

UCLA

UCLA Previously Published Works

Title

Perceptual consequences of changes in epilaryngeal area and shape.

Permalink

<https://escholarship.org/uc/item/0cf8k9gz>

Journal

The Journal of the Acoustical Society of America, 136(5)

ISSN

0001-4966

Authors

Samlan, Robin A
Kreiman, Jody

Publication Date

2014-11-01

DOI

10.1121/1.4896459

Peer reviewed

Perceptual consequences of changes in epilaryngeal area and shape

Robin A. Samlan^{a)} and Jody Kreiman

Department of Head and Neck Surgery, UCLA School of Medicine, 31-24 Rehab Center,
1000 Veteran Avenue, Los Angeles, California 90095-1794

(Received 14 July 2014; accepted 9 September 2014)

The influence of epilaryngeal area on glottal flow and the acoustic signal has been described [Titze, *J. Acoust. Soc. Am.* **123**, 2733–2749 (2008)], but it is not known how (or whether) changes in epilaryngeal area influence perceived voice quality. This study examined these relationships in a kinematic vocal tract model. Epilaryngeal constrictions and expansions were simulated at the levels of the aryepiglottic folds and the ventricular folds in the context of four glottal configurations representing normal vibration to severe vocal fold paralysis, for the three corner vowels /a/, /i/, and /u/. Minimum and maximum glottal flow, maximum flow declination rate, spectral slope, cepstral peak prominence, and the harmonics-to-noise ratio were measured, and listeners completed a perceptual sort-and-rate task for all samples. Epilaryngeal constriction and expansion caused salient differences in voice quality. The location of constriction was also perceivable. Vowels simulated with aryepiglottic constriction demonstrated lower maximum airflow and less noise than the other epilaryngeal shapes, and listeners consistently perceived them as distinct from other stimuli. Acoustic differences decreased with increasing severity of simulated paralysis. Results of epilaryngeal constriction and expansion were similar for /a/ and /i/, and produced slightly different patterns for /u/. © 2014 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4896459>]

PACS number(s): 43.70.Mn, 43.70.Gr, 43.71.Bp [ADP]

Pages: 2798–2806

I. INTRODUCTION

This study applied a computational model to examine how modifying epilaryngeal cavity shape in different ways alters glottal flow, the acoustic signal, and voice quality for normal and disordered voices. Previous studies (Rothenberg, 1981; Fant, 1986; Titze and Story, 1997; Titze, 2006) have documented the effects of changes in epilaryngeal shape on glottal flow and the acoustic signal, but the nature and extent of the influence of such changes on perceived quality remain unknown.

Several factors predict that changes in quality should occur with changes in epilaryngeal configuration. Theoretical studies suggest that epilaryngeal configuration can alter vocal fold vibration; constriction increases inertance, which improves the impedance match with the glottis and decreases oscillation threshold pressure (Titze and Story, 1997; Titze, 2006). Increased inertive reactance also results in skew of the glottal flow pulse, which increases the maximum flow declination rate (MFDR), harmonic amplitudes, and loudness (Rothenberg, 1981; Fant, 1986; Titze, 2006). A narrowed epilarynx is the hypothetical source of the “singer’s formant,” which increases spectral energy around 2500–3000 Hz (Sundberg, 1974; Titze and Story, 1997). In one computational study, these factors were associated with increased perception of “ring,” although listeners were highly variable in their responses (Bergan *et al.*, 2004).

To provide a more detailed examination of the perceptual consequences of epilaryngeal modification, we used a computational vocal tract model to generate /a/ vowels with epilaryngeal constrictions and expansions at the levels of the aryepiglottic folds and the ventricular folds, in the contexts of varying glottal gap sizes (Experiment 1) or vocal tract configurations for different vowels (Experiment 2). Previously described associations led to the expectation that the two constriction locations would be perceived differently; ventricular fold constriction has been associated with muscle tension dysphonia (Morrison *et al.*, 1986), laryngealization, and harsh voice (Edmondson and Esling, 2006), while aryepiglottic fold constriction has been associated with enhanced voice, such as the singer’s formant (Sundberg, 1974) and the “ringing” quality of twang, belt, and operatic singing (Bergan *et al.*, 2004; Yanagisawa *et al.*, 1989). Thus, we anticipated, for Experiment 1, that the glottal flow, MFDR, and source spectral changes would occur with epilaryngeal constriction in both locations and cause perceivable changes in voice quality. Previous research does not strongly support predictions about the perceptual consequences of epilaryngeal expansion, although the Estill Voice Training System (Estill, 1997) teaches flexibility between constricted and retracted/expanded postures, suggesting that the effects of epilaryngeal expansion may also be perceivable.

If listeners can perceive differences in normal voice quality with changes in epilaryngeal configuration, it is important to determine whether that is also the case when a voice is disordered because such effects are assumed to underlie a variety of common voice therapeutic techniques. For example, a narrowed epilarynx underlies “resonant voice therapy” (Titze and Story, 1997; Titze, 2001, 2006; Titze

^{a)} Author to whom correspondence should be addressed. Currently at the Department of Speech, Language, and Hearing Sciences, University of Arizona, Tucson, AZ. Electronic mail: rsamlan@email.arizona.edu.

and Verdolini Abbott, 2012). Resonant voice is defined clinically as “any voice production that is both easy to produce and vibrant in the facial tissues” (Titze and Verdolini Abbott, 2012, p. 286) and is described perceptually as “neither pressed nor breathy” (Titze and Verdolini Abbott, 2012, p. 286). The acoustic and aerodynamic changes reported with epilaryngeal constriction theoretically result from nonlinear source-filter coupling, the degree of which is proportional to the ratio of glottal area to epilaryngeal area (Titze, 2008). A large persistent glottal gap in the context of a small epilaryngeal area should, therefore, increase source-filter coupling. In previous modeling studies, increased coupling, as the result of decreased epilaryngeal area, was found to increase inertive reactance and create a closed phase in the glottal flow waveform even when the minimum glottal area remained >0 (no glottal closure; Titze, 2008). It is not known, though, whether the strong nonlinear coupling in a context where the glottis never closes (i.e., in cases of severe unilateral vocal fold paralysis) will provide any perceptual benefit given the limited harmonics and strong turbulence resulting from a large glottal gap. To examine this issue, glottal configuration was covaried with epilaryngeal configuration to determine how degree of maximum glottal closure mediates the perceptual consequences of this coupling.

In Experiment 2, additional simulations simultaneously varied epilaryngeal cavity shapes and vocal tract area functions (Story, 2008). This experiment was included to determine whether the effects of epilaryngeal shape from Experiment 1 are specific to the /a/ vowel, or whether generalization to other vowels can be assumed. It is possible that the resonances associated with different vowels might differentially interact with the epilaryngeal resonances and modulate the perceptual effects of the epilaryngeal configurations.

II. EXPERIMENT ONE

A. Methods

1. Stimuli

For modeling purposes, the epilarynx was defined as a 2.4 cm segment of the vocal tract just superior to the vocal folds. A kinematic vocal tract model (Story, 2005; Samlan et al., 2013) was used to create five epilaryngeal cavity shapes by orthogonally modifying epilaryngeal area (“constricted” vs “expanded”) and location of the constriction/expansion within the epilarynx (approximate locations of the ventricular folds and aryepiglottic folds), relative to a fifth “neutral” condition based on a naturally produced vowel. The ventricular folds were modeled as a 0.4 cm section beginning just superior to the vocal folds,¹ and the aryepiglottic folds were similarly modeled as a 0.4 cm section at a location ~ 2.0 cm superior to the vocal folds. Vocal tract area for each condition, as a function of distance from the glottis, is shown in Fig. 1. Epilaryngeal area for the imposed segments was smoothly varied upstream and downstream from the target values using a Gaussian interpolation function (see Story, 2005). Epilaryngeal area equaled 0.2 cm^2 for both constriction locations and 1.2 cm^2 for both expansion locations, consistent with the ranges of modification reported in Story et al. (2001).

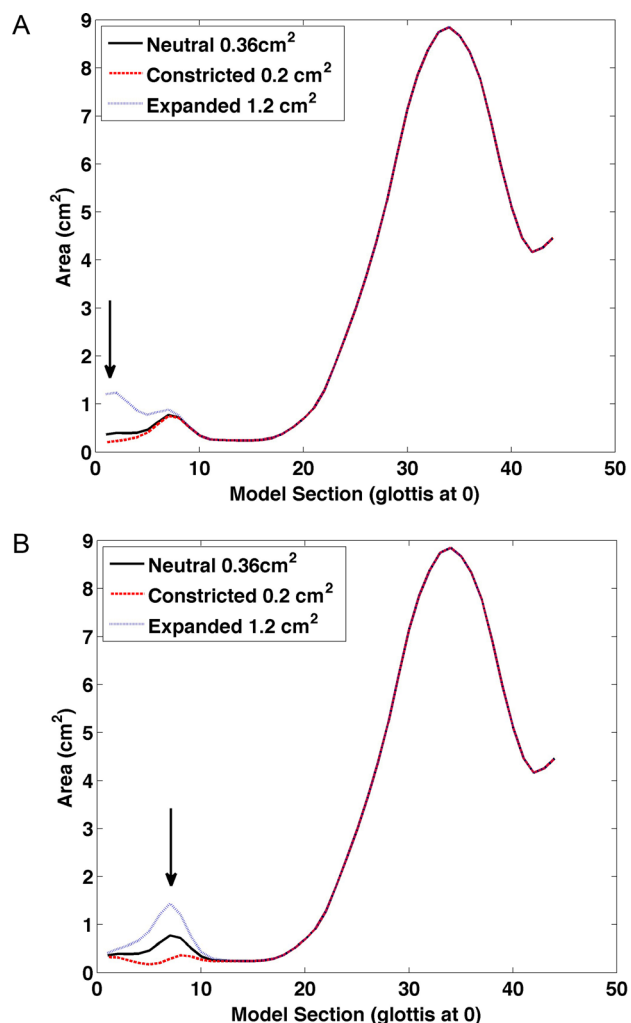


FIG. 1. (Color online) Area functions for the neutral and modified epilaryngeal cavity shapes paired with the /a/ vowel. The arrow indicates the distance from the glottis at which the area was specified. (a) With constriction and expansion at the level of the ventricular folds and (b) with constriction and expansion at the level of the aryepiglottic folds.

In the neutral condition, epilaryngeal area was 0.36 cm^2 at the level of the ventricular folds and 0.62 cm^2 at the level of the aryepiglottic folds. Settings for the vocal tract area function downstream from the epilarynx were identical for every simulation and produced the vowel /a/ (Story, 2008). The effect of the pyriform sinuses was not included in this study.

A kinematic vocal fold model based on Titze (1984, 2006b) was used to create four sets of model parameters by specifying values for adduction, bulging, nodal point ratio, amplitude, and phase. The parameter sets sequentially introduced postural and vibratory asymmetries between the left and right vocal folds, analogous to increasingly severe left vocal fold paralysis, and resulting in normal, mildly, moderately, and severely asymmetric glottal configurations. Adduction is defined as the distance, in mm, of the superior aspect of the vocal process from midline during vibration. Bulging is the curvature of the medial surface of the vocal fold in the vertical dimension. Nodal point is the pivot point around which the rotational mode changes phase, and was represented in this study as the ratio of the nodal point to the thickness of the vocal fold. The amplitude of vibration was

defined in this study as the amplitude of the right surface relative to the left, and phase refers to the difference in timing between the lateral movements of the two vocal fold surfaces. The parameter settings used to generate each configuration are shown in Table I, and were selected based on a previous study demonstrating that similar settings simulate increasingly impaired voice qualities (Samlan, 2012). The fundamental frequency (F_0) was set at 100 Hz for all simulations and did not vary over the duration of the vowel. Each of the four sets of vocal fold parameters was combined with all five epilaryngeal conditions to create a set of 20 stimuli.

2. Vocal function analysis

Measures of minimum glottal flow, maximum flow, and MFDR were made on each simulated voice sample to assess variations in airflow and MFDR with changes in epilaryngeal cavity shape. Low frequency spectral slope was assessed using $H1^* - H2^*$ (where the asterisks indicate that harmonic amplitudes were corrected for the effect of vowel formants; Iseli and Alwan, 2004), measured from the output pressure signal with VoiceSauce (Shue et al., 2011). High frequency spectral slope (B0 - B2) was measured by subtracting the mean spectral amplitude in the range 2000–5000 Hz (B2) from the mean amplitude from 60–400 Hz (B0) (de Krom, 1995; Hartl et al., 2003). Cepstral peak prominence (CPP; Hillenbrand et al., 1994) and the harmonics-to-noise ratio for 0–3500 Hz (HNR) were measured from the output pressure signal using VoiceSauce, to determine whether constriction caused increased prominence of harmonic energy and decreased prominence of noise.

3. Perceptual task

To determine whether listeners heard differences in the acoustic stimuli resulting from these manipulations, 14 normal-hearing listeners completed a sort and rate task (Granqvist, 2003) in which the 5 /a/ vowels for a single glottal configuration were represented on a computer screen as differently shaped and colored icons. Stimuli were presented over Etymotic ER-1 insert earphones (Etymotic Research, Inc., Elk Grove Village, IL), which simulate a free field listening environment. Listeners played the vowels by clicking each icon and then dragged the icons to a line on the screen so that the physical distance between the icons on the line represented how similar or different the vowels sounded (farther apart corresponded to a larger difference in quality; vowels that sounded exactly the same were to be stacked). Each glottal configuration (normal, mild, moderate, severe paralysis) was presented in a different screen so that listeners completed four trials, each

including the five samples (neutral plus four epilaryngeal shape manipulations) for one glottal configuration. The order in which listeners completed the trials was randomized. Testing was combined with the tasks for Experiment 2; the combined tests lasted between 10 and 20 min. All procedures were approved by the UCLA Institutional Review Board.

B. Results

1. Vocal function analysis

Figure 2 shows the glottal flow and flow spectrum for each epilaryngeal shape paired with the normal laryngeal configuration, and results for individual acoustic measures of the stimuli are shown in Fig. 3. In this section, we describe how each measure varied across epilaryngeal shapes, relative to the neutral epilarynx condition, with changing glottal configuration. In the neutral condition, not surprisingly, minimum and maximum glottal flow increased with increasing severity of paralysis, while MFDR, B0–B2, CPP, and HNR all decreased; $H1^* - H2^*$ first increased, and then decreased for severe paralysis. Altering epilaryngeal area at the level of the aryepiglottic folds influenced glottal flow and acoustic measures relative to this pattern more than altering area at the ventricular folds, but the specific effects at both locations depended on the glottal configuration. Note that, because measurements were obtained from modeling, all reported differences are “reliable” in the sense that identical parameter sets will always produce identical results. Thus, statistical analysis is not relevant (because there is no within-category variability), and the importance of the observed differences derives from their perceptual salience. We turn to this topic in Sec. II B 2.

a. Aryepiglottic fold expansion and constriction. Generally speaking, effects on vocal function of epilaryngeal expansion at the level of the aryepiglottic folds were smaller than and opposite to those of aryepiglottic constriction, although the precise effects depended somewhat on glottal configuration. As expected, maximum glottal flow increased with expansion and decreased with constriction; but, contrary to expectations, MFDR also increased with expansion and decreased with constriction. Minimum flow was affected only in the context of a severe glottal paralysis when it increased with expansion and decreased with constriction. When paralysis was present, $H1^* - H2^*$ increased with aryepiglottic expansion. With aryepiglottic constriction $H1^* - H2^*$ increased in the mild paralysis simulation and decreased when paralysis was moderate or severe. B0–B2 increased with constriction (reflecting decreased high frequency energy), but was unchanged with expansion. In

TABLE I. Settings of the kinematic vocal fold model parameters for the four glottal configurations.

Glottal configuration	Adduction (cm)		Bulging (cm)		Nodal point ratio		Amplitude <i>R</i> (relative to <i>L</i>)	Phase (radians)	
	<i>R</i>	<i>L</i>	<i>R</i>	<i>L</i>	<i>R</i>	<i>L</i>		<i>R</i>	<i>L</i>
Normal	0.08	0.08	0.1	0.1	0.8	0.8	1	0	0
Mild paralysis	0.1	0.1	0.1	0	0.8	0.6	1	0	1.4
Moderate paralysis	0.1	0.2	0.1	0	0.8	0.4	1	0	1.8
Severe paralysis	0.1	0.3	0.1	0	0.8	0.4	0.5	0	1.8

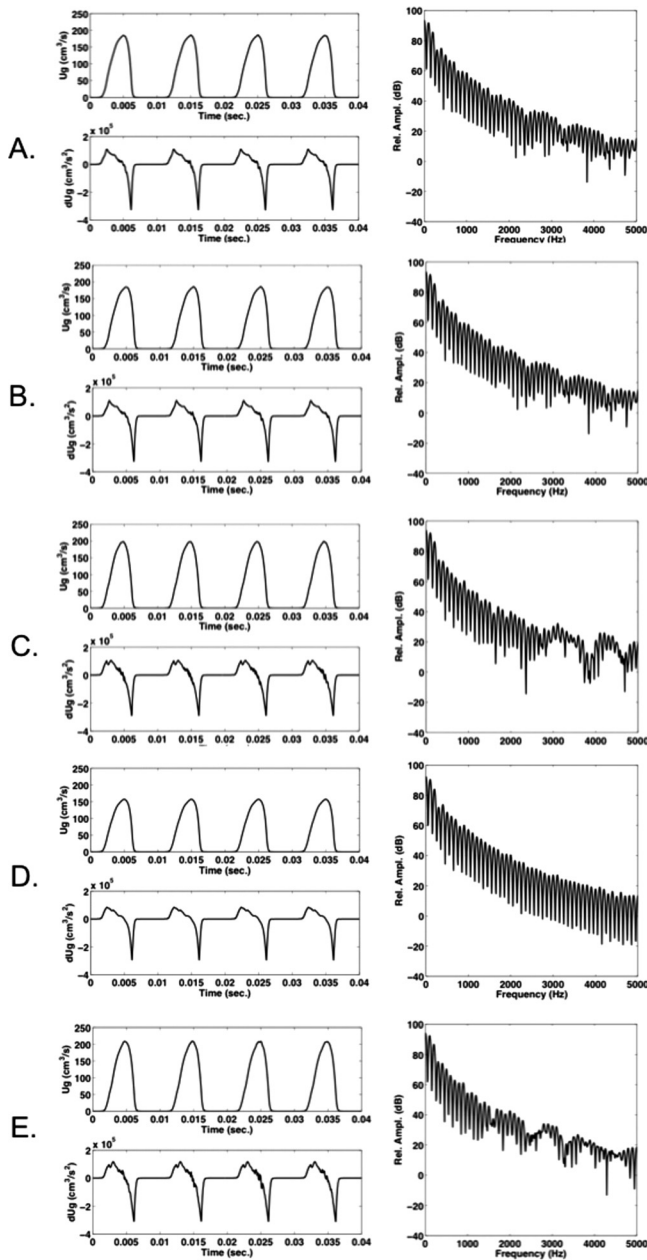


FIG. 2. Glottal flow in top left panel, flow derivative in bottom left panel, and spectrum of the glottal flow in the right panel for the normal laryngeal configuration and (a) neutral epilarynx, (b) ventricular fold constriction, (c) ventricular fold expansion, (d) aryepiglottic fold constriction, and (e) aryepiglottic fold expansion.

the absence of paralysis, however, B0–B2 increased with expansion and decreased with constriction, while H1*–H2* was unchanged by expansion and increased with constriction. CPP and HNR showed the most complex interactions with glottal configuration. For a normal glottis or mild paralysis, both measures decreased with aryepiglottic expansion, while for moderate or severe paralysis, both were unchanged relative to the neutral epilaryngeal configuration. Aryepiglottic constriction resulted in increases in CPP and HNR relative to a neutral glottis except when paralysis was severe, in which case no change was observed.

b. Ventricular fold expansion and constriction. In contrast to the modifications at the level of the aryepiglottic

folds, much smaller changes were measured in the voice samples associated with epilaryngeal modification at the level of the ventricular folds. Ventricular fold expansion did not affect minimum flow, but increased maximum flow relative to the neutral epilarynx. Ventricular fold expansion lowered MFDR slightly in the normal and mild paralysis conditions while increasing it for the moderate and severe conditions. This condition resulted in an H1*–H2* decrease and an increase in B0–B2 in the normal and mild paralysis glottal configurations, and in increases in both H1*–H2* and B0–B2 in the moderate and severe paralysis configurations. Decreased noise (i.e., higher HNR) was evident with ventricular expansion for the three paralyzed glottal configurations, although CPP did not change.

Epilaryngeal constriction at the level of the ventricular folds did not affect minimum flow, and decreased maximum flow and MFDR relative to the neutral epilarynx for all glottal configurations. Ventricular fold constriction increased H1*–H2* for the normal and mild paralysis conditions, and decreased it for the moderate and severe paralysis conditions. B0–B2 decreased with ventricular fold constriction in the normal and severe paralysis conditions, and increased with constriction in the mild and moderate paralysis conditions. Ventricular fold constriction led to noise increase (lower HNR) for moderate and severe paralysis. There were no differences in CPP.

2. Perceptual analyses

To determine the perceptual importance of differences in epilaryngeal cavity shape for each glottal configuration, scores from the sort and rate task were analyzed via separate three-way (individual differences) multidimensional scaling (MDS) analyses. The MDS solutions explained listener variability well: R^2 for the complete group of listeners was 0.57 for the normal glottal configuration, 0.71 for mild paralysis, 0.75 for moderate paralysis, and 0.72 for severe paralysis. Figure 4 shows the scaling solutions for the normal, mildly, moderately, and severely paralyzed glottal configurations. Relative to the neutral vocal tract configuration (plotted with a circle), the perceptual importance of an epilaryngeal expansion (“x” and inverted triangle) depended on glottal configuration context, but not on the location of the expansion; the acoustic characteristics generated by expansions were easier to hear for the normal or mildly paralyzed glottal configurations, independent of the location of the epilaryngeal expansion, and harder to hear as the overall extent of modeled paralysis worsened. In contrast, the perceptual importance of a constriction (star and upright triangle) relative to the neutral configuration depended on both its location and the glottal configuration; aryepiglottic constrictions always caused perceptible changes in quality, but ventricular constrictions had little or no perceptual impact when glottal configuration reflected a normal glottis or mild paralysis (although their perceptual importance increased for moderate to severe modeled paralyzes).

C. Discussion

Modifying the epilaryngeal area caused perceivable differences in voice quality for both normal and disordered

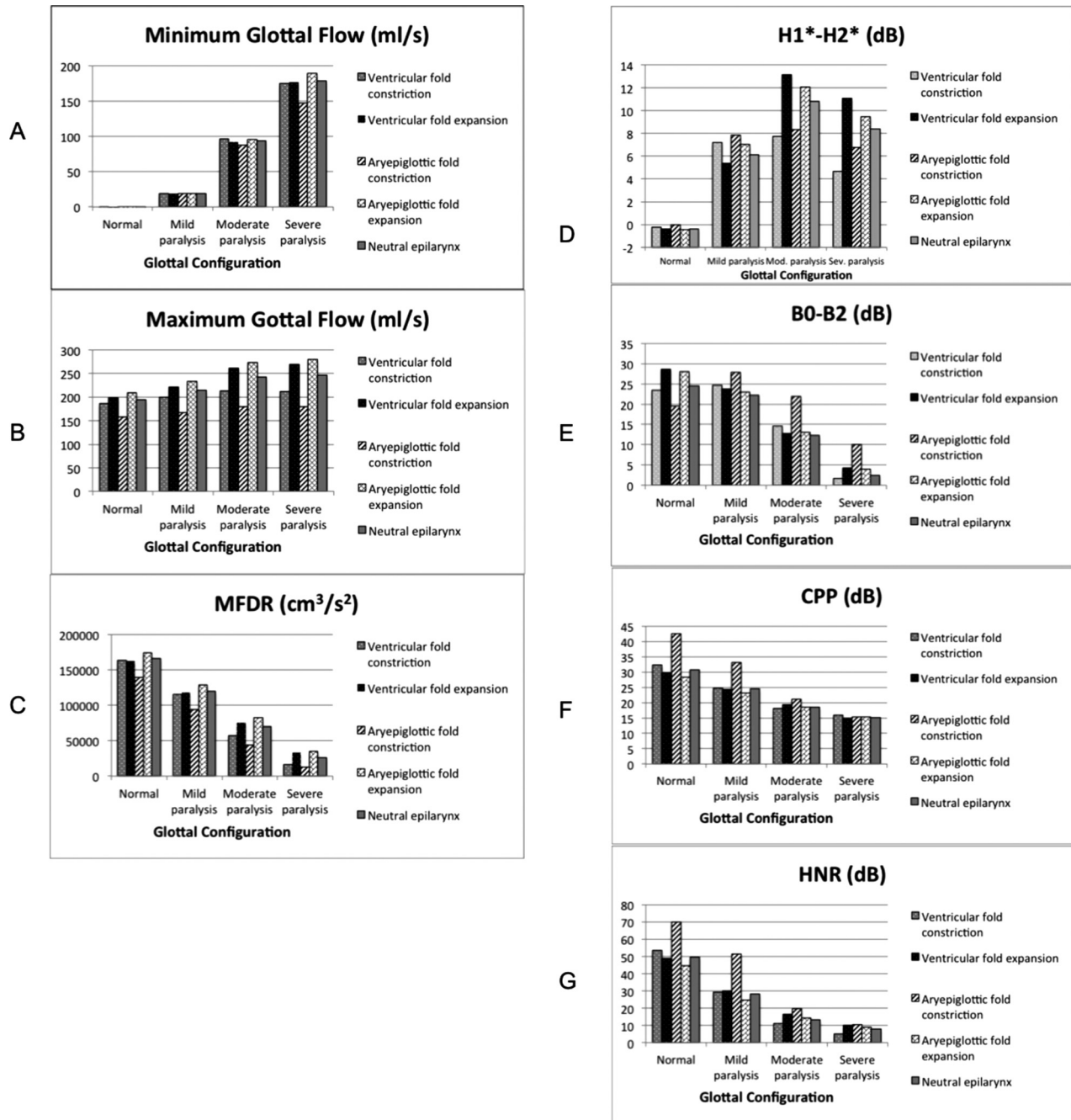


FIG. 3. (a) Minimum glottal flow, (b) maximum glottal flow, (c) MFDR, (d) $H1^*-H2^*$, (e) B0-B2, (f) CPP, and (g) HNR for each epilyrnygeal cavity shape for the four glottal configurations.

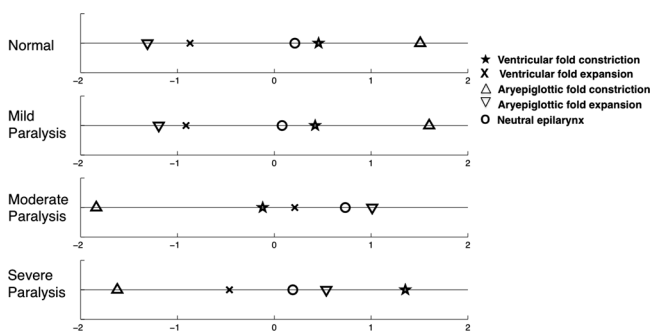


FIG. 4. MDS solutions for (a) normal, (b) mild paralysis, (c) moderate paralysis, and (d) severe paralysis.

voices. Further, listeners perceived each of the location and area modifications differently, indicating that both kinds of change affected voice quality.

Within a particular glottal configuration, the vowel produced with aryepiglottic fold constriction was always perceived as most distinct from the other stimuli. Although listeners did not assess voices on a “better” or “worse” basis, the acoustic and aerodynamic effects of these constrictions (reduced minimum and maximum glottal flow for all voice qualities, and increased CPP and HNR for normal, mildly, and moderately impaired voices) are all associated with improved voice quality. Not all measures for epilyrnygeal constriction were consistent with improved voice quality;

MFDR increased and measures of spectral slope were inconsistent across glottal configuration. Aryepiglottic constriction caused inconsistent changes in low frequency spectral slope among the five epilaryngeal configurations. Higher frequency spectral slope, as assessed using B0–B2, was flatter with aryepiglottic fold constriction for the normal laryngeal configuration and steeper for all three disordered configurations. Together, the results indicate that aryepiglottic fold constriction did not increase harmonic amplitudes in the disordered voices, yet decreased noise generation secondary to decreased airflow. Note that the decrease in minimum flow with aryepiglottic constriction occurred even though there was no antecedent decrease in minimum glottal area, demonstrating the effect of vocal tract interaction with glottal area in common vocal function measures.

Aerodynamic and acoustic measures of ventricular fold and aryepiglottic fold constriction typically differed from the neutral epilarynx in the same direction, though with differing magnitudes. While these data do not provide evidence that ventricular fold constriction worsens voice quality in disordered voices, neither do they support ventricular fold constriction as effective compensation or improvement for disordered voices.

These results are consistent with the hypothesis that the severe paralysis configuration leads to large source-filter interactions. More changes in glottal flow occurred as a result of this coupling than in the normal or mildly impaired glottal configuration, and spectral slope continued to vary with epilaryngeal shape. CPP and HNR differences with epilaryngeal constriction, however, were much smaller in the case of severe paralysis than for the normal or mild glottal configurations. This is presumably because the lack of closure resulted in fewer, lower amplitude harmonics, and the gap created higher noise levels that masked any gain in harmonic energy contributed by the downstream constriction (Shrivastav and Sapienza, 2003; Kreiman and Gerratt, 2012).²

The finding that poor glottal closure mitigates functional improvements resulting from increased aryepiglottic constriction serves as a reminder that clinicians must exercise caution when using information about how a technique affects an essentially normal voice to motivate therapy for disordered voices. Based on the current study, limited improvement in voice quality would be expected from epilaryngeal constriction (and, thus, from resonant voice therapy or semi-occluded voice therapy) in patients with moderate or severe vocal paralysis. This is not to say that voice therapy cannot improve voice quality in such situations. In this study, we did not assess whether other mechanisms might facilitate voice quality improvement in patients with moderate or severe paralysis, such as “level 2” source–vocal tract interaction (whereby pressures above and below the glottis influence vocal fold vibration; Titze, 2008), or altering multiple levels of production simultaneously.

Contrary to theoretical expectations and assumptions underlying resonant voice and semi-occluded vocal tract therapy, MFDR decreased from the neutral values with epilaryngeal constriction, and the presumed effect of higher MFDR (i.e., increased amplitude of high frequency harmonics) did not occur. It is likely that the MFDR decrease with

constriction occurred because F_0 was fixed at 100 Hz for these voices and, by definition, period length was also constant. MFDR, for this particular kinematic model, was primarily dependent on the amplitude of the flow waveform, and the amplitude decreased with aryepiglottic and ventricular fold constriction.

The time varying glottal area was identical within a glottal configuration and the pharyngeal and oral cavity vocal tract area functions were held constant. The differences in glottal flow, the output pressure spectrum, and voice quality were, therefore, fully the result of changes to the epilaryngeal cavity shape, demonstrating that not all changes in what we consider the “voice source” (i.e., glottal flow) should be attributed to properties of vocal fold vibration. We extended our investigation of this kind of interaction to the supraglottal configuration in Experiment 2.

III. EXPERIMENT 2

A. Methods

The five epilaryngeal cavity shapes and the vocal fold settings for normal voice quality described for Experiment 1 were also used in Experiment 2. For comparison to the /a/ vowels generated in Experiment 1, the vocal tract area function downstream from the epilarynx was set to produce /i/ and /u/ vowels, based on Story’s (2008) area functions. The three area functions are shown in Fig. 5. The epilaryngeal area at the ventricular fold segment was 0.28 cm^2 for the neutral /i/ and 0.37 cm^2 for the neutral /u/. Epilaryngeal area at the aryepiglottic fold segment was 0.87 cm^2 for the neutral /i/ and 1.55 cm^2 for the neutral /u/. Note that the area at the ventricular fold segment for /a/ (0.36 cm^2) was close to that for /u/ and aryepiglottic segment (0.62 cm^2) closer to that for /i/. Epilaryngeal area at the level of the aryepiglottic folds for the neutral /u/ was larger than the setting for the expanded condition (1.2 cm^2). Thus, though the term “expanded” is still used in the descriptions that follow, the “expanded” setting for the /u/ vowel represents a constriction from the Magnetic Resonance Imaging-based area function used for the “neutral” /a/ in Experiment 1.

The vocal function analyses (minimum glottal flow, maximum flow, MFDR, CPP, and HNR) and perceptual methods were identical to those used in Experiment 1.

B. Results and discussion

1. Vocal function analysis

Consistent with the results of Experiment 1, with a normal glottal configuration, minimum glottal flow was always $\leq 0 \text{ ml/s}$ for all vowels. Maximum glottal flow and MFDR are shown in Figs. 6(a) and 6(b). Across epilaryngeal shapes, the highest maximum flows occurred for the vowel /u/, which had the largest epilaryngeal area across all sections. Within the /a/ and /i/ vowels, the highest maximum flows occurred for the two expanded epilaryngeal conditions and the lowest maximum flow occurred with aryepiglottic fold constriction for all three vowels. Aryepiglottic fold constriction generated the highest $H1^* - H2^*$ values and lowest B0–B2 for /a/ and /i/. MFDR, within a vowel, showed the

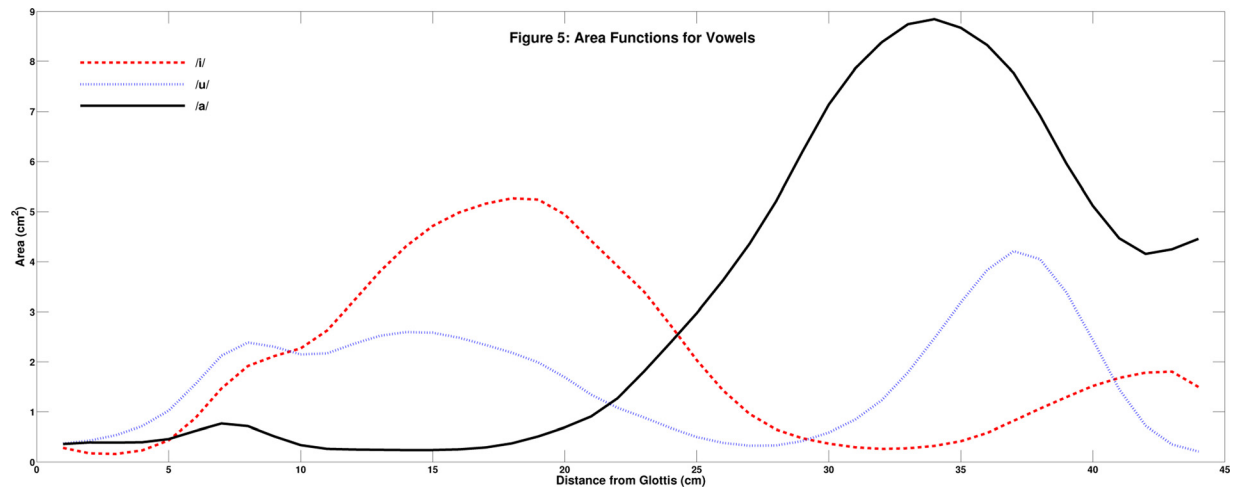


FIG. 5. (Color online) Vocal tract area as a function for distance from the glottis for /i/, /u/, and /a/ vowels.

same pattern of changes as did maximum glottal flow and, across vowels, the highest MFDR values occurred for the /u/ vowels.

As in the asymmetric glottal configurations, the lowest noise (highest CPP and HNR) for each of the vowels occurred with aryepiglottic constriction. For /a/ and /i/, aryepiglottic fold expansion increased noise relative to the neutral epilarynx, while the opposite occurred for the /u/ vowel, where the “expanded” condition was a constriction from neutral. The /u/ vowels generated the most spectral noise: CPP was lowest for /u/ across epilaryngeal shapes [Fig. 6(e)] and HNR was lowest for /u/ in all conditions except ventricular fold expansion, where it was lowest for /i/ [Fig. 6(f)].

2. Perceptual analysis

For the /i/ and /u/ vowels, one-dimensional MDS solutions explained 75% of the variance in the underlying data. Figures 7(a), 7(b), and 7(c) show these solutions for the /a/, /i/, and /u/ vowels, respectively. As for the disordered glottal configurations [Figs. 3(b)–3(d)], the production generated with aryepiglottic constriction (plotted with an upright triangle) was distinctly different from the others for all three vowels. MDS solutions were quite similar to one another for the /a/ and /i/ vowels, with the ventricular fold constriction condition (star) most similar to that of the neutral epilarynx (circle) and the two expanded epilaryngeal shapes (“x” and inverted triangle) most similar to one another. The similarity grouping was slightly different for the /u/ vowel in that the aryepiglottic expansion was more similar to the neutral epilarynx than to ventricular expansion and more similar to neutral than was ventricular constriction. The difference in grouping with the neutral condition is consistent with the similarity of the epilaryngeal area at the aryepiglottic folds in the neutral and expanded conditions in the /u/ vowel.

Note that /a/ and /i/ (which were more similar in area at the level of the aryepiglottic folds) were more similar in vocal function and in how listeners perceived differences in epilaryngeal shape modification than either was to the /u/ vowel. The difference in /u/ was the large epilaryngeal area at the level of the aryepiglottic folds, consistent with the

Experiment 1 findings that changes in aryepiglottic fold area created more perceptually salient differences in the acoustic signal than did ventricular fold area changes.

IV. GENERAL DISCUSSION

We expected that any epilaryngeal constriction would limit maximum glottal flow and enhance harmonic structure, while expansion would increase maximum glottal flow and enhance spectral levels of noise. Simulation findings generally fit these expectations regarding glottal flow and spectral noise levels, particularly for /a/ and /i/ with normal voice or mild paralysis. The effect of constriction was perceptually, aerodynamically, and acoustically larger for the aryepiglottic folds than for the ventricular folds, with the exception that the HNR was lowest for ventricular constriction in the context of a moderate or severe vocal fold paralysis. Effects were limited for the /u/ vowel, likely because /u/ has a larger underlying epilaryngeal area than either /a/ or /i/. In the normal laryngeal configuration, B0–B2 did indeed decrease with constriction in both locations, though it typically increased with constriction for the disordered voices. The other spectral slope measure did not show as clear of a pattern in response to constriction at either location.

Beyond their role in artistic effects in singing (Yanagisawa *et al.*, 1989; Estill, 1997; Titze, 2001), epilaryngeal cavity shape changes are of interest because of their suspected role as an underlying mechanism in resonant voice therapies (Titze, 2001; Titze and Verdolini Abbott, 2012). Increases in CPP and HNR were evident with aryepiglottic constriction in the mild paralysis case, but less so in the moderate and severe paralysis cases. Even though the acoustic changes were small when paired with the severe paralysis glottal configuration, cases with aryepiglottic fold constriction were distanced from the other samples in the sort and rate tasks. Further study is indicated to better understand the conditions under which such constriction will be beneficial. It is also important to recognize that endoscopic evidence of epilaryngeal constriction as an underlying mechanism for resonant voice therapies is equivocal. Endoscopic findings have

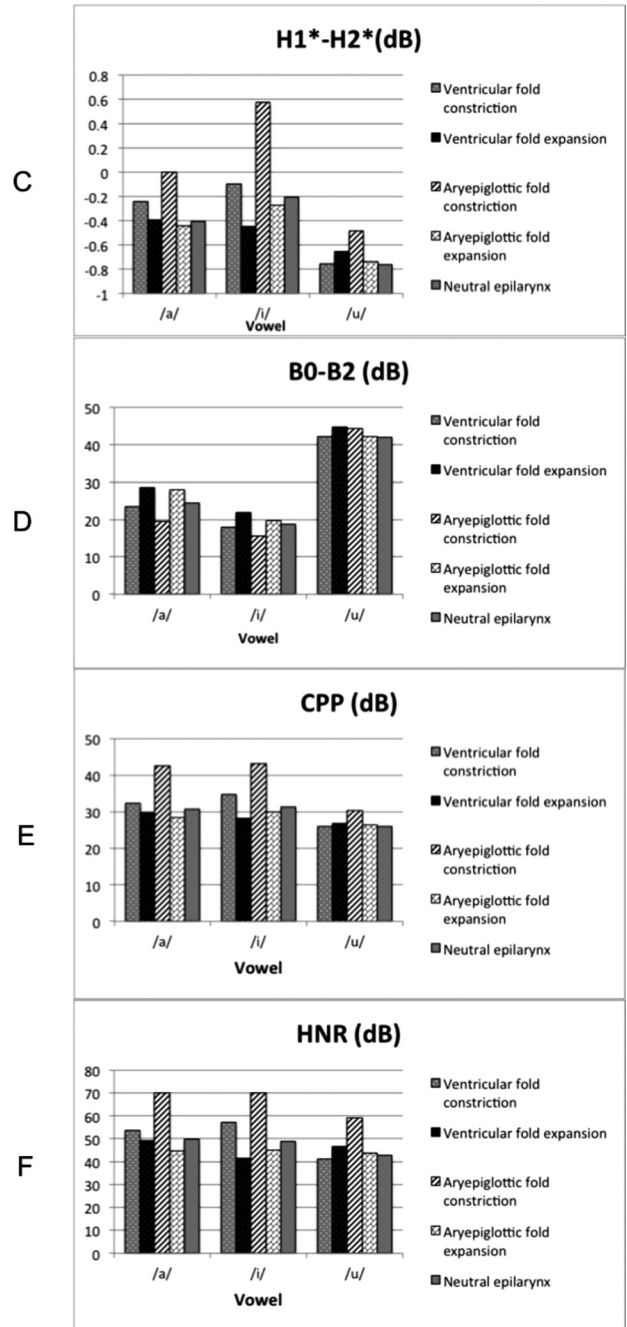
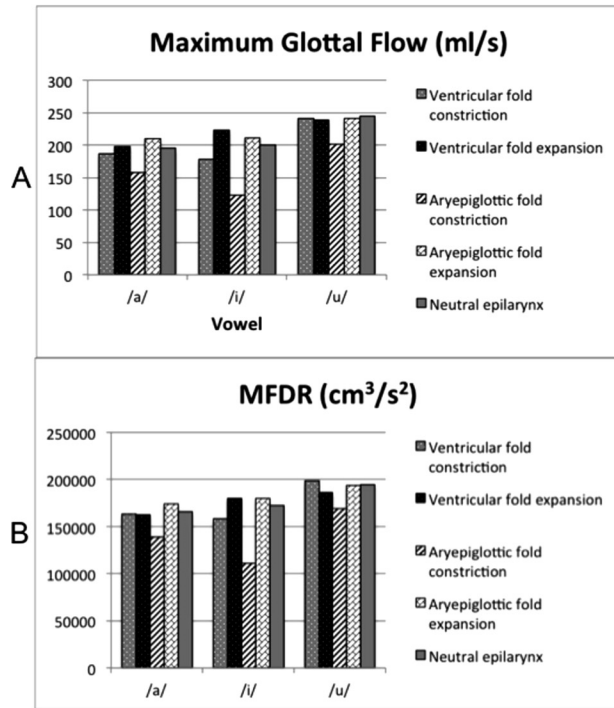


FIG. 6. (a) Maximum glottal flow, (b) MFDR, (c) H1*-H2*, (d) B0-B2, (e) CPP, and (f) HNR for each epilaryngeal cavity shape for the three vowels using the “normal” glottal configuration.

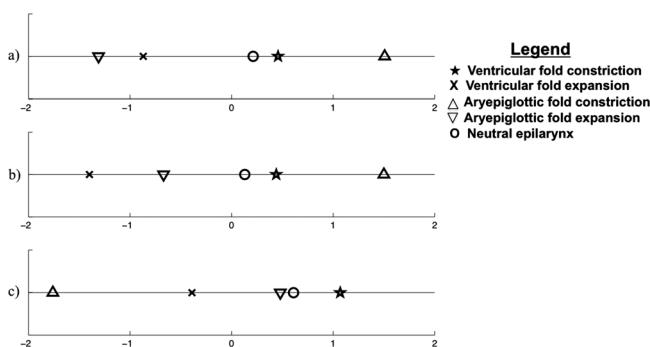


FIG. 7. MDS solutions for (a) normal /a/, (b) normal /i/, and (c) normal /u/.

demonstrated aryepiglottic constriction for twang, belting, and opera singing voice qualities (Yanagisawa *et al.*, 1989), yet constriction was not observed during resonant spoken voice productions for six speakers (Smith *et al.*, 2005).

V. CONCLUSIONS

To return to the original questions:

- (1) Epilaryngeal constriction and expansion caused perceivable differences in voice quality, and listeners readily perceived differences between the constriction locations.
- (2) Vocal function measures demonstrated that constriction at the aryepiglottic location generated vowels with lower maximum flow and less noise.

- (3) The vowels simulated with aryepiglottic constriction were perceived as different than those with other epilaryngeal shapes in normal and disordered voices.
- (4) Acoustic differences with aryepiglottic fold constriction were minimal for severely disordered voices.
- (5) Vocal function and perceptual results from epilaryngeal area constriction and expansion were similar for /a/ and /i/. Differences among shapes were smaller for /u/ and followed a different pattern.

ACKNOWLEDGMENTS

Thank you to Brad Story for providing the model used in the study and sharing insights regarding parameter selection and data interpretation and to Bruce Gerratt for his helpful comments on earlier versions of this manuscript. This work was supported by research Grant No. DC01797 from the National Institute of Health/National Institute on Deafness and Other Communication Disorders (NIH/NIDCD).

¹The wave propagation algorithm used in the model (Liljencrants, 1985; Story, 1995) led to a segment length (Δ) = $c/2(F_s)$ where c = speed of sound = 3500 cm/s and F_s = sampling frequency = 44 100 Hz. Therefore, Δ = 0.396825 cm (Story, 2005).

²It should be noted that changing the epilaryngeal configuration, even with glottal and vocal tract areas downstream from the epilarynx remaining constant, might also influence formant bandwidths.

Bergan, C. C., Titze, I. R., and Story, B. (2004). "The perception of two vocal qualities in a synthesized vocal utterance: ring and pressed voice," *J. Voice* **18**, 305–317.

de Krom, G. (1995). "Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments," *J. Speech Hear. Res.* **38**, 794–811.

Edmondson, J. A., and Esling, J. H., (2006). "The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies," *Phonology*, **23**, 157–191.

Estill, J. (1997). *Level One: Primer of Basic Figures* (Estill Voice Training Systems, Santa Rosa, CA), pp. 55–63.

Fant, G. (1986). "Glottal flow: Models and interaction," *J. Phonetics* **14**, 393–399.

Granqvist, S. (2003). "The visual sort and rate method for perceptual evaluation in listening tests," *Logoped. Phoniatr. Vocol.* **28**, 109–116.

Hartl, D. M., Hans, S., Vaissière, J., and Brasnu, D. F. (2003). "Objective acoustic and aerodynamic measures of breathiness in paralytic dysphonia," *Eur. Arch. Oto-Rhino Laryngol.* **260**, 175–182.

Hillenbrand, J., Cleveland, R., and Erickson, R. (1994). "Acoustic correlates of breathy vocal quality," *J. Speech Hear. Res.* **37**, 769–778.

Iseli, M., and Alwan, A. (2004). "An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient

estimation," in *Proceedings of ICASSP*, Montreal (May 2004), pp. 669–672.

Kreiman, J., and Gerratt, B. R. (2012). "Perceptual interaction of the harmonic source and noise in voice," *J. Acoust. Soc. Am.* **131**, 492–500.

Liljencrants, J. (1985). "Speech synthesis with a reflection-type line analog," Doctoral dissertation, Royal Institute of Technology, Stockholm, Sweden.

Morrison, M. D., Nichol, H., and Rammage, L. A. (1986). "Diagnostic criteria in functional dysphonia," *Laryngoscope* **94**, 1–8.

Rothenberg, M. (1981). "Acoustic interaction between the glottal source and the vocal tract," in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo), pp. 305–328.

Samlan, R. A. (2012). "Kinematic modeling of asymmetric vocal fold vibration," Ph.D. dissertation, University of Arizona, Tucson, AZ.

Samlan, R. A., Story, B. H., and Bunton, K. (2013). "Relation of perceived breathiness to laryngeal kinematics and acoustic measures based on computational modeling," *J. Speech Hear. Res.* **56**, 1209–1223.

Shrivastav, R., and Sapienza, C. M. (2003). "Objective measures of breathy voice quality obtained using an auditory model," *J. Acoust. Soc. Am.* **114**, 2217–2224.

Shue, Y., Keating, P., Vicenik, C., and Yu, K. (2011). "VoiceSauce: A program for voice analysis," *Proc. ICPhS* **17**, 1846–1849.

Smith, C. G., Finnegan, E. M., and Karnell, M. P. (2005). "Resonant voice: Spectral and nasoendoscopic analysis," *J. Voice* **19**, 607–622.

Story, B. H. (1995). "Speech simulation with an enhanced wave-reflection model of the vocal tract," Ph.D. dissertation, University of Iowa, Iowa City, IA.

Story, B. H. (2005). "A parametric model of the vocal tract area function for vowel and consonant simulation," *J. Acoust. Soc. Am.* **117**, 3231–3254.

Story, B. H. (2008). "Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002," *J. Acoust. Soc. Am.* **123**, 327–335.

Story, B. H., Titze, I. R., and Hoffman, E. A. (2001). "The relationship of vocal tract shape to three voice qualities," *J. Acoust. Soc. Am.* **109**, 1651–1667.

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

Titze, I. R. (1984). "Parameterization of the glottal area, glottal flow, and vocal fold contact area," *J. Acoust. Soc. Am.* **75**, 570–580.

Titze, I. R. (2001). "Acoustic interpretation of resonant voice," *J. Voice* **15**, 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**, 448–459.

Titze, I. R. (2006b). *The myoelastic aerodynamic theory of phonation* (National Center for Voice and Speech, Iowa City, IA), pp. 197–214.

Titze, I. R. (2008). "Nonlinear source-filter coupling in phonation: Theory," *J. Acoust. Soc. Am.* **123**, 2733–2749.

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.

Titze, I. R., and Verdolini Abbott, K. (2012). *Vocology: The Science and Practice of Voice Habilitation* (National Center for Voice and Speech, Salt Lake City, UT), pp. 286–310.

Yanagisawa, E., Estill, J., Kmucha, S. T., and Leder, S. B. (1989). "The contribution of aryepiglottic constriction to 'ringing' voice quality—A video-laryngoscopic study with acoustic analysis," *J. Voice* **3**, 342–350.