JSLHR

Research Note

Formant-Estimated Vocal Tract Length and Extrinsic Laryngeal Muscle Activation During Modulation of Vocal Effort in Healthy Speakers

Matti D. Groll,[a,b] Victoria S. McKenna,[c] Surbhi Hablani,[b] and Cara E. Stepp[a,b,d]

Purpose: The goal of this study was to explore the relationships among vocal effort, extrinsic laryngeal muscle activity, and vocal tract length (VTL) within healthy speakers. We hypothesized that increased vocal effort would result in increased suprahyoid muscle activation and decreased VTL, as previously observed in individuals with vocal hyperfunction.

Method: Twenty-eight healthy speakers of American English produced vowel–consonant–vowel utterances under varying levels of vocal effort. VTL was estimated from the vowel formants. Three surface electromyography sensors measured the activation of the suprahyoid and infrahyoid muscle groups. A general linear model was used to investigate the effects of vocal effort level and surface electromyography on VTL. Two additional general linear models were used to investigate the effects of vocal effort on suprahyoid and infrahyoid muscle activities.

Results: Neither vocal effort nor extrinsic muscle activity showed significant effects on VTL; however, the degree of extrinsic muscle activity of both suprahyoid and infrahyoid muscle groups increased with increases in vocal effort.

Conclusion: Increasing vocal effort resulted in increased activation of both suprahyoid and infrahyoid musculature in healthy adults, with no change to VTL.

Vocal hyperfunction is a common symptom of voice disorders, defined as "abuse and/or misuse of the vocal mechanisms due to excessive and/or 'imbalanced' muscular forces" (Hillman et al., 1989). Excessive vocal effort is a primary component of vocal hyperfunction (Roy et al., 2007, 1996). By extension, excessive vocal effort is also thought to be a correlate of increased tension in the intrinsic and extrinsic laryngeal muscles (Boone & McFarlane, 1988; Laver, 1980; Roy et al., 1997). It has been suggested that the presence of vocal hyperfunction and the accompanying increase in vocal effort and laryngeal tension result in a change in vocal tract length (VTL), primarily due to laryngeal elevation and reduced hyolaryngeal space (Aronson & Bless, 2009; Roy & Ferguson, 2001). Indeed, radiographic measures show that individuals with vocal hyperfunction exhibit a significantly higher laryngeal position than individuals with healthy voices (Lowell et al., 2012). Raising of the larynx is a consequence of increased extrinsic laryngeal muscle activation; specifically, activation of the thyrohyoid, digastric, stylohyoid, mylohyoid, geniohyoid, hyoglossus, and/or genioglossus muscles can all elevate the larynx (Hixon et al., 2008; Ludlow et al., 2007).

Changes in VTL cause a change to all formant frequencies, with a shorter VTL corresponding to increased formant frequencies (Roy & Ferguson, 2001; Shipp, 1987; Stevens, 1998). A simple relationship between VTL and formant frequency can be derived by modeling the vocal tract as a uniform tube that is closed at one end and open at the other, which exhibits odd quarter-wave harmonic resonances (Resnick & Halliday, 1966). The vocal tract is assumed to be open at its proximal end (lips) and effectively closed at its distal end (phonating vocal folds). Equation 1 shows the resulting inverse relationship between VTL and formant frequencies (Stevens, 1998), where *n* is

---

[a] Department of Biomedical Engineering, Boston University, MA
[b] Department of Speech, Language, & Hearing Sciences, Boston University, MA
[c] Department of Speech, Language, & Hearing Sciences, Purdue University, West Lafayette, IN
[d] Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, MA

Correspondence to Matti D. Groll: mgroll@bu.edu

Editor-in-Chief: Bharath Chandrasekaran
Editor: Jack J. Jiang

Received September 23, 2019
Revision received December 16, 2019
Accepted January 28, 2020
https://doi.org/10.1044/2020_JSLHR-19-00234

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

the formant number, f_n is the formant frequency, and c is the speed of sound.

$$f_n = \frac{(2n-1)\times c}{4\times \mathrm{VTL}} \tag{1}$$

Several studies have explored the relationship between formants and VTL. For example, a study by Lindblom and Sundberg (1971) showed that simulated lowering of the laryngeal position by 10 mm decreased all formant frequencies. Another study by Sundberg and Nordström (1976) showed that the first formant (F1) of close front vowels (/i/) remained unaffected, whereas the second (F2), third (F3), and fourth (F4) formants increased when the larynx was raised. Thus, increases in formant frequencies, particularly in higher formants, may be used to approximate a shortening of the VTL. Furthermore, higher formant frequencies in a given vocal tract shape tend to be regularly spaced and do not greatly deviate from higher formant frequencies of a uniform tube with the same length, therefore maintaining the relationship in Equation 1 (Lammert & Narayanan, 2015; Wakita, 1977). Thus, Equation 1 can be used to estimate VTL by measuring higher formant frequencies.

Manual circumlaryngeal massage has been used as a therapeutic technique to lower the larynx in order to relieve laryngeal tension and decrease excessive vocal effort. A study by Roy and Ferguson (2001) observed that the F1, F2, and F3 of speakers with vocal hyperfunction significantly decreased following manual circumlaryngeal massage. The authors reasoned this to be due to a lowering of the larynx from an elevated position to an optimal position as a result of a decrease in laryngeal tension (Roy & Ferguson, 2001). However, this lowering could have occurred by a decrease in suprahyoid muscle activation or a rebalancing of activation between the suprahyoid and infrahyoid musculature for a neutral laryngeal position.

A commonly used methodology to noninvasively measure extrinsic laryngeal muscle activity is through surface electromyography (sEMG). Previous studies have shown that individuals with vocal hyperfunction (and presumably increased vocal effort) exhibit increased levels of sEMG at the larynx during phonation when compared to individuals with healthy voices (Milutinovic et al., 1988; Redenbaugh & Reich, 1989). Similarly, Hocevar-Boltezar et al. (1998) found that individuals with vocal hyperfunction exhibit greater levels of sEMG at the supralaryngeal muscles during phonation when compared to individuals with healthy voices. However, more recent studies have not observed differences in the neck sEMG of individuals with vocal hyperfunction (Stepp et al., 2010; Van Houtte et al., 2013). Thus, the relationship between vocal hyperfunction and laryngeal muscle tension measured by sEMG remains unknown.

Recent work by McKenna et al. (2019) observed changes in sEMG when individuals with healthy voices modulated their vocal rate and vocal effort to simulate laryngeal tension. The level of activation of an sEMG sensor placed over the extrinsic suprahyoid muscles (i.e., digastric, stylohyoid, mylohyoid, geniohyoid, hyoglossus, and genioglossus) was a significant predictor of the degree of self-reported vocal effort, whereas the level of activation of sEMG sensors placed over the left and right infrahyoid muscles (i.e., sternohyoid and omohyoid) were not. One interpretation of these results from McKenna et al. (2019) is that the increased activation of the suprahyoid muscles during increased vocal effort could indicate a heightening of the larynx, similarly to the elevated laryngeal position observed in individuals with vocal hyperfunction (Lowell et al., 2012; Roy & Ferguson, 2001). However, an increase in muscle activity does not explicitly indicate that the larynx has been raised. Indeed, it is possible that muscle activation may be independent of laryngeal position. For example, the co-contraction of contradictory muscles (e.g., the suprahyoid and infrahyoid muscles) could result in increased muscle activation without a net change in laryngeal position. This could indicate a difference in the mechanisms by which vocal effort manifests in healthy individuals and individuals with vocal hyperfunction.

This study expands on the analysis of acoustic data collected in the study by McKenna et al. (2019) in order to investigate the relationships between VTL, extrinsic laryngeal muscle activity, and increased vocal effort, thereby providing insight into the mechanisms behind vocal hyperfunction. Specifically, we estimated VTL from vowel formants and measured suprahyoid and infrahyoid muscle activity during phonation at different levels of vocal effort. We explored the following research questions: (a) Is VTL related to vocal effort and sEMG of extrinsic laryngeal muscle activity, and (b) does sEMG change with increased vocal effort? Our hypotheses were that VTL would be inversely related to both the level of vocal effort and suprahyoid muscle activity and that suprahyoid muscle activity would increase with increased vocal effort.

Method

Participants

This study included all 26 participants from the previous study described by McKenna et al. (2019). An additional two participants, who were excluded from the previous study due to poor laryngeal imaging data, were included in this study, because laryngeal imaging data were not used in the present analysis. In total, 28 healthy adult speakers (12 men, 16 women) aged 18–29 years (M = 20.8 years, SD = 2.76 years) participated. All participants were native speakers of American English, were nonsmokers, did not have any trained singing experience beyond grade school, and reported no history of neurological, speech, or hearing disorders. All participants were screened for healthy vocal function by a certified speech-language pathologist (SLP). Prior to starting the experimental protocol, all participants completed written consent in compliance with the Boston University Institutional Review Board.

Experimental Design

Details of the full experimental protocol and setup can be found in the study by McKenna et al. (2019). In brief, participants were seated in a sound-treated booth and produced a series of vowel–consonant–vowel (VCV) utterances during flexible laryngoscopy. Videoendoscopic images, acoustic signals, sEMG, and neck surface vibrations via an accelerometer were collected across six experimental conditions consisting of different vocal rates and vocal efforts. In this study, only the acoustic, accelerometric, and sEMG data during a regular vocal rate (referred to as no effort in this study) and at increased levels of vocal effort were analyzed.

For each experimental condition, participants produced a set of eight uniform /ifi/ utterances. There were four experimental conditions of interest in this study: no effort, mild effort, moderate effort, and maximum effort. Each experimental condition was repeated at least twice to obtain 16 /ifi/ utterances per condition. For all experimental conditions, participants were instructed to maintain a comfortable speaking rate and loudness. For the no effort condition, no additional instructions were given. For increased effort levels, participants were asked to modulate their vocal effort using the following instructions: “Now we would like you to increase your effort during your speech as if you are trying to create tension in your voice as if you are trying to push your air out. Try to maintain the same volume while increasing your effort.” Mild effort was described as “mildly more effort than your regular speaking voice.” Moderate effort was described as “more effort than your mild effort,” and maximal effort was described as “as much effort as you can, while still having a voice.” All participants practiced each vocal effort condition with a certified SLP for approximately 10 min to ensure they understood the task and were comfortable producing different levels of vocal effort.

Instrumentation and Calibration

Prior to recording, three Bagnoli sEMG differential sensors (Delsys) were placed on the anterior neck using medical-grade adhesive. The electrodes of each sEMG sensor were 10 mm × 1 mm diameter and spaced 10 mm apart. Prior to sensor placement, the skin surface of the anterior of the neck was abraded and exfoliated to enhance skin-electrode contact and reduce artifacts and noise in the sEMG signal (Stepp, 2012). A certified SLP manually palpated the submandibular region (mandible, soft tissue) during a swallowing maneuver. A single differential sensor was placed at the midline, posterior to the mandible in the submandibular region, in order to capture suprahyoid muscle activation (specifically, the mylohyoid muscles and, to a lesser extent, the geniohyoid and the anterior belly of the digastrics). The single differential sensor was used to simultaneously capture activity from the left and right suprahyoid muscles and minimize error resulting from activity comparisons between the two sites. Next, the SLP palpated the thyroid cartilage and located the thyroid prominence during tasks such as humming and swallowing. Two double-differential sensors were placed approximately 1 cm to the left and right sides of the thyroid prominence. This placement captured the activity of the external infrahyoid muscles (i.e., the thyrohyoids, omohyoids, and sternohyoids), without capturing the activity from the more lateral sternocleidomastoids. Double-differential electrodes were used here to reduce electrical signals that were common across all electrodes (Rutkove, 2007). To account for environmental noise, as well as physiological noise such as heartbeat, a ground electrode was placed on the acromion of the right shoulder. Participants then completed a set of maximal voluntary contraction (MVC) tasks, including neck flexion, saliva swallows, throat clears, and isometric contraction against resistance, which were later used to normalize sEMG voltage measurements. All sEMG sensors are also likely to capture activation of the platysma muscle due to its superficial location compared to surrounding musculature (Stepp, 2012).

Acoustic data were collected through a directional headset microphone (Shure SM35 XLR) placed at 45° from the midline of the vermilion and 7 cm from the corner of the lips. Accelerometric data were obtained via a BU-21771-000 accelerometer (Knowles Electronic) placed on the anterior of the neck with double-sided adhesive tape, superior to the thyroid notch and inferior to the cricoid cartilage. Neck sEMG signals were acquired using a 16-channel Delsys Bagnoli EMG System (DS-160) and analog bandpass filtered with roll-off frequencies of 20 and 450 Hz and a gain of 1,000. Microphone and accelerometer signals were first preamplified (Xenyx Behringer 802 Preamplifier). All signals were simultaneously time-aligned and digitized at 30 kHz via a data acquisition board (National Instruments USB-6312).

Data Analysis

Formants

Author S. H. manually calculated formants of /ifi/ utterances using acoustic recordings. For each VCV utterance, a wide-band grayscale spectrogram was generated, and the first four formants were identified for both vowel segments using Praat (Boersma, 2002). For each vowel segment, the accuracy of the automated formant tracking was inspected visually, and the default settings of the formant-tracking tool were adjusted to ensure that formant tracking aligned with spectral representation on the grayscale spectrogram (Hillenbrand et al., 1995). Formants were estimated over the central stable portion of the vowel, omitting both the onset and offset of the vowel, because the vocal tract movements needed to transition from voiced to voiceless speech result in rapidly changing formants (Lehiste & Peterson, 1961; Öhman, 1966; Stevens & House, 1955). If an individual formant could not be confidently identified in the spectrogram, the formant was rejected from analysis. An entire vowel segment was rejected from analysis if the sample was aperiodic, glottalized, misarticulated, or lacking steady-state voicing. Mean formant values for each experimental condition were computed in hertz by averaging across all usable vowel segments for that condition.

Vocal Tract Length

VTL was estimated from formant values using the relationship in Equation 1. Estimates of VTL based on the third formant and the fourth formant were averaged to compute a mean VTL estimate per participant for each experimental condition. The third and fourth formants were used because higher formants are more stable and a better estimate of VTL using Equation 1 (Wakita, 1977).

Neck sEMG

The sEMG signals were digitally band passed with a second-order Butterworth filter between 20 and 500 Hz in order to remove movement artifacts at low frequencies and electrode and equipment noise at high frequencies (Stepp, 2012). Strings of /ifi/ utterances were segmented into individual VCV productions. Voicing was identified from the accelerometer signal and an additional 250 ms of prephonation was added to account for muscle activity prior to the start of voicing in the acoustic signal (Shipp, 1975; Stepp, Heaton, Stadelman-Cohen, et al., 2011). The percent activation (amplitude compared to each sensor MVC) was calculated. The root-mean-square of each individual /ifi/ utterance was divided by the root-mean-square of the MVC tasks as determined for each sensor during the calibration process. The entire /ifi/ utterance, as well as the preceding 250 ms, was chosen for analysis, because muscle activation has been shown to increase at both the initiation and termination of voicing as a result of quick dynamic laryngeal movements for speech (Hirose & Gay, 1973; Sawashima et al., 1973). The resulting activation value was a percentage of the maximum possible value for each sensor for each /ifi/ segment. Due to a strong correlation (Pearson correlation of $r = .82$), the percent activation values for the left and right infrahyoid muscles were averaged together, resulting in a single activation value for the infrahyoid muscles per /ifi/ segment (McKenna et al., 2019). These values were then averaged within an experimental condition resulting in two final average percent activation values per condition—one corresponding to suprahyoid activity and the other corresponding to infrahyoid activity—per participant.

Statistical Analysis

All statistical analyses were completed using Minitab Statistical Software (Version 17; Minitab, Inc.). Significance for all statistical testing was set a priori at $p < .05$.

To determine interrater reliability for formant measurements, author M. G. measured formants for four randomly selected participants. A two-way intraclass correlation (ICC) analysis was conducted for F3 and F4 measurements between the two raters, with an ICC(2, 1) of .95 and .93, respectively. Two weeks after initial analysis, author S. H. reanalyzed three randomly selected participants, and ICC analysis was performed for F3 and F4 measurements in order to determine intrarater reliability. The ICC values for intrarater reliability for F3 and F4 measurements were ICC(2, 1) = .99 and .99, respectively. These values all indicated excellent reliability for formant estimates (Koo & Li, 2016).

To address the first research question, we performed a general linear model (GLM) using VTL as the outcome measure, with effort level as a fixed factor, participant as a random factor, and percent activation of the suprahyoid and infrahyoid muscle groups as covariates. Upon initial analysis, muscle activity was found to have high collinearity (variance inflation factors > 10) with the participant factor. In order to reduce collinearity, Z scores were computed for both muscle groups within individual participants.

To address the second research question, we performed two separate GLMs for the percent activation of the suprahyoids and infrahyoids, with effort level as a fixed factor and participant as a random factor. Partial eta squared (η_p^2) was calculated to determine effect size for significant factors (Witte & Witte, 2010). Post hoc Dunnett's tests were used to evaluate significant differences between no effort and effortful conditions. Cohen's d was used to calculate effect sizes for significant differences (Witte & Witte, 2010).

Results

Across all participants, a total of 1,643 /ifi/ utterances were recorded, resulting in 3,286 individual vowels for formant estimations. In order to maintain a relationship between formants and the muscle activation calculated across an entire VCV utterance, an /ifi/ utterance was only accepted for analysis if F3 and F4 were successfully calculated in both /i/ vowels. In total, 392 third formants and 522 fourth formants could not be confidently identified in the spectrogram, and 241 vowels were rejected due to utterances that were aperiodic, glottalized, misarticulated, or lacking steady-state voicing. Thus, 541 /ifi/ utterances were rejected (32.9% of all data), resulting in an average of 9.8 usable /ifi/ utterances per condition per participant. The average number of usable /ifi/ utterances was not meaningfully different across conditions, with averages of 10.4, 10.6, 9.7, and 8.7 for no, mild, moderate, and maximum effort conditions, respectively.

Vocal Tract Length

The average and range of VTL values for male and female participants across all effort conditions are shown in Table 1, with no effort means of 14.20 and 15.92 cm for female and male participants, respectively. The results of the first GLM, shown in Table 2, revealed that neither effort level nor muscle activity was significant predictors of VTL.

Vocal Effort and Laryngeal Muscle Activation

Table 3 shows the results of the two ANOVAs performed on the muscle activation of the suprahyoid and infrahyoid muscles. Effort level had a significant effect on the percent activation of the suprahyoid muscles ($p < .001$) with a η_p^2 of .30, indicating a large effect size. Effort level also had a significant effect on the percent activation of the infrahyoid muscles ($p = .027$), with a η_p^2 of .11 indicating

Table 1. Mean (range) of vocal tract length estimates for male and female participants for each effort condition.

	Vocal tract length (cm)			
Gender	**No effort**	**Mild effort**	**Moderate effort**	**Maximum effort**
Female	14.20 (12.79–15.52)	14.14 (12 .91–15.50)	14.20 (13.08–15.41)	14.15 (13.11–15.53)
Male	15.92 (14.70–16.81)	16.04 (14.60–17.54)	16.03 (14.83–17.32)	16.01 (14.15–17.58)

Note. Vocal tract length was estimated from measures of the third and fourth formants.

a medium effect size. Mean values and confidence intervals for the percent activation of the suprahyoid and infrahyoid muscles are plotted per experimental condition in Figure 1. Post hoc Dunnett's tests were used to compare muscle activity at increased effort levels to muscle activity at no effort. For the percent activation of the suprahyoid muscles, moderate effort ($p_{\text{adj}} = .015$) and maximum effort ($p_{\text{adj}} < .001$) conditions were significantly higher than at no effort, with Cohen's *d* values of 0.44 and 0.75, indicating small and medium effect sizes, respectively. For the percent activation of the infrahyoid muscles, only the maximum effort condition was significantly higher ($p_{\text{adj}} = .032$) than at no effort, with a Cohen's *d* value of 0.21, indicating a small effect size.

Discussion

Adult vocal tracts can range from 13 to 20 cm in length (Lammert & Narayanan, 2015). Average female and male VTLs have been reported across literature in the ranges of 13.88–15.14 and 15.54–17.04 cm, respectively (Fitch & Giedd, 1999; Stevens, 1998; Vorperian et al., 2011). The average VTL for female (14.20 cm) and male (15.92 cm) participants were both near previously reported means. However, it is important to note that VTL estimates from formants taken during a VCV utterance are not directly comparable to those during a sustained isolated vowel, as the consonant preceding or following the vowel would likely influence vocal tract configuration as well as formant stability. These effects were minimalized by rejecting unstable utterances, only selecting the central part of each vowel, and using the same utterance type across all conditions.

We predicted that an increase in activation of the suprahyoid muscles with increased effort would result in a shortening of the VTL (McCabe & Titze, 2002; Roy & Ferguson, 2001). However, neither vocal effort nor sEMG had a significant effect on VTL, which did not support our first hypothesis. Nevertheless, we did find that vocal effort had a significant effect on suprahyoid muscle activity with a large effect size, supporting our second hypothesis. Thus, increased suprahyoid activation at increased vocal effort was not accompanied by a shortening of the VTL. Given that the suprahyoid muscles can elevate the larynx, these results were unexpected (Hixon et al., 2008; Shipp, 1975).

Vocal effort also had a significant effect on the sEMG of the infrahyoid muscles with a small effect size. In contrast to the suprahyoid muscles, the infrahyoid muscles are largely responsible for depressing the larynx (Hixon et al., 2008). This co-contraction of opposing muscles is consistent with the theoretical framework that vocal hyperfunction is due to inefficient and imbalanced levels of laryngeal muscle activity (Hillman et al., 1989). When individuals with healthy voices are asked to speak with increased vocal effort, they may be engaging in hyperfunctional behavior. Thus, an increase in activity of both the suprahyoid and infrahyoid muscles could be seen as an increase in general muscular activation. This may explain why the sEMG of the extrinsic laryngeal muscles increased with increased vocal effort even though VTL remained unchanged.

The present findings are inconsistent with previous work in individuals with vocal hyperfunction. Roy and Ferguson (2001) showed that formant frequencies decreased in a population of individuals with vocal hyperfunction following a session of manual laryngeal massage. The authors argued that this decrease in formant frequencies indicated a lengthening of the vocal tract caused by the larynx being lowered

Table 2. Analysis of variance model results for vocal tract length with participant as a random factor, effort level as a fixed factor, and the percent activation of suprahyoid and infrahyoid muscles as covariates.

Effect	*df*	*F*	*p*	Effect size (η_p^2)	Qualitative effect size
Participant	27	64.29	< .001	.96	Large
Effort level	3	0.15	.93	—	—
Suprahyoid muscle activation	1	0.02	.88	—	—
Infrahyoid muscle activation	1	2.41	.13	—	—

Note. Significance was set a priori at $p < .05$. Dashes indicate nonsignificant findings.

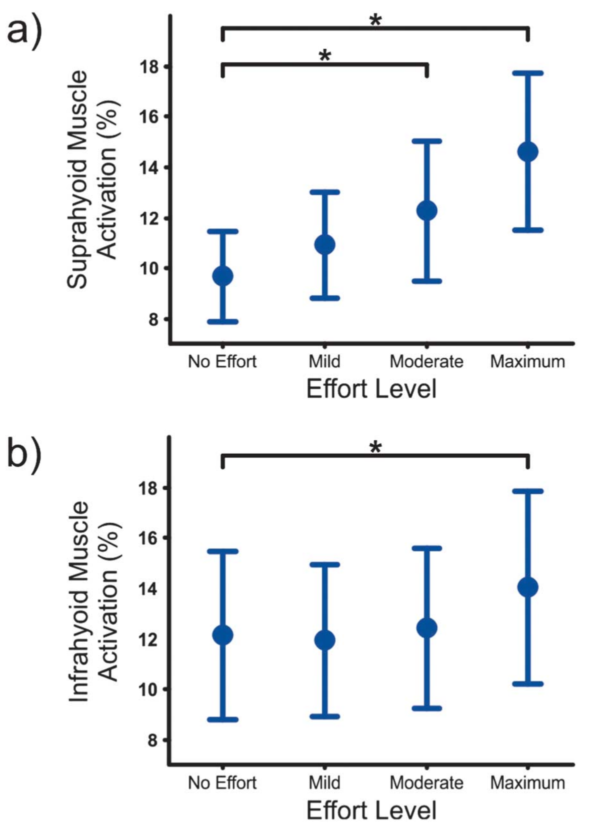
Table 3. Separate analysis of variance model results for the percent activation of suprahyoid and infrahyoid muscles with participant and effort level as model factors.

Suprahyoid muscle activation					
Effect	***df***	***F***	***p***	**Effect size (η_p^2)**	**Qualitative effect size**
Participant	27	13.37	< .001	.82	Large
Effort level	3	11.21	< .001	.30	Large
Infrahyoid muscle activation					
Effect	***df***	***F***	***p***	**Effect size (η_p^2)**	**Qualitative effect size**
Participant	27	31.87	< .001	.92	Large
Effort level	3	3.23	.027	.11	Medium

Note. Effect sizes reported for factors with significant effects at a significance level of $p < .05$.

from an atypically high position to an optimal, lower one following therapy (Aronson & Bless, 2009; Roy & Ferguson, 2001). More recently, Lowell et al. (2012) used radiography to show that individuals with muscle tension dysphonia have an elevated laryngeal position during phonation when compared to individuals with healthy voices. However, unlike prior research on VTL, studies of anterior neck sEMG in individuals with vocal hyperfunction have resulted in conflicting findings. Several studies using single sensors placed on the anterior of the neck reported increased activation in individuals with vocal hyperfunction (Milutinovic et al., 1988; Redenbaugh & Reich, 1989). Another study observed a six- to eightfold increase in sEMG placed at the perioral and supralaryngeal muscles in individuals with vocal hyperfunction (Hocevar-Boltezar et al., 1998). However, more recent studies observed no differences in the anterior neck sEMG of individuals with vocal hyperfunction (Stepp et al., 2010; Van Houtte et al., 2013).

Figure 1. Mean and 95% confidence intervals of the level of muscle activation (measured as a percentage of the maximum voluntary contraction) across each effort level for (a) the sensor corresponding to the suprahyoid muscle placement and (b) the average of the two sensors corresponding to the left and right infrahyoid muscle placement. Black bars indicate significant differences between effort levels at a value of $p_{adj} < .05$.



An elevated hyolaryngeal complex accompanied by potentially no change in sEMG suggests that the mechanisms resulting in increased vocal effort in individuals with vocal hyperfunction may be different than those used by individuals with healthy voices. One clear difference between these two populations is the timing and duration of vocal effort. In this study, individuals with healthy voices were instructed to make brief, unsustained modulations to their vocal effort. In contrast, individuals with vocal hyperfunction experience prolonged and unrelenting increases in vocal effort that are accompanied by chronic vocal fatigue (Solomon, 2008). Though it is unclear whether vocal hyperfunction precedes or is a result of vocal fatigue, this interplay is not present during the shorter modulations of vocal effort in healthy individuals. Thus, laryngeal elevation may be a compensatory mechanism to alleviate chronic vocal effort and vocal fatigue, whereas increased extrinsic laryngeal muscle activity may be a response to acute increases in vocal effort. Future work should investigate the relationship between extrinsic laryngeal muscle activity and VTL during chronic increases of vocal effort in individuals with healthy voices.

Although the relationship between VTL and formant values in Equation 1 is well established, acoustic measures can only estimate VTL (Lammert & Narayanan, 2015). However, as this study was concerned with changes in VTL across experimental conditions as opposed to the exact measure of each participant's VTL, we reasoned that an acoustic estimate of VTL was appropriate. Nevertheless, changes in acoustic estimates of VTL are not necessarily an indication of a repositioning of the larynx. Roy and Ferguson (2001) comment on these problems, indicating that several changes in articulators such as tongue height, degree of mouth opening, posterior–anterior position of the tongue, epilarynx configurations, and constriction at the level of the vocal folds may affect individual formants, and alterations in lip position can also cause a change in VTL (Roy & Ferguson, 2001). In this study, we used higher formants (F3 and F4), which are less sensitive to changes from articulatory movements (Lammert & Narayanan, 2015; Wakita, 1977), yet it is possible that higher formants are still affected by small changes to the cross-sectional area of the vocal tract or the position of the articulators. Because it is unclear how exactly individuals produce increased vocal effort, it is important to note that a combination of oral and laryngeal changes could occur in conjunction with laryngeal elevation, resulting in a constant VTL despite a raising of the larynx. For example, suprahyoid muscles are part of the complex musculature that can depress the mandible, particularly during mastication (Luschei & Goldberg, 1981). When the hyoid is fixed, the digastric muscle depresses the mandible (Möller, 1966). When combined with additional suprahyoid muscle activation as a result of increased effort, this could result in no overall change to VTL. Direct imaging of the vocal tract via magnetic resonance imaging (MRI) could provide a better understanding of the true nature of changes in VTL during increased vocal effort (Baer et al., 1991; Lowell et al., 2012). However, the use of MRI for VTL estimations has several limitations as well. The complex shape of the vocal tract can make it difficult to get a reliable length measurement from MRI during articulator movements. Additionally, MRI is incompatible with the sEMG equipment used in this study, making it impossible to observe the real-time relationships between VTL, vocal effort, and the activation of separate laryngeal muscle groups. One possible solution to this could be to use MRI to perform calibrations prior to sEMG placement in order to determine and validate a direct relationship between VTL and formant values. Due to the retrospective nature of the current investigation, this was not completed in this study. MRI calibration should be implemented in future studies exploring the relationship between sEMG and VTL.

There are limitations to this study that should be considered. A challenge that routinely comes with measurement of extrinsic laryngeal sEMG is contamination from the platysma muscle, which extends over the anterolateral aspect of the neck, from the mandible to the supraclavicular regions. As a result, platysma activation is not associated with changes to the laryngeal structure. However, the platysma lies superficial compared to surrounding muscles and therefore dominates the electrical signal when active (Stepp, 2012). Thus, increased sEMG at increased effort levels may be attributed to increased activation of the platysma instead of extrinsic laryngeal muscles.

Formants were only analyzed during the vowel portion of each VCV utterance, whereas sEMG was calculated over the entire utterance. Therefore, the relationship between sEMG measures and formants is not strictly observing the same task. However, sEMG was calculated over the entire utterance, as well as the preceding 250 ms, because muscle activation has been shown to increase at both the initiation and termination of voicing as a result of quick dynamic laryngeal movements for speech (Hirose & Gay, 1973; Sawashima et al., 1973; Stepp, Heaton, Braden, et al., 2011). Thus, recording sEMG over more than just the vowel portions is necessary for capturing the activation of laryngeal muscles that may be affected by increased vocal effort and laryngeal tension.

Additionally, VTL measures were estimated from /ifi/ utterances during laryngoscopy. The purpose of this was to capture additional potential correlates of laryngeal tension, which have been discussed in previous articles (McKenna et al., 2019; McKenna & Stepp, 2018). It is possible that, under endoscopy, participants may employ altered physiological mechanisms of vocal effort. Thus, it may be useful to examine VTL during increased levels of vocal effort in a more typical and comfortable speaking environment. Relatedly, /ifi/ utterances were selected because the /i/ vowel provides an optimal view of the vocal folds during endoscopy (McKenna et al., 2016). Future studies should investigate the relationship between vocal effort, muscle activity, and VTL in different phonetic contexts, such as in the vowel /u/, during which lip rounding is known to increase VTL and lower formants (Roy & Ferguson, 2001).

Conclusions

Formant values were used to estimate VTL during modulation of vocal effort. Suprahyoid and infrahyoid muscle activation measured via sEMG significantly increased when vocal effort was increased. VTL did not change as a result of increased vocal effort. These findings are seemingly in contrast with the shortening of VTL previously observed in individuals with vocal hyperfunction, indicating that there may be a difference in physiological mechanisms between how vocal effort manifests in these two populations.

Acknowledgments

This work was supported by the National Institutes of Health through Grants R01DC015570 (Cara E. Stepp), T32DC013017 (Christopher A. Moore), and T32DC000030 (Elizabeth A. Strickland) from the National Institute on Deafness and Other Communication Disorders. It was also supported by a Sargent College Dudley Allen Research Grant (Victoria S. McKenna) from Boston University. Thanks to Nicole Enos for sEMG data processing.

References

Aronson, A. E., & Bless, D. M. (2009). *Clinical voice disorders* (4th ed.). Thieme.

Baer, T., Gore, J. C., Gracco, L. C., & Nye, P. W. (1991). Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *The Journal of the Acoustical Society of America, 90*(2 Pt. 1), 799–828. https://doi.org/10.1121/1.401949

Boersma, P. (2002). Praat: A system for doing phonetics by computer. *Glot International, 5*(9/10), 341–345.

Boone, D. R., & McFarlane, S. (1988). *The voice and voice disorders*. Prentice Hall.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America, 106*(3), 1511–1522. https://doi.org/10.1121/1.427148

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America, 97*(5), 3099–3111. https://doi.org/10.1121/1.411872

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech and Hearing Research, 32*(2), 373–392. https://doi.org/10.1044/jshr.3202.373

Hirose, H., & Gay, T. (1973). Laryngeal control in vocal attack: An electromyographic study. *Folia Phoniatrica et Logopaedica, 25*(3), 203–213. https://doi.org/10.1159/000263688

Hixon, T. J., Weismer, G., & Hoit, J. D. (2008). *Preclinical speech science: Anatomy physiology acoustics perception*. Plural.

Hocevar-Boltezar, I., Janko, M., & Zargi, M. (1998). Role of surface EMG in diagnostics and treatment of muscle tension dysphonia. *Acta Oto-Laryngologica, 118*(5), 739–743. https://doi.org/10.1080/00016489850183287

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lammert, A. C., & Narayanan, S. S. (2015). On short-time estimation of vocal tract length from formant frequencies. *PLOS ONE, 10*(7), Article e0132193. https://doi.org/10.1371/journal.pone.0132193

Laver, J. (1980). *The phonetic description of voice quality: Cambridge studies in linguistics*. Cambridge University Press.

Lehiste, I., & Peterson, G. E. (1961). Transitions, glides, and diphthongs. *The Journal of the Acoustical Society of America, 33*(3), 268–277. https://doi.org/10.1121/1.1908638

Lindblom, B. E., & Sundberg, J. E. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *The Journal of the Acoustical Society of America, 50*(4B), 1166–1179. https://doi.org/10.1121/1.1912750

Lowell, S. Y., Kelley, R. T., Colton, R. H., Smith, P. B., & Portnoy, J. E. (2012). Position of the hyoid and larynx in people with muscle tension dysphonia. *The Laryngoscope, 122*(2), 370–377. https://doi.org/10.1002/lary.22482

Ludlow, C. L., Humbert, I., Saxon, K., Poletto, C., Sonies, B., & Crujido, L. (2007). Effects of surface electrical stimulation both at rest and during swallowing in chronic pharyngeal dysphagia. *Dysphagia, 22*(1), 1–10. https://doi.org/10.1007/s00455-006-9029-4

Luschei, E. S., & Goldberg, L. J. (1981). Neural mechanisms of mandibular control: Mastication and voluntary biting. In V. B. Brooks (Ed.), *Handbook of physiology. The nervous system* (Vol. 2, Part 2, pp. 1237–1274). American Physiological Society.

McCabe, D. J., & Titze, I. R. (2002). Chant therapy for treating vocal fatigue among public school teachers. *American Journal of Speech-Language Pathology, 11*(4), 356–369. https://doi.org/10.1044/1058-0360(2002/040)

McKenna, V. S., Diaz-Cadiz, M. E., Shembel, A. C., Enos, N. M., & Stepp, C. E. (2019). The relationship between physiological mechanisms and the self-perception of vocal effort. *Journal of Speech, Language, and Hearing Research, 62*(4), 815–834. https://doi.org/10.1044/2018_JSLHR-S-18-0205

McKenna, V. S., Heller Murray, E. S., Lien, Y.-A. S., & Stepp, C. E. (2016). The relationship between relative fundamental frequency and a kinematic estimate of laryngeal stiffness in healthy adults. *Journal of Speech, Language, and Hearing Research, 59*(6), 1283–1294. https://doi.org/10.1044/2016_JSLHR-S-15-0406

McKenna, V. S., & Stepp, C. E. (2018). The relationship between acoustical and perceptual measures of vocal effort. *The Journal of the Acoustical Society of America, 144*(3), 1643–1658. https://doi.org/10.1121/1.5055234

Milutinovic, Z., Lastovka, M., Vohradnik, M., & Janosevic, S. (1988). EMG study of hyperkinetic phonation using surface electrodes. *Folia Phoniatrica et Logopaedica, 40*(1), 21–30. https://doi.org/10.1159/000265880

Möller, E. (1966). The chewing apparatus. An electromyographic study of the action of the muscles of mastication and its correlation to facial morphology. *Acta Physiologica Scandinavica. Supplementum, 280,* 1–229.

Öhman, S. E. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America, 39*(1), 151–168. https://doi.org/10.1121/1.1909864

Redenbaugh, M. A., & Reich, A. R. (1989). Surface EMG and related measures in normal and vocally hyperfunctional speakers. *Journal of Speech and Hearing Disorders, 54*(1), 68–73. https://doi.org/10.1044/jshd.5401.68

Resnick, R., & Halliday, D. (1966). *Physics: Part II*. Wiley.

Roy, N., Bless, D. M., Heisey, D., & Ford, C. N. (1997). Manual circumlaryngeal therapy for functional dysphonia: An evaluation of short- and long-term treatment outcomes. *Journal of Voice, 11*(3), 321–331. https://doi.org/10.1016/S0892-1997(97)80011-2

Roy, N., & Ferguson, N. A. (2001). Formant frequency changes following manual circumlaryngeal therapy for functional dysphonia: Evidence of laryngeal lowering. *Journal of Medical Speech-Language Pathology, 9*(3), 169–176.

Roy, N., Ford, C. N., & Bless, D. M. (1996). Muscle tension dysphonia and spasmodic dysphonia: The role of manual laryngeal tension reduction in diagnosis and management. *Annals of Otology, Rhinology & Laryngology, 105*(11), 851–856. https://doi.org/10.1177/000348949610501102

Roy, N., Mauszycki, S. C., Merrill, R. M., Gouse, M., & Smith, M. E. (2007). Toward improved differential diagnosis of adductor spasmodic dysphonia and muscle tension dysphonia. *Folia Phoniatrica et Logopaedica, 59*(2), 83–90. https://doi.org/10.1159/000098341

Rutkove, S. B. (2007). Introduction to volume conduction. In *The clinical neurophysiology primer* (pp. 43–53). Springer.

Sawashima, M., Kakita, Y., & Hiki, S. (1973). Activity of the extrinsic laryngeal muscles in relation to Japanese word accent. *Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo), 7,* 19–25.

Shipp, T. (1975). Vertical laryngeal position during continuous and discrete vocal frequency change. *Journal of Speech and Hearing Research, 18*(4), 707–718. https://doi.org/10.1044/jshr.1804.707

Shipp, T. (1987). Vertical laryngeal position: Research findings and application for singers. *Journal of Voice, 1*(3), 217–219. https://doi.org/10.1016/S0892-1997(87)80002-4

Solomon, N. P. (2008). Vocal fatigue and its relation to vocal hyperfunction. *International Journal of Speech-Language Pathology, 10*(4), 254–266. https://doi.org/10.1080/14417040701730990

Stepp, C. E. (2012). Surface electromyography for speech and swallowing systems: Measurement, analysis, and interpretation. *Journal of Speech, Language, and Hearing Research,* *55*(4), 1232–1246. https://doi.org/10.1044/1092-4388(2011/11-0214)

Stepp, C. E., Heaton, J. T., Braden, M. N., Jette, M. E., Stadelman-Cohen, T. K., & Hillman, R. E. (2011). Comparison of neck tension palpation rating systems with surface electromyographic and acoustic measures in vocal hyperfunction. *Journal of Voice, 25*(1), 67–75. https://doi.org/10.1016/j.jvoice.2009.08.001

Stepp, C. E., Heaton, J. T., Jetté, M. E., Burns, J. A., & Hillman, R. E. (2010). Neck surface electromyography as a measure of vocal hyperfunction before and after injection laryngoplasty. *Annals of Otology, Rhinology & Laryngology, 119*(9), 594–601. https://doi.org/10.1177/000348941011900905

Stepp, C. E., Heaton, J. T., Stadelman-Cohen, T. K., Braden, M. N., Jette, M. E., & Hillman, R. E. (2011). Characteristics of phonatory function in singers and nonsingers with vocal fold nodules. *Journal of Voice, 25*(6), 714–724. https://doi.org/10.1016/j.jvoice.2010.06.003

Stevens, K. N. (1998). *Acoustic phonetics*. MIT Press.

Stevens, K. N., & House, A. S. (1955). Development of a quantitative description of vowel articulation. *The Journal of the Acoustical Society of America, 27*(3), 484–493. https://doi.org/10.1121/1.1907943

Sundberg, J., & Nordström, P.-E. (1976). Raised and lowered larynx-the effect on vowel formant frequencies. *Speech Transmission Laboratory. Quarterly Progress and Status Reports, 17*(2–3), 35–39.

Van Houtte, E., Claeys, S., D'Haeseleer, E., Wuyts, F., & Van Lierde, K. (2013). An examination of surface EMG for the assessment of muscle tension dysphonia. *Journal of Voice, 27*(2), 177–186. https://doi.org/10.1016/j.jvoice.2011.06.006

Vorperian, H. K., Wang, S., Schimek, E. M., Durtschi, R. B., Kent, R. D., Gentry, L. R., & Chung, M. K. (2011). Developmental sexual dimorphism of the oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of Speech, Language, and Hearing Research, 54*(4), 995–1010. https://doi.org/10.1044/1092-4388(2010/10-0097)

Wakita, H. (1977). Normalization of vowels by vocal-tract length and its application to vowel identification. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 25*(2), 183–192. https://doi.org/10.1109/TASSP.1977.1162929

Witte, R. S., & Witte, J. S. (2010). *Statistics*. Wiley.