JSLHR

Research Article

The Quantal Larynx: The Stable Regions of Laryngeal Biomechanics and Implications for Speech Production

Scott Reid Moisik[a,b] and Bryan Gick[c,d]

Purpose: Recent proposals suggest that (a) the high dimensionality of speech motor control may be reduced via modular neuromuscular organization that takes advantage of intrinsic biomechanical regions of stability and (b) computational modeling provides a means to study whether and how such modularization works. In this study, the focus is on the larynx, a structure that is fundamental to speech production because of its role in phonation and numerous articulatory functions.
Method: A 3-dimensional model of the larynx was created using the ArtiSynth platform (http://www.artisynth.org). This model was used to simulate laryngeal articulatory states, including inspiration, glottal fricative, modal prephonation, plain glottal stop, vocal–ventricular stop, and aryepiglotto–epiglottal stop and fricative.
Results: Speech-relevant laryngeal biomechanics is rich with "quantal" or highly stable regions within muscle activation space.
Conclusions: Quantal laryngeal biomechanics complement a modular view of speech control and have implications for the articulatory–biomechanical grounding of numerous phonetic and phonological phenomena.

The term *quantal* has often been applied to a subset of nonlinear effects in speech—traditionally those that underlie robust auditory–perceptual goals (e.g., Stevens, 1972, 1989; Stevens & Keyser, 2010). The present article uses computational biomechanical modeling to examine whether and to what extent different postures of the larynx exhibit quantal-like nonlinear behavior in biomechanical space. It is hoped that this examination of quantal biomechanics of the larynx will simultaneously (a) provide insight into the nature of speech motor control, particularly with respect to the larynx, and (b) aid in understanding the factors that shape the emergence and organization of speech sound systems more generally. *Quantality* and the closely related concept of *saturation* in biomechanics (see, e.g., Perkell, 2012) are terms that have been used to describe aspects of biomechanical robustness. In understanding the central role of quantal biomechanics in motor control and the emergence of speech sounds, it is necessary to consider the importance of biomechanical robustness in dimensionality reduction of motor systems.

The human vocal tract is endowed with seemingly innumerable degrees of freedom, raising the question of how a finite nervous system copes with the task of generating movement (e.g., Bernstein, 1967). A large and growing body of work in neurophysiology and other fields supports the long-standing notion that the human nervous system reduces dimensionality of these many degrees of freedom by using a "library" of neuromuscular modules (for a review, see d'Avella, Giese, Ivanenko, Schack, & Flash, 2015; Safavynia & Ting, 2012; Ting et al., 2015), each built to perform a specific function. Modularization, broadly speaking, is the solution Bernstein himself proposed, and it remains one that has shown continued success in explaining how nervous systems can deal with the degrees of freedom problem (e.g., Berger, Gentner, Edmunds, Pai, & d'Avella, 2013). Gick and Stavness (2013) proposed a model for speech production incorporating embodied modules of this kind as speech movement primitives and argued that biomechanical modeling will be essential in revealing such structures. This approach proposes that widely attested speech movements are the outputs

[a]Division of Linguistics and Multilingual Studies, Nanyang Technological University, Singapore
[b]The Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
[c]Department of Linguistics, University of British Columbia, Vancouver, Canada
[d]Haskins Laboratories, New Haven, CT
Correspondence to Scott Reid Moisik: scott.moisik@ntu.edu.sg

Editor: Julie Liss
Associate Editor: Nelson Roy
Received January 14, 2016
Revision received July 18, 2016
Accepted August 28, 2016
https://doi.org/10.1044/2016_JSLHR-S-16-0019

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

of discrete, functionally independent neuromuscular modules, selected for use in speech because they take advantage of biomechanical properties intrinsic in specific body structures.

Because any group of muscles could in principle be combined into such a module, it becomes essential to consider the question of which particular groupings are likely to emerge and why. It is natural to assume that those muscle groupings will be selected that correspond with the most effective body structures for reliably performing important tasks. Key properties of optimally effective motor structures have been identified through computational approaches to motor control, such as that outlined in Todorov and Jordan's (2003) minimal intervention principle. A logical implication of minimal intervention is that there should be an inverse relationship between the need for intervention (i.e., correction of the feed-forward command set; note that *feed-forward* here refers to operation without correction that is based on immediate sensory feedback) and the range of error allowable in achieving successful task performance. That is, all else being equal, a motor system should always prefer structures that require less intervention to achieve their tasks, such that the operation of a perfectly optimal system would be entirely feed-forward. It is expected that body structures selected for speech would be ones that can achieve their tasks even under noisy everyday operating conditions; such structures should allow a large range of error, optimizing for feed-forward function. Speech production mechanisms should thus exhibit "quantal" properties of this kind, where a quantal region in some function is a region in which large variation (error) in one dimension produces little response in some other (task) dimension. From the nervous system's point of view, because body structures with this property will be more likely to succeed in producing reliable outcomes, such structures are more likely to be selected for repeated use, reinforcing neural pathways that lead to these mechanically robust muscle groupings.

Given that optimal control should always favor modules that exhibit quantal effects (biomechanical or otherwise), such effects should be pervasive in the modules used in speech motor control. Indeed, quantal behavior in the biomechanical–articulatory domain of speech has long been presumed to be an important factor governing speech sound production and thought to play a key role in shaping the types of sounds found in language (Fujimura, 1989; Schwartz, Boë, Vallée, & Abry, 1997; Stevens, 1989). Recent modeling attempts have begun to show evidence of such biomechanical quantality in supralaryngeal subsystems (Buchaillard, Perrier, & Payan, 2009; Gick et al., 2014; Gick, Stavness, Chiu, & Fels, 2011; Nazari, Perrier, Chabanas, & Payan, 2011). Although some three-dimensional models of the larynx do exist (e.g., Hunter, Titze, & Alipour, 2004; Moisik, 2008; Moisik & Gick, 2013), the larynx has not yet been examined in the context of a biomechanical model sophisticated enough to demonstrate quantality in laryngeal articulation. The present work attempts to achieve this goal by simulating a range of widely attested laryngeal states in a substantially improved model that is based on the model described by Moisik and Gick (2013).

The present article is not the first to consider quantal aspects of laryngeal articulation. Stevens (1989, pp. 27–28) speculated that the vocal fold abduction–adduction continuum is quantized into breathy, modal, and pressed phonatory states. Stevens' exploration of laryngeal quantal effects is revisited here, but the focus of the present work is on biomechanical–articulatory relations. In addition to testing the quantal proposal in the context of a new and highly realistic model of laryngeal biomechanics, the present study goes well beyond previous approaches in the following respects: (a) by giving full consideration to the role played by the supraglottal portion of the larynx, thus taking a whole-larynx approach to laryngeal phonetics (Moisik & Esling, 2011); (b) by offering qualitative validation of the model by comparison with representative laryngoscopic images; (c) by providing a means for quantifying quantality using a numeric index, enabling a more objective characterization of quantal effects; and (d) by linking quantality to a modular approach to neuromuscular organization (e.g., Gick & Stavness, 2013), predicting that each laryngeal posture needed for successful speech may be generated by varying a single parameter—the activation of an appropriate fixed ratio of muscles—and that each such combination will show evidence of robust output over a wide range of muscle activation levels, thus rooting quantality in biomechanical and computational principles.

Method

Biomechanical Model of the Larynx

The biomechanical model of the larynx described here, called QL2, was developed using the ArtiSynth biomechanical modeling toolkit (http://www.artisynth.org; e.g., Lloyd, Stavness, & Fels, 2012). The model was designed to replicate larynx anatomy as accurately as possible. Any operative properties the model may have are a direct function of the structures themselves. Note that this model simulates only laryngeal biomechanics, not aerodynamics or acoustics. A precursor to this model (QL1) appeared in Moisik and Gick (2013), but QL2 has undergone substantial changes and improvements from this earlier state. QL2 consists of a three-dimensional fully hexahedral finite-element model (FEM) mesh representing the mucosa and soft tissues of the larynx (see Figure 1) and rigid bodies for the cartilaginous and skeletal framework of the larynx (see Figure 2) and several major components of the vocal tract (see Figure 3), such as the mandible and upper jaw (maxilla and palatine bones). See the Appendix for a detailed description of the methods used to create the model.

Model Evaluation

QL2 was used to study seven articulatory and postural states of the larynx (referred to collectively as the *articulatory states* simulations): inspiration, glottal fricative, modal prephonation, glottal stop, vocal–ventricular (VV) stop, aryepiglotto–epiglottal (AE) stop, and AE fricative.

Figure 1. Superficial (top row; A) and ""X-ray"" (bottom row; B) views of the laryngeal mucosa and the embedded laryngeal cartilages and finite-element model–intrinsic musculature. Structures (bold italics): a = arytenoid; aef = aryepiglottic fold; c = cuneiform; ct = cuneiform tubercle; e = epiglottis; et = epiglottic tubercle; ff = ventricular (false) fold; gfm = glossoepiglottic fold medial; gfl = glossoepiglottic fold lateral; kt = corniculate tubercle; pf = piriform fossa; tf = vocal (true) fold; val = valleculae; vent = ventricle (space). Muscles: 1 = thyroepiglottic; 2 = thyroarytenoid vocalis; 3 = thyroarytenoid muscularis; 4 = ventricularis anterolateral; 5 = ventricularis anteromedial; 6 = ventricularis posterolateral.

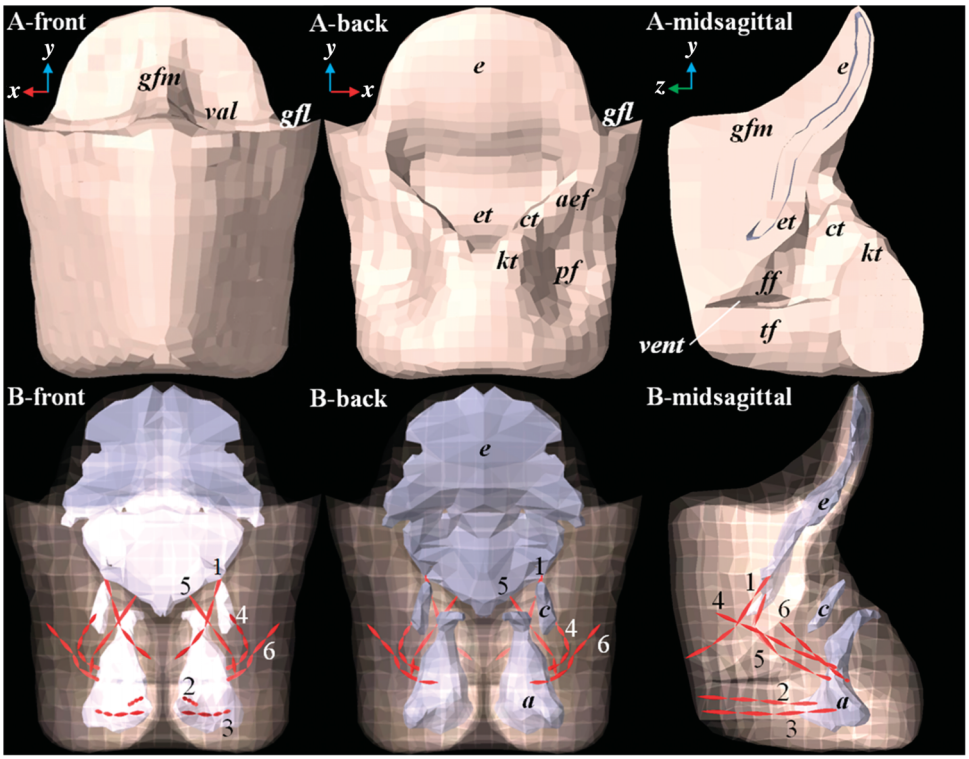


Figure 2. Laryngeal cartilages and the axial musculoelastic framework of QL2 (the biomechanical model of the larynx described here): (A) back, (B) right side (midsagittal cut), and (C) top (with transparent epiglottis). Rigid bodies (bold italics): a = arytenoid; c = cuneiform; cr = cricoid; e = epiglottis; h = hyoid bone; t = thyroid. Connective tissues: 1 = cricothyroid joint; 2 = lateral glossoepiglottic fold; 3 = medial hyoepiglottic ligament; 4 = lateral thyrohyoid ligament; 5 = thyrohyoid membrane; 6 = vocal ligament. Muscles: 7 = interarytenoid transverse superior; 8 = interarytenoid oblique; 9 = interarytenoid transverse inferior; 10 = lateral cricoarytenoid; 11 = posterior cricoarytenoid oblique; 12 = posterior cricoarytenoid horizontal. 13 = cricothyroid (pars recta; CTr)

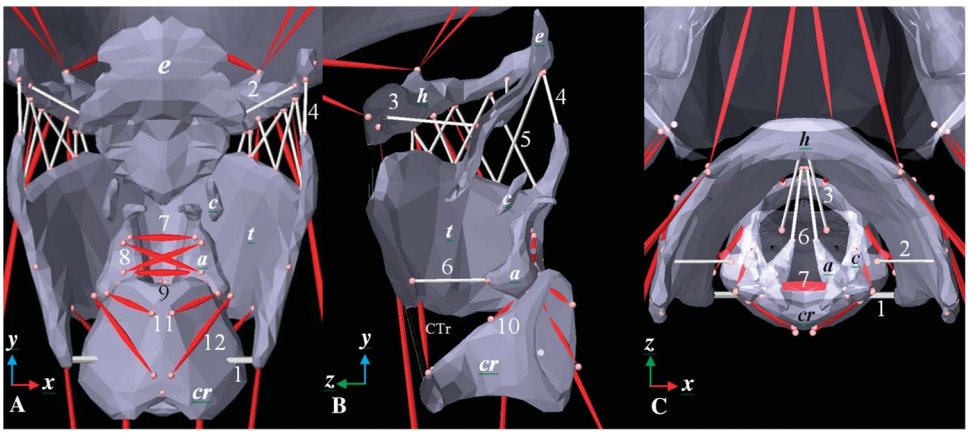
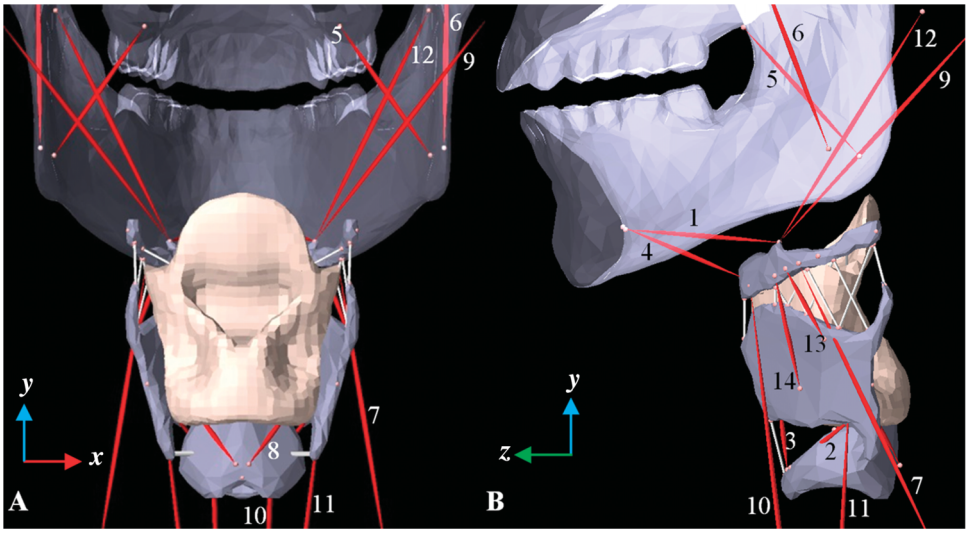


Figure 3. Major model components: (A) back side; (B) left side. Muscles: 1 = anterior digastric; 2 = cricothyroid (pars oblique); 3 = cricothyroid (pars recta); 4 = geniohyoid; 5 = internal pterygoid; 6 = masseter; 7 = omohyoid; 8 = posterior cricoarytenoid (oblique); 9 = posterior digastric; 10 = sternohyoid; 11 = sternothyroid; 12 = stylohyoid; 13 = thyrohyoid (oblique); 14 = thyrohyoid (vertical).



These seven states were selected because they represent a wide range of laryngeal behavior, from fully open (inspiration) to massively constricted (AE stop), and include all of the major laryngeal structures. All simulation targets were designed on the basis of phonetic criteria using primarily laryngoscopy and magnetic resonance imaging and are grounded in the literature on laryngeal muscle physiology (e.g., Faaborg-Andersen, 1957; Faaborg-Andersen & Buchthal, 1956; Hillel, 2001; Hirano & Ohala, 1969; Zemlin, 1998).

A heuristic process was used to find suitable combinations of muscle activation levels (in most cases, identified within the confines delimited by the literature on laryngeal muscle function) in a forward simulation mode (i.e., specifying muscle activation and solving for the resulting movements) that would result in the closest possible approximations to the intended phonetic targets. Once obtained, the fixed ratio of muscle activation was defined and controlled by means of a single, master muscle activation parameter. (For instance, if an articulatory state has a 2:1 ratio of muscle A and B, then at 50% muscle activation, muscle A would be excited twice as much as muscle B, but both would be only at half of the maximum excitation specified for that state, with the exact forces exerted determined by the muscle force scaling.) The combinations used for each articulatory state are provided in Table 1 (expressed as proportions of the total muscle force scaling).

For each articulatory state, a simulation series was created by activating the relevant set of muscles (the specifics of which are found in the appropriate Results sections) from no (0%) activation to full (100%) activation. Each simulation in the series was run for 1.00 s with a maximal time step of 0.01 s (less if numerical stability issues were encountered). In each individual run, muscle activation was set to climb linearly to the target level from 0.00 s to 0.50 s and then remain constant until the end of the simulation at 1.00 s. Every simulation was then summarized by taking the average value observed between 0.50 s and 1.00 s for the response variable in question (see the discussion below about measured distances).

The larynx does not remain in a stable vertical orientation when we speak (Esling, 1999; Ewan & Krones, 1974; Honda, Hirai, Masaki, & Shimada, 1999; Moisik, Lin, & Esling, 2014) but rather shows considerable variation in height and its relationship to nearby structures, especially the hyoid bone. To examine extrinsic larynx posture, wherever possible, each simulation target was conducted in five fixed hyo-laryngeal contexts; these are listed in Table 2 along with the plot markers used to represent them. The default target (black dot) has no modification to extrinsic larynx posture independent of the articulatory objective (none, in this case). The raised target (solid red upward-pointing triangle) involves elevation of the larynx and elevation and advancement of the hyoid bone; in this context, the hyoid bone tends to move away from the thyroid cartilage. The lowered target (solid blue downward-pointing triangle) results in lowering of the hyo-laryngeal complex. The constricting target (hollow red upward-pointing triangle) features approximation of the hyoid bone and thyroid cartilage. The expanding target (hollow blue downward-pointing triangle) exhibits separation of the hyoid bone from the thyroid cartilage by hyoid elevation and advancement and thyroid cartilage lowering. Due to instability of the simulation with increasing tissue contacts, it was not possible to run simulations for the four larynx height settings in the case of the AE states.

A set of measurements was defined to obtain response variables used to evaluate the behavior of QL2. These measurements are listed in Table 3 and illustrated in Figure 4.

Table 1. Peak muscle activations for the states simulations (expressed as proportions).

Simulation	PCAh	PCAo	LCA	IAti	IAts	IAo	TAv	TAm	VTal	VTpl	VTam	TE	TH
Inspiration	**1.00**	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Glottal fricative	0.00	**0.60**	0.00	0.00	**0.65**	**0.35**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Modal prephonation	0.00	0.00	**0.50**	**0.50**	**0.75**	**0.50**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Plain glottal stop	0.00	0.00	**0.50**	**0.75**	**0.75**	**0.75**	**0.60**	**0.60**	0.00	0.00	0.00	0.00	0.00
VV stop	0.00	0.00	**0.50**	**0.50**	**0.75**	**0.50**	**0.25**	**0.30**	**0.25**	**0.25**	**0.25**	**1.00**	0.00
AE stop	0.00	0.00	**0.40**	**0.90**	**0.90**	**1.00**	**0.20**	**0.25**	**0.80**	**0.35**	**0.25**	**0.35**	**0.60**
AE fricative	**1.00**	**1.00**	0.00	0.00	**0.80**	**0.80**	0.00	0.00	**0.20**	**0.65**	**0.20**	**0.25**	**0.60**

Note. Bold indicates non-zero numbers; PCAh = posterior cricoarytenoid horizontal; PCAo = posterior cricoarytenoid oblique; LCA = lateral cricoarytenoid; IAti = interarytenoid transverse inferior; IAts = interarytenoid transverse superior; IAo = interarytenoid oblique; TAv = thyroarytenoid vocalis; TAm = thyroarytenoid muscularis; VTal = ventricularis anterolateral; VTpl = ventricularis posterolateral; VTam = ventricularis anteromedial; TE = thyroepiglottic; TH = thyrohyoid (oblique and vertical); VV = vocal–ventricular; AE = aryepiglotto–epiglottal.

They were selected to capture information about tissue proximity and contact for the key internal structures of the larynx. The measurement was the distance between selected nodes on the laryngeal mucosa. Selection of representative nodes was based on the centrality of these within the surface of the structures of interest. The response variables are as follows: Apposition of the medial surfaces of the vocal folds (true fold [TF] distance) was measured by measurement 1, and apposition of the medial surfaces of the ventricular folds (false fold [FF] distance) was determined by measurement 2. Ventricle height (vocal fold to ventricular fold [VV] distance) was gauged using measurement 3. Anteroposterior narrowing of the upper epilarynx (AE distance) was judged by measurement 4, which measures the distance between the anterior surface of the right cuneiform tubercle and the epiglottic tubercle.

Quantality Score

In previous work dealing with the notion of quantal effects in speech production, quantality has been described qualitatively but has not been quantified. Responding to this gap, a numeric index has been developed, called the quantality (Q-) score, to allow for characterization of these quantal effects more objectively (rather than just making qualitative observations that certain signals appear to exhibit quantal effects).

Table 2. Peak muscle activations (expressed as proportions) for extrinsic larynx posture.

Configuration	Marker	AD	PD	GH	ST	STH	THo	THv
Default	·	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
Raised	▲	1.0	**0.0**	**0.0**	**0.0**	1.0	**0.0**	**0.0**
Lowered	▼	**0.0**	**0.0**	**0.0**	**1.0**	**0.0**	0.0	0.0
Constricting	Δ	**0.0**	**0.0**	**0.0**	0.0	**0.0**	**1.0**	**1.0**
Expanding	∇	0.3	0.3	0.3	0.4	**0.0**	**0.0**	**0.0**

Note. AD = anterior digastric; PD = posterior digastric; GH = geniohyoid; ST = sternothyroid; STH = stylohyoid; THo = thyrohyoid oblique; THv = thyrohyoid vertical.

With the Q-score, a higher score means more quantal behavior of the system, which is biomechanical in this case. The Q-score is defined by the following equation:

$$Q\text{-}score(f_{norm}(x)) = \ln\left(\frac{\sum_{i=1}^{n} w_i |f_{norm}'(x_i)|}{\sum_{i=1}^{n} w_i}\right) - \ln\left(\frac{\sum_{i=1}^{n} z_i |f_{norm}'(x_i)|}{\sum_{i=1}^{n} z_i}\right), \quad (1)$$

where n = number of samples in f_{norm}', $w_i = 1 - z_i$ and $z_i = \frac{i-1}{n-1}$.

The concept behind this formula is to examine the behavior of the first derivative, f_{norm}', of a given normalized response variable f_{norm} (e.g., TF distance) for a given simulation series (ranging from 0%–100% muscle activation). The response variable signals were normalized to make the Q-score independent of absolute signal magnitude and thus to better reflect the shape of the signal; this normalization was made relative to the maximum absolute value observed. Highly "quantal" articulation should show an initial period of rapid change (absolute value of the derivative is large) at low muscle activation levels but tend to stabilize (derivative tends to 0) at higher levels of muscle activation. To capture this intuition, two weighting functions are used: w(i) and z(i), which place emphasis on the early and later parts of the derivative function, respectively. The first term, being weighted for the early part of the derivative function, favors higher absolute values early on. (Low values would indicate that the response variable is not changing and is thus stable without any muscular intervention, suggesting that the muscle activation set has little effect on the structure.) The second term, being weighted for the later part of the signal, favors lower absolute values of the derivative later on (which indicates stability despite high muscle activation). Note that a perfectly constant signal would give a zero derivative and thus be undefined. In practice, no such signals were produced by the model.

Figure 5 shows Q-scores computed for an illustrative function (at selected values of the parameter σ). This function is simply meant to emulate the types of response

Table 3. Measurements used as response variables.

No.	Name	Description	Indicator of
1	TF distance	Distance between *x*-components of FEM nodes of the midway medial surface of the vocal folds	Vocal fold constriction
2	FF distance	Distance between *x*-components of FEM nodes of the medial region of the ventricular (false) folds	Ventricular fold constriction, lower epilaryngeal constriction
3	VV distance	Distance between *y*-components of medially located FEM nodes of the inferior aspect of the right-side ventricular fold and the ipsilateral nodes on the upper surface of the vocal fold	Available ventricular space, vocal–ventricular fold contact
4	AE distance	Anteroposterior distance between the *z*-components of the right-side epiglottic and cuneiform tubercles	Upper epilaryngeal constriction

Note. See Figure 4 for interpretation of coordinates. TF = true (vocal) fold; FEM = finite-element model; FF = false (ventricular) fold; VV = vocal fold to ventricular fold; AE = aryepiglotto-epiglottal.

signals found for the actual simulations in relation to muscle activity (for the purposes of illustration, just x values 0–100). Higher Q-scores indicate model behavior that exhibits significant initial change that rapidly settles into a stable pattern. Because nothing guarantees that the first term in Equation 1 will be higher than the second term, it is possible for the Q-score to be negative. A approximately linear relationship gives a Q-score of 0.0 (bold black line, Figure 5); 100,000 simulated response-variable signals randomly sampled from the uniform distribution (on the interval [0, 1]) yield a mean Q-score close to 0.0 (-5.8×10^{-4}; $SD = 0.22$). In practice, Q-scores above 0.0 appear more and more quantal. The illustrative function at $\sigma \approx 3.05$ (blue line, Figure 5) becomes stable at high values of muscle activation (at 90%) and gives a Q-score of 0.54. At $\sigma \approx 1.70$ (green line, Figure 5), the function stabilizes at about the 50% mark and results in a Q-score of 1.38. Q-scores higher than this indicate response functions that settle very quickly even at low levels of muscle activity (e.g., at 10% muscle activity for the red line, Figure 5). To ease interpretation, values of Q-scores on the following intervals are labeled as such: $[-\infty, 0.00]$ is nonquantal, [0.00, 0.54] is mildly quantal (blue area), [0.54, 1.38] is moderately quantal (green area), and $[1.38, +\infty]$ is strongly quantal (red area).

Results: Simulations of Laryngeal Articulatory States

This section presents results for the simulations of the seven laryngeal articulatory states. For each state a plot is given showing the effect of increasing muscle activation on the response variables, TF distance, VV distance, FF distance, and, in the case of the AE simulations, AE distance. As a reminder, each data point represents, for a given simulation run in a given simulation series, the average of the value of the response variable during the constant muscle activation phase (the period from 0.50 to 1.00 s), when the model assumes a static configuration. Alongside these plots are selected frames obtained from laryngoscopic videos associated with the target state. These laryngoscopic images were obtained from videos produced by John Esling (with his permission) and available on the Internet for viewing (http://web.uvic.ca/ling/research/phonetics/SOG/index.htm; Esling & Harris, 2003). Similar images and states are documented in Esling and Harris (2005). Also shown are five frames (selected at equal intervals starting at 0.05 s and stopping at 0.45 s) that demonstrate the appearance of QL2 from top-down (for all plots) and midsagittal (for some plots) views as it appeared in the neutral larynx height series

Figure 4. Measurement vectors corresponding to Table 3. 1 = vocal (true) fold distance; 2 = ventricular (false) fold distance; 3 = vocal–ventricular fold distance; 4 = aryepiglotto–epiglottal distance.

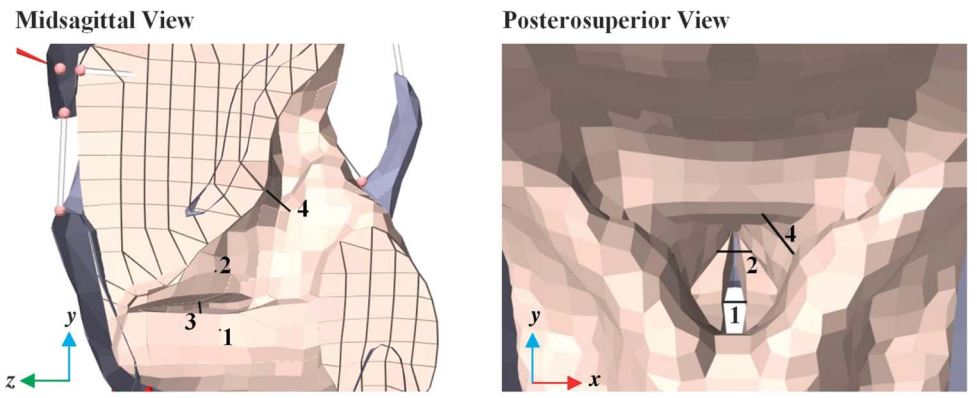
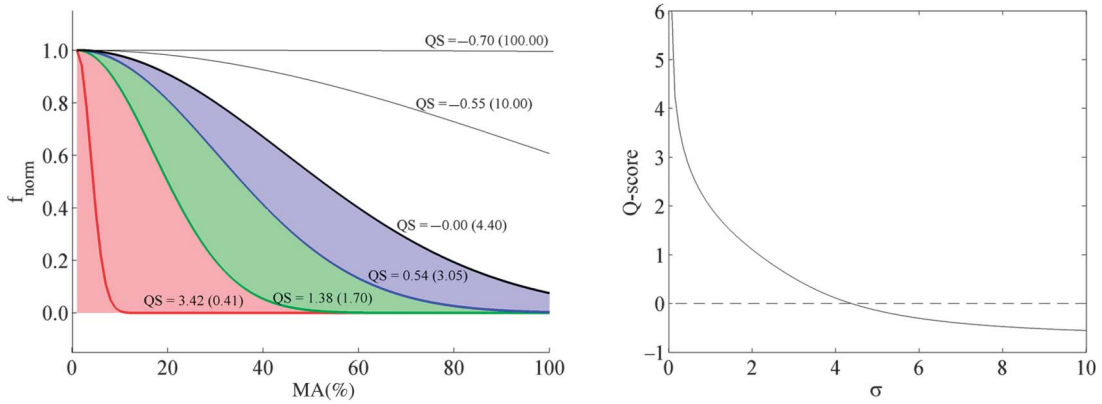


Figure 5. Illustration of Q-scores (QS) for $f_{norm} = e^{\wedge}(-x^2/2\sigma^2)$. Left: Function shape illustration and Q-scores for selected values of σ (given in parentheses). f_{norm} = a given normalized response variable; MA = muscle activation. Bold lines indicate appearance of the illustrative function at critical Q-scores. Colored regions are interpretive zones (white = nonquantal; blue = mildly quantal; green = moderately quantal; red = strongly quantal). Right: Q-score as a function of σ. (Note that σ values higher than approximately 10 produce smaller and smaller changes in the shape of the function.)



at 75% muscle activation (denoted by the dashed line in the corresponding response variable plots). The Q-scores (see the Quantality Score section) are used to characterize the articulatory state simulations. Q-score values are given for each response variable associated with the default hyo-laryngeal configuration (dotted black lines) in the respective sections. As a visual aid, gray regions in the response variable plots have been manually added to highlight stabilization behavior.

Inspiration

Although inspiration is not normally used in the production of the segmental content of an utterance, it precedes and follows speech utterances and thus forms an integral component of the sequence of motor behaviors characterizing speech. At the peak of muscle activity, the arytenoids are widely separated and the interarytenoid mucosa is plainly visible (arrow 1, Figure 6). Laryngoscopic and X-ray cinematography (Moisik, 2013, p. 276) show that the larynx typically descends and the tongue root advances during inspiration (consistent with electromyographic [EMG] recordings of the genioglossus muscle showing elevated levels during inspiration—e.g., Sauerland & Harper, 1976; also see Schwab, Gefter, Pack, & Hoffman, 1993, p. 1513), especially when forceful or deep. These actions very likely assist in increasing the overall patency of the airway to reduce airflow resistance. Fink (1974a) demonstrated that the larynx-lowering activity in inspiration is associated with a wide separation of the thyroid cartilage and hyoid bone; this corresponds with a wide vertical spacing of the ventricle and a lateral displacement of all soft tissues of the laryngeal airway, all of which would benefit airflow resistance reduction.

Concordant with Fink's observations, an increase in VV distance occurs with larynx lowering; the lateral traction on the vocal folds was not apparent (TF distance is largely unaffected by larynx height condition, except slightly for the "raised" case). There is a sharp inflection in the observed response variables (e.g., TF distance; arrow 2 in Figure 6) that happens at a low level of muscle activation. This indicates a nonlinear shift in the posturing of the arytenoids as they are rocked backward and outward. As a consequence, Q-scores for all variables are in the moderate to strong range. However, after this inflection point, the response for TF and FF increases linearly as a function of muscle activation, and VV distance indicates that the vocal and ventricular folds maintain constant separation. Q-scores for TF, VV, and FF computed on the response signals (braces, Figure 6) after the inflection point (arrow 2) are all less than zero (–0.0018, –0.3300, and –0.0650, respectively), supporting the interpretation of a nonquantal, largely linear response.

Glottal Fricative (Also Aspiration and Expiration)

Unlike inspiration, this laryngeal state is responsible for the generation of a commonly occurring speech sound, namely the glottal fricative [h]. However, it also serves in the production of aspiration, which can occur in connection with stops (e.g., the [t^h] in English [t^hɑp] *top*), voiceless fricatives, and, with sufficient airflow to drive vocal fold vibration, breathy voice (Esling & Harris, 2005; in this work, the authors refer to the nonvibrating state as *breath*). The persistent interarytenoid gap allows for continuous airflow and the anterior ligamentous glottis is variably abducted (but not so widely as in inspiration), adding more airway resistance. This state also plays a fundamental physiological role in the respiratory cycle because the additional resistance during expiration increases the time available for gas exchange to occur (Bartlett, Remmers, & Gautier, 1973; England, Bartlett, & Daubenspeck, 1982; Negus, 1949, pp. 63–64; cf. Hillel, 2001, p. 23). Unlike the inspiration simulations, which show a linear increase in TF distance past the initial nonlinear inflection point (arrow 2, Figure 6), the glottal fricative state exhibits clear biomechanical stability (gray regions, Figure 7). This interpretation is supported by the Q-scores, which are all in the moderately quantal range.

Figure 6. Simulation of inspiration. Data points show individual simulation runs (with values averaged from 0.5 to 1.0 s in each case). See Table 1 for muscles activated, and see Table 2 for interpretation of plot markers, colors, and extrinsic larynx posture muscle parameter settings. MA = muscle activation; QS = quantality score for the default extrinsic larynx posture. Distances: TF = vocal (true) folds; VV = vocal–ventricular; FF = ventricular (false) folds. t = time. Dashed line indicates 75% muscle activation visualized with model screen shots. Arrow 1 shows visible interarytenoid mucosa, and arrow 2 shows inflection point in distance measurements indicating sudden change in vocal fold (and ventricular fold) behavior. The laryngoscopic still frame (upper right) is the property of John Esling, Copyright 2016. Adapted with permission from http://web.uvic.ca/ling/research/phonetics/SOG/index.htm (Esling & Harris, 2003).

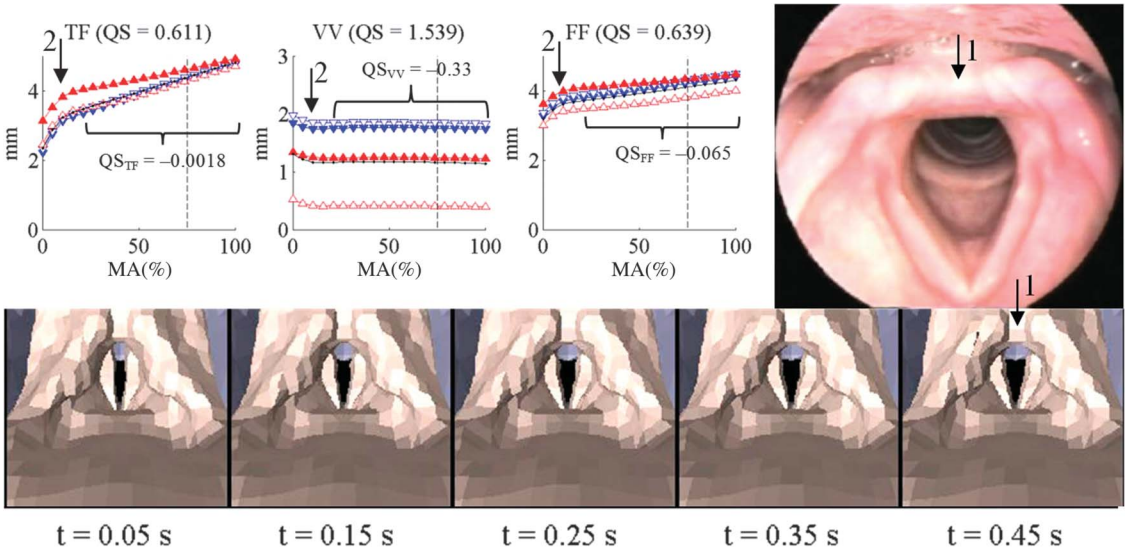
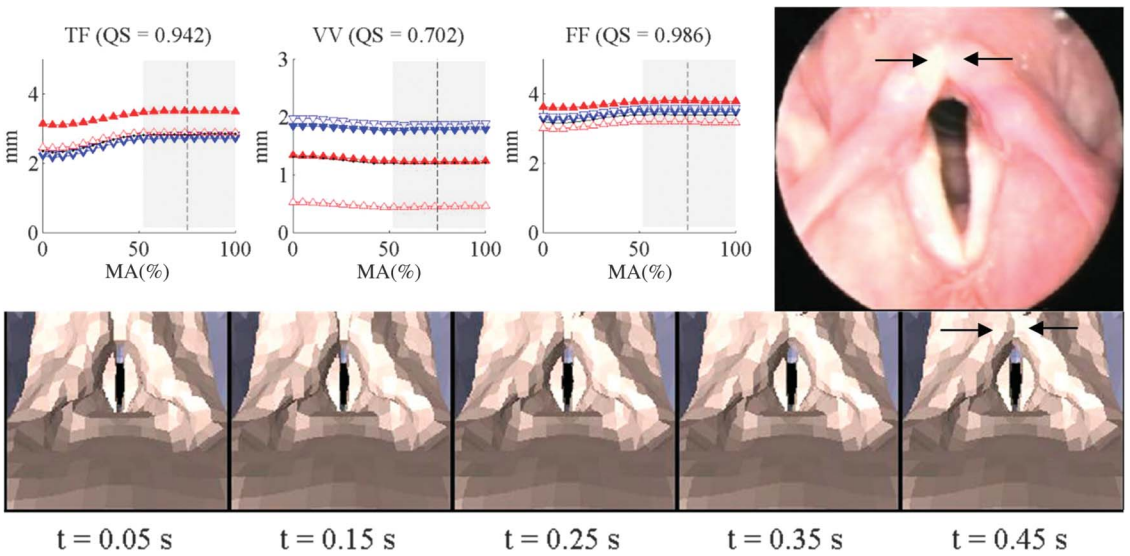


Figure 7. Simulation of glottal fricative [h] (also aspiration and expiration). Data points show individual simulation runs (with values averaged from 0.5 to 1.0 s in each case). See Table 1 for muscles activated, and see Table 2 for interpretation of plot markers, colors, and extrinsic larynx posture muscle parameter settings. MA = muscle activation; QS = quantality score for the default extrinsic larynx posture. Distances: TF = vocal (true) folds; VV = vocal–ventricular; FF = ventricular (false) folds. t = time. Dashed line indicates 75% muscle activation visualized with model screen shots. Gray rectangles indicate visual indication of stabilized movement. Inward-pointing arrows indicate corniculate tubercle contact. The laryngoscopic still frame (upper right) is the property of John Esling, Copyright 2016. Adapted with permission from http://web.uvic.ca/ling/research/phonetics/SOG/index.htm (Esling & Harris, 2003).



Q-scores for TF and FF distance are higher here than the corresponding values observed for the inspiration simulations. The contact of the corniculate tubercles is the key factor underlying this stability (inward-facing arrows, Figure 7). This accords well with what is observed in the laryngoscopic appearance of these sounds (the still frame selected is from a video of aspiration during [p^h], but expiration and glottal fricative [h] appear almost identical to this).

Important in the formation of this state is that the action of the interarytenoid (IA) muscles be primarily confined to the superior transverse fibers, which rock the arytenoid–corniculate complexes together without forcing too much adduction of the cartilaginous glottis. Engagement of the lower section of this muscle achieves such closure, which is the case for modal prephonation (see the next section). At the same time, the posterior cricoarytenoid muscles can effect a rotation of the vocal processes to drive the abduction of the vocal folds without separating the arytenoids as during inspiration (Zemlin, Davis, & Gaza, 1984). Such specific activation of the superior fibers of the transverse interarytenoid muscle in relation to expiration (or similar adjustments) has not been reported in the literature: In fact, it seems to be the case that no physiological study has been able to resolve even the transverse and oblique portions of the IA muscle. Nevertheless, such local differences in activation are indeed possible in principle for individual muscles (e.g., Wickham & Brown, 1998; also see Knudson, 2007, pp. 57–58). Physiological measurements (e.g., using EMG) suitable for capturing differential activation of different portions of the IA muscle would be difficult to obtain, but the present result would be grounds for empirically investigating the matter further. The issue of intramuscular contraction aside, Kagaya and Hirose (1975) demonstrated with EMG evidence that, although diminished somewhat, the IA muscles are still active during aspirated stops, such as [p^h].

Modal Prephonation

As the name suggests, modal prephonation occurs immediately prior to modal phonation characterized by a smooth attack or onset. It is the static state through which sufficient airflow can set the vocal folds in motion for modal phonation. Thus, it represents one of the most important phonetic states because of the essential role played by modal phonation in speech production. Although the glottal fricative state (see the section Glottal Fricative [Also Aspiration and Expiration]) was identified as corresponding to breathy phonation, reduction of medialization in the prephonation state would possibly also produce somewhat breathy phonation. This is thus a second form of breathiness, but one that shows no interarytenoid gap, as can be observed for the Bai speaker presented in Edmondson and Esling (2006, p. 174, figure 10b). This setting is also said to occur in the context of voiceless unaspirated stops (Esling & Harris, 2005, p. 350).

To achieve the simulation of this state, all subcomponents of the IA muscle were activated and complemented by lateral cricoarytenoid (LCA) activity. Hillel (2001) demonstrated that modal prephonation is characterized by a specific temporal sequencing of LCA muscle activity prior to IA engagement; the former primes the adducted state, and the IAs then maintain adduction during phonation. No attempt was made to simulate this level of temporal detail (although it is possible in principle); thus, QL2 may appear to be overadducted compared with the laryngoscopic image because LCA activity is never relaxed.

Similar to glottal fricative, modal prephonation exhibits contact of the corniculate tubercles, but, as is evident in the laryngoscopic image, the extent of this contact is greater than what occurs in the simulation: There is still some space between the cuneiform tubercles. QL2 nearly replicates this contact, but, owing to the rigidity of the arytenoid cartilages (in QL2), it cannot completely emulate the compression of these structures into each other. Rather, upon contact, QL2 allows some rigid anterior rolling of the arytenoids as muscle excitation increases, causing the corniculate tubercles to separate somewhat. Thus, the response variables never fully saturate, as the comparatively lower Q-scores in the high–mild to low–moderate quantality range indicate. This is attributable to a known limitation of the model: the lack of deformable arytenoid cartilages. Some moderate stabilization is evident in TF and FF distances (see Figure 8), and such stabilization would almost certainly be more pronounced were deformable arytenoid cartilages to be modeled.

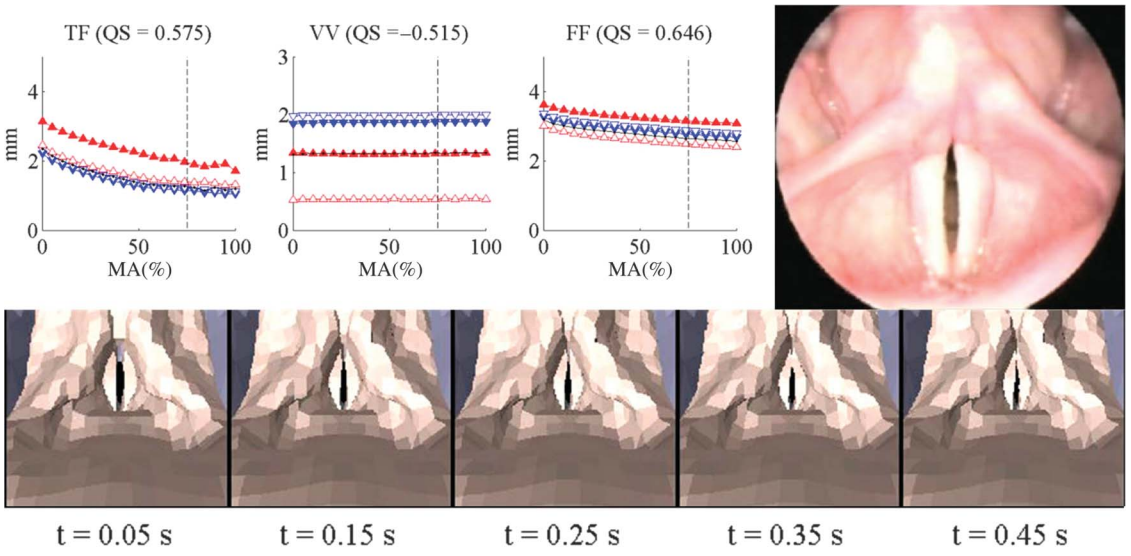
Plain Glottal Stop

Plain glottal stop refers to a prolonged arrest of the vocal folds caused by their medialization and contact. The use of *plain* is intended to emphasize that only the vocal folds are involved. It is possible for the ventricular folds to become engaged in the production of glottal stop, and this is the topic of the following section. In either case, such vocal fold arrest can occur at the onset (or offset) of modal phonation, giving the quality of a hard phonatory onset or attack (or termination), but glottal stop also constitutes a speech sound in its own right and has phonemic status in many languages (Ladefoged & Maddieson, 1996). Furthermore, this method can accompany supralaryngeal consonants, in which case it is referred to as *glottalization*; such glottalized stops are widespread and are even common in certain varieties of English (Roach, 1979).

Low dimensional vocal fold vibration models suggest that vocal fold adduction should be sufficient to achieve glottal stop (this is verified in Moisik & Esling, 2014). To model plain glottal stop in QL2, moderate thyroarytenoid activity (cf. Hirano & Ohala, 1969) was added to the muscle set used in modal prephonation. This causes intrinsic stiffening and a medial bulging of the vocal folds, which aids in closure (see Figure 9).

Note that the laryngoscopic still frame does not represent an end-point stricture and actually comes from a video of a glottal stop produced with accompanying adduction of the ventricular folds (the frame in question comes some short moments after the point of full stricture when

Figure 8. Simulation of modal prephonation. Data points show individual simulation runs (with values averaged from 0.5 to 1.0 s in each case). See Table 1 for muscles activated, and see Table 2 for interpretation of plot markers, colors, and extrinsic larynx posture muscle parameter settings. MA = muscle activation; QS = quantality score for the default extrinsic larynx posture. Distances: TF = vocal (true) folds; VV = vocal–ventricular; FF = ventricular (false) folds. t = time. Dashed line indicates 75% muscle activation visualized with model screen shots. The laryngoscopic still frame (upper right) is the property of John Esling, Copyright 2016. Adapted with permission from http://web.uvic.ca/ling/research/phonetics/SOG/index.htm (Esling & Harris, 2003).



the ventricular folds are beginning to abduct). From a wide cross-linguistic laryngoscopic data set (Edmondson & Esling, 2006; Edmondson, Esling, Harris, & Wei, 2004; Esling, Fraser, & Harris, 2005; Esling & Harris, 2003), it would seem that glottal stop with such ventricular fold reinforcement (a VV stop; see the section VV Stop [or "Reinforced" Glottal Stop]) is actually much more commonly encountered than the plain glottal stop. There are, however, reports of endoscopic evidence for plain glottal stop (Edmondson, Chang, Hsieh, & Huang, 2011; Garellek, 2013; Iwata, Sawashima, Hirose, & Niimi, 1979). Nonetheless, even in such cases of apparent plain glottal stops, it is still possible that the ventricular folds come into contact with the vocal folds, even if the ventricular folds do not completely adduct (Moisik, Esling, Crevier-Buchman, Amelot, & Halimi, 2015).

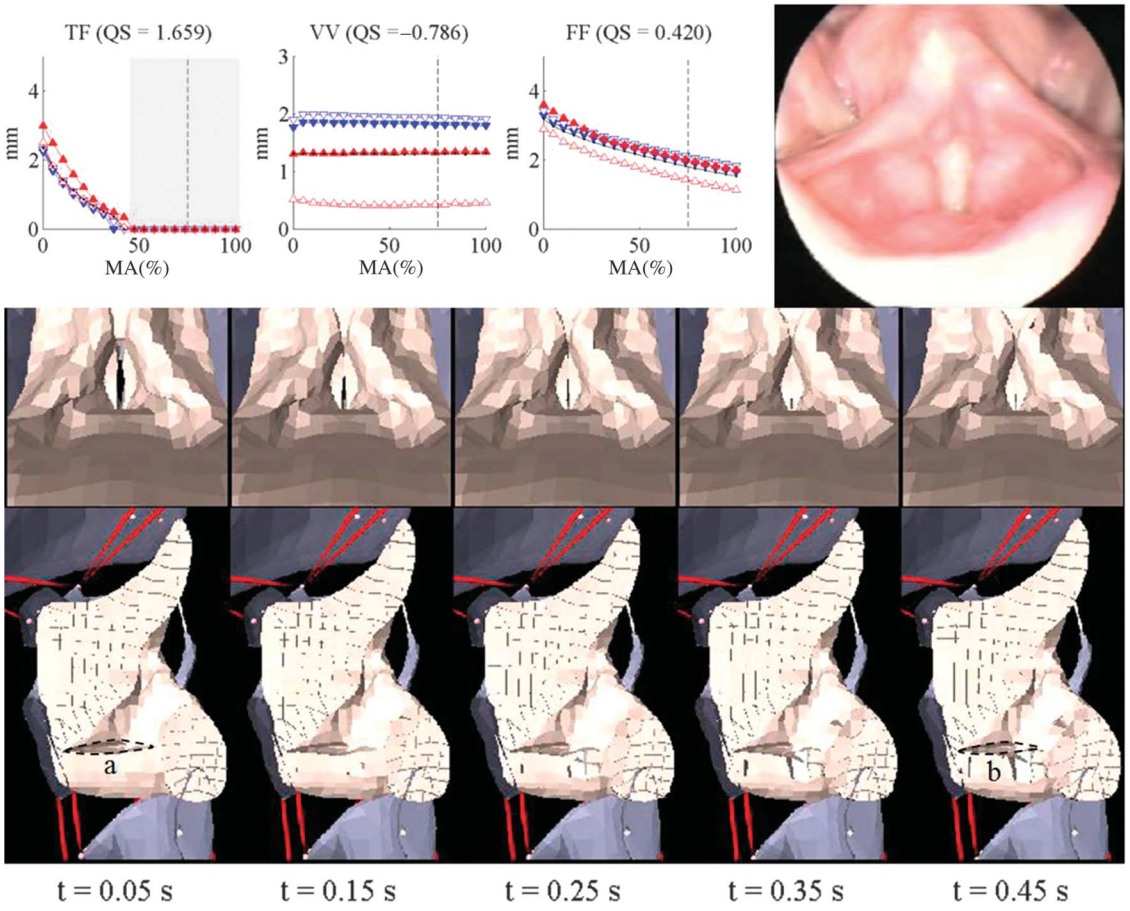
In the simulations, no such contact occurs throughout the simulation (see Figure 9; compare dashed outlines *a* and *b* in the midsagittal view, showing that the ventricle stays open), and the vertical VV distance is not much different from that in the preceding cases examined (and contributes nothing to quantality, as the negative Q-score indicates). What is certain is that vocal fold contact constitutes a biomechanical stabilization event (with a commensurately high strong-range Q-score for TF distance) that occurs in addition to the contact of the posterior cartilages during glottal stop; noteworthy, however, is the fact that the FF distance still decreases approximately linearly as a function of increasing muscle activation (as indicated by the comparatively low Q-score indicating only mild quantality). Plain glottal stop, though moderately quantal for some measures, thus exhibits epilaryngeal instability, which may account for the greater likelihood of observing VV (or reinforced glottal) stop.

VV Stop (or "Reinforced" Glottal Stop)

As discussed in the previous section, plain glottal stop is possible but is not commonly observed; instead, glottal stop typically occurs with reinforcement from the ventricular folds. Laryngoscopic imaging has never been able to show definitively whether, when this occurs, there is VV fold contact (VVFC; i.e., the ventricular folds descending upon and compressing into the vocal folds). VVFC has been demonstrated for glottal stop and creaky voice with laminography (Hollien, 1974), laryngeal ultrasound (Esling & Moisik, 2012), and MR imaging (Moisik et al., 2015).

For VVFC to be achieved, it seems necessary to engage the supraglottal musculature of the larynx, including the thyroepiglottic (TE) and ventricularis (VT) muscles (Reidenbach, 1998b). The observation of epiglottis movements for phonetically related postures (Brunelle, Nguyễn, & Nguyễn, 2010) supports the interpretation of TE involvement. The simulations here confirm that VVFC can be achieved by activation of the TE and all three branches of the VT muscles: VV distance goes to 0 mm (see Figure 10; compare dashed outlines *a* and *b* in the midsagittal view; note also that the thicker dashed outlines in the laryngoscopic still frame and the model images at 0.45 s indicate that the upper epilarynx is still open despite being narrowed). The simulations show that ventricular fold midline contact also occurs but at higher muscle activation levels (compare the zeroing point in VV distance with that in FF distance;

Figure 9. Simulation of plane glottal stop. Data points show individual simulation runs (with values averaged from 0.5 to 1.0 s in each case). See Table 1 for muscles activated, and see Table 2 for interpretation of plot markers, colors, and extrinsic larynx posture muscle parameter settings. MA = muscle activation; QS = quantality score for the default extrinsic larynx posture. Distances: TF = vocal (true) folds; VV = vocal–ventricular; FF = ventricular (false) folds. t = time. Dashed line indicates 75% muscle activation visualized with model screen shots. Dashed outline a shows ventricle opening at start; dashed outline b shows ventricle opening late in the MA increase phase. Gray rectangles are a visual indication of stabilized movement. The laryngoscopic still frame (upper right) is the property of John Esling, Copyright 2016. Adapted with permission from http://web.uvic.ca/ling/research/phonetics/SOG/index.htm (Esling & Harris, 2003).



arrows, Figure 10). Thyroid–hyoid approximation associated with the constricting extrinsic larynx posture makes VVFC occur even sooner (VV distance, hollow red upward-pointing triangles, Figure 10).

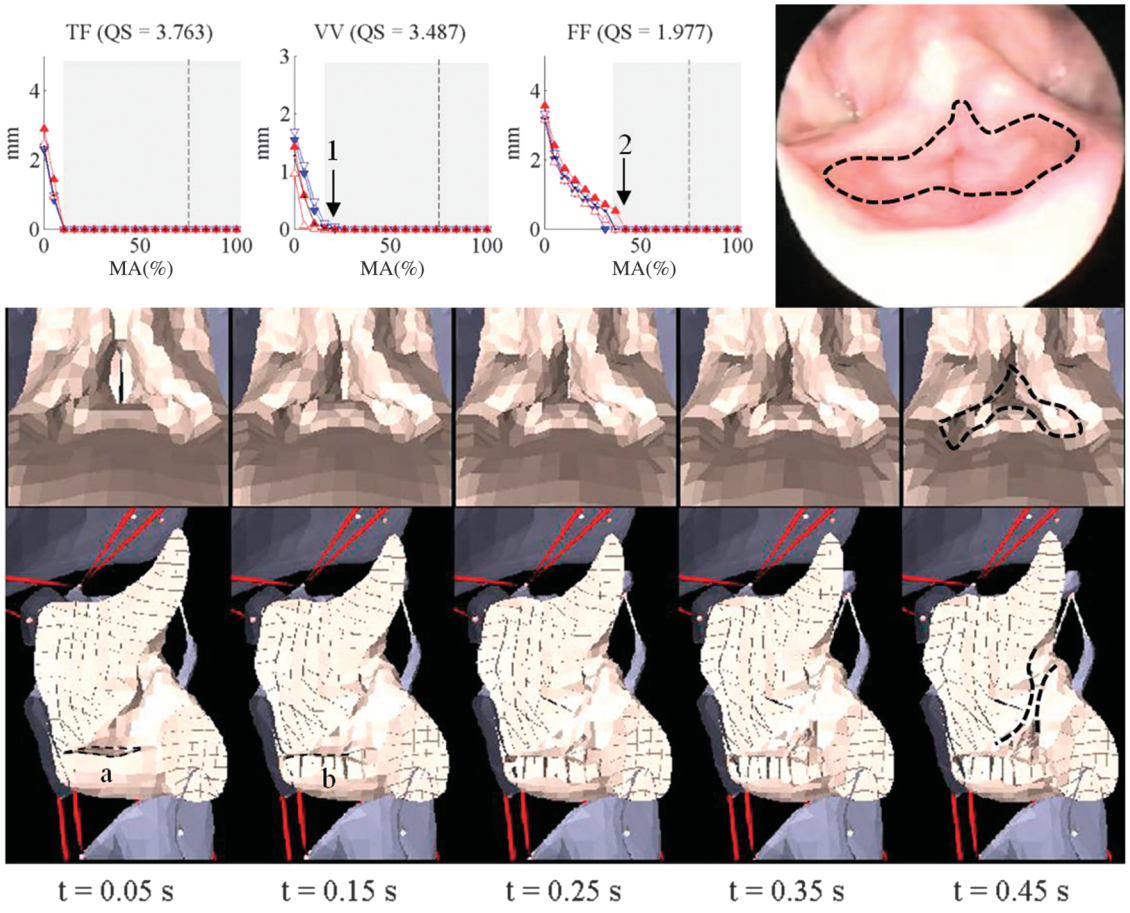
Thus, VV stop adds two extra stabilization events—VVFC and ventricular fold midline contact—on top of those occurring in plain glottal stop. This would serve to enhance articulatory stability. The high Q-scores (all in the strongly quantal range) support this interpretation, but the higher Q-score for VV distance than for FF distance indicates that VVFC has an earlier occurring (arrow 1, Figure 10) stabilization compared with complete ventricular fold medialization and contact (arrow 2, Figure 10). This last result suggests that VVFC might occur even if there is only partial ventricular fold medialization (such that the vocal folds are still visible).

If the posture between the muscle activation levels demarcated by arrow 1 and arrow 2 (see Figure 10) were to occur, such that there is VV contact but with a slight gap between the ventricular folds (corresponding to QL2 appearance at t = 0.15 s in Figure 10) and the vocal folds set into vibration (not simulated here), it is expected that such vibration would be characteristically perturbed as in creaky voice (Moisik & Esling, 2014). VVFC should disrupt the normal transmission of the mucosal wave of the vocal folds and possibly alter the effective vibrating mass, both of which should yield vibratory patterns associated with phonatory qualities such as creakiness or harshness (depending on factors such as subglottal pressure).

AE Stop

Laryngeal constriction beyond VV stop leads to reduction of the laryngeal vestibule and ultimately to collapse of this upper epilaryngeal airspace as the cuneiform tubercles come into apposition with the epiglottic tubercle. Such a state is found phonetically in the context of both glottal stop (Lindqvist-Gauffin, 1972) and pharyngeal/epiglottal

Figure 10. Simulation of vocal–ventricular stop ("reinforced" glottal stop). Data points show individual simulation runs (with values averaged from 0.5 to 1.0 s in each case). See Table 1 for muscles activated, and see Table 2 for interpretation of plot markers, colors, and extrinsic larynx posture muscle parameter settings. MA = muscle activation; QS = quantality score for the default extrinsic larynx posture. Distances: TF = vocal (true) folds; VV = vocal–ventricular; FF = ventricular (false) folds. t = time. Dashed line indicates 75% muscle activation visualized with model screen shots. Gray rectangles are a visual indication of stabilized movement. Dashed outline a shows ventricle opening at start; dashed outline b shows ventricle closure early in the MA increase phase. Thick dashed outlines are an indication of the epilaryngeal tube aperture. The laryngoscopic still frame (upper right) is the property of John Esling, Copyright 2016. Adapted with permission from http://web.uvic.ca/ling/research/phonetics/SOG/index.htm (Esling & Harris, 2003).



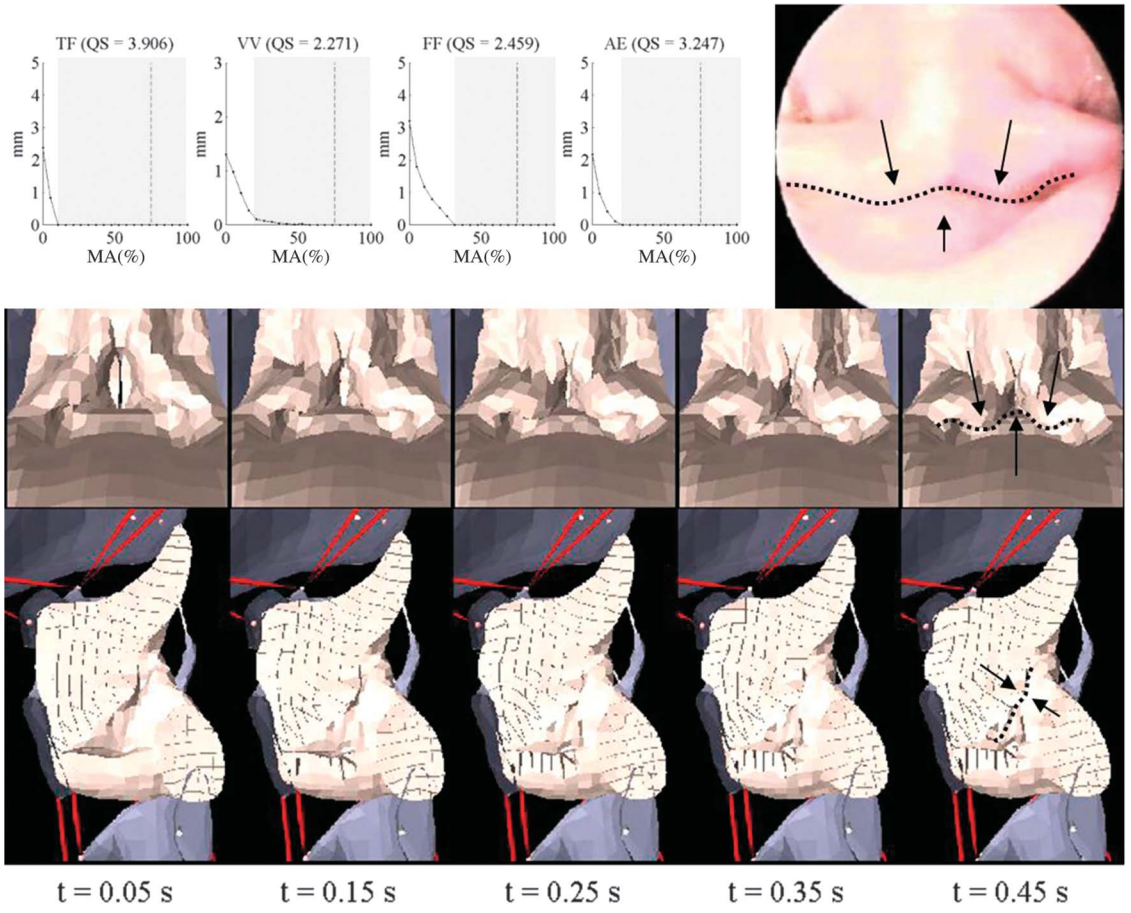
stop [ʔ] (Edmondson & Esling, 2006; Esling, 1996), although the latter typically also have accompanying tongue retraction, as in Arabic 'ayn (Heselwood, 2007). This state also occurs in effort closure and is prior (and fundamental) to swallowing (Fink, 1974a). AE contact is achieved in the simulation (indicated by the dotted contour line and arrows in Figure 11) with a combination of muscle activations similar to VV stop and the addition of, strong thyrohyoid engagement. Aryepiglottic muscles are not represented in QL2, and this simulation demonstrates that such constriction does not require these muscles, mostly consistent with both Fink's (1974a) and Reidenbach's (1997, 1998a, 1998b) theoretical descriptions of the general closure mechanism (i.e., that anteroposterior laryngeal vestibule closure occurs under thyroid–hyoid approximation, the action of the VT and TE musculature, and the buckling and medializing effects of tissue contacts).

The AE stop state represents a culmination of biomechanical stabilization events beginning with corniculate tubercle contact (as seen in expiration) and leading to extensive compression and contact of most of the vocal fold and epilaryngeal surfaces. These simulation results suggest that this configuration is highly stable, producing very high Q-scores across all of the response variables. Tongue retraction tends to accompany this state (Edmondson & Esling, 2006) but was not included in this simulation, demonstrating that tongue retraction is not necessary for full AE contact to be achieved.

AE Fricative (Also Whisper)

Like AE stop, AE fricative is characterized by contact between the cuneiform and epiglottic tubercles (thick dashed contour lines and arrows in Figure 12). The key

Figure 11. Simulation of aryepiglotto-epiglottal stop. Data points show individual simulation runs (with values averaged from 0.5 to 1.0 s in each case). See Table 1 for muscles activated, and see Table 2 for interpretation of plot markers, colors, and extrinsic larynx posture muscle parameter settings. MA = muscle activation; QS = quantality score for the default extrinsic larynx posture. Distances: TF = vocal (true) folds; VV = vocal–ventricular; FF = ventricular (false) folds; AE = aryepiglotto–epiglottic. t = time. Dashed line indicates 75% muscle activation visualized with model screen shots. Gray rectangles are a visual indication of stabilized movement. Dotted line and arrows show anteroposterior contact of the epilarynx. The laryngoscopic still frame (upper right) is the property of John Esling, Copyright 2016. Adapted with permission from http://web.uvic.ca/ling/research/phonetics/SOG/index.htm (Esling & Harris, 2003).



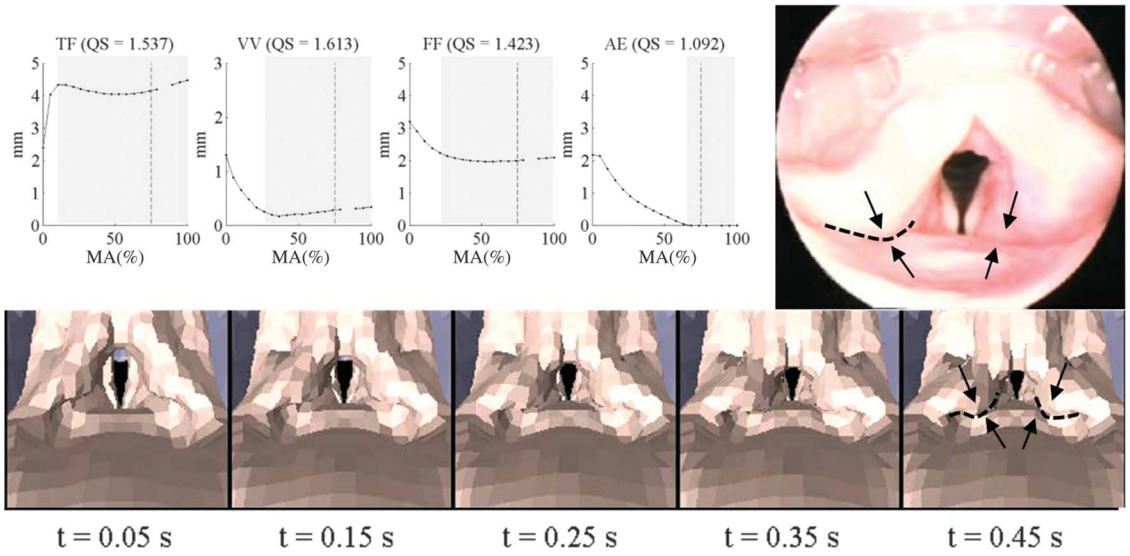
difference is that the vocal folds must abduct, particularly along their cartilaginous extent, to allow for the airflow necessary to generate a noisy sound source. The configuration used to produce this sound occurs during ordinary whisper (Esling & Harris, 2005; Honda et al., 2010) but also occurs in the context of pharyngeal/epiglottal fricatives [ħ], and these latter sounds often feature substantial tongue retraction (Esling, 1996; Wilson, 2007). In addition, and like in modal prephonation, this state involves abduction of the cartilaginous glottis but with simultaneous contact of the corniculate tubercles.

As with AE stop, simulation of this state presents a considerable modeling challenge given the large number of tissue contacts involved (requiring the computation of self-collisions for the FEM of the laryngeal mucosa); one simulation run failed because of laryngeal mucosa FEM element inversion (missing data in Figure 12). (The simulations presented here represent the muscle activation combination that produced the largest number of stable simulations.) The simulation series demonstrates that AE fricative is less quantal in nature than AE stop, having lower Q-scores, although still in the strongly quantal range. (Note that Q-scores were computed for response-variable signals after their missing values were supplied by means of a linear-interpolation gap-filling algorithm.) This might be attributable to lack of contact between the corniculate tubercles.

Discussion and Conclusions

The present study has simulated a wide range of phonetic behaviors using a physiologically highly detailed model of the larynx (QL2), validating the resulting simulated states against laryngoscopic image data. The range of articulatory states simulated spans the continuum of laryngeal constriction from those states that are characterized by a relatively open larynx (e.g., inspiration, glottal fricative,

Figure 12. Simulation of aryepiglotto–epiglottal fricative ("whisper"). Data points show individual simulation runs (with values averaged from 0.5 to 1.0 s in each case). See Table 1 for muscles activated, and see Table 2 for interpretation of plot markers, colors, and extrinsic larynx posture muscle parameter settings. MA = muscle activation; QS = quantality score for the default extrinsic larynx posture. Distances: TF = vocal (true) folds; VV = vocal–ventricular; FF = ventricular (false) folds; AE = aryepiglotto–epiglottic. t = time. Dashed line indicates 75% muscle activation visualized with model screen shots. Gray rectangles are a visual indication of stabilized movement. Thick dashed line and arrows show anteroposterior contact (or approximation) of the epilarynx; contact is unilateral in the laryngoscopic still frame. The laryngoscopic still frame (upper right) is the property of John Esling, Copyright 2016. Adapted with permission from http://web.uvic.ca/ling/research/phonetics/SOG/index.htm (Esling & Harris, 2003).



and modal prephonation) to those with a more narrowed and constricted larynx (e.g., glottal stop and epiglottal stop). These states are recruited phonetically for many phonological functions, including the production of laryngeal (e.g., /h/ and /ʔ/) and pharyngeal (e.g., /ħ/ and /ʕ/) consonants and secondary laryngeal and pharyngeal articulations (as in aspiration, glottalization, and pharyngealization), and they can be related to phonatory states with common settings. Related to this final point, even though vibration was not simulated, a plausible set of close postural correspondences hold between static (simulated) articulatory states and vibratory (not simulated) phonatory states. Catford (1964) originally presented this line of reasoning and posited that corresponding states involve minimal adjustments; the near visual parity between nonvibratory and vibratory states in Esling and Harris (2005) further supports this interpretation. Thus, in the model, if sufficient subglottal pressure were to be applied in the glottal fricative posture, breathy voiced phonation would result; in the modal prephonation state, modal phonation would occur (with increasingly breathy phonation as less medialization is applied); and with the more constricted states, creakiness or harshness would occur with or without possible epilaryngeal vibration (all depending on the amount of applied subglottal pressure). The results of all of these simulations are discussed in turn in the remainder of this section.

One of the primary uses of a biomechanical model of a vocal tract structure is to reveal aspects of that structure's biomechanics that are difficult or impossible to measure directly. The property of interest in the present article is quantality, exhibited in regions of biomechanical space where structural configuration is relatively insensitive to changes in muscle activity. Quantality is one property that can help increase the controllability of body structures given realistic limitations on the capacity of the central nervous system. To more objectively characterize this property in different laryngeal structures, a numeric index of quantality called the *Q-score* was developed. Using this score, quantal effects were found to be evident in simulations of all laryngeal configurations, although some were stronger than others. A summary for each simulation type is given in Table 4.

As the table details indicate, quantality in laryngeal posturing manifests primarily as tissue-on-tissue contacts. These include contact between the arytenoid apices/corniculate tubercles, contact between the body of the arytenoids, vocal fold contact, VV fold contact, ventricular fold contact, epiglottis–ventricular fold contact, and cuneiform–epiglottic tubercle contact. Each contact can be considered the locus of a potential quantal effect in relation to different combinations of muscle activations, and these in turn can be thought of as forming some of the key articulatory actions of the larynx, which either occur independently as speech sounds or are coproduced with supralaryngeal configurations to form more complex articulatory possibilities. As noted above, each of the states examined corresponds with a phonatory state, which, in addition to being subject to quantallike transitions in a

Table 4. Summary of simulations and quantal scores (rounded to the nearest hundredth) measured for the four response variables in the default extrinsic larynx posture.

Category	TF	VV	FF	AE	Description
Inspiration (full sequence)	**0.61**	**1.54**	**0.64**	—	Sudden transition of arytenoid orientation followed by linear vocal fold abduction
Inspiration (posttransition[a])	–0.00	–0.33	–0.07	—	Linear vocal fold abduction
Glottal fricative	**0.94**	**0.70**	**0.99**	—	Contact of corniculate tubercles
Modal prephonation	**0.58**	–0.52	**0.65**	—	Contact of corniculate tubercles and vocal processes
Plain glottal stop	**1.66**	–0.79	*0.42*	—	Contact of corniculate tubercles, vocal processes, and medial vocal fold surfaces; linear movement of ventricular folds
VV stop	**3.76**	**3.49**	**1.98**	—	Contact of corniculate tubercles, vocal processes, medial vocal fold surfaces, medial ventricular fold surfaces, vocal folds and ventricular folds, and ventricular folds and epiglottis
AE stop	**3.91**	**2.27**	**2.46**	**3.25**	Contact of corniculate tubercles, vocal processes, medial vocal fold surfaces, medial ventricular fold surfaces, between the vocal folds and ventricular folds, ventricular folds and epiglottis, and between the cuneiform tubercles and the epiglottic tubercle
AE fricative	**1.54**	**1.61**	**1.42**	**1.09**	Contact of corniculate tubercles and between the cuneiform tubercles and the epiglottic tubercle

Note. Q-score numbers: underlined bold font = strongly quantal; bold font = moderately quantal; italic font = mildly quantal; plain font = non-quantal. TF = true (vocal) fold; VV = VV = vocal fold to ventricular fold; FF = false (ventricular) fold; AE = aryepiglotto-epiglottal em dashes = not applicable.

[a]Values were obtained by computing quantality scores after the sudden transition of arytenoid orientation.

phonatory/vibratory state, may be presumed to inherit the quantality of the related static (i.e., nonphonatory) articulatory state.

A point of discussion is that the biomechanical quantal effects discussed here and in articles such as that by Gick et al. (2011) appear one sided compared with the sigmoidal patterns used by Stevens (1989) in his illustration of quantal theory. The simulations presented here range only over a single bundle of muscle activation ratios. It would be expected (indeed, predicted) that, were two different such ratios to be systematically varied in increments, the more familiar sigmoidal pattern (i.e., with two plateaus separated by a phase of rapid transition) would be observed. Likewise, if one were to observe an articulatory–acoustic quantal effect of the sort discussed by Stevens from only one end of the parameter range, the effect would remain quantal in nature. After all, quantality relates to the basic idea of nonlinearity in the mapping between input parameter and response. Some regions of the input show rapid response, and other regions exhibit stability. This one-sidedness is, therefore, just an artefact of the method used to examine the behavior in the model as individual simulations.

Judging by the Q-scores (and the associated interpretive regions), those phonetic states with comparatively lower scores, such as glottal fricative and modal prephonation, are actually much more commonly attested in the world's languages than those with the highest scores, particularly the AE states. The one exception here was noted for glottal stop, which seems to occur much more commonly with some degree of adduction of the ventricular folds, often to the point of their complete medialization. What this suggests is that phonological systems clearly do not optimize solely for quantality. A more plausible scenario is that multiple factors are involved, not all of which are biomechanical—for instance, perceptual distinctiveness. Keeping the discussion limited to biomechanical factors, it seems reasonable that properties, such as those associated with the notion of ease of articulation (Napoli, Sanders, & Wright, 2014), including speed and metabolic demands, might contribute to the cost of diachronic maintenance of a particular speech sound and contribute to the likelihood of it arising through sound change. Stavness, Gick, Derrick, and Fels (2012) argued that biomechanical factors such as volume displacement, relative strain, and relative muscle stress (which all could be considered aspects of articulatory ease) help account for the selection of preferential North American English /r/ variants used in particular vowel contexts.

Similar reasoning might be applied to the cases examined in the present article. Those articulatory states with lower Q-scores are still at least mildly quantal in nature but also benefit from requiring less overall muscular action and less mass displacement. The greater frequency of occurrence of VV stop might demonstrate that quantality is preferable when other biomechanical cost differences are marginal, and, furthermore, perceptual factors are not likely to differ much between plain and ""reinforced" (VV)" glottal stop. The cost boundary between the relatively frequent VV stop and the rather uncommon AE stop or fricative (in association with pharyngeal consonants) may relate to the added costs arising from tongue retraction (not simulated here but known to commonly occur) and hyolaryngeal approximation associated with these latter states.

The fact that some measures exhibit nonquantal responses to muscle activity is worth considering. Similar results were found for simulations of the oropharyngeal isthmus (Gick et al., 2014), indicating that only certain combinations of muscle activations gave rise to quantal effects. Such findings demonstrate that quantal effects are not a trivial finding. That is, it is not simply that quantal effects are observable for any given arbitrary action of a

computational biomechanical model. Rather, they arise only in specific circumstances and for reasons that are typically manifest (e.g., contact between opposing structures). It is also acknowledged that graded posturing governed by the relative activation of agonist–antagonist muscle groups can further operate above the quantal effects to provide more nuanced control. Moreover, the quantal effects observed here are associated with very specific phonetic articulatory states of the larynx, most of which are well known to have important linguistic functions. Biomechanical quantal effects are thus interpretable as providing a firm articulatory foundation upon which further refinements are always possible but may invoke muscle activity compositions beyond what was simulated here.

In any case, simulations with more linear behavior illustrate that some aspects of speech-related biomechanics might benefit from having more linear action, possibly requiring the more delicate balancing of agonist–antagonist muscle forces and perhaps with the assistance of cortical feedback. Along these lines, Buchaillard et al. (2009) suggested that, despite the stabilizing effect of palatal contact in the production of /i/, genioglossus anterior (in opposition to genioglossus posterior) action provides control over lingual grooving that operates without a saturation effect and requires delicate control. For the larynx, it seems plausible that pitch might operate similarly and be influenced by agonist–antagonist relationships for delicate and nuanced control. In the context of the simulations presented here, if inspiration is compared to expiration (which was associated with glottal fricative, a phonetic function), it is evident that inspiration has more linear than quantal operation. The simulation results indicate that although the mechanism of initiating inspiration is quantal, the variation in the degree of opening is linear. Such differentiation in the biomechanics of these two states (inspiration vs. expiration) possibly reflects phylogenetically deep aspects of laryngeal physiology. Inspiration requires varying quantities of air in relation to the widely varying metabolic demands of different activities, but because gas exchange occurs at a fixed rate, expiration requires a stable configuration to check the flow of air and maintain the respiratory cycle (Negus, 1929). Expiration is used as a basis for a phonetic state with phonological applications (e.g., glottal fricative), whereas inspiration is not; the interpretation here then is that the physical posture of the (deep or forced) inspiration state (ignoring airflow direction) does not provide a stable biomechanical basis for speech sound production.

Apart from identifying those aspects of laryngeal biomechanics that are well suited to modular control by exhibiting stable biomechanics (i.e., quantal effects), QL2 also allows another important aspect of speech biomechanics to be explored—that of functional interactions. Examples of such interactions are segmental coarticulation and the effect that voice quality (i.e., long-term articulatory settings of the vocal tract, in the sense of Laver, 1980) is thought to have on the articulation of individual segments. This aspect of articulation should not be overlooked because it relates to how speech sounds might be influenced by—and might even come to change over time in—certain phonetic contexts. The QL2 simulations demonstrate that hyo-laryngeal configuration influences the degree of laryngeal constriction. A good example is found with the VV distance measure in relation to the constricting and expanding extrinsic larynx postures: The constricting adjustment causes a reduction in the height of the ventricle, and expanding has the opposite effect. In articulatory state simulations I to V, the raised posture consistently showed the greatest amount of vocal fold and ventricular fold separation (as indicated by TF and FF distances). Thus, QL2 demonstrates that laryngeal constriction is facilitated through hyo-laryngeal approximation. On the other hand, laryngeal anticonstriction (expansion) is facilitated by the opposite effects of hyo-laryngeal separation. Such patterns are consistent with observations from theoretical models (Esling, 2005; Moisik, 2013) concerning laryngeal articulation and have implications for phonological patterns associated with laryngeal and pharyngeal sounds.

In the spirit of Moisik and Esling (2011), the QL2 model underscores that the speech functioning of the larynx must be considered from a whole-larynx perspective, meaning that the reductive view of viewing laryngeal function as mediated by essentially a one-dimensional glottis (Halle & Stevens, 1971; Ladefoged, 1971) is simply not tenable. What is needed is a more holistic approach to speech-related laryngeal biomechanics that sees the larynx—both its intrinsic and extrinsic components—as intimately connected to those surrounding structures, including the tongue, jaw, pharynx, trachea, and cervical spine. QL2, however, represents only one small advancement toward this larger whole-larynx goal of understanding how the larynx interacts with neighboring components of the vocal tract. The next step will be to simulate laryngeal behavior with a tongue, a pharynx, a mobile cervical spine, and a trachea and then to simulate laryngeal vibration under conditions of an expanded anatomical representation. These projects are currently underway.

Although this is the second iteration of the larynx model in ArtiSynth, it still is subject to a number of limitations. First, although the parameter settings have been determined to fall within reasonable, physiologically appropriate ranges, the model represents a vast number of choices that needed to be made to set these parameters. Even though these parameters were set with great care in ensuring that they produced a model with sufficient but not excessive flexibility, in many cases they represent informed estimations at best. Future versions of the model need to continue the effort to find better approximations to these parameters. This is especially the case for the ligaments (and membranes) and muscles. Although this does not greatly influence the broad details of the simulation of static configurations, especially where the muscles are concerned, such parameters do have a large effect on the temporal response of the model. Another major issue is the handling of collision. In particular, no collision handling is specified for the cartilages. This is particularly important with regard to the behavior of the arytenoids and the cricoarytenoid joints. Another weakness concerns the general coarseness of the laryngeal mucosa and the lack

of deformability of the laryngeal cartilages. This made it difficult to simulate in fine detail the contact occurring between the corniculate tubercles, which in real life undergo considerable deformation—a property that has been attributed considerable functional significance by Fink (1974b). Last, it must be acknowledged that although ArtiSynth represents a significant achievement in terms of numerical stability combined with efficient computation of such complex dynamics, there are still many cases that simply cannot be simulated because of numerical instability or outright crashes. Such issues were discovered when attempting to collide several subparts of the laryngeal mucosa together (as in the AE simulations). Further advancements in stability must be in place before the prospect of studying even more complex interactions (e.g., the interaction between the tongue and the larynx or between the tongue, larynx, and pharynx) becomes fully feasible.

The present study strongly depends on Stevens' (1989) seminal work on quantal theory, which identified numerous laryngeal quantalities primarily focusing on the nonlinear relationships between glottal aperture and phonatory state. QL2 indicates that there is a wide range of quantal biomechanical–articulatory effects that are advantageous to a modular neuromuscular system for speech production, both for laryngeal control and more generally. This is because such quantal effects help reduce variability in positioning of articulators and thereby reduce the need for feedback-based control.

Acknowledgments

We gratefully acknowledge the funding support of an Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant to the second author. We acknowledge John Esling for generously giving us permission to use the States of the Glottis videos and for his conceptual and experimental contributions to the field that have made this research possible. We also offer our thanks to Sid Fels, Ling Tsou, Ian Stavness, Peter Anderson, and John Lloyd for their support with ArtiSynth and in conducting this research.

References

Bartlett, D., Remmers, J. E., & Gautier, H. (1973). Laryngeal regulation of respiratory airflow. *Respiration Physiology, 18,* 194–204. doi:10.1016/0034-5687(73)90050-9

Berger, D. J., Gentner, R., Edmunds, T., Pai, D. K., & d'Avella, A. (2013). Differences in adaptation rates after virtual surgeries provide direct evidence for modularity. *Journal of Neuroscience, 33,* 12384–12394. doi:10.1523/JNEUROSCI.0122-13.2013

Bernstein, N. (1967). *The co-ordination and regulation of movements.* London, United Kingdom: Pergamon.

Berry, D. A., Montequin, D. W., Chan, R. W., Titze, I. R., & Hoffman, H. T. (2003). An investigation of cricoarytenoid joint mechanics using simulated muscle forces. *Journal of Voice, 17,* 47–62. doi:10.1016/S0892-1997(03)00026-2

Brunelle, M., Nguyễn, D. D., & Nguyễn, K. H. (2010). A laryngographic and laryngoscopic study of Northern Vietnamese tones. *Phonetica, 67,* 147–169.

Buchaillard, S., Perrier, P., & Payan, Y. (2009). A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. *The Journal of the Acoustical Society of America, 126,* 2033–2051.

Catford, J. C. (1964). Phonation types: The classification of some laryngeal components of speech production. In D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott, & J. L. M. Trim (Eds.), *In Honour of Daniel Jones* (pp. 26–37). London, United Kingdom: Longmans, Green & Co.

d'Avella, A., Giese, M., Ivanenko, Y. P., Schack, T., & Flash, T. (2015). Modularity in motor control: From muscle synergies to cognitive action representation. *Frontiers in Computational Neuroscience, 9,* 126. doi:10.3389/fncom.2015.00126

Edmondson, J. A., Chang, Y., Hsieh, F., & Huang, H. J. (2011). Reinforcing voiceless finals in Taiwanese and Hakka: Laryngoscopic case studies. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 627–630). Hong Kong: City University of Hong Kong.

Edmondson, J. A., & 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.

Edmondson, J. A., Esling, J. H., Harris, J. G., & Wei, J. (2004). A phonetic study of the Sui consonants and tones. *Mon-Khmer Studies, 34,* 47–66.

England, S. J., Bartlett, D., & Daubenspeck, J. A. (1982). Influence of human vocal cord movements on airflow and resistance during eupnea. *Journal of Applied Physiology, 52,* 773–779.

Esling, J. H. (1996). Pharyngeal consonants and the aryepiglottic sphincter. *Journal of the International Phonetic Association, 26,* 65–88.

Esling, J. H. (1999). The IPA categories "pharyngeal" and "epiglottal": Laryngoscopic observations of pharyngeal articulations and larynx height. *Language and Speech, 42,* 349–372.

Esling, J. H. (2005). There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics, 50,* 13–44.

Esling, J. H., Fraser, K. E., & Harris, J. G. (2005). Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics, 33,* 383–410.

Esling, J. H., & Harris, J. G. (2003). *States of the glottis: An articulatory phonetic model based on laryngoscopic observations*. Retrieved from http://web.uvic.ca/ling/research/phonetics/SOG/index.htm

Esling, J. H., & Harris, J. G. (2005). States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In W. J. Hardcastle & J. M. Beck (Eds.), *A figure of speech: A Festschrift for John Laver* (pp. 347–383). Mahwah, NJ: Erlbaum.

Esling, J. H., & Moisik, S. R. (2012). Laryngeal aperture in relation to larynx height change: An analysis using simultaneous laryngoscopy and laryngeal ultrasound. In D. Gibbon, D. Hirst, & N. Campbell (Eds.), *Rhythm, melody and harmony in speech: Studies in honour of Wiktor Jassem* (Vol. 14/15, pp. 117–127). Poznań, Poland: Polskie Towarzystwo Fonetyczne.

Ewan, W. G., & Krones, R. (1974). Measuring larynx movement using the thyroumbrometer. *Journal of Phonetics, 2,* 327–335.

Faaborg-Andersen, K. (1957). Electromyographic investigation of intrinsic laryngeal muscles in humans. *Acta Physiologica Scandinavica, 41*(Suppl. 140), 1–150.

Faaborg-Andersen, K., & Buchthal, F. (1956). Action potentials from internal laryngeal muscles during phonation. *Nature, 177,* 340–341. doi:10.1038/177340a0

Fink, B. R. (1974a). Folding mechanism of the human larynx. *Acta Oto-Laryngologica, 78,* 124–128.

Fink, B. R. (1974b). Spring mechanisms in the human larynx. *Acta Oto-Laryngologica, 77,* 295–304.
Fujimura, O. (1989). Comments on "On the quantal nature of speech," by K. N. Stevens. *Journal of Phonetics, 17,* 87–90.
Garellek, M. (2013). *Production and perception of glottal stops* (Doctoral dissertation). University of California, Los Angeles.
Gick, B., Anderson, P., Chen, H., Chiu, C., Kwon, H. B., Stavness, I., . . . Fels, S. (2014). Speech function of the oropharyngeal isthmus: A modelling study. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2,* 217–222.
Gick, B., & Stavness, I. (2013). Modularizing speech. *Frontiers in Cognitive Science*. doi:10.3389/fpsyg.2013.00977
Gick, B., Stavness, I., Chiu, C., & Fels, S. S. (2011). Categorical variation in lip posture is determined by quantal biomechanical-articulatory relations. *Canadian Acoustics, 39,* 178–179.
Halle, M., & Stevens, K. N. (1971). A note on laryngeal features. *MIT Quarterly Progress Report, 101,* 198–212.
Heselwood, B. (2007). The "tight approximant" variant of the Arabic 'ayn. *Journal of the International Phonetic Association, 37*(1), 1–32.
Hillel, A. D. (2001). The study of laryngeal muscle activity in normal human subjects and in patients with laryngeal dystonia using multiple fine-wire electromyography. *The Laryngoscope, 111*(Suppl. 97), 17–47.
Hirano, M., & Ohala, J. J. (1969). Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. *Journal of Speech, Language, and Hearing Research, 12,* 362–373.
Hirano, M., & Sato, K. (1993). *Histological color atlas of the human larynx*. San Diego, CA: Singular.
Hollien, H. (1974). On vocal registers. *Journal of Phonetics, 2,* 125–143.
Honda, K., Hirai, H., Masaki, S., & Shimada, Y. (1999). Role of vertical larynx movement and cervical lordosis in f0 control. *Language and Speech, 42,* 401–411.
Honda, K., Kitamura, T., Takemoto, H., Adachi, S., Mokhtari, P., Takano, S., . . . Dang, J. (2010). Visualisation of hypopharyngeal cavities and vocal-tract acoustic modelling. *Computer Methods in Biomechanics and Biomedical Engineering, 13,* 443–453.
Honda, K., Takemoto, H., Kitamura, T., Fujita, S., & Takano, S. (2004). Exploring human speech production mechanisms by MRI. *IEICE Transactions on Information and Systems, 87–D,* 1050–1058.
Hunter, E. J., Titze, I. R., & Alipour, F. (2004). A three-dimensional model of vocal fold abduction/adduction. *The Journal of the Acoustical Society of America, 115,* 1747–1759. doi:10.1121/1.1652033
Iwata, R., Sawashima, M., Hirose, H., & Niimi, S. (1979). Laryngeal adjustment of Fukienese stops: Initial plosives and final applosives. *Annual Bulletin Research Institute of Logopedics and Phoniatrics, 13,* 61–81.
Kagaya, R., & Hirose, H. (1975). Fiberoptic, electromyographic and acoustic analyses of Hindi stop consonants. *Annual Bulletin Research Institute of Logopedics and Phoniatrics, 9,* 27–46.
Knudson, D. (2007). *Fundamentals of biomechanics*. New York, NY: Springer.
Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. Chicago, IL: University of Chicago Press.
Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Cambridge, MA: Blackwells.
Laver, J. (1980). *The phonetic description of voice quality*. Cambridge, MA: Cambridge University Press.
Lindqvist-Gauffin, J. (1972). *A descriptive model of laryngeal articulation in speech*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.447.9782&rep=rep1&type=pdf
Lloyd, J. E., Stavness, I., & Fels, S. S. (2012). ArtiSynth: A fast interactive biomechanical modeling toolkit combining multibody and finite element simulation. In Y. Payan (Ed.), *Soft tissue biomechanical modeling for computer assisted surgery* (pp. 355–394). Berlin, Germany: Springer.
Moisik, S. R. (2008). *A three-dimensional model of the larynx and the laryngeal constrictor mechanism: Visually synthesizing pharyngeal and epiglottal articulations observed in laryngoscopy* (Master's thesis). University of Victoria, British Columbia, Canada.
Moisik, S. R. (2013). *The epilarynx in speech* (Doctoral dissertation). University of Victoria, British Columbia, Canada.
Moisik, S. R., & Esling, J. H. (2011). The "whole larynx" approach to laryngeal features. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1406–1409). Hong Kong: City University of Hong Kong.
Moisik, S. R., & Esling, J. H. (2014). Modeling the biomechanical influence of epilaryngeal stricture on the vocal folds: A low-dimensional model of vocal–ventricular fold coupling. *Journal of Speech, Language, and Hearing Research, 57,* S687–S704.
Moisik, S. R., Esling, J. H., Crevier-Buchman, L., Amelot, A., & Halimi, P. (2015). Multimodal imaging of glottal stop and creaky voice: Evaluating the role of epilaryngeal constriction. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0247.pdf. Glasgow, United Kingdom: The University of Glasgow.
Moisik, S. R., & Gick, B. (2013). The quantal larynx revisited. *The Journal of the Acoustical Society of America, 133,* 3522. doi:10.1121/1.4806322
Moisik, S. R., Lin, H., & Esling, J. H. (2014). A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS). *Journal of the International Phonetic Association, 44,* 21–58.
Napoli, D. J., Sanders, N., & Wright, R. (2014). On the linguistic effects of articulatory ease, with a focus on sign languages. *Language, 90,* 424–456. doi:10.1353/lan.2014.0026
National Institutes of Health and U.S. National Library of Medicine. (2009). *The Visible Human Project*. Retrieved from http://www.nlm.nih.gov/research/visible/visible_human.html
Nazari, M. A., Perrier, P., Chabanas, M., & Payan, Y. (2010). Simulation of dynamic orofacial movements using a constitutive law varying with muscle activation. *Computer Methods in Biomechanics and Biomedical Engineering, 13,* 469–482. doi:10.1080/10255840903505147
Nazari, M. A., Perrier, P., Chabanas, M., & Payan, Y. (2011). Shaping by stiffening: A modeling study for lips. *Motor Control, 15,* 141–168.
Negus, V. E. (1929). *The mechanism of the larynx*. London, United Kingdom: Heinemann.
Negus, V. E. (1949). *The comparative anatomy and physiology of the larynx*. London, United Kingdom: Heinemann.
Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics, 25,* 382–407.
Reidenbach, M. M. (1997). Anatomical considerations of closure of the laryngeal vestibule during swallowing. *European Archives of Oto-Rhino-Laryngology, 254,* 410–412.
Reidenbach, M. M. (1998a). Aryepiglottic fold: Normal topography and clinical implications. *Clinical Anatomy, 11,* 223–235.
Reidenbach, M. M. (1998b). The muscular tissue of the vestibular folds of the larynx. *European Archives of Oto-Rhino-Laryngology, 255,* 365–367.

Roach, P. J. (1979). Laryngeal-oral coarticulation in glottalized English plosives. *Journal of the International Phonetic Association, 9,* 2–6.

Safavynia, S. A., & Ting, L. H. (2012). Task-level feedback can explain temporal recruitment of spatially fixed muscle synergies throughout postural perturbations. *Journal of Neurophysiology, 107,* 159–177. doi:10.1152/jn.00653.2011

Sauerland, E. K., & Harper, R. M. (1976). The human tongue during sleep: Electromyographic activity of the genioglossus muscle. *Experimental Neurology, 51,* 160–170. doi:10.1016/0014-4886(76)90061-3

Schwab, R. J., Gefter, W. B., Pack, A. I., & Hoffman, E. A. (1993). Dynamic imaging of the upper airway during respiration in normal subjects. *Journal of Applied Physiology, 74,* 1504–1514.

Schwartz, J.-L., Boë, L.-J., Vallée, N., & Abry, C. (1997). The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics, 25,* 255–286. doi:10.1006/jpho.1997.0043

Stavness, I., Gick, B., Derrick, D., & Fels, S. S. (2012). Biomechanical modeling of English /r/ variants. *The Journal of the Acoustical Society of America, 131,* EL355–EL360.

Stavness, I., Lloyd, J. E., Payan, Y., & Fels, S. S. (2011). Coupled hard-soft tissue simulation with contact and constraints applied to jaw-tongue-hyoid dynamics. *International Journal for Numerical Methods in Biomedical Engineering, 27,* 367–390.

Stavness, I., Nazari, M. A., Perrier, P., Demolin, D., & Payan, Y. (2013). A biomechanical modeling study of the effects of the orbicularis oris muscle and jaw posture on lip shape. *Journal of Speech, Language, and Hearing Research, 56,* 878–890. doi:10.1044/1092-4388(2012/12-0200)

Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In P. B. Denes & E. E. David, Jr. (Eds.), *Human communication: A unified view* (pp. 51–66). New York, NY: McGraw Hill.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17,* 3–45.

Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics, 38,* 10–19.

Thumfart, W. F., Platzer, W., Gunkel, A. R., Mauer, H., & Brenner, E. (1999). *Surgical Approaches in otorhinolaryngology*. Stuttgart, Germany: Thieme.

Ting, L. H., Chiel, H. J., Trumbower, R. D., Allen, J. L., McKay, J. L., Hackney, M. E., & Kesar, T. M. (2015). Neuromechanical principles underlying movement modularity and their implications for rehabilitation. *Neuron, 86,* 38–54. doi:10.1016/j.neuron.2015.02.042

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

Todorov, E., & Jordan, M. I. (2003). A minimal intervention principle for coordinated movement. *Advances in Neural Information Processing Systems, 15,* 27–34.

Von Leden, H., & Moore, P. (1961). The mechanics of the cricoarytenoid joint. *Archives of Otolaryngology, 81,* 616–625.

Wickham, J. B., & Brown, J. M. (1998). Muscles within muscles: The neuromotor control of intra-muscular segments. *European Journal of Applied Physiology and Occupational Physiology, 78,* 219–225. doi:10.1007/s004210050410

Wilson, I. (2007). The effects of post-velar consonants on vowels in Nuu-chah-nulth: Auditory, acoustic, and articulatory evidence. *The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique, 52,* 43–70.

Zander, T., Rohlmann, A., & Bergmann, G. (2004). Influence of ligament stiffness on the mechanical behavior of a functional spinal unit. *Journal of Biomechanics, 37,* 1107–1111. doi:10.1016/j.jbiomech.2003.11.019

Zemlin, W. R. (1998). *Speech and hearing science: Anatomy and physiology* (4th ed.). Boston, MA: Allyn & Bacon.

Zemlin, W. R., Davis, P., & Gaza, C. (1984). Fine morphology of the posterior cricoarytenoid muscle. *Folia Phoniatrica, 36,* 233–240.

Appendix (p. 1 of 2)

Biomechanical Model Details

The laryngeal cartilages and hyoid bone models were developed specifically for QL2 and fitted to a pre-existing model of the jaw–tongue–hyoid complex (see Nazari, Perrier, Chabanas, & Payan, 2010; Stavness, Lloyd, Payan, & Fels, 2011; Stavness, Nazari, Perrier, Demolin, & Payan, 2013). The point of registration between the two models was thus the hyoid, and the pre-existing hyoid was replaced with the one developed for QL2. These structures were connected together with point-to-point axial connections (i.e., connectors that transmit force only along their axis) representing both musculature and membranous and ligamentous attachments (see Figure 2), all of which was determined by reference to various anatomical sources (e.g., Hirano & Sato, 1993; Thumfart, Platzer, Gunkel, Mauer, & Brenner, 1999; Zemlin, 1998).

Ligaments and membranes behave as simple springs, passively generating a restoring tension force when stretched or compressed beyond optimum length and a damping force in response to change in length. The scaling of the stiffness and damping of these structures was accomplished primarily experimentally, with initial values determined and then fine-tuned to provide appropriate levels of constraint on structural movements under exploratory muscle contractions on the basis of sources found in the literature (Buchaillard et al., 2009; Honda, Takemoto, Kitamura, Fujita, & Takano, 2004; Titze, 2006; Zemlin, 1998; also see Moisik, 2008). It was not possible to map empirical measurements of stiffness and damping to all ligaments and membranes used in the model, but the values fall within physiological ranges (for stiffness, 1–5000 N m^{-1}; for damping, ζ ranges from 0.1–1.0) for the structures in question (e.g., for the cricoarytenoid joint, see Berry, Montequin, Chan, Titze, & Hoffman, 2003) and for ligaments and membranes in general (e.g., Zander, Rohlmann, & Bergmann, 2004). Muscles are similar but can also generate an active contractile force (more details are provided below). The thyrohyoid membrane was treated as a lattice work of axial connections spanning the upper edge of the thyroid lamina and the lower edge of the greater cornua of the hyoid bone. Special attention was given to the cricoarytenoid joints, which were modeled using a collection of axial connections approximating the structure of the connective tissue joining the cricoid and arytenoid cartilages (Von Leden & Moore, 1961). The cricothyroid joint was modeled as a revolute joint and set to be compliant enough to allow some small translational displacement. Full planar (midsagittal) constraints were applied to the epiglottis, mandible, and hyoid rigid bodies.

Medical image segmentation software (Amira; FEI, Hillsboro, OR) was used to segment the mucosa and cartilages of the larynx and the hyoid bone taken from a set of axial-plane images of a cryosectioned neck of an adult male in the Visible Human data set (obtained with permission from the National Institutes of Health and U.S. National Library of Medicine, 2009). The lower extent of the mucosa mesh is found a short distance below the apex of the blade of the cricoid cartilage, and the upper extent is found at the inferior rim of the hyoid bone and includes the epiglottic mucosa. The segmentations were then converted using Amira into a set of surface meshes. Blender (http://blender.org) was then used to refine and symmetrize the raw meshes, and these were exported into ArtiSynth. Within ArtiSynth, the larynx mucosa surface mesh was converted into a hexahedral FEM mesh (see Figure 1, top row) using an in-house semiautomatic gridding-and-projection algorithm with a resolution of 2.5×10^{-3} m; manual adjustments were made following this to further improve the quality of the mesh in order to increase simulation stability. This coarseness of the mucosa mesh was chosen to provide a satisfactory tradeoff between deformability of the model and computational cost. Finer discretizations can yield more agile, deformable models but quickly become so computationally demanding that they are unfeasible for simulating on conventional systems.

Structural mass was determined by specifying tissue density and multiplying by volume of the structure in question (handled internally by the ArtiSynth engine). The density used for the laryngeal cartilages was estimated to be 1900 kg m^{-3}, a small amount below that used in Stavness et al. (2011; also see Buchaillard et al., 2009), which was used for bone. The mucosa had a density of 1040 kg m^{-3}, following the value used for the tongue in those sources just mentioned.

Collision computation is expensive and often inaccurate because of its discontinuous nature. Thus, the computation of collision physics is handled as sparsely as possible. ArtiSynth provides a collision manager object that enables collision behavior among its interacting components to be specified in detail, including specifying whether collisions are processed for a given component, which components can collide, and, by means of collision submeshes (descriptors of which surface faces and nodes of an FEM are eligible for collision computation), what parts of deformable components can collide. Collision submeshes were defined over the vocal folds, ventricular folds, and aryepiglottic folds and in the region of the epiglottic tubercle. Stavness et al. (2011) gives more details on how collision is computed within ArtiSynth.

An approximate-containment approach was used to connect the laryngeal cartilages to the mucosa. FEM nodes of the mucosa were set to be rigidly attached to these cartilages if they were either contained within the corresponding meshes or very close to the surface (no more than 2 mm away, although this varies by cartilage). The epiglottic, cuneiform, and arytenoid cartilages (rigid bodies) are all entirely embedded within the mucosa mesh. The larynx mucosa FEM uses an incompressible Mooney–Rivlin material with parameters similar to those used for the face (Nazari et al., 2011) and tongue (Buchaillard et al., 2009).

The FEM laryngeal mucosa is influenced by a set of muscles that include point-to-point muscles connected to the rigid-body framework (the interarytenoids, lateral cricoarytenoids, and posterior cricoarytenoids) and a set of FEM-intrinsic muscles (see Figure 1, bottom row), which includes the thyroarytenoid, thyroepiglottic, and ventricularis muscles. It was deemed necessary to represent these using FEM-intrinsic muscles, which can enable use of the active stiffening feature of ArtiSynth FEM models but which come at the cost of increased computational complexity. Furthermore, the lateral cricoarytenoid muscles and oblique fibers of the posterior cricoarytenoid muscles were not entirely contained by the laryngeal mucosa, and thus these were represented using FEM-extrinsic muscles. The thyroarytenoid and thyroepiglottic muscles were developed with reference to Zemlin (1998, pp. 128–129). The ventricularis (sometimes referred to as the *external thyroarytenoid*) muscles were developed with reference to the work of Reidenbach (1997, 1998a, 1998b). Aryepiglottic muscles were omitted on the basis of the lack of clear histological evidence for their existence (Reidenbach, 1998a, p. 233).

Appendix (p. 2 of 2)

Biomechanical Model Details

Model movement is generated through muscle activation, which requires sending an excitation signal (ranging from 0% for no contraction to 100% for maximum contraction) to a given muscle, informing it to begin to contract. All muscles are point-to-point axial muscles, which generate an active contractile force, a passive recoil force when extended past the optimum length, and a damping force in response to change in length, all of which is directed along the muscle axis. This approach to modeling muscles follows the approach taken with other models developed in ArtiSynth (e.g., Gick et al., 2014; Stavness et al., 2011). Because no attempt was made to segment individual muscles from the Visible Human data, it was not possible to estimate muscle force scaling directly from the cross-sectional areas of muscles. Instead, muscle force scaling was heuristically determined by gauging the stability of QL2 under exploratory contraction. Force scaling varies by muscle, but the values are within a physiologically normal range (0.5–5.0 N). This solution to muscle force scaling was deemed acceptable for three reasons. First, because a forward approach to simulation was used (meaning that muscle action was explicitly adjusted to visually achieve certain target configurations rather than computed via inverse model), exactness in muscle scaling was less important than the flexibility in posturing afforded by a given muscle force scaling. Second, the force scaling primarily influences the behavior over time of a given muscle; however, the simulations do not demand a high degree of accuracy in the temporal behavior of muscle contraction. Last, although every effort was made to make the model as anatomically and physiologically realistic as possible, the model is still a highly idealized representation of the actual system of the hyo-laryngeal complex, and the model should be interpreted as such.

References

Berry, Montequin, Chan, Titze, & Hoffman (2003)
Buchaillard, Perrier, & Payan (2009)
Gick et al. (2014)
Hirano & Sato (1993)
Honda, Takemoto, Kitamura, Fujita, & Takano (2004)
Moisik (2008)
National Institutes of Health and U.S. National Library of Medicine (2009)
Nazari, Perrier, Chabanas, & Payan (2010)
Nazari, Perrier, Chabanas, & Payan (2011)
Reidenbach (1997)
Reidenbach (1998a)
Reidenbach (1998b)
Stavness, Lloyd, Payan, & Fels (2011)
Stavness, Nazari, Perrier, Demolin, & Payan (2013)
Thumfart, Platzer, Gunkel, Mauer, & Brenner (1999)
Titze (2006)
Von Leden & Moore (1961)
Zander, Rohlmann, & Bergmann (2004)
Zemlin (1998)