Published in final edited form as:

J Voice. 2012 July ; 26(4): 403–411. doi:10.1016/j.jvoice.2011.07.002.

Measures of Spectral Slope Using an Excised Larynx Model

Fariborz Alipour¹, Ronald C. Scherer², and Eileen Finnegan¹

¹Department of Communication Sciences & Disorders, The University of Iowa, Iowa City, Iowa 52242

²Department of Communication Sciences and Disorders, Bowling Green State University, Bowling Green, Ohio 43403

Abstract

Spectral measures of the glottal source were investigated using an excised canine larynx model for various aerodynamic and phonatory conditions. These measures included spectral harmonic difference H1-H2 and spectral slope that are highly correlated with voice quality but not reported in a systematic manner using an excised larynx model. It was hypothesized that the acoustic spectra of the glottal source were significantly influenced by the subglottal pressure, glottal adduction, and vocal fold elongation, and the resulting vibration pattern. Canine larynges were prepared, mounted on the bench with and without false vocal folds, and made to oscillate with flow of heated and humidified air. Major control parameters were subglottal pressure, adduction, and elongation. Electroglottograph (EGG), subglottal pressure, flowrate, and audio signals were analyzed using custom software. Results suggest that an increase in subglottal pressure and glottal adduction may change the energy balance between harmonics by increasing the spectral energy of the first few harmonics in an unpredictable manner. It is suggested that changes in the dynamics of vocal fold motion may be responsible for different spectral patterns. The finding that the spectral harmonics do not conform to previous findings was demonstrated through various cases. Results of this study may shed light on phonatory spectral control when the larynx is part of a complete vocal tract system.

INTRODUCTION

Spectral measures of the glottal source have been correlated with voice quality by many investigators.¹⁻⁴ Two of the most reported measures are the spectral difference between the first and second harmonics (H1-H2) and the spectral slope. The common assumption concerning the glottal source assumes a uniformly decaying spectral envelope with a typical spectral slope value of -12 (dB/octave) for normal chest phonation.⁵ Similarly, using a glottal pulse model, Childers and Lee reported spectral value of -12 (dB/octave) for modal voice and a value of -18 (dB/octave) for falsetto and breathy voices. Other investigators have reported spectral measures based on the inverse filtering of the human glottal flow signal.^{2, 6-8}

Some links between the glottal source spectra and phonatory mechanics are well known. For example, Gauffin and Sundberg⁹ studied the spectral correlate of the glottal source for six subjects (singers and non-singers). They used hardware inverse filtering and acquired the flow with a Rothenberg flow mask system. They obtained spectral information via a B&K condenser microphone without the presence of the flow mask. Subglottal pressure was estimated from the oral pressure during /p/ occlusion for /pæ/ repetitions. They also modeled

the estimated peak glottal area (EPA) from the subglottal pressure and the peak volume velocity using the lossless Bernoulli equation relating area to transglottal pressure and flow. They found that adduction change from pressed to normal to flow and breathy was associated with EPA increase and subglottal pressure decrease; SPL increased from pressed to flow and then decreased from flow to breathy. In addition, for a given register, SPL corresponded to the amplitude of the differentiated flow except at low intensity levels.

Hillenbrand *et al.*¹⁰ investigated the acoustic characteristics of breathy voice. They measured spectral information from 15 subjects (8 men and 7 women) phonating 4 vowels at 3 conditions of normal, moderately breathy, and very breathy. They selected the one second of the maximally stable segment from their audio recordings for analysis. Using six different acoustic measures, they compared their data and the effects of phonation type, vowel, and gender on breathiness ratings.

Holmberg *et al.*¹¹ compared aerodynamic, EGG, and spectral measures for the vowel /æ/ and the syllable /pæ/ from 20 female subjects. Their aerodynamic analyses included glottal airflow parameters, transglottal pressure, DC flow, and flow adduction quotient (defined as glottal closed time divided by the period), and the acoustic measures SPL, fundamental frequency, and the amplitude differences between the first two spectral harmonics. They found that there were no significant differences in parameter values for the vowel in repeated /pæ/ syllables versus sustained phonation of /æ/. They indicated that H1-H2 is directly related to the degree to which the glottal waveform has a sinusoidal shape and inversely related to adduction. Also, based on their statistical analysis, they showed that the correlation between the adduction quotient from the EGG signal and the glottal waveform was weak.

It is noted that fundamental frequency is related to subglottal pressure^{12,13}, both increasing together for the same adduction, typically, and this should be reflected primarily in the source alone (not the vocal tract). Also, MFDR tends to increase with peak-to-peak glottal flow as well as with the closed quotient (for softer phonation) and the skewing quotient. Sundberg *et al.*¹⁴ also showed that the ratio of peak glottal flow by the subglottal pressure (flow “permittance”) consistently separated the phonatory modes of normal, pressed, and “flow”, suggesting a combination of adductory and possibly vocal fold tension relationships.

Although the reviewed literature points to consistent trends for change of glottal acoustics, measuring the acoustic effects directly appears to be an important step. If the vocal tract and its corresponding resonance and interaction effects are not involved, as is the case for excised larynges where real or artificial vocal tracts are not attached, then the results should reflect the acoustics of the glottal flow (plus radiation) directly. Although inverse filtering might have provided a different view of the glottal source, directly measuring the subglottal pressure and microphone signal near the source is more suitable for the *in vitro* setup. In the experimental setup that was used for the current study, there was a sufficient signal-to-noise ratio to permit spectral analysis of the acoustic output, where the signal was taken to be the signal-plus-noise during excised larynx phonation, and the noise was taken to be the sounds picked up by the microphone when the larynx was not phonating but while the rest of the equipment was running. Unlike the human larynx, the excised larynx as a source of sound is directly accessible and one can measure and study the changes in the output phonation spectra while the subglottal pressure, adduction, and vocal fold length conditions are specifically controlled. It is noted that the excised larynges were not innervated, so the results of this study reflect passive conditions.

Alipour *et al.*¹⁵ studied the aerodynamic and acoustic effects of sudden phonatory changes (chest to falsetto and vice versa) in ten excised canine larynges and reported that continuous

changes in subglottal pressure and flow rate alone can trigger mode changes in the excised canine larynx during passive phonation (no innervation). Also, their spectral analysis of the microphone signals indicated major differences in the spectral slope and harmonic structure of the chest and falsetto modes of phonation. This work is an extension of that previous study.

In a study to determine which glottal source measures are meaningful and robust, Kreiman *et al.*⁴ identified four independent factors that may be important determinants of voice quality. These included: the difference between the first and second harmonic intensities (H1-H2), overall spectral slope, high frequency noise excitation, and H2-H4. To be consistent with that study, the present study reports H1-H2, the spectral slope, and the closed quotient using EGG, as well as the overall sound pressure level (SPL).

The purpose of this study was to demonstrate that: (1) increasing subglottal pressure increases SPL due to an increase in the energy of the harmonics but not necessarily the fundamental, (2) increasing adduction or elongation can change the distribution of spectral harmonics such that the spectral slope may not conform to the general known values and overall sound pressure level may decrease, (3) the amplitude of the first harmonic is not always greater than that of the second harmonic, and (4) a distinct change in vibratory pattern will occur under certain tension conditions for a small increase or decrease in subglottal pressure.

METHODS

Seven excised canine larynges were obtained following cardiovascular research experiments at the University of Iowa Hospitals and Clinics. The canines ranged in weight from 17 kg to 25 kg with vocal fold length ranging from 12 mm to 16 mm (see Table 1). Excised larynges were mounted and operated according to previous work^{15, 16}. Subglottal pressure, flow rate, adduction, and vocal fold elongation were the major control variables. Each excised larynx experiment started with pressure-flow sweeps at specific vocal fold lengths to evaluate the operating ranges of the larynx. Then, a series of sustained phonation runs were made within the working range of pressure and flow to record and observe oscillation of the vocal folds in slow motion visualized with a strobe light.

For each sustained oscillation, the 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 (Gilmont rotameter, J197). The subglottal pressure signal was recorded using a pressure transducer (Microswitch 136PC01G1) mounted perpendicular to the flow in the tracheal tube 10-12 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). Two electrode plates from a Synchrovoice Electroglottograph (EGG) were placed on the thyroid lamina at the level of the vocal folds and a third electrode was placed on the posterior side of the larynx; they all were then secured with duct tape to obtain the maximum EGG signal during phonation. The audio signal was obtained with a microphone (Sony ECM-MS907) at a distance of 15-20 cm from the larynx and recorded on a digital audio tape recorder (Sony PCM-M1). The sound intensity (SPL) was measured with a sound level meter with “A” weighting (Extec model 407738), placed about 15 cm from the larynx. Because this weighting attenuates intensities at lower frequencies, a correction was applied to the readings using the following formulae

$$C = 117.9 - 30.359 \ln(F0) + 1.918[F0]^2$$

where C is the correction to be added to the SPL reading, F_0 is the fundamental frequency, and $\ln$ stands for the natural logarithm. This formula was obtained by curve fitting for the correction.

Adduction and Elongation

Adduction was controlled either by approximating the arytenoid cartilages against metal shims of various thicknesses (0.3 – 1.0 mm) or 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 arytenoid adduction. Medial vocal fold adduction corresponding to vocal fold bulging due to TA contraction could not be achieved due to the lack of muscle innervation. The adduction levels were the weights (50-200 grams) that pulled the sutures attached to the muscular process of the arytenoid cartilages. The vocal folds were elongated by pulling the anterior aspect of the thyroid cartilage with a micrometer-controlled alligator clip attached to the middle of the thyroid cartilage, or by weights pulling posteriorly on sutures attached to the arytenoid cartilages. The elongation levels in millimeters were the amount of the pull on either the thyroid cartilage or arytenoid to elongate the vocal folds.

Data collection and processing

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 into a computer by using an A/D (14 bit) board and software (DATAQ Instruments). The signals were then converted to calibrated physical quantities in a MATLAB routine and used for the aerodynamic and acoustic analyses.

Spectral analyses of the signals were obtained with a Fast Fourier Transform (FFT) of the microphone signal in the MATLAB computing environment. Only those cases with a relatively high signal-to-noise ratio ($\text{SNR} > 20$) were included in the spectral analysis. SNR was calculated in TF32 Basic level program (Paul Milenkovic, Madison WI). Each FFT was calculated with at least 4096 data points for adequate resolution. The amplitude of at least four consecutive harmonics (H1 through H4) was estimated with a cursor from the FFT plot. In addition, the spectral slope was calculated using the harmonics between 0 and 2000 Hz. When the spectral harmonics did not decay in logarithmic fashion from the first harmonic, the slope was computed from the largest harmonic. A peak picking computer program was used to locate and determine the frequency and amplitude of the harmonics and a MATLAB program was used to compute the spectral slope. By linear fitting of the amplitude in dB and $\log_{10}$ of the frequency (Figure 1), the slope in dB per octave was calculated by multiplying the slope of the fitted line by $\log_{10}$ of 2.

To obtain sound pressure levels during sweeps, the RMS of the microphone signal was calibrated against measured sound pressure levels for sustained phonation of the same larynx at various adduction and subglottal pressure values, and a regression equation was established between the RMS of the microphone signal and the sound pressure level.

To calculate the fundamental frequency (F_0), the EGG signal was low-pass filtered at 150% of its estimated F_0 value seen from the spectrogram or an oscilloscope. The fundamental frequency was then calculated with a zero crossing method. First the signal DC offset was removed and then the periods of all the cycles in the selected segment were calculated from consecutive zero crossings and averaged. Closed quotient (CQ) was estimated using the differentiated EGG (original) signal. The point of the maximum peak in the differentiated signal marked the estimated start of the glottal closed time and the point of the minimum

peak marked the estimated end of the glottal closed time. The closed quotient was obtained by dividing this closed time by the cycle period. The pressure-flow sweep data were averaged according to the procedure outlined in Alipour and Scherer,¹³ where the sweep data were divided into 50-100 segments of 10-20 phonatory cycles and the mean subglottal pressure, mean flow rate, and RMS values of the microphone signal were obtained for each segment.

Human glottal flow

The cycles of the glottal flow create sound that is radiated into the room. The microphone essentially records the acoustics of the differentiated glottal flow due to the radiation effect.¹⁷ As an orientation to the correspondence between spectral shape and phonatory conditions and understanding the relation between excised phonation and human phonation, human (RS: adult male) glottal flow is illustrated in Figure 2, where a normal phonation (inverse filtered) glottal flow signal is contrasted with a very breathy (hypoadducted) and a highly pressed (hyperadducted) phonation. These signals were obtained by inverse-filtering of the flow signal obtained from the Glottal Enterprises flow system. The program used was TF32 Basic level (Paul Milenkovic, Madison WI). The glottal flow signal associated with breathy phonation, with an average flow rate of 995 ml/s, appears nearly sinusoidal, while normal phonation with an average flow of 239 ml/s and pressed phonations with an average of 90 ml/s have flow pulses that are skewed more to the right. The strong negative airflow in this case is probably due to the inverse filtering software. The corresponding FFT spectra for these (undifferentiated) signals are shown in Figure 3. It is seen that the spectral slope over the first two octaves is approximately -20 dB/octave for breathy phonation, -10 dB/octave for normal phonation, and -9.6 dB/octave for the pressed phonation for the cases shown, a wide range of slopes. Thus, in the study here, there should also be the expectation that a wide range of glottal source spectral characteristics should be found as phonatory adduction is varied over a wide range.

While it is difficult to categorize the sound of excised larynx phonation as pressed, normal, or breathy qualities as one might be able to do for human phonation, phonatory parameters shared by both excised and live human phonation, such as adduction, elongation, and subglottal pressure can be studied. This control results in important shared variables, such as the closed quotient of the glottal flow, which may change depending upon glottal adduction, length, and subglottal pressure. Thus, trends for the alteration of spectral characteristics of the source should be similar for both excised and human phonation, which in turn should correspond to the voice quality ultimately produced by the speaker.

RESULTS

Figure 4 (A & B) shows sound pressure level variations during series of sustained phonation of excised larynx CL37 at various adduction and elongation levels (hypothesis 1). The data support the common findings that sound pressure level increases with subglottal pressure (here approximately 0.7 dB/cm H₂O for both figures combined). Figure 4A represents the cases for adduction levels of 50-200 gram weights on arytenoid sutures with no vocal fold elongation and Figure 4B shows sound pressures at 1, 2, and 3 millimeter elongation levels for an adduction level of 150 grams. The data suggest that, at high pressures (over 16 cm H₂O), the sound pressure level is almost insensitive to adduction (no consistent trend). On the other hand, for the medium adduction level, an increase of elongation tends to lower SPL. Also, the results suggest that, in general, a particular SPL is achieved at higher subglottal pressure values for conditions of less adduction or greater elongation.

Figure 5 compares the EGG and subglottal pressure waveforms of the excised larynx CL66 at low and medium adductions (100 & 200 grams of weight on sutures, respectively) of the

same glottal length of 1.2 cm. In the low adduction condition (top two traces), the vocal folds oscillated at about 195 Hz with mean subglottal pressure of 16 cm H₂O and mean flow rate of 900 ml/s. The EGG signal indicates only a short period of contact (CQ=0.178). The video image indicated a posterior gap with harmonic anterior-posterior contact (almost half) of the membranous vocal folds. The subglottal pressure appears sinusoidal, similar to what one might expect for the breathy signal of human phonation. The medium adducted folds (bottom two traces) oscillated at about 210 Hz with approximately the same subglottal pressure of 15 cm H₂O and a much reduced mean flow rate of 300 ml/s. The EGG signal shows longer contact (CQ=0.37). The video image indicated no posterior gap and full anterior-posterior contact. The pressure signal indicates sufficient energy to have excited a subglottal resonance of the system, as indicated by the dent in the subglottal pressure signal. Because the pressure measures were obtained from a transducer placed 10-12 cm below the vocal folds, deviations in the Ps from a regular sinusoidal shape are due to subglottal resonances. Figure 6 shows the FFT spectrum of the microphone signal corresponding to the above conditions of Figure 5. It should be noted that the strongest harmonic is of the radiated source without a vocal tract. In low adduction (top trace) the second and third harmonics are about 23 dB weaker than the first harmonic. However, the bottom trace for the higher adduction indicates that the difference between the first and second harmonics, and first and third harmonics, is considerably less, approximately 5 and 10 dB, respectively. Thus, in this case, the adduction increase changed the energy balance between harmonics by increasing the spectral energy of the first few harmonics and reducing the difference between H1 and H2 and between H3 and H4. In addition, the relative energy of the 6th and 8th harmonics is higher, with a “missing” 7th harmonic, reminiscent of the complexity of spectra for different duty cycles. Similar results were obtained for four other larynges under contrasting adduction conditions.

Of particular interest is the case in which a small increase of mean pressure caused a mode change for a larynx with the vocal folds under tension (elongated). This is demonstrated in Figure 7 which shows the change in phonation that results from a gradual increase in subglottal pressure from 26 to 27 cm H₂O. The top two traces correspond to the EGG and subglottal pressure waveforms of sustained oscillation of excised larynx CL33 at mean subglottal pressure of 26 cm H₂O and mean flow rate of 1000 ml/s. The oscillation is at a high frequency of about 468 Hz. The EGG signal appears nearly sinusoidal with very small amplitude due to lack of vocal fold closure. A few moments later in the pressure sweep a significant change in oscillation occurs. The bottom two traces of Figure 7 are for a subglottal pressure of 27 cm H₂O (just slightly higher during the pressure sweep) and a flowrate of 900 ml/s. There is a sudden drop in frequency to about 166 Hz and the waveforms are significantly different from sinusoidal. The EGG signal has larger amplitude with a closed quotient of about 0.19. These differences are translated in their audio spectra shown in Figure 8. In the top trace, the spectrum shows 6 harmonics with a strong fundamental and a second harmonic 14 dB weaker which can result in a well-defined spectral slope. The average spectral slope is -10 dB/octave (over the first four harmonics). However, the lower frequency oscillation has more harmonics of higher relative intensity, but interestingly the second harmonic is 9 dB stronger than the first harmonic, despite the absence of a vocal tract or supraglottic laryngeal structures. This may be due to the large oscillation amplitude of the vocal folds at this chest mode which excites many harmonics with considerable energy. The spectral slope is not well defined in this case due to the lack of logarithmic energy decay. However, one can define the spectral slope starting from the largest harmonic, which was used in this study for those cases for which the fundamental was not the largest harmonic.

Figure 9 shows FFT plots of acoustic signals from three excised larynges CL72, CL73, and CL74 at relatively high flow rates. In the top panel, CL72 oscillated at a frequency of 259

Hz, sound pressure level of 82 dB, subglottal pressure of 22 cm H₂O, and flowrate of 1.3 l/s. The pressure and microphone signals were approximately sinusoidal, the EGG signal was weak, and there was no vocal fold contact. The spectral shape conforms to breathy or falsetto mode with three major harmonics and a sharp decline of energy (H1-H2=24 dB). In the second panel case (CL73), the pressure, flow and sound intensity are similar in values to the top panel case (22 cm H₂O, 1.4 l/s, and 92 dB). However, the spectral pattern is completely different. The larynx oscillated at 110 Hz with vocal fold contact, a strong EGG signal, and a closed quotient of 0.14. The pressure and microphone signals were complex waveforms with many more spectral harmonics. The second harmonic is about 9 dB stronger than first harmonic. This larynx probably oscillated in chest mode. In the bottom panel, excised larynx CL74 oscillated with 18 cm H₂O pressure, 0.8 l/s flowrate, 73.3 dB, and at 187 Hz with asymmetric motion (the two vocal folds moved in the same direction most of the time, with little contact). The pressure and microphone signals were complex waveforms, yet the EGG signal was relatively sinusoidal. The first three harmonics have similar strength with a stronger second harmonic (H1-H2= -1.5 dB).

Figure 10 is another example of spectral changes, this time for excised larynx CL64 at various vocal fold elongations (tensions) in falsetto mode at the same adduction level. In the top panel, the vocal folds were stretched 15% (low tension) and the larynx oscillated with 12 cm H₂O pressure, 0.56 l/s flow at 209.4 Hz. The spectrum shows a large peak at the fundamental, with a large drop to the second harmonic (H1-H2=24.5 dB). The video indicated a large amplitude of vocal fold oscillation with minor tissue contact. The second panel shows the spectrum of the same larynx with 30% vocal fold elongation (medium tension) that oscillated with 16 cm H₂O pressure, 0.45 l/s flowrate at 368.1 Hz. The first three harmonics are almost equally strong with 5 dB from first to second. The vocal folds oscillated with low amplitude and no contact and was similar to the falsetto mode, but surprisingly the spectrum is not. The third panel corresponds to the 45% elongation (high tension) with 29.8 cm H₂O pressure, 0.76 l/s flowrate, and oscillation frequency of 452.8 Hz. Again, the video image suggested a falsetto-like low amplitude oscillation with no contact. The second harmonic for this case is 1.1 dB stronger than its fundamental. This is unlike what is typically expected from falsetto phonation in the human.

Figure 11 shows the spectral ranges for seven canine larynges including the first-second harmonic difference (H12), second-fourth harmonic difference (H24), and spectral slope (SS) across a wide variety of phonatory conditions. On the X-axis larynges are identified and the Y-axis represents the values in the corresponding unit in the key. The mean and standard deviation for each parameter are represented with their bar chart and error bar. It is noted that H12 for larynges CL33, CL72, and CL73 took on large negative values. The H24 difference was always positive for all seven larynges. As can be seen in this graph, the spectral peak differences have large variability across larynges, but the spectral slope shows the least variability (except for CL73). The average spectral slope was approximately -10 dB. Given that the larynges produced both falsetto and chest phonation, the slope was not as steep as might be expected.

To examine the general trend of these parameters for adduction and subglottal pressure across all larynges, a grouping of data was made based on the range of adduction levels of low (50 g), medium (100 g), and high (>100 g), also based on subglottal pressure ranges of low (<12 cm H₂O), medium (between 12 to 20 cm H₂O) and high (>20). Figure 12 represents grouped average and standard deviations for all seven larynges for the same parameters as in Figure 11. The two-digit group number indicates the range of the adduction (first digit) and pressure (second digit) from low to high values. It is interesting to note that H12 has positive average values for groups, but still contains large variability. In the medium adduction group (21-23), there is an increasing trend of these parameters with

subglottal pressure. In all adduction levels, the higher pressure values resulted in steeper spectral pattern. The H24 has the highest values for highest adduction or highest pressure values. The spectral slope has the lowest variability compared to the other variables.

DISCUSSION

Acoustic spectra of the glottal source have been the focus of many investigations because of the direct effect on the perception of vocal quality. This study investigated the effects of subglottal pressure, glottal adduction, and vocal fold tension on acoustic measures of the glottal source using excised canine larynges. At any given level of adduction and elongation, sound pressure level was shown to increase with subglottal pressure. In Figure 4B, the longer elongation conditions have lower SPL values for the same pressures, especially above 20 cm H₂O. This suggests that the control of SPL is strongly related to subglottal pressure, and is sensitive to both adduction and elongation, with greater sensitivity to adduction at lower subglottal pressure values, and more sensitive to elongation at higher subglottal pressures.

The effects of vocal fold dynamics on acoustic correlates are interpreted through glottal contact or EGG closed quotient and through the aerodynamic parameters. For example as indicated in **Figures 5 & 6**, an increase of adduction at similar subglottal pressures, which is accompanied by the decrease of flowrate and increase of closed quotient, feeds more energy into the second and third harmonics. This may be due to the increased dynamic activity of the vocal folds with more contact. This is consistent with findings of Sundberg and Gauffin¹⁸, where they observed that with greater adduction and higher levels of SPL, higher harmonics achieve greater strength than lower harmonics. This can also be observed in an enhanced condition of a mode change from falsetto-like to chest-like as shown in **Figures 7 & 8**. For the falsetto mode the waveforms look sinusoidal with little or no contact and the energy is mostly in the first two harmonics. On the other hand, the chest mode (bottom traces) is rich in harmonics which is consistent with the findings of other investigators such as Colton¹⁹. The EGG waveform indicates a sharper closure that feeds more energy into the higher harmonics.

As suggested by Alipour *et al.*¹⁵, the subglottal pressure is a major mechanism of controlling the modes of phonation besides adduction and it is seen again here in Figure 8 that a mode shift due to a slight increase in subglottal pressure can have major effects on the increase in intensity of higher harmonics. This can be due to the changes in the dynamics of the vocal folds which can be explained through the eigenmode analysis discussed by Berry²⁰, where the nonlinearities in the system, such as stress-strain relationships or pressure-flow relationships, may facilitate the mode changes. When conditions are right, mode change can happen with a small change in one parameter such as subglottal pressure. The vocal fold dynamics selects a vibration mode that is closer to its eigenmode for the given condition of pressure, flow, and adduction. The excitation of the different vibration modes may result in emphasis of different spectral harmonics.

Excised phonation with high flow rates may resemble the source of breathy phonation and should have a sharp contrast between its fundamental and second harmonic spectral values (large spectral slope). Figure 9 demonstrated a discrepancy in this idea with examples that have completely different spectral patterns for three different excised larynges oscillation in similar aerodynamic conditions. While the spectral pattern in one (CL72) conforms to breathy phonation with a spectral slope of -16 dB/octave, the other larynges do not. For example, the second panel shows the larynx (CL73) with flowrate and pressure patterns similar to chest mode (109.6 Hz) with the second harmonic as its strongest harmonic. The third panel has a similar pattern to the second at a higher fundamental frequency of 186.7

Hz. It appears that the dynamics of the vocal fold motion is the main reason for the spectral harmonics and not just the aerodynamic mode.

Similarly, we noticed in Figure 10 that different falsetto conditions of the excised larynx CL64 at 15%, 30%, and 45% elongation rates demonstrated different spectral behaviors that were not expected. The first panel showed a falsetto mode with spectral slope of -14.5 dB/octave which conforms to values reported in the literature. However, the spectral patterns of the second and third panels do not. This suggests that the distribution of the energy between harmonics is not defined by the modes of phonation *per se*, but by the dynamics of vocal folds. Besides the elongations, the subglottal pressure was increased from 12 cm H₂O in panel one to 16 cm H₂O in panel two to almost 30 cm H₂O in panel three. This might be another reason for the greater energy in the higher harmonics of the second and third panels.

The comparative spectral data of all excised larynges shown in Figure 11 indicate that the relative consecutive harmonic difference between the first and second harmonics (H1-H2) has a large variability across the different larynges. This suggests that H1-H2 may not be a consistent measure to indicate spectral slope but may be useful in characterizing other important characteristics of the signal. This is shown in Figure 11 also, where the grouped data have positive average H1-H2 values, but the large variability makes it hard to use this parameter alone. The second-fourth harmonic difference (H2-H4) is always positive with similar ranges except for larynx CL72. Despite the inconsistency in spectral pattern, when spectral slope is defined from the largest harmonic as described in our methodology, the range for spectral slope falls within expected values (-12 to -9 dB/octave) excluding CL73. Also, when these parameters are estimated for similar adduction and pressure ranges (Figure 12), a better result can be expected.

CONCLUSIONS

The findings here support observations of other investigators that sound pressure level increases with subglottal pressure at any adduction or elongation level. The limited data so far suggest that increase of subglottal pressure is a primary means to increase the intensity of all harmonics, with the first harmonic generally changing the least, and the second harmonic gaining greater intensity than the first harmonic for higher P_s values. For a given P_s value, greater adduction tends to increase SPL and greater elongation tends to decrease SPL. However, the spectral pattern does not always follow any particular trend; in particular, the spectral difference H1-H2 has large variability and may be an inconsistent measure of glottal source.

Acknowledgments

The project described was supported by Award Number R01DC009567 from the National Institute on Deafness and other Communication Disorders. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Deafness and other Communication Disorders or the National Institutes of Health.

REFERENCES

- Childers DG, Lee CK. Vocal quality factors: analysis, synthesis, and perception. *J Acoust Soc Am*. 1991; 90(5):2394–2410. [PubMed: 1837797]
- Holmberg EB, Perkell JS, Hillman RE, Gress C. Individual variation in measures of voice. *Phonetica*. 1994; 51(1-3):30–37. [PubMed: 8052674]
- Shrivastav R, Sapienza CM. Objective measures of breathy voice quality obtained using an auditory model. *J Acoust Soc Am*. 2003; 114(4):2217–2224. [PubMed: 14587619]

4. Kreiman J, Gerratt BR, Antonanzas-Barroso N. Measures of the glottal source spectrum. *J Speech Hear Res.* 2007; 50:595–610.
5. Titze IR. Acoustic interpretation of the voice range profile (Phonetogram). *J Speech Hear Res.* 1992; 35:21–34. [PubMed: 1735970]
6. Childers DG, Ahn C. Modeling the glottal volume-velocity waveform for three voice types. *J Acoust Soc Am.* 1995; 97(1):505–519. [PubMed: 7860829]
7. Eskenazi L, Childers DG, Hicks DM. Acoustic correlates of vocal quality. *J Speech Hear Res.* 1990; 33(2):298–306. [PubMed: 2359270]
8. Laukkanen AM, Bjorkner E, Sundberg J. Throaty voice quality: subglottal pressure, voice source, and formant characteristics. *J Voice.* 2006; 20(1):25–37. [PubMed: 15990272]
9. Gauffin J, Sundberg J. Spectral correlates of glottal voice source waveform characteristics. *J Speech Hear Res.* 1989; 32:556–565. [PubMed: 2779199]
10. Hillenbrand J, Cleveland RA, Erickson RL. Acoustic correlates of breathy vocal quality. *J Speech Hear Res.* 1994; 37(4):769–778. [PubMed: 7967562]
11. Holmberg EB, Hillman RE, Perkell JS, Guid PC, Goldman SL. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *J Speech Hear Res.* 1995; 38(6):1212–1223. [PubMed: 8747815]
12. Titze IR. On the relation between subglottal pressure and fundamental frequency in phonation. *J Acoust Soc Am.* 1989; 85(2):901–906. [PubMed: 2926005]
13. Alipour F, Scherer RC. On pressure-frequency relations in the excised larynx. *J Acoust Soc Am.* 2007; 122(4):2296–2305. [PubMed: 17902865]
14. Sundberg J, Titze I, Scherer R. Phonatory control in male singing: A study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source. *J Voice.* 1993; 7(1):15–29. [PubMed: 8353616]
15. Alipour F, Finnegan EM, Scherer RC. Aerodynamic and acoustic effects of abrupt frequency changes in excised larynges. *J Speech Hear Res.* 2009; 52:465–481.
16. Alipour F, Scherer RC. Pulsatile airflow during phonation: an excised larynx model. *J Acoust Soc Am.* 1995; 97(2):1241–1248. [PubMed: 7876445]
17. Fant, G. KTH Quarterly Report. Vol. 1. The Royal Institute of Technology; Stockholm, Sweden: 1979. Glottal source and excitation analysis; p. 85-107.
18. Sundberg, J.; Gauffin, J. Waveform and spectrum of the glottal voice source. In: Lindblom, B.; Ohman, S., editors. *Frontiers of Speech Communication Research*. Academic Press; London: 1979.
19. Colton RH. Spectral characteristics of the chest and falsetto registers. *Folia Phoniatica.* 1972; 24:337–344. [PubMed: 4670039]
20. Berry DA. Mechanisms of modal and nonmodal phonation. *J Phonetics.* 2001; 29:431–450.

Figure 1.

Spectral peaks and a fitted line for calculation of the spectral slope. Note that $(\text{dB1} - \text{dB2}) / [(\log_{10}f_1 - \log_{10}f_2) \log_{10}2]$ is the slope in dB/octave.

Figure 2. Inverse-filtered glottal flow waveforms for breathy, normal, and pressed human phonation from an adult male human (RS).

Figure 3.
Spectral patterns corresponding to Figure 2.

Figure 4. Sound pressure variations during pressure-flow sweeps in excised larynx CL37. Figure 4A represents adduction cases with different weights (grams) pulling the arytenoid sutures.

Figure 4B shows cases with 0, 1, and 2 mm elongations from a nominal length for a constant adduction.

Figure 5.
EGG and subglottal pressure waveforms of excised larynx CL66 at low (upper two traces) and medium adduction (lower two traces) levels.

Figure 6.
Spectral patterns corresponding to Figure 5.

Figure 7.
EGG and pressure waveforms of excised larynx CL33 with vocal fold under tension before and after mode changes.

Figure 8.
Spectral patterns corresponding to Figure 7.

Figure 9.
Spectral patterns of three excised larynges CL72, CL73, and CL74 for similar aerodynamic conditions.

Figure 10. Spectral patterns of excised larynx CL64 at elongation rates of 15, 30, and 45% resulting in three falsetto oscillation modes.

Figure 11. Spectral ranges of individual canine larynges, including average and standard deviation values of first-second harmonic difference (H12), second-fourth harmonic difference (H24), and spectral slope (SS).

Figure 12.

Spectral ranges of the canine excised larynges based on the adduction and subglottal pressure grouping. The two-digit group number indicates the range of the adduction (first digit) and pressure (second digit) from low to high values. The adduction range was 50 [1], 100 [2], 150 [3] grams weight and pressure range was 4-12 cm H₂O [1], 12-20 cm H₂O [2], and 20-35 cm H₂O [3].

Table 1

Information on the canine larynges used in this study. Vocal fold length (VFL) was measured at rest position from the anterior commissure to the tip of vocal processes.

Larynx	Gender	Weight	VFL
CL33	M	22	16
CL37	M	N/A	12
CL64	M	18	13
CL66	F	17	12
CL72	F	26	12
CL73	M	17	14
CL74	F	18	14