JASA ARTICLE

Register transitions in an *in vivo* canine model as a function of intrinsic laryngeal muscle stimulation, fundamental frequency, and sound pressure level

Patrick Schlegel,[a] David A. Berry, Clare Moffatt, Zhaoyan Zhang, and Dinesh K. Chhetri
Department of Head and Neck Surgery, David Geffen School of Medicine at the University of California-Los Angeles, Los Angeles, California 90095, USA

ABSTRACT:
Phonatory instabilities and involuntary register transitions can occur during singing. However, little is known regarding the mechanisms which govern such transitions. To investigate this phenomenon, we systematically varied laryngeal muscle activation and airflow in an *in vivo* canine larynx model during phonation. We calculated voice range profiles showing average nerve activations for all combinations of fundamental frequency (F0) and sound pressure level (SPL). Further, we determined closed-quotient (CQ) and minimum-posterior-area (MPA) based on high-speed video recordings. While different combinations of muscle activation favored different combinations of F0 and SPL, in the investigated larynx there was a consistent region of instability at about 400 Hz which essentially precluded phonation. An explanation for this region may be a larynx specific coupling between sound source and subglottal tract or an effect based purely on larynx morphology. Register transitions crossed this region, with different combinations of cricothyroid and thyroarytenoid muscle (TA) activation stabilizing higher or lower neighboring frequencies. Observed patterns in CQ and MPA dependent on TA activation reproduced patterns found in singers in previous work. Lack of control of TA stimulation may result in phonation instabilities, and enhanced control of TA stimulation may help to avoid involuntary register transitions, especially in the singing voice.
 https://doi.org/10.1121/10.0025135
(Received 12 October 2023; revised 9 January 2024; accepted 16 February 2024; published online 18 March 2024)
[Editor: James F. Lynch] Pages: 2139–2150

I. INTRODUCTION

In singing, especially with the high tenor voice, there is always a risk of phonatory instabilities and involuntary register transitions. That is, sudden, unexpected jumps or drops in the fundamental frequency (F0) may occur, along with associated changes in the phonation quality, due to a small change in some phonatory input parameters (Titze *et al.*, 1994). While a singer naturally strives to avoid these unexpected anomalies during performance, this can sometimes be challenging, as there are multiple contributing factors to this phenomenon.

One potential mechanism contributing to vocal instability is source-filter coupling (Titze, 2008; Herbst *et al.*, 2024). For example, one factor which can affect sound source stability is the configuration of the supraglottal tract. Titze *et al.* (2008) investigated the occurrence of voice source instabilities in human subjects performing singing tasks. During systematic increasing or lowering of F0, a higher rate of these instabilities was observed when the F0 crossed over the first formant frequency (F1) of the vocal tract. This effect was explained with source–tract coupling. A similar study was performed by Wade *et al.* (2017), where singers shifted F1 across constant F0 or F0 across constant F1. They found that instabilities were less frequent when singers were allowed to change lip geometry to alter F1, avoiding resonance crossovers.

Another factor are subglottal resonances. It has been hypothesized that many instances of partially chaotic vocal fold oscillatory behavior, or sudden changes in vibratory mode in *ex vivo* experiments (Berry *et al.*, 1996), may be caused by a strong coupling between subglottal tract and vocal fold oscillations (Zhang *et al.*, 2006). Zhang *et al.* (2006) varied subglottal tract length in a physical model of vocal fold vibration between 17 and 325 cm. They found a pattern of instabilities at onset, with stable and unstable onsets alternating with increasing vocal tract lengths. This pattern persisted even in the longest vocal tracts.

Naturally, laryngeal anatomy and morphology of the larynx itself are also of great importance. In particular, vocal fold medial surface thickness has long been considered an important factor in regulating voice quality at different vocal registers (Zhang, 2024), with chest and falsetto voices often produced with thick and thin vocal folds, respectively. For instance, some singers may have a morphology that favors a deeper and more "chest-like" voice, e.g., due to a thicker vocal fold mucosa (Herbst, 2020). In general, variation in laryngeal morphology can be considerable and correlated with different types of the voice produced (Herbst, 2020; Zarachi *et al.*, 2021).

Thus, in addition to laryngeal anatomy, register transitions can also be affected by laryngeal muscles modulating

[a]Email: patrickschlegel93@yahoo.de

the laryngeal configuration (Herbst *et al.*, 2024; Zhang, 2024): among others, the cricothyroid muscle (CT), the lateral cricoarytenoid and interarytenoid muscles (LCA+IA), and the thyroarytenoid muscle (TA). The CT is activated by the superior laryngeal nerve and elongates the vocal folds. By activating the CT during phonation, higher fundamental frequencies (F0) can be achieved. LCA+IA are activated by the recurrent laryngeal nerve (RLN) and adduct the posterior cartilaginous portion of the vocal folds, thus creating a standard phonatory posture. The TA muscle is also activated by a branch of the RLN and constitutes the bulk of the vocal folds. Its activation leads to medial surface bulging of the vocal folds which facilitates vocal fold thickening and closure of the membranous portion of the vocal folds. All these muscles counterbalance each other to some extent. For example, activation of the TA leads to medial surface bulging and hence shortening of the vocal folds, and counteracts CT activation that elongates the vocal folds (Zhang, 2016; Soriano and Gupta, 2021; Luegmair *et al.*, 2014). Also, the aforementioned bulging of the TA leads to mostly medial and anterior adduction, as opposed to the posterior adduction occasioned by the LCA+IA muscles. Hence complex interactions of these muscles may contribute to phonatory mode stability or instability (Chhetri and Park, 2016; Schlegel *et al.*, 2022).

It is in general premised that humans possess some degree of separate volitional control over cartilaginous and membranous vocal fold adduction (Herbst *et al.*, 2011). However, as the TA muscle is also attached to the arytenoid cartilage (Zhang, 2016) and the LCA appears to be able to rock the arytenoid cartilage anteriorly, inferiorly, and medially, it is not entirely clear to which degree this separate control is based on separate TA or LCA activation (Yin and Zhang, 2014). Herbst *et al.* investigated voluntary cartilaginous and membranous vocal fold adduction in singers by introducing "aBducted" and "aDducted" falsetto and chest registers, referring to the ab- or adducted posterior glottis. Using this approach they were able to show that voluntarily manipulating cartilaginous or membranous adduction was possible after some practice. However, measuring the underlying activations of TA or LCA that resulted in the respective posture changes was not possible (Herbst *et al.*, 2011).

There are multiple, somewhat conflicting, definitions of vocal registers (Herbst *et al.*, 2024; Hollien 1974; Roubeau *et al.*, 2009). Further, existing registers have been subdivided into subgroups as for example in the study by Herbst *et al.* in abducted and adducted sub-registers (Herbst *et al.*, 2011). For simplicity, in this work we will refer to the higher-pitched type of phonation that we observe in our *in vivo* canine model as falsetto-like and to the lower pitched phonation as chest-like.

In a previous study performed on the same dataset as this work, the effect of muscle activation level and subglottal pressure (pSub) on different phonation parameters such as F0, intensity, vocal efficiency, and sound pressure level was investigated (Chhetri and Park, 2016). In this previous work, phonatory stability was only assessed in terms of the duration of the first stable mode, from phonation onset until the first significant change in F0 at various combinations of muscle activations of the TA (5 levels), LCA/IA (8 levels), and CT (8s levels). Looking at the interaction between CT versus TA, phonatory instability was observed once TA was activated beyond threshold activation (level 1). Phonatory stability decreased with increasing TA levels, which was antagonistic to the effects of CT activation levels, such that as TA activation level increased, increased CT activation level was needed to improve stability of phonation. However, no high-speed video data were evaluated, and the nature of this counterbalancing effect was not further explored.

To help remedy this deficiency, in this work we use high-speed video data recorded from an *in vivo* canine model to investigate the occurrence of register transitions from falsetto-like to chest-like phonation as a function of CT, LCA+IA, and TA muscle activation levels. We investigate potential effects of subglottal resonances on the distribution of "stable" oscillation states. By systematically quantifying phonatory output across a wide range of neuromuscular conditions, with an intact subglottal tract, the relative independence of source and tract in differentiating vocal registers will be investigated. Further, we relate our muscle-activation based findings to the abducted and adducted falsetto and chest registers, as described in the work of Herbst *et al.* (2011) to substantiate the roles of LCA+IA and TA in cartilaginous and membranous vocal fold adduction.

Thus, with this work we want to answer the following questions: (1) Which laryngeal phonatory conditions favor specific combinations of F0 and SPL? (2) Do subglottal resonances influence this map of muscle activation states? (3) Does our systematic variation of laryngeal muscle activation reproduce the abducted and adducted falsetto and chest registers, as described by Herbst *et al.* (2011)?

II. MATERIALS AND METHODS

In this study, data collected in a previous study (Chhetri and Park, 2016) was used. While the previous study focused on the effect of laryngeal muscle activation on acoustic and aerodynamic parameters, the current study analyzes previously omitted high-speed video (HSV) data with a focus on phonatory mode stability and register transitions.

A. Canine larynx model

The setup of the experiment performed in this study has been described in detail before (Chhetri and Park, 2016; Chhetri *et al.*, 2010). Hence, only a summary is given here:

The canine was anesthetized and placed in a supine position on an operating table. The larynx was surgically exposed, eliminating any effects of the supraglottal tract on vocal fold vibration. Selective nerve branches to TA, LCA/IA, and CT muscles were connected to tripolar cuff electrodes (Ardiem Medical, Indiana, PA) for independent graded stimulation. To avoid cross-stimulation, the nerve branches to the posterior

https://doi.org/10.1121/10.0025135

cricoarytenoid (PCA) muscles and Galen's anastomosis were cut. Therefore, three muscles could be activated independently: CT, LCA+IA (LCA and IA are grouped functionally together), and TA. To achieve phonation, humidified airflow was provided from a subglottal tube connected to an airflow controller (MCS Series Mass Flow Controller, Alicat Scientific, Tucson, AZ). The subglottal tube had a diameter of 2.5 cm and a length of approximately 11 cm ranging from the vocal folds upstream to an expansion chamber. The expansion chamber had a width of 8.9 cm and a height of 10.16 cm with its inner walls lined with foam, only omitting a central path of about 2.5 cm width for the airflow. Phonation was recorded with an HSV camera (Phantom v210, Vision Research Inc., Wayne, NJ) at 3000 fps with a resolution of 512×512 pixels. A probe microphone (model 4128, Bruel & Kjaer North America, Norcross, GA) and a pressure transducer (MKS Baratron 220D, MKS Instruments, Andover, MA) mounted to the inner wall of the subglottic tube were used for recording of the acoustic and aerodynamic signals.

B. Laryngeal muscle activation combinations

In total, 320 combinations of CT, LCA+IA, and TA muscle activation were investigated (i.e., 320 different phonatory conditions). For each muscle, stimulation was performed gradually from threshold activation (muscle barely reacts) to maximum activation (muscle achieved maximal change in posture and strain). TA was stimulated at 5 levels, including no stimulation (TA0), and 4 levels from minimum activation (TA1) to maximum activation (TA4) all equidistant from each other. For each stimulation condition of TA, 64 combinations of CT and LCA+IA activation were investigated in which CT and LCA+IA were stimulated independently at 8 different levels. Similar to the TA muscle, these levels included no stimulation (CT0 / LCA+IA0) and 7 levels of minimum to maximum activation (CT1-7 / LCA+IA1-7), all equidistant from each other. For each condition, stimulation duration was 1.5 s during which airflow rate was increased linearly from 300 to 1600 ml/s. The investigated phonatory conditions are explicitly enumerated in Table I.

C. High-speed video analysis

Phonation onset was determined manually using acoustic and video data for all 320 recorded HSV videos (one for each phonatory condition). Very early onset and offset of phonation were excluded by discarding the first and last 10 ms (30 frames) of all phonation segments. Further, all datasets that did not achieve 13.3 ms of phonation (at least 10 cycles of glottis opening/closing at the lowest F0s) were discarded. This resulted in 255 datasets remaining for analysis.

TABLE I. Investigated phonatory conditions.

TA activation	CT activation	LCA+IA activation	Combinations
0	0–7	0–7	64
1	0–7	0–7	64
2	0–7	0–7	64
3	0–7	0–7	64
4	0–7	0–7	64
0–4	0–7	0–7	320 (total)

For these remaining videos, the glottal area between the vocal folds was segmented using the software tool GLOTTIS ANALYSIS TOOLS (version 2020, Friedrich Alexander University Erlangen- Nürnberg, Germany). The segmentation process (Schlegel *et al.*, 2020) and the software (Kist *et al.*, 2021) have already been described in detail and are therefore only briefly outlined, emphasizing relevant steps:

The glottis was divided in an anterior half and a posterior half. For all the videos in this study, the anterior and posterior halves were segmented separately. Segmentation conditions were kept constant in all videos, and the resultant anterior and posterior glottal area waveforms were calculated (GAW) (Soron and Lieberman, 1963) (changing anterior/posterior glottal area over time). This separation was made because for most of the subsequently calculated parameters [with the exception of the minimum posterior area (MPA)], only the anterior part of the glottis was of interest, as the posterior part partially did not oscillate, but stayed constantly open throughout the glottal cycle. The acoustic signal and pSub were recorded synchronously with the GAW. In Fig. 1(A), different states of the glottis, before and during oscillation, are shown. Figure 1(B) shows where the different signals were recorded and illustrates the segmentation process.

D. Cycle detection and parameter calculation

Cycle based parameters were calculated from different signals, but all parameters were extracted based on anterior GAW cycles, as signals were recorded synchronously, and anterior GAW based cycles were the most straightforward to detect. Maximum based GAW cycles (maximum open – closed – maximum open) were detected using a custom MATLAB algorithm. Based on these cycles the following parameters were calculated:

Local fundamental frequency (LF0): The inverse of the length of each cycle, smoothed with a moving median filter using 10 neighboring cycles.

dB-SPL: Calculated from the high-pass (50 Hz) filtered acoustic signal using the custom-made MATLAB function "spl" (Greene, 2016) with a window length of 0.1 s. The dB-SPL curve was then resampled so that one value for each GAW-based cycle could be extracted.

MPA: Calculated from the minimum of the posterior GAW for each cycle and normalized to the highest MPA value.

Closed Quotient (CQ): Calculated form the anterior GAW as closed time/cycle length. A video frame was determined as "closed" if the glottal area was equal to or less than 5% of the maximum area of the video. All parameters for all conditions and all cycles are given in the supporting information (File S1).

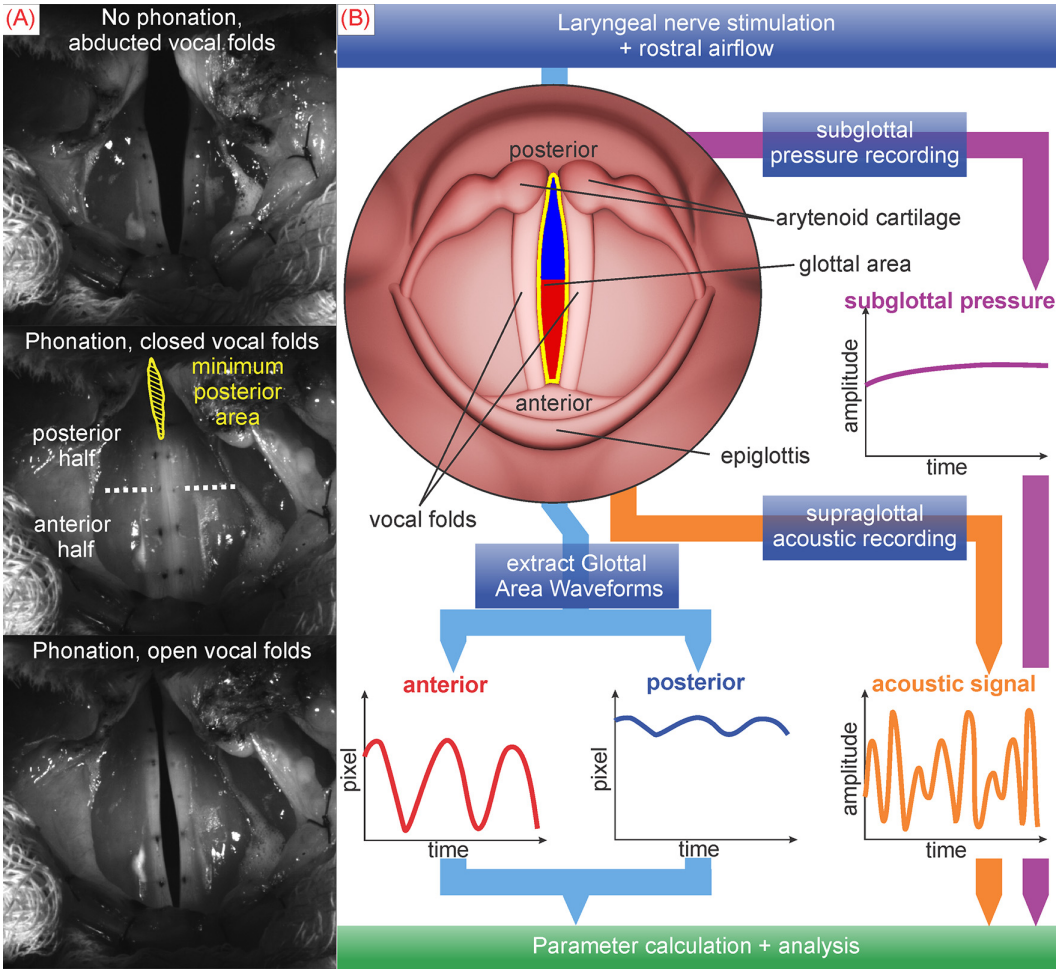

FIG. 1. (Color online) Illustration of (A) the recorded glottis during different states of vocal fold oscillation (phonatory condition TA2 LCA+IA6 CT7) and (B) the glottis segmentation separated into anterior and posterior area as well as the extracted signals.

Register transitions from falsetto-like to chest-like—i.e., changes of oscillatory mode—were detected based on fast and sudden jumps or drops in LF0 by at least 50 Hz within ten cycles. A threshold of 50 Hz was chosen as a balanced option between including clearly audible changes in the signal and excluding smaller local and temporary instabilities just as in previous work (Schlegel *et al.*, 2022). Changes at the very beginning or end of the signal sections were ignored as they could also be side effects of onset or offset. Register changes were detected semi-automatically with manual removal of outliers after automatic detection.

E. Statistical analysis

As we stimulate multiple laryngeal muscles, changes in parameters may not be easily attributable to a single laryngeal muscles or muscle groups. For this reason, when measuring correlations we use Kendall's partial correlation for ordinal variables (Kendall, 1942) (as observed trends were approximately linear). This allows to measure the association between one muscle and one parameter while removing the effect of the remaining measured muscles. When comparing three data vectors the correlation coefficient is defined as given in Eq. (1), where τ_{12} is the Kendall rank correlation coefficient between data vectors 1 and 2, τ_{13} the coefficient for vectors 1 and 3, and τ_{23} the coefficient for vectors 2 and 3,

$$\tau_{12-3} = \frac{\tau_{12} - \tau_{13}\tau_{23}}{\sqrt{1-(\tau_{12})^2}\sqrt{1-(\tau_{23})^2}}. \quad (1)$$

This concept can be expanded to four data vectors (Wang and Chen, 2020) as shown in Eq. (2). In our work we use this definition to assess the correlation between one muscle and one parameter whilst eliminating the effects of the two remaining stimulated muscles/muscle groups,

$$\tau_{12-34} = \frac{\tau_{123} - \tau_{143}\tau_{243}}{\sqrt{1-(\tau_{142})^2}\sqrt{1-(\tau_{243})^2}}. \quad (2)$$

III. RESULTS

Parameters were calculated for all vibratory cycles detected in the 255 HSV datasets. In some videos small numbers of consecutive video frames (always below 20) were damaged and hence cycles that contained such damaged frames were excluded. Further, cycles within register

https://doi.org/10.1121/10.0025135

changes (fast F0 drops) were excluded. This resulted in 36 to 678 cycles detected per dataset/phonatory condition with 69 572 cycles in total. This large variation of cycle numbers is due to onset only occurring relatively late at high airflow in some conditions (as airflow was increased as a ramp over time during recording of each dataset). Respectively, this yielded 69 572 values for all parameters at unique combinations of muscle activations and flow, allowing for the exploration of the parameter space based on all relevant input signals.

A. Muscle activation voice range profiles

For each level of TA activation, voice range profiles based on the corresponding 64 combinations of LCA+IA and CT activations were generated (see Table I). They reflect which combinations of SPL and F0 were achieved as a function of TA and CT or LCA+IA nerve stimulation. Figure 2 illustrates how these plots were generated. Figure 2(a) shows two examples of measured anterior GAWs vs time and airflow for phonatory conditions TA2 LCA+IA0 CT1 and TA2 LCA+IA5 CT5, as well as the extracted LF0 from GAW cycles. A register change is marked in the second phonatory condition. In Fig. 2(b) extracted parameter values for all 64 CT and LCA+IA phonatory conditions for TA2 are plotted in 3D space (F0 vs SPL vs CT activation), featuring one value per cycle. Values from both conditions in Fig. 2(a) are highlighted in color. Figure 2(c) shows the corresponding voice range profile for TA level 2. The plot area is divided into 27 × 45 squares (each square has a width of 0.33 milliseconds, i.e., 1 HSV video-frame and a height of 1 dB SPL) and all CT values within each square are averaged, color-coding the respective square. Squares containing less than 5 values are excluded. F0 and SPL values from the two phonatory conditions presented in Fig. 2(a) are marked.

In Figs. 3(a)–3(e) voice range profiles for all five TA levels are shown illustrating average CT levels for all combinations of F0 and SPL. The smaller plots show the number of included samples per square area (top right) and the standard deviation of CT activation in each square area (bottom right). Due to the logarithmic nature of SPL, sample density is higher for higher SPL values. Standard deviation is low overall.

It is apparent that higher CT activation results in higher F0. The diagonal "stripes" at F0s lower than 250 Hz indicate that each level of CT also seems to favor a specific ratio of SPL and F0. At F0s higher than 250 Hz, high CT levels are split into two modes at about 300 Hz and 600 Hz with little to no phonation occurring between these frequencies. With no TA activation [Fig. 3(a)], no phonation occurs at high CT stimulation, then, with rising TA [Fig. 3(b)] the mode at 600 Hz first appears to be the preferred oscillatory condition at high CT.

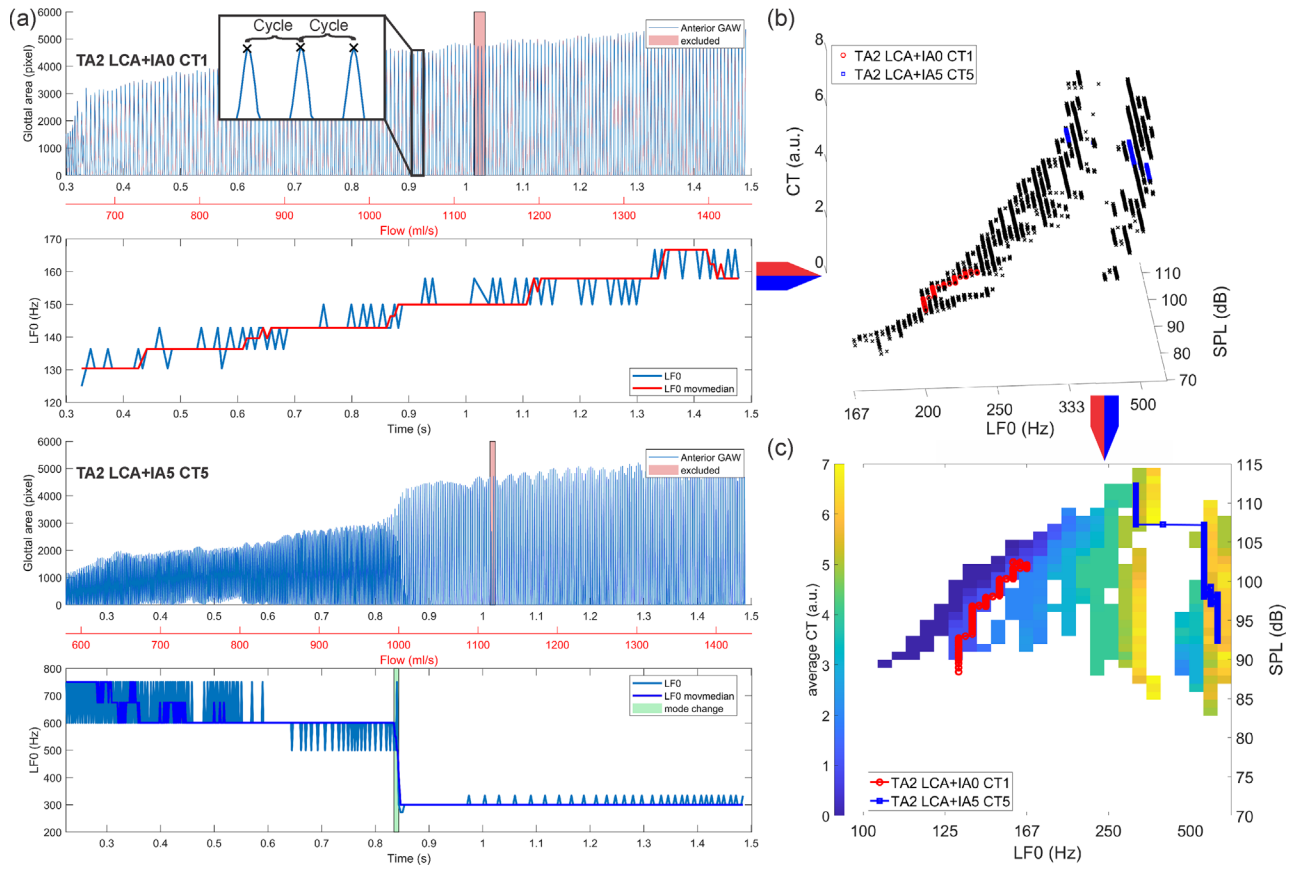

FIG. 2. (Color online) Illustration of the generation of voice range profiles. In (a) GAWs and LF0s of two example phonatory conditions (TA2 LCA+IA0 CT1 and TA2 LCA+IA5 CT5) are shown. As depicted in (b), for each TA level all parameters of 64 phonatory conditions of CT and LCA+IA are extracted. (c) Muscle activation levels are then averaged to create a color coded map of average activation conditions.

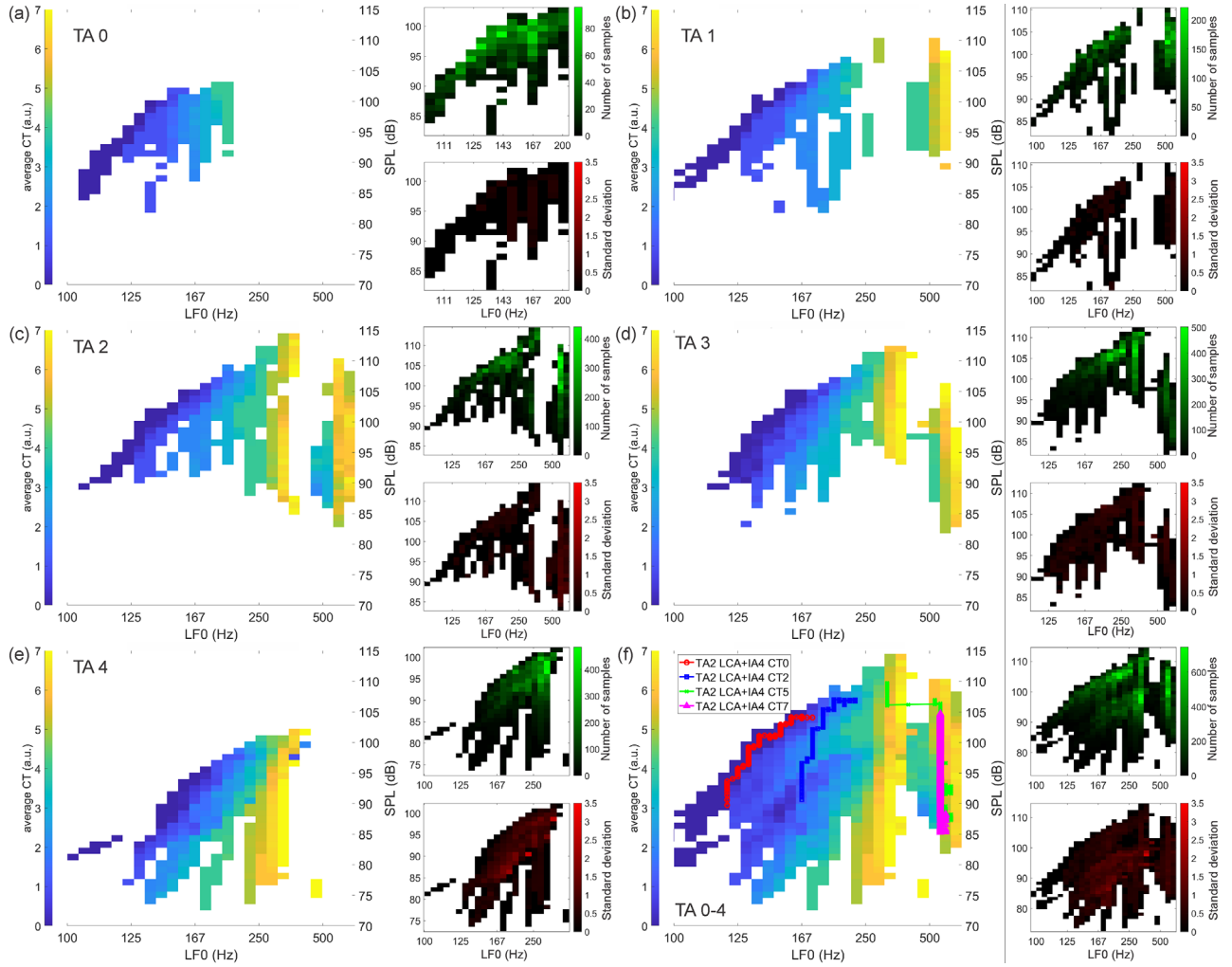

FIG. 3. (Color online) Voice range profiles for TA levels 0 to 4 (a)–(e) showing average CT levels vs F0 and SPL as well as sample density (green gradient plots, top right) and CT standard deviation (red gradient plots, bottom right). In (f) all data combined with some example phonatory conditions are depicted.

With further rising TA [Figs. 3(c) and 3(d)] this changes as the 300 Hz mode appears and the 600 Hz mode starts to disappear at higher SPLs. For high TA activation [Fig. (3e)] only the 300 Hz mode remains. Figure 3(f) depicts all 320 phonatory conditions with some example conditions marked to illustrate how values changeover time/with increasing airflow in different phonatory conditions. It is evident that there are only very few cycles with a LF0 between the 300 Hz and 600 Hz modes [Fig. 3(F), number of samples] indicating a region of instability between these modes.

Analogously, in Figs. 4(a)–4(e) voice range profiles for all five TA levels are shown illustrating average LCA+IA levels for all combinations of F0 and SPL. For these conditions standard deviations of LCA+IA levels (small plots bottom right) are much larger than for CT levels with LCA+IA activation levels often covering the entire spectrum of 0–7 for each square. It is apparent that high LCA+IA activation in general supports higher SPL. Further, the 600 Hz mode in Fig. 4(b) seems to depend on high LCA+IA activation. However, overall trends are not as clear as for CT activation. Figure 4(f) again shows all 320 combinations with some examples marked.

An analogous figure was also created for pSub as a control (see supplementary material Fig. S1) and shows the expected high correlation between pSub and SPL.

B. Abducted and adducted falsetto and chest

In 2011, Herbst *et al.* (2011) proposed that singers would be able to consciously control TA activation to produce two different subtypes of voice registers: abducted and adducted falsetto and chest. To objectively differentiate between these sub-registers, they calculated CQ (based on kymographs) and mimimum posterior glottal area for the different phonation types. To illustrate that these sub-registers are indeed mainly dependent on TA activation, we calculated the same types of parameters (CQ and MPA) from all detected cycles and separated them by register (falsetto-like and chest-like) and either the level of TA activation (from 0 to 4) or the level of LCA+IA activation (from 0 to 7). As threshold for the differentiation between falsetto-like and chest-like, 400 Hz was chosen, as this was in the middle of the region of instability (1/600–1/400 = 1/400–1/300) which most observed mode-changes crossed. This simple but

https://doi.org/10.1121/10.0025135

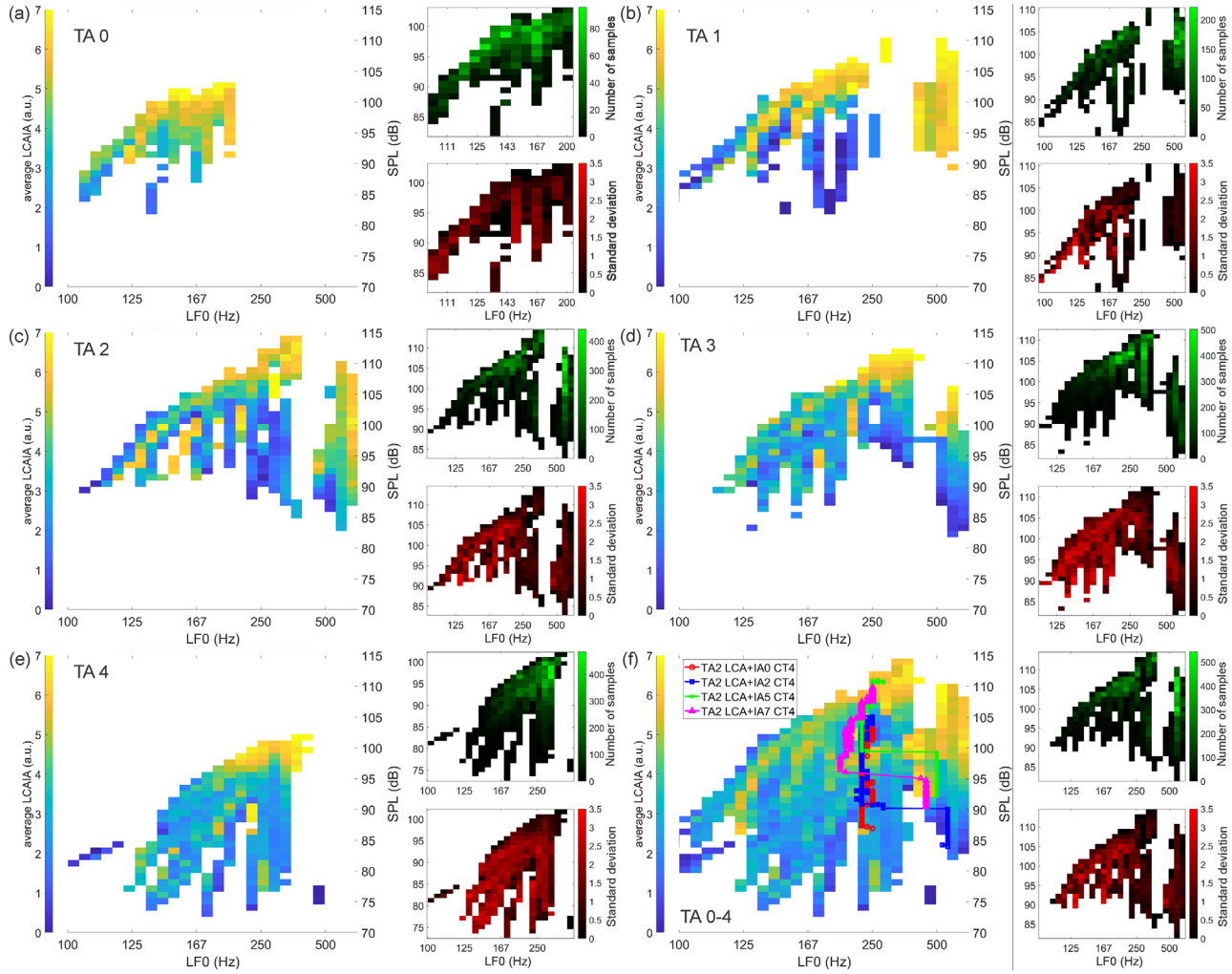

FIG. 4. (Color online) Voice range profiles for TA levels 0 to 4 (a)–(e) showing average LCA+IA levels vs F0 and SPL as well as sample density (green gradient plots, top right) and LCA+IA standard deviation (red gradient plots, bottom right). In (f) all data combined with some example phonatory conditions are depicted.

objective separation was preferred over more subjective approaches.

In Figs. 5(a)–5(e) CQ vs normalized MPA for all cycles and TA activations from 0 to 4 are depicted, separated by registers. Each plot also contains the outline of the area where most points were located. Figure 5(f) shows all the outlines from all TA levels for chest-like and falsetto-like registers. Overall, normalized MPA decreases and CQ increases with increasing TA activation, following the expected trend for a transition from abducted to adducted falsetto/chest registers.

The change from abducted to adducted registers is continuous with muscle stimulation and no clear gaps occur that are not caused by discretization artifacts. However, for minimum (TA0) and maximum (TA4) TA levels abducted and adducted chest-like sub-registers are clearly separable [Fig. 5(f), dark and light blue outlines for TA0 and TA4]. This is not as clearly the case for falsetto-like sub-registers as the overlap for different TA activations is higher.

Analogously in Figs. 6(a)–6(h) CQ vs normalized MPA for all cycles are depicted, separated by LCA+IA activation levels (0–7) and registers. Figure 6(i) shows the outlines from five selected LCA+IA levels. In this figure trends are less clear, with the chest-like outlines following the trend from abducted to adducted less consistently than they did when depicted dependent on TA. Falsetto-like areas show no clear trend from abducted to adducted at all when depicted dependent on LCA+IA.

For quantification of these effects, Kendall's partial correlation coefficients between one muscle and CQ and MPA are given in Table II. For each muscle the effect of the remaining two other stimulated muscles was eliminated. Of all muscles TA had the largest positive correlation with CQ and the largest negative one with MPA in both registers. The effects of LCA+IA and CT were considerably weaker.

IV. DISCUSSION

In this work we investigated register changes occurring in a canine model during different levels of symmetric laryngeal muscle stimulation and airflow. Using video data and the GAW, accurate cycle length measurement for each

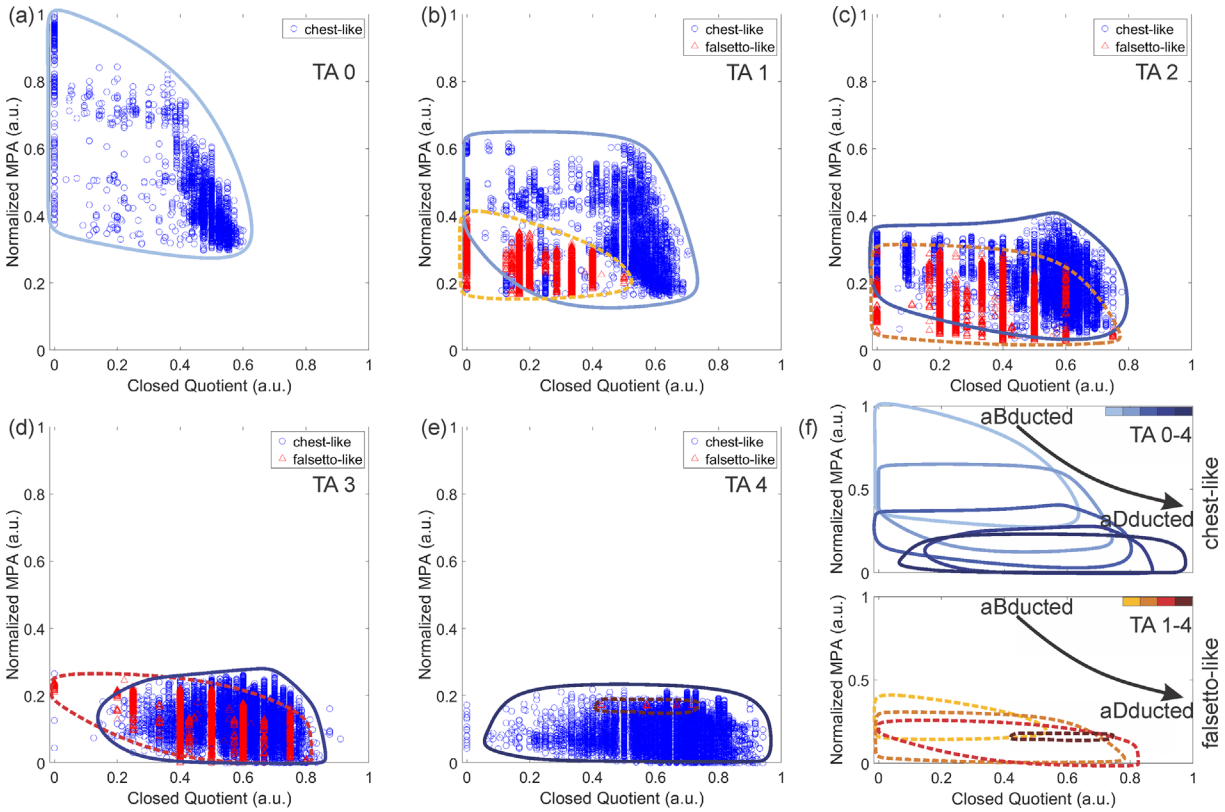

FIG. 5. (Color online) CQ and normalized MPA for TA levels 0 to 4 (a)–(e) separated by falsetto-like and chest-like registers. In (f) the general trend with increasing TA activation is shown based on point cloud outlines for chest-like and falsetto-like registers.

vocal fold oscillation cycle was possible, resulting in cycle-specific LF0 calculation for all signals. By increasing airflow for each condition with synchronous acoustic recording, intensity levels in SPL were also captured. With these data, voice range profiles were generated to help visualize average muscle activation maps as a function of LF0 and SPL levels.

A. Muscle activation voice range profiles

Notably, no register changes occurred in any phonatory conditions with very low CT stimulation (CT0-1) or without TA stimulation (TA0). This is consistent with a previous study on asymmetric laryngeal nerve activation in which it was also found that CT activation was necessary for register changes to occur (Schlegel *et al.*, 2022) and with a study by Geng *et al.* (2021) in a computational canine larynx model, where sudden drops in F0 only occurred with TA activity. However, in a canine hemilarynx experiment some early register changes did occur briefly after phonation onset with no CT activation (Schlegel *et al.*, 2024), although this could be an effect of the hemi-larynx setup used in that study.

As Fig. 3 shows, for no TA activation and low CT activation the region of instability around a LF0 of 400 Hz was never reached, whilst the voice range profile was split between high CT activation states. Respectively, the stability or instability of the signal appears to be mostly dependent on laryngeal muscle activation; i.e., if muscle activation results in low LF0 levels far apart from the region of instability at 400 Hz, the resulting signal is stable (displaying no sudden frequency changes) and vice versa.

At low LF0s until the first mode at 300 Hz, LF0 increases with increasing SPL (Fig. 3). This effect grows weaker with increasing CT activation until the 300 Hz mode at high CT levels is reached and LF0 stays relatively constant (except for potential sudden register changes) regardless of the increase in SPL and related increase in airflow over time. Notably, the slope of the LF0 increase stays also relatively constant and does not change with intensity but changes dependent on CT activation.

The region of instability was always in approximately the same fundamental frequency range around 400 Hz. As only the supraglottal tract of the canine was removed, the question remains if this instability may be a subglottal resonance phenomenon, since subglottal resonances in general cannot easily be avoided (Zhang *et al.*, 2006) and are also part of normal phonation in humans (Wade *et al.*, 2017). The dimensions of the subglottal tract upstream would suggest minimum resonance frequencies of approximately 800 Hz (assuming a speed of sound of 352.95 m/s at 37 °C) which is slightly higher than the 600 Hz mode we observe. This would suggest no relevant effect of subglottal resonances on the observed instabilities. To further confirm this, we re-investigated old datasets from two different canines that underwent similar muscle stimulation experiments, as the canine used in this study, and had approximately the same subglottal tract configurations [for details of the experiments see Schlegel *et al.* (2022) and Schlegel and Chhetri (2021)].

https://doi.org/10.1121/10.0025135

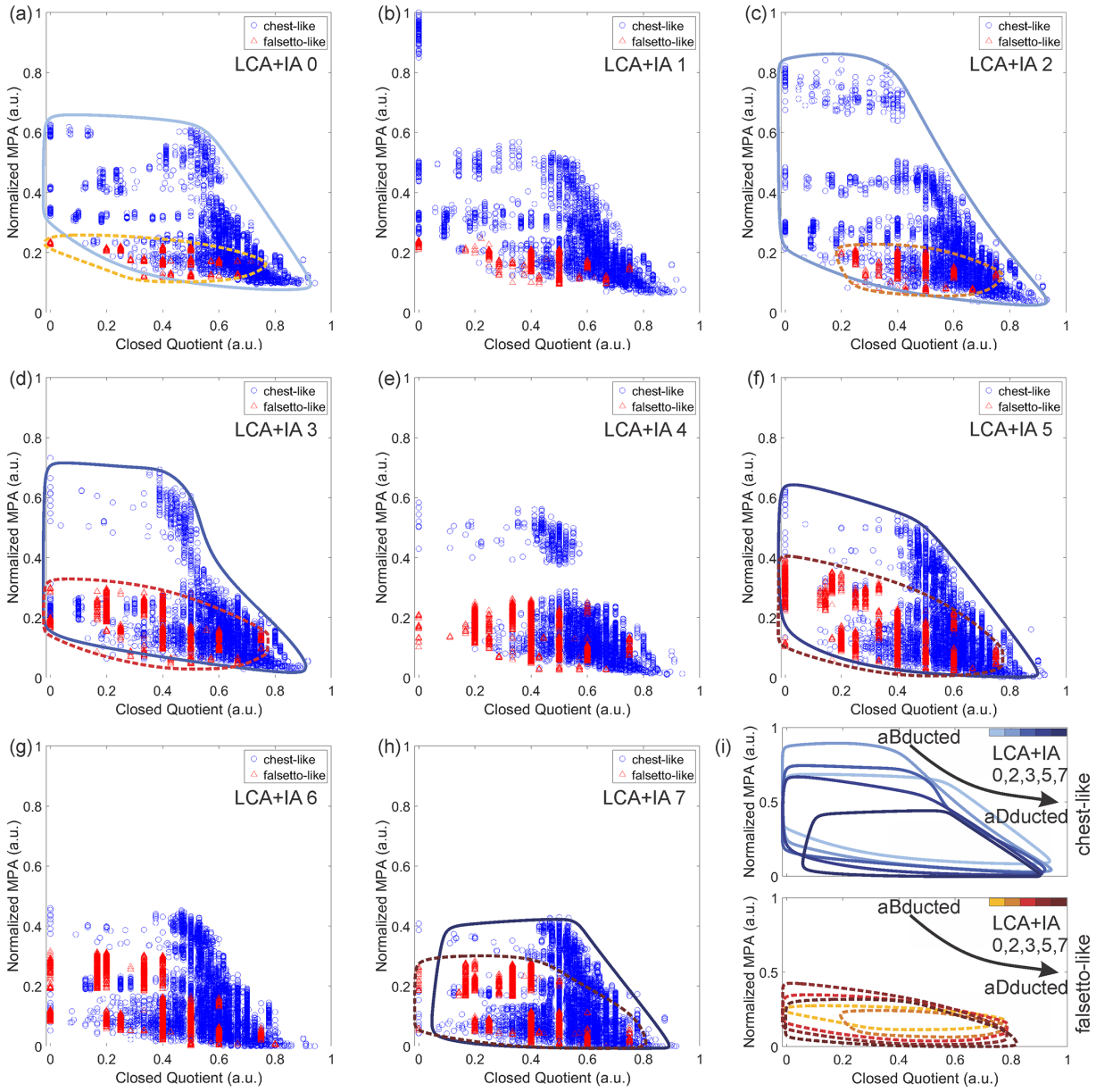

FIG. 6. (Color online) CQ and normalized MPA for LCA+IA levels 0 to 7 (a)–(h) separated by falsetto-like and chest-like registers. In (i) the general trend with increasing LCA+IA activation is shown based on point cloud outlines for chest-like and falsetto-like registers.

Figure 7 depicts the frequency of cycles with specific LF0 values for (a) the canine investigated in this study and (b), (c) for the previous experiments on different canines as well as the subglottal setup related to experiments (a) and (c) in (a_S) and (c_S). As can be seen in the figure, each canine has a region of instability, however, this region differs for each canine with neighboring peaks also differing in dominant frequency. As the subglottal tract configurations for these three canines were approximately the same, this also indicates that the observed phenomenon may be less likely significantly affected by the subglottal tract alone. Instead, it could be a product of interactions between subglottal acoustics and intrinsic properties of the larynx itself varying between individuals, such as slight variations in muscle microanatomy. However, it has to be noted that the type of register changes observed in this study do also occur in computational models that do not simulate complex subglottal acoustics (Geng *et al.*, 2020).

As observed in previous work (Schlegel *et al.*, 2022), high CT activation resulted in LF0 potentially suddenly dropping to lower LF0 levels [see, e.g., Fig. 2(a) condition TA2 LCA+IA5 CT5 and Figs. 3(b)–3(f)]. These register changes were almost exclusively LF0 drops with a LF0 "jump" occurring in only a single case. This type of behavior could be interpreted as caused by acoustical coupling between the subglottal tract and the larynx (Zhang *et al.*, 2006). In the context of this study, this would imply that the existence of the "region of instability" observed in multiple larynges could be due to a source-filter-interaction between larynx and subglottal tract. Alternatively, it could also be a

TABLE II. Kendall's partial correlation coefficients between TA, LCA+IA and CT and CQ and MPA corrected for the effect of the remaining laryngeal muscles or muscle groups (e.g., in case of TA the effect of LCA+IA and CT was eliminated).

Laryngeal muscle	CQ	MPA
	Chest-like	
TA	0.346	−0.632
LCA+IA	−0.112	−0.385
CT	0.180	0.031
	Falsetto-like	
TA	0.516	−0.617
LCA+IA	0.184	−0.472
CT	−0.001	0.250

purely morphological effect of CT and TA activation changing oscillatory behavior (Geng *et al.*, 2021). In any case, the location of the region of instability is likely mostly dependent on morphological properties of the larynx such as laryngeal muscle composition.

Last, TA activation had the expected effect of determining which of the two regions neighboring the region of instability was preferred (Titze *et al.*, 1989). Lower TA activation resulting in the higher mode being more stable and vice versa. Hence, whilst phonation in the region of instability cannot be achieved by changing muscle activation conditions, muscle adjustments lead to favoring the higher or lower pitched mode neighboring the region of instability. Translated to singers, this means that a slight "aberration" in TA activation could induce a sudden and significant frequency drop.

B. Abducted and adducted falsetto and chest

By calculating CQ and MPA for different levels of TA activation (Fig. 5) we observed that with increasing levels of TA, normalized MPA would decrease and CQ would slightly increase for both registers. Separating the plots by LCA+IA activation levels instead (Fig. 6), resulted in a less and less consistent decrease in normalized MPA with increasing LCA+IA in the chest-like register, and an increase in the range of both parameters in the falsetto-like register. A partial explanation for this difference may have been a stronger TA activation overall, as the graded nerve stimulation approach is subject to certain inaccuracies (see limitations). Nevertheless, these observations suggest that the abducted and adducted separation that Herbst *et al.* (2011) observed may be mainly caused by TA activation.

Further, abducted and adducted chest sub-registers are separable for very low (TA0) and very high (TA4) TA activation with a less clean separation for falsetto-like sub-registers [Fig. 5(f)]. Consequently, the impression of distinct sub-registers may arise when observing human subjects that achieve these phonation types by just crudely varying TA activation from very weak to very strong. However, the transition between these states appears to be more continuous dependent on the degree of TA activation.

Kendall's partial correlation coefficients (Table II) confirm these subjective assessments as TA had the strongest positive correlation with CQ and the strongest negative correlation with MPA. As both parameters are associated with changing sub-registers, this respectively suggests that the effect of TA on these was the strongest. The slight negative

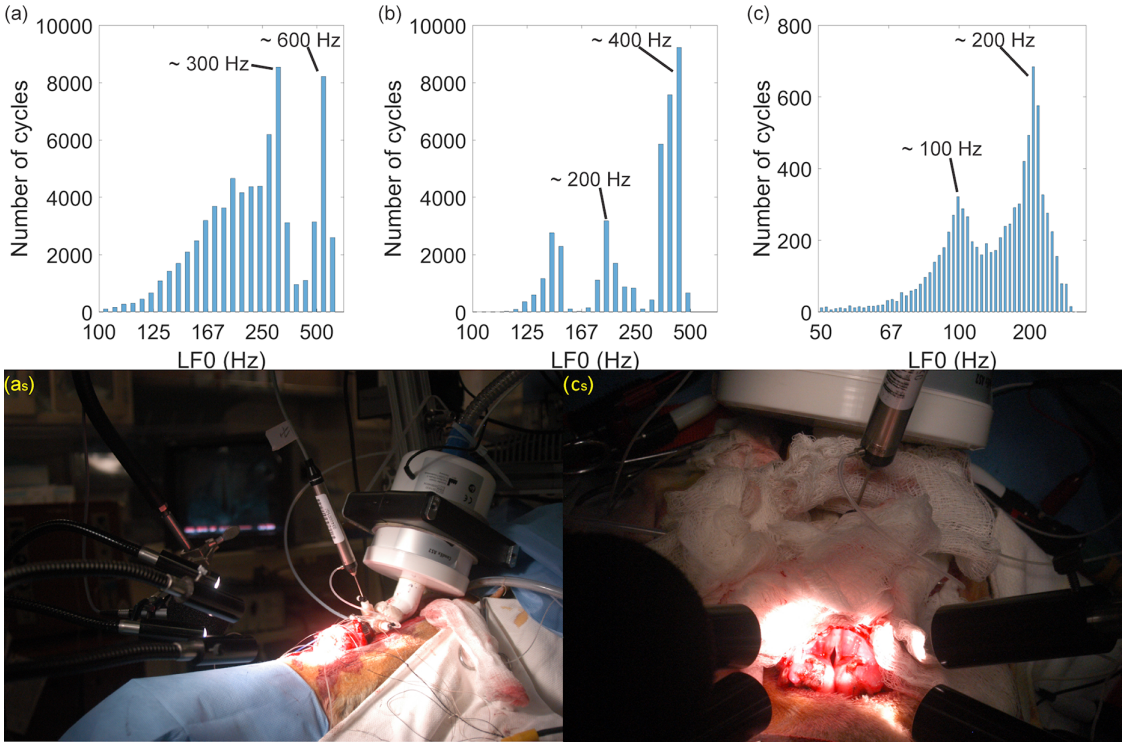

FIG. 7. (Color online) Histograms depicting the frequency of cycles with specific LF0 lengths in canines in (a) this and (b), (c) two previous studies. The subglottal tract setup for the experiments related to (a) and (c) is depicted in (a_S) and (c_S).

https://doi.org/10.1121/10.0025135

correlation between LCA+IA and CQ in chest-like register is explainable by a higher variation between cycles and some cycles with an artificially elongated open phase in conditions with high TA and LCA+IA. This in turn induces a slight negative association between muscle activation and CQ. However, this effect only becomes visible in the LCA+IA correlation as the positive association between CQ and LCA+IA was weaker than for CQ and TA. The minor correlation between CT and CQ in the chest-like register may be explainable by CT increasing F0 and hence leading to shorter cycles that then cause differences in CQ due to discretization artifacts. The minor correlation between CT and MPA in the falsetto-like register indicates that the CT-induced stretching of the vocal folds slightly diminishes the open posterior area as captured by our stationary HSV camera.

C. Summary

We answer our research questions as follows: We show in Figs. 3 and 4 the specific combinations of F0 and SPL favored by different laryngeal phonatory conditions. They visualize the interplay between CT and TA that leads to the falsetto-like or chest-like register being more stable, dependent on muscle stimulation with a LF0 gap between the registers. The existence of this frequency gap may be caused by a source-filter-interaction between larynx and subglottal tract but current computational models suggest that it may more likely be a purely laryngeal event (Geng *et al.*, 2020; Zhang, 2024). Further, we were able to reproduce the abducted and adducted falsetto and chest registers as described by Herbst *et al.* (2011) with systematic TA activation in our canine model. Respectively, this suggests that voluntary control of TA also played a major role in the work of Herbst *et al.* to produce abducted and adducted registers. This in turn may indicate that the extent of voluntary control that humans have over TA activation, or can achieve with sufficient training, may be greater than previously thought.

D. Limitations

In this study, only a single canine was used for the main investigation, as we focused on systematically investigating 320 different nerve activation conditions and in total 1.44×10^6 video frames to create a map of stable and unstable phonatory states for this larynx. Further, the canine model had no vocal tract and the subglottal tract, whilst chosen to have nearly physiological length (Zhang *et al.*, 2006), differs from a living animal as airflow is produced by a machine and not from its lungs. Additionally, the graded nerve stimulation approach relies on visual detection of minimal and maximal muscle activation, which can lead to inaccuracies if the stimulation current, at which minimum or maximum muscle activation is achieved, is set too low or too high. For example, if the point of maximum activation is overestimated, the resulting muscle activation may be stronger than intended. As PCA was cut for this experiment, its possible effects on oscillatory behavior were not assessed.

V. CONCLUSION

In this work, we visualized interrelationships between laryngeal airflow, muscle input, and voice output by mapping laryngeal muscle activation vs LF0 and SPL using voice range profiles. We found a region of instability that persisted at a specific LF0 level regardless of muscle activation. This region appears to be larynx-specific, as a comparison with data from previous studies showed. The higher pitched region neighboring this LF0 level can display sudden frequency drops to the lower pitched neighboring region, whilst otherwise being stable with rising airflow. Which of the neighboring modes is preferred, appears to be related primarily to TA activation. Further, we could confirm trends in CQ and MPA based on TA activation observed by Herbst *et al.* (2011), which indicates that abducted and adducted presentations of registers in their work were primarily achieved by voluntary control of TA. Our study emphasizes the importance of holistic approaches, looking at the interplay of the different laryngeal muscles, and suggests that training exercises centered on consciously manipulating TA activation may be beneficial for singers to avoid involuntary register changes. More investigations with varying lengths and properties of the subglottal tract and larynx are necessary to better capture the parameter space of human phonation.

SUPPLEMENTARY MATERIAL

See the supplementary material for a table containing all parameters for all conditions and all cycles (file S1) and a figure depicting voice range profiles for TA levels 0 to 4 (a)–(e) showing average pSub vs F0 and SPL as well as sample density (green gradient plots, top right) and pSub standard deviation (red gradient plots, bottom right) (file S2). In (f) all data combined with some example phonatory conditions are depicted.

ACKNOWLEDGMENTS

D.K.C. received funding from the National Institutes of Health (NIH)/National Institute on Deafness and Other Communication (NIDCD) under Grant No. R01 DC11300. (Funder website: https://www.nidcd.nih.gov/.) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR DECLARATIONS

Conflict of Interest

The authors of this manuscript have no conflicts of interest to disclose.

Ethics Approval

An *in vivo* canine larynx experiment was performed. This experiment was approved by the Animal Research Committee (ARC) of the University of California, Los Angeles (Protocol No. ARC-2010-021). The canine model

was chosen for its similarity to the human larynx, in terms of size, neuromuscular anatomy, and histopathology of the vocal folds (Chhetri *et al.*, 2014; Garrett *et al.*, 2000).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

Berry, D. A., Herzel, H., Titze, I. R., and Story, B. H. (**1996**). "Bifurcations in excised larynx experiments," J. Voice **10**, 129–138.

Chhetri, D. K., Neubauer, J., and Berry, D. A. (**2010**). "Graded activation of the intrinsic laryngeal muscles for vocal fold posturing," J. Acoust. Soc. Am. **127**, EL127–EL133.

Chhetri, D. K., Neubauer, J., and Sofer, E. (**2014**). "Influence of asymmetric recurrent laryngeal nerve stimulation on vibration, acoustics, and aerodynamics," Laryngoscope **124**, 2544–2550.

Chhetri, D. K., and Park, S. J. (**2016**). "Interactions of subglottal pressure and neuromuscular activation on fundamental frequency and intensity," Laryngoscope **126**, 1123–1130.

Garrett, C. G., Coleman, J. R., and Reinisch, L. (**2000**). "Comparative histology and vibration of the vocal folds: Implications for experimental studies in microlaryngeal surgery," Laryngoscope **110**, 814–824.

Geng, B., Movahhedi, M., Xue, Q., and Zheng, X. (**2021**). "Vocal fold vibration mode changes due to cricothyroid and thyroarytenoid muscle interaction in a three-dimensional model of the canine larynx," J. Acoust. Soc. Am. **150**, 1176–1187.

Geng, B., Pham, N., Xue, Q., and Zheng, X. (**2020**). "A three-dimensional vocal fold posturing model based on muscle mechanics and magnetic resonance imaging of a canine larynx," J. Acoust. Soc. Am. **147**, 2597–2608.

Greene, C. (**2016**). "Sound pressure level calculator," MATLAB Central File Exchange, https://www.mathworks.com/matlabcentral/fileexchange/35876-sound-pressure-level-calculator (Last viewed October 5, 2023).

Herbst, C. T. (**2020**). "Registers-the snake pit of voice pedagogy Part 1: Proprioception, perception, and laryngeal mechanisms," J. Singing **77**, 175–190.

Herbst, C. T., Elemans, C. P. H., Tokuda, I. T., Chatziioannou, V., and Švec, J. G. (**2024**). "Dynamic system coupling in voice production," J. Voice (in press).

Herbst, C. T., Qiu, Q., Schutte, H. K., and Švec, J. G. (**2011**). "Membranous and cartilaginous vocal fold adduction in singing," J. Acoust. Soc. Am. **129**, 2253–2262.

Hollien, H. (**1974**). "On vocal registers," J. Phon. **2**, 125–143.

Kendall, M. G. (**1942**). "Partial rank correlation," Biometrika **32**, 277–283.

Kist, A. M., Gómez, P., Schlegel, P., Kunduk, M., Echternach, M., Patel, R., Semmler, M., Bohr, C., Dürr, S., Schützenberger, A., and Döllinger, M. (**2021**). "A deep learning enhanced novel software tool for laryngeal dynamics analysis," J. Speech. Lang. Hear. Res. **64**, 1889–1903.

Luegmair, G., Chhetri, D. K., and Zhang, Z. (**2014**). "The role of the thyroarytenoid muscle in regulating glottal airflow and glottal closure in an in vivo canine larynx model," Proc. Mtgs. Acoust. **22**(1), 060007.

Roubeau, B., Henrich, N., and Castellengo, M. (**2009**). "Laryngeal vibratory mechanisms: The notion of vocal register revisited," J. Voice **23**, 425–438.

Schlegel, P., Berry, D. A., and Chhetri, D. K. (**2022**). "Analysis of vibratory mode changes in symmetric and asymmetric activation of the canine larynx," PLoS One **17**, e0266910.

Schlegel, P., and Chhetri, D. (**2021**). "Asymmetric superior laryngeal nerve stimulation: Influence on acoustic cepstral peak prominence and glottis symmetry and perturbation measures," in *The 14th International Conference on Advances in QuantitativeLaryngology, Voice and Speech Research (AQL),* Bogotá, Colombia June 7–10.

Schlegel, P., Chung, H. R., Döllinger, M., and Chhetri, D. K. (**2024**). "3D reconstruction of vocal fold medial surface trajectories: Effects of neuromuscular stimulation and airflow," Laryngoscope **134**(3), 1249–1257.

Schlegel, P., Kniesburges, S., Dürr, S., Schützenberger, A., and Döllinger, M. (**2020**). "Machine learning based identification of relevant parameters for functional voice disorders derived from endoscopic high-speed recordings," Sci. Rep. **10**, 10517.

Soriano, R. M., and Gupta, V. (**2021**). *Anatomy, Head and Neck, Larynx Nerves* (StatPearls, Treasure Island, FL).

Soron, H. I., and Lieberman, P. (**1963**). "Some measurements of the glottal-area waveform," J. Acoust. Soc. Am. **35**, 1876–1877.

Titze, I. R. (**2008**). "Nonlinear source–filter coupling in phonation: Theory," J. Acoust. Soc. Am. **123**(4), 2733–2749.

Titze, I. R., Luschei, E. S., and Hirano, M. (**1989**). "Role of the thyroarytenoid muscle in regulation of fundamental frequency," J. Voice **3**, 213–224.

Titze, I. R., Mapes, S., and Story, B. (**1994**). "Acoustics of the tenor high voice," J. Acoust. Soc. Am. **95**, 1133–1142.

Titze, I. R., Riede, T., and Popolo, P. (**2008**). "Nonlinear source–filter coupling in phonation: Vocal exercises," J. Acoust. Soc. Am. **123**, 1902–1915.

Wade, L., Hanna, N., Smith, J., and Wolfe, J. (**2017**). "The role of vocal tract and subglottal resonances in producing vocal instabilities," J. Acoust. Soc. Am. **141**, 1546–1559.

Wang, J.-H., and Chen, Y.-H. (**2020**). "Interaction screening by Kendall's partial correlation for ultrahigh-dimensional data with survival trait," Bioinformatics **36**, 2763–2769.

Yin, J., and Zhang, Z. (**2014**). "Interaction between the thyroarytenoid and lateral cricoarytenoid muscles in the control of vocal fold adduction and eigenfrequencies," J. Biomech. Eng. **136**, 111006.

Zarachi, A., Tafiadis, D., Ziavra, N., Kastanioudakis, I., Dimakis, E., Liontos, A., Argyropoulou, M., and Exarchakos, G. (**2021**). "Morphometric correlations of the voice category (VC) in professional singers with oropharyngeal and laryngeal anatomy using stroboscopy and cervical posteroanterior radiography," Int. J. Otolaryng. Head Neck Surg. **10**, 277–299.

Zhang, Z. (**2016**). "Mechanics of human voice production and control," J. Acoust. Soc. Am. **140**, 2614–2635.

Zhang, Z. (**2024**). "Vocal fold vertical thickness in human voice production and control: A review," J. Voice. (in press).

Zhang, Z., Neubauer, J., and Berry, D. A. (**2006**). "The influence of subglottal acoustics on laboratory models of phonation," J. Acoust. Soc. Am. **120**, 1558–1569.