnostic tool to confirm supraglottic activity. Future work should include use of a human excised larynx and manipulation of supraglottic activity such as FVF medial compression

and anterior–posterior compression for which acoustic and spectrographic characteristics will be examined.

V. CONCLUSIONS

The excised canine larynx model was used to study the acoustic effects of the FVFs and laryngeal ventricle. The major control parameters were subglottal pressure, glottal adduction, and existence or lack of SGSs. With pressure-flow sweeps, the range of stable and sustained oscillations was identified and then, for a series of sustained oscillations in the determined range, acoustic data were recorded and spectral information was processed. It was found that:

- (1) The excised larynx with FVFs had a limited range of frequency change, as a function of subglottal pressure while the larynx without FVF had a larger frequency range.
- (2) The excised canine larynx with FVFs and epiglottis oscillated louder (with a significantly higher SPL, $p = 0.0266$ for 210 observations).
- (3) Generally, the canine larynx without FVFs and epiglottis oscillated at a significantly lower PTP ($p = 0.0303$ for 23 observations).

ACKNOWLEDGMENTS

The project described was supported by Award No. R01DC009567 from the National Institute on Deafness and other Communication Disorders. The authors would like to thank Frances E. Kell for assistance in data collection.

Alipour, F., Jaiswal, S., and Finnegan, E. (2007). "Aerodynamic and acoustic effects of false vocal folds and epiglottis in excised larynx models," *Ann. Otol. Rhinol. Laryngol.* **116**, 135–144.

Alipour, F., and Scherer, R. C. (2007). "On pressure-frequency relations in the excised larynx," *J. Acoust. Soc. Am.* **122**, 2296–2305.

Arnold, G. E., and Pinto, S. (1960). "Ventricular dysphonia: New interpretation of an old observation," *Laryngoscope* **70**, 1608–1627.

Bailly, L., Henrich, N., and Pelorson, X. (2010). "Vocal fold and ventricular fold vibration in period-doubling phonation: Physiological description and aerodynamic modeling," *J. Acoust. Soc. Am.* **127**, 3212–3222.

Bailly, L., Pelorson, X., Henrich, N., and Rutu, N. (2008). "Influence of a constriction in the near field of the vocal folds: Physical modeling and experimental validation," *J. Acoust. Soc. Am.* **124**, 3296–3308.

Boessenecker, A., Berry, D. A., Lohscheller, J., Eysholdt, U., and Doellinger, M. (2007). "Mucosal wave properties of a human vocal fold," *Acta. Acust. Acust.* **93**(5), 815–823.

Deem, J. F., Manning, W. H., Knack, J. V., and Matesich, J. S. (1989). "The automatic extraction of pitch perturbation using microcomputers: Some methodological considerations," *J. Speech Lang. Hear. Res.* **32**, 689–697.

Drechsel, J. S., and Thomson, S. L. (2008). "Influence of supraglottal structures on the glottal jet exiting a two-layer synthetic, self-oscillating vocal fold model," *J. Acoust. Soc. Am.* **123**(6), 4434–4445.

Finnegan, E. M., and Alipour, F. (2009). "Phonatory effects of supraglottal structures in excised canine larynges," *J. Voice* **23**, 51–61.

Fuks, L., Hammarberg, B., and Sundberg, J. (1998). "A self-sustained vocal-ventricular phonation mode: Acoustical, aerodynamic and glottographic evidences," *KTH TMH-QPSR* **3**, 49–59.

Kitzing, P., and Sonesson, B. (1967). "Shape and shift of the laryngeal ventricle during phonation," *Acta Oto-laryngol.* **63**, 479–488.

Kucinschi, B. R., Scherer, R. C., Dewitt, K. J., and Ng, T. T. (2006). "Flow visualization and acoustic consequences of the air moving through a static model of the human larynx," *J. Biomech. Eng.* **128**(3), 380–390.

Kurita, S., Nagata, K., and Hirano, M. (1983). "A comparative study of the layer structure of the vocal fold," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College Hill Press, San Diego, CA), pp. 3–21.

Kutta, H., Steven, P., Kohla, G., Tillmann, B., and Paulsen, F. (2002). "The human false vocal folds—an analysis of antimicrobial defense mechanisms," *Anat. Embryol.* **205**(4), 315–323.

Lindestad, P. A., Blixt, V., Pahlberg-Olsson, J., and Hammarberg, B. (2004). "Ventricular fold vibration in voice production: A high-speed imaging study with kymographic, acoustic and perceptual analyses of a voice patient and a vocally healthy subject," *Logoped. Phoniatr. Vocol.* **29**, 162–170.

Lindestad, P. A., Sodersten, M., Merker, B., and Granqvist, S. (2001). "Voice source characteristics in Mongolian 'throat singing' studied with high-speed imaging technique, acoustic spectra, and inverse filtering," *J. Voice* **15**, 78–85.

Maryn, Y., De Bont, M. S., and Van, C. P. (2003). "Ventricular dysphonia: Clinical aspects and therapeutic options," *Laryngoscope* **113**(5), 859–866.

McGowan, R. S., and Howe, M. S. (2010). "Influence of the ventricular folds on a voice source with specified vocal fold motion," *J. Acoust. Soc. Am.* **127**, 1519–1527.

Nemetz, M. A., Pontes, P. A., Vieira, V. P., and Yazaki, R. K. (2005). "Vestibular fold configuration during phonation in adults with and without dysphonia," *Braz. J. Otorhinolaryngol.* **71**, 6–12.

Olthoff, A., Schiel, R., and Kruse, E. (2007). "The supraglottic nerve supply: An anatomic study with clinical implications," *Laryngoscope* **117**(11), 1930–1933.

Pepinsky, A. (1942). "The laryngeal ventricle considered as an acoustical filter," *J. Acoust. Soc. Am.* **13**, 32–35.

Pinho, S. M., Pontes, P. A., Gadelha, M. E., and Biasi, N. (1999). "Vestibular vocal fold behavior during phonation in unilateral vocal fold paralysis," *J. Voice* **13**, 36–42.

Stager, S. V., Bielamowicz, S., Gupta, A., Marullo, S., Regnell, J. R., and Barkmeier, J. (2001). "Quantification of static and dynamic supraglottic activity," *J. Speech Lang. Hear. Res.* **44**(6), 1245–1256.

Stager, S. V., Bielamowicz, S. A., Regnell, J. R., Gupta, A., and Barkmeier, J. M. (2000). "Supraglottic activity: Evidence of vocal hyperfunction or laryngeal articulation," *J. Speech Lang. Hear. Res.* **43**, 229–238.

Stager, S. V., Neubert, R., Miller, S., Regnell, J. R., and Bielamowicz, S. A. (2003). "Incidence of supraglottic activity in males and females: A preliminary report," *J. Voice* **17**, 395–402.

Sundberg, J. (1974). "Articulatory interpretation of the 'singing formant,'" *J. Acoust. Soc. Am.* **55**, 838–844.

Svec, 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**, 1397–1407.

Svec, J. G., and Schutte, H. K. (1996). "Videokymography: High-speed line scanning of vocal fold vibration," *J. Voice* **10**, 201–205.

Titze, I. R. (1989). "On the relation between subglottal pressure and fundamental frequency in phonation," *J. Acoust. Soc. Am.* **85**(2), 901–906.

Titze, I. R., and Story, B. H. (1997). "Acoustic interactions of the voice source with the lower vocal tract," *J. Acoust. Soc. Am.* **101**, 2234–2243.

van den Berg, J. (1955). "On the role of the laryngeal ventricle in voice production," *Folia Phoniatr.* **7**, 57–69.

Zhang, C., Zhao, W., Frankel, S. H., and Mongeau, L. (2002). "Computational aeroacoustics of phonation, Part II: Effects of flow parameters and ventricular folds," *J. Acoust. Soc. Am.* **112**, 2147–2154.

Zheng, X., Bielamowicz, S., Luo, H., and Mittal, R. (2009). "A computational study of the effects of false vocal folds on glottal flow and vocal fold vibration during phonation," *Ann. Biomed. Eng.* **37**, 625–642.