ASHA
American
Speech-Language-Hearing
Association

JSLHR

Research Article

Impact of Vocal Effort on Respiratory and Articulatory Kinematics

Defne Abur,[a] Joseph S. Perkell,[a,b] and Cara E. Stepp[a,c,d]

[a] Department of Speech, Language and Hearing Sciences, Boston University, MA [b] Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge [c] Department of Biomedical Engineering, Boston University, MA [d] Department of Otolaryngology-Head & Neck Surgery, Boston University School of Medicine, MA

ARTICLE INFO

Article History:
Received June 10, 2021
Revision received July 27, 2021
Accepted August 24, 2021

Editor-in-Chief: Bharath Chandrasekaran
Editor: Kate Bunton

https://doi.org/10.1044/2021_JSLHR-21-00323

ABSTRACT

Purpose: The goal of this study was to examine the effects of increases in vocal effort, without changing speech intensity, on respiratory and articulatory kinematics in young adults with typical voices.
Method: A total of 10 participants completed a reading task under three speaking conditions: baseline, mild vocal effort, and maximum vocal effort. Respiratory inductance plethysmography bands around the chest and abdomen were used to estimate lung volumes during speech, and sensor coils for electromagnetic articulography were used to transduce articulatory movements, resulting in the following outcome measures: lung volume at speech initiation (LVSI) and at speech termination (LVST), articulatory kinematic vowel space (AKVS) of two points on the tongue dorsum (body and blade), and lip aperture.
Results: With increases in vocal effort, and no statistical changes in speech intensity, speakers showed: (a) no statistically significant differences in LVST, (b) statistically significant increases in LVSI, (c) no statistically significant differences in AKVS measures, and (d) statistically significant reductions in lip aperture.
Conclusions: Speakers with typical voices exhibited larger lung volumes at speech initiation during increases in vocal effort, paired with reduced lip displacements. To our knowledge, this is the first study to demonstrate evidence that articulatory kinematics are impacted by modulations in vocal effort. However, the mechanisms underlying vocal effort may differ between speakers with and without voice disorders. Thus, future work should examine the relationship between articulatory kinematics, respiratory kinematics, and laryngeal-level changes during vocal effort in speakers with and without voice disorders.
Supplemental Material: https://doi.org/10.23641/asha.17065457

Vocal effort is the perceived exertion as experienced by the speaker and not the listener (Banister, 1979; Borg, 1982; Hunter et al., 2020), and previous work implicates both psychological and physiological factors (McKenna et al., 2019; Van Mersbergen et al., 2008) as contributors during speech production. Vocal effort is one of the most prevalent symptoms reported by speakers with voice disorders (Altman et al., 2005; Bach et al., 2005; Roy et al., 2005; Smith et al., 1998), speakers who are high-voice users (Bermúdez de Alvear et al., 2011; Sampaio et al., 2012; Smith et al., 1997), and approximately 10% of older adults with no history of voice disorders (Merrill et al., 2013); thus, many people are affected. It is important to characterize contributors to vocal effort, as it can negatively affect crucial aspects of daily life such as social interactions and job performance and attendance (Roy et al., 2005). Physiological coordination across all speech subsystems (laryngeal, respiratory, and articulatory) is necessary for speech production and indirect evidence indicates that all speech subsystems may be affected by increases in vocal effort.

Physiological components of vocal effort have been extensively examined in various speaker populations with

---

Correspondence to Defne Abur: dabur@bu.edu. ***Disclosure:*** *Cara E. Stepp has received consulting fees from Altec, Inc./Delsys, Inc., companies focused on developing and commercializing technologies related to human movement. Stepp's interests were reviewed and are managed by Boston University in accordance with their conflict of interest policies. The other authors have declared that no other competing financial or nonfinancial interests existed at the time of publication.*

regard to the laryngeal subsystem. When speakers with typical voices increase their vocal effort, laryngeal muscle tension, supraglottal compression, and subglottal pressure show deviations from values during typical speech (Lien et al., 2015; McKenna et al., 2019; Rosenthal et al., 2014). These same laryngeal-level changes have been found in populations with voice disorders who report increased vocal effort, such as speakers with hyperfunctional voice disorders (HVDs; Espinoza et al., 2017; Hillman et al., 1989; Redenbaugh & Reich, 1989; Stager et al., 2000; Stepp et al., 2010), and spasmodic dysphonia (Ludlow, 2009, 2011; McCall et al., 1973).

The anatomical connections between the laryngeal and respiratory subsystems provide a theoretical rationale for physiological interactions between respiration and voicing during increased vocal effort, which is supported by empirical evidence. Specifically, when lung volume increases during inhalation, the laryngeal structures, including the cricoid cartilage, are pulled downward (Fink, 1978; Macklin, 1925; Sundberg et al., 1989). As the space between the cricoid and thyroid cartilages grows wider, the vocal folds shorten and become more relaxed; these changes result in compensatory increases in cricothyroid muscle tension (Sundberg et al., 1989; Zenker & Zenker, 1960). Thus, when speakers with typical voices produce speech at higher intensities, the concurrent lung volume increases exert a greater pull downward on the laryngeal structures and a greater increase in cricothyroid muscle tension during voicing. Given that increases in vocal effort and intensity have some shared underlying mechanisms, that is, both co-occur with increases in subglottal pressure (Espinoza et al., 2017; Fryd et al., 2016), an interplay between respiratory and laryngeal features during increased vocal effort is also expected.

Studies in individuals with HVDs reporting increased vocal effort have implicated differences in respiratory features compared with speakers with typical voices (Dastolfo et al., 2016; Holmberg et al., 2003; Sapienza & Stathopoulos, 1994; Sapienza et al., 1997), suggesting that the laryngeal features associated with increased vocal effort interact with respiratory patterns. Compared with controls, speaker populations who report increased vocal effort have demonstrated lower lung volume at speech initiation (LVSI), lower lung volume at speech termination (LVST), and lower total lung volume excursions (LVEs), measured with respiratory inductance plethysmography and similar instrumentation (Hixon & Putnam, 1983; Iwarsson & Sundberg, 1999; Lowell et al., 2004, 2008). However, speech intensity may interact with these observed relationships. One study found that, compared with controls, speakers with HVDs showed lower LVSI and LVST during loud speech but not during typical speech (Iwarsson & Sundberg, 1999). In speakers with typical voices, the opposite trend has been reported: Compared with typical speech, loud speech results in increases in both LVSI and LVST (Dromey & Ramig, 1998; Huber et al., 2005; Stathopoulos & Sapienza, 1997). A pressed voice, associated with vocal features of HVDs (Berry et al., 2001; Grillo & Verdolini, 2008; Jiang & Titze, 1994), contains more energy at higher frequencies and can also contribute to the production and perception of increased loudness (Södersten et al., 2005). Thus, speech intensity appears to interact with respiratory and laryngeal dynamics differently in speakers with and without voice disorders.

Respiratory changes have been observed in speakers with typical voices when they modulate vocal effort, but the impact of speech intensity remains unclear. In one study, speech produced before and after a vocal loading task (40 min of reading over background noise) resulted in lower LVSI, greater LVST, and higher self-perceptions of vocal effort in speakers with typical voices (Sundarrajan et al., 2017). The observed patterns in LVSI and LVST may suggest a trend for decreased total LVE (equivalent to LVSI–LVST). Additionally, these respiratory kinematic changes were accompanied by changes in speech intensity: Speakers demonstrated lower speech intensity after vocal loading. In contrast, McKenna et al. (2017) trained speakers with typical voices to increase their vocal effort and found that effortful productions were associated with greater mean speech intensity as well as greater LVE. Taken together, these observed relationships suggest that respiratory patterns during increased vocal effort are related to speech intensity (likely due to changes in breathing demands) in speakers with typical voices. In order to fully clarify how vocal effort impacts respiratory kinematics during typical speech, the relationship between vocal effort and respiratory kinematics needs to be examined while speech intensity is held constant. This is particularly relevant to speakers with HVDs, who demonstrate increased vocal effort without increased speech intensity (Espinoza et al., 2017; Friedman et al., 2013; Hillman et al., 1989; Zietels et al., 2008).

Anatomical connections between the laryngeal and articulatory subsystems and articulatory–acoustic evidence from speakers with HVDs point to a potential influence of vocal effort on the articulatory subsystem of speech as well. Speakers with HVDs demonstrate an elevated position of the larynx and hyoid bone (Lowell, Kelley, Colton, et al., 2012), presumably due to dysregulations in extrinsic laryngeal muscle tension (Aronson, 1990; Morrison et al., 1983), which would alter vocal tract length and articulatory–acoustic features during voicing (Rosner & Pickering, 1994). Increased extrinsic laryngeal muscle tension in HVDs may also restrict movements of the tongue base, which would reduce the tongue's physiological range of motion and articulatory working space. In support of this, articulatory–acoustic measures have shown that speakers with HVDs demonstrate increased vowel space area (reflecting greater range of tongue motion) and reduced third formant frequency (reflecting a lower laryngeal position and increased vocal tract length) following voice therapy and muscle tension relief (Roy & Ferguson, 2001; Roy et al., 2009).

Thus, prior to voice therapy and symptom relief, speakers with HVDs appear to have smaller vowel space area and higher third formant frequencies. Given this articulatory–acoustic evidence for restricted tongue movement and elevated laryngeal height with increased vocal effort, articulatory kinematics could be impacted by increased vocal effort.

Additional support for the influence of vocal effort on the articulatory system comes from studies that report a relationship between laryngeal height changes during phonation and speech articulation with increases in speech intensity, which shares commonalities with vocal effort (as reviewed earlier). During increases in speech intensity, driven by both the respiratory and laryngeal subsystems (Finnegan et al., 2000), concurrent increases in articulatory kinematic measures of tongue and lip movements have been found when using electromagnetic articulography (EMA; e.g., Mefferd, 2015; Mefferd & Dietrich, 2020; Perkell et al., 1992; Schönle et al., 1987; Tiede et al., 2010; Wang et al., 2016). In speakers with typical voices, prior work has found larger tongue displacements for loud intensities compared with conversational speech intensities (Mefferd & Green, 2010; Whitfield et al., 2018), presumably because the mouth is opened wider in loud speech to more effectively project speech sounds. Measures of vertical lip displacement increase as speech intensity increases in speakers with typical voices (Dromey & Ramig, 1998; Huber & Chandrasekaran, 2006), which is consistent with a wider mouth opening. Furthermore, speakers with spasmodic dysphonia (Shoffel-Havakuk et al., 2019) demonstrate prolonged lowering of the lower lip and reduced coordination of the upper and lower lips (Dromey et al., 2007). Taken together, these collective findings suggest that increased vocal effort could influence both tongue and lip kinematics.

In summary, the mechanisms involved in increased vocal effort at the laryngeal level have been studied extensively, but the respective roles of the respiratory and articulatory subsystems during vocal effort have not been explored as thoroughly and warrant investigation. In speakers without voice disorders, the influence of vocal effort, alone, on speech subsystems is not fully clear due to concurrent changes in speech intensity in prior investigations. Because speakers with HVDs demonstrate self-perception of increased vocal effort at typical speech intensities, it is important to elucidate the mechanisms underlying vocal effort without concurrent changes to speech intensity.

Objectives of This Study

The purpose of this work was to examine how respiratory kinematic and articulatory kinematic features are influenced by increased vocal effort when speakers with typical voices are instructed to change vocal effort while maintaining a steady level of speech intensity. Speakers with typical voices were chosen because they have the capacity to increase vocal effort and can serve as their own controls (their typical speech patterns can be used as a baseline to compare to effortful productions). Ten participants read an experimental passage aloud under three speaking conditions: typical (or baseline), with mild vocal effort, and with maximal vocal effort. Respiratory inductance plethysmography and EMA were used to examine changes in lung volume and kinematics of the tongue and lips as a function of speaking condition.

The first aim of the study was to investigate how increasing vocal effort, without changing speech intensity, affects LVSI and termination. We hypothesized that, during increases in vocal effort without changes in speech intensity, speakers with typical voices would exhibit higher LVSI and lower lung volume at speech termination (indicating larger overall LVEs) due to increased inspiratory demands under increased vocal effort. The second aim of this study was to examine articulatory kinematics during increases in vocal effort with minimal changes in speech intensity. Given the high prevalence of vocal effort in speakers with HVDs and the evidence that they demonstrate reduced vowel space areas (Roy et al., 2009), we hypothesized that articulatory kinematics of the tongue would show a reduction in measures of articulatory kinematic working space (AKVS) and that maximum lip aperture (LA) would reduce with increased vocal effort.

Method

Participants

A total of 10 speakers with typical voices participated in the study. Five speakers were assigned female at birth (four cisgender and one nonbinary) and five speakers were assigned male at birth (all cisgender). Speakers were aged 17–29 years (*M* = 21.3, *SD* = 3.4) with no history of neurological, speech, language, hearing, or respiratory disorders. None of the speakers had a history of speech-language therapy. Given that vocal training and experience playing wind instruments can influence respiratory kinematics during speech (Mendes et al., 2006; Zuskin et al., 2009), participant recruitment excluded those with any history of formal vocal training or experience playing wind instruments. Furthermore, no participants reported a history of smoking, which can affect laryngeal parameters (Awan, 2011). All participants were native speakers of American English, to control for possible effects of accent on articulatory–acoustic measures (e.g., vowel formant features or articulatory patterning; Robb & Chen, 2008; Vatikiotis-Bateson & Kelso, 1993). All participants completed written consent in compliance with the Boston University Institutional Review Board. For the 17-year-old participant, additional consent was obtained from the participant's legal guardian per Boston University Institutional Review Board guidelines.

Hearing Screening

To control for possible effects of hearing impairment on speech subsystems (Higgins et al., 1994), a hearing screening was completed using 3M E-A-RTONE Gold 3A earphones and a Grason-Stadler GSI 18 Screening Audiometer. All participants passed the screening (a threshold of 25 dB or less at 125, 250, 1000, 2000, 4000, and 8000 Hz; American Speech-Language-Hearing Association [ASHA], 1997).

Vocal Effort Training

Prior to data collection, speakers were provided with training on how to produce speech with increased vocal effort. First, participants were given a printed version of the experimental speech material and instructed to read it aloud in their typical speaking voice. The speech material consisted of one paragraph with four sentences and a total of 100 words (see Figure 1). The sentences were designed to be long (> 22 words) in order to make the reading task challenging for the respiratory system. Specifically, using longer sentences ensured that each speaker would be induced to a large range in lung volume initiation and termination across all speaking conditions. The experimental speech corpus was designed as part of a larger study to include all the corner vowels (/ɑ/, /i/, /u/, and /æ/), diphthong transitions from low to high vowels (e.g., light contains the transition from /ɑ/ to /i/), and vowel–fricative–vowel utterances (e.g., Reef eats contains the utterance /ifi/). Reading the text in a typical speaking voice was considered the "baseline" condition throughout the study.

Next, each participant was given the definition of vocal effort as "*how easy or difficult it is to talk in terms of how much effort, strain, discomfort, and/or fatigue you perceive when using your voice*" (McKenna & Stepp, 2018). Additionally, speakers were told "*Think only of vocal effort [experimenter pointing to larynx], not mental effort or concentration it takes to produce an effortless voice*" (van Leer & van Mersbergen, 2017). Following this definition, participants were trained on how to increase vocal effort using experiential anchoring (adapted from van Leer & van Mersbergen, 2017). This design choice was used in order to examine unbiased, speaker-created targets of vocal effort.

To induce increased vocal effort to a "mild" level, the following instructions were given to participants "*Think of a time when you were aware of slight effort, strain, discomfort and/or fatigue when talking. For example, think of a time when you talked for a long time, without losing your voice, and felt mild amounts of effort, strain, discomfort, and/or fatigue when talking. I would like you to read the next set using that amount of vocal effort. When increasing your vocal effort, you want to make sure you are still using your voice, so please do not whisper or speak in voice breaks.*" Subsequently, participants were asked to practice reading the passage aloud with mild vocal effort until they reported that it matched their experiential anchor.

To increase vocal effort to a "maximal" level, speaking through severe laryngitis was used as the experiential anchor (based on the maximum in the scale used by van Leer & van Mersbergen, 2017). Participants received the following instructions: "*Think of a time when you had to talk through laryngitis, and talking was extremely effortful, strained, uncomfortable, and/or fatiguing. I would like you to read the next set using that amount of vocal effort. When increasing your vocal effort, you want to make sure you are still using your voice, so please do not whisper or speak in voice breaks.*" Participants practiced reading the passage with maximal vocal effort until they reported that it matched their experiential anchor.

All participants demonstrated changes in voice quality during vocal effort conditions, subjectively determined during by the experimenter, and reported an awareness of increasing their vocal effort after training. This was confirmed offline via auditory-perceptual ratings by a voice-specializing speech-language pathologist blinded to the paradigm (see *Auditory Perceptual Ratings*).

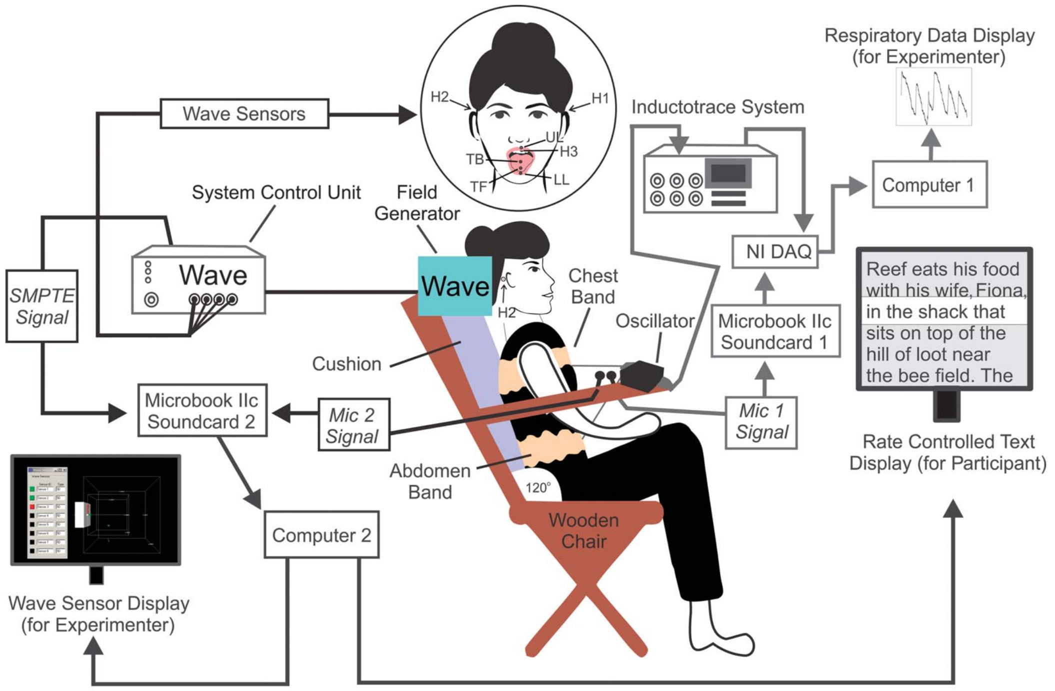
Data Collection

Simultaneous respiratory, articulatory, and acoustic data were collected in a sound-attenuated booth. This was achieved by using respiratory inductance plethysmography and EMA with hardware and software detailed below (see Figure 2 for a schematic). To avoid disturbing the electromagnetic field of the EMA system, participants were asked to remove all metal objects on their body outside the sound booth to prevent measurement artifacts. Participants were seated in a wooden chair in front of a monitor displaying the experimental text. To control for the effects of speaking rate on articulatory kinematics (Dromey & Ramig, 1998; Mefferd & Green, 2010) and on vocal effort (McKenna & Stepp, 2018), reading the text was rate controlled

Figure 1. The text used for the experimental prompt.

"Reef eats his food with his wife, Fiona, in the shack that sits on top of the hill of loot near the bee fields. The couple brew fruity hops together and have to wear heavy black boots because of the messy process, but the wee fees make it worth it. They have a cute black dog named Tia with a small cough who acts sweet when she gets a happy feast of treats. Tia loves to snoop around for the sack of toffee tied tightly in the kitchen or the night light that is behind Fiona and Reef's fancy suits."

Figure 2. Schematic of experimental setup. Gray lines indicate connections related to the respiratory data collection, whereas black lines indicate connections related to articulatory kinematic data collection. Participants were seated in a wooden chair in a slightly reclined position to avoid disturbing the electromagnetic field of the electromagnetic articulography system (NDI Wave Speech Research System). Participants were fitted with two respiratory bands, one around the chest (below the axilla) and one around the abdomen (at the level of the navel), that connected to the oscillator of an Inductotrace respiratory inductive plethysmograph system. Seven sensors were placed on the head as labeled in the circle at the top of the schematic: H1 = head correction sensor 1 (placed over the left mastoid); H2 = head correction sensor 2 (placed over the right mastoid); H3 = head correction sensor 3 (placed on the upper incisor); TF = tongue front sensor, placed on the tongue blade (5 mm from tongue apex); TB = tongue body sensor (10–20 mm from TF); UL = upper lip sensor (placed on the upper lip vermilion border at midline); LL = lower lip sensor (placed on the lower lip vermilion border at midline). The text for the speech production task was displayed on a screen with a white bar moving down the screen at a controlled rate during each experimental trial. SMPTE = Society of Motion Picture and Television Engineers.



during the experiment: A white bar moved horizontally down the screen at a steady rate, and participants were asked to follow the bar while speaking (see Figure 2).

Each experimental trial lasted 110 s and contained three repetitions of the passage always in the same order: a baseline (typical speech), a mild vocal effort, and a maximal vocal effort condition. The order was kept consistent to ensure that speakers could clearly differentiate their anchors for the mild and maximal effort productions. Each of five trials contained all three speaking conditions to control for fatigue across the session (e.g., if all baseline productions were completed first, respiratory and articulatory kinematic changes observed during the later vocal effort conditions could have resulted from vocal fatigue).

Prior to the five consecutive trials (15-passage repetitions; five for each vocal effort condition), all speakers were given instructions to maintain the same volume while increasing vocal effort (McKenna & Stepp, 2018). Before beginning each trial, the experimenter used an electrolarynx (held at the corner of the speaker's mouth) to generate a tone, record the acoustic signal with the microphone, and note the concurrent sound pressure level (in dB SPL) value on a CM-160 (Galaxy Audio, Wichita, KS) sound-level meter held level with the microphone (15 cm from the mouth). This information was later used to determine corresponding dB SPL values from the acoustic signal amplitude in Praat. Throughout the experiment, speech intensity was monitored using the sound-level meter held level with the microphone. The measurement error of speech intensity is thought to be at least 2.5 dB (based uncertainties in microphone distance and the inherent error of sound-level meters; Švec & Granqvist, 2018). For differences greater than 2.5 dB, visual feedback was provided by the experimenter pointing upward or downward to indicate that they needed to increase or decrease their speech intensity, respectively.

Respiratory Data Collection

Respiratory kinematic data were collected using an Inductotrace respiratory inductive plethysmograph system (Ambulatory Monitoring, Inc, Ardsley, New York) connected to wire coils embedded in two flexible bands, one placed around the rib cage and one around the abdomen of each participant. The Inductotrace equipment was configured according the manufacturer's guidelines (Ambulatory Monitoring, Inc, Ardsley, New York). Consistent with prior work (e.g., Sundarrajan et al., 2017), the first band was placed at a level just beneath the axilla and the second band was placed at the level of the navel (see Figure 2). The coils in the two bands measure changes in band inductance (due to expansion and compression) of the rib cage and abdomen, reflecting lung volume during inhalation and exhalation. Additionally, the acoustic (sound pressure) signal was collected with an omnidirectional microphone (MX153, Shure, Niles, IL) located 15 cm from the participant's mouth. The microphone signal was preamplified using a Microbook IIc soundcard (MOTU, Cambridge, MA) and acquired using a NI-DAQ (NI USB-6212; National Instruments, Austin, TX) to time align the respiratory and articulatory data. Respiratory inductance plethysmography and microphone signals were acquired together with the NI-DAQ at a sampling rate of 11025 Hz. All signals were displayed in real time for an experimenter throughout the study to ensure no movement artifact was present in the recording (Computer 1; see Figure 2).

Participants were seated with a slight recline to assure a constant reproducible posture across speakers over time and to avoid interference of the respiratory band coils with the electromagnetic field of the EMA system. The chair was set to a 120° reclined position (i.e., reclined 30° from the upright position; see Figure 2). A cushion was placed on the chair for postural stability and comfort, which slightly reduced the incline. As a result, each participant was seated at approximately a 110° reclined position (i.e., reclined 20° from the upright position). Although seating position can influence respiratory kinematics, prior work has demonstrated typical vital capacity and respiratory kinematic measures when speakers are sitting in a 120° supine position (Heller Murray et al., 2018).

Once participants were fitted with the respiratory bands and seated in a comfortable position on the chair, they were asked to relax for a few minutes. During this time, the experimenter recorded tidal breathing for 1–2 min. Next, participants were asked to perform three consecutive maximal displacement (MD) maneuvers (based on Rochet-Capellan & Fuchs, 2013). The MD maneuver was described as "inhaling maximally then exhaling maximally," which is similar to the instructions used for a vital capacity measurement. Lung volume calibrations, which allow for comparisons of lung volume estimates in liters, were not preferred in this study due to time constraints with articulatory sensor setup and because only within-speaker changes were examined. Hence, lung volume was quantified as the percent of MD of the weighted sum of the Inductotrace band signals (in volts), rather than vital capacity (in liters), to represent a speaker's full range of inhalation and exhalation.

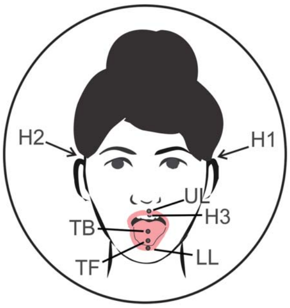
Articulatory Data Collection

An NDI Wave Speech Research System (Northern Digital Inc., Waterloo, Ontario, Canada) was used to collect EMA data at a sampling rate of 100 Hz using the NDI WaveFront Software v1.1.1. An omnidirectional microphone (MX153, Shure, Niles, IL), located 15 cm from the participant's mouth (at a 45° angle from the midline), was used to collect an acoustic signal simultaneously with the EMA data (see Figure 2). The microphone signal was acquired using a Microbook IIc soundcard (MOTU, Cambridge, MA) with a sampling rate of 22050 Hz. This acoustic signal was used for determining acoustic measures of speech, time aligning the respiratory kinematic and articulatory kinematic data, confirming steady speech intensity across conditions, and for auditory-perceptual ratings of speech.

Seven small sensor coils (inductors) were attached to the head to capture data in three dimensions: x (lateral), y (vertical), and z (anteroposterior). The coils were connected with fine insulated wires to the signal processing electronics of the wave system. During sensor placement and the experimental task, positional data were displayed for the experimenter to ensure that sensors remained in place (see Figure 2).

First, three sensors were attached to the head to provide data to define a head-based, three-dimensional coordinate system (Cler et al., 2017): head correction sensor 1 (H1) was placed on the left mastoid with medical tape, head correction sensor 2 (H2) was placed on the right mastoid with medical tape, and head correction sensor 3 (H3) was placed on the mucosa between the upper central incisors with stoma adhesive (ConvaTec, Bridgewater, NJ). Next, the location of each speaker's maxillary occlusal plane was obtained with three sensors attached to a 54 mm × 86 mm polyvinyl chloride card. One sensor was at the front of the card (approximately behind the diastema of the front teeth), and two sensors were at the back corners of the card. Each participant was instructed to align the back corners of the card with their back molars and bite lightly on the card while positional data were recorded. Lastly, the three sensors were removed from the card and attached to the (a) tongue front (TF; 5 mm from tongue tip, i.e., the tongue blade), (b) tongue body (TB; 10–20 mm from TF), and (c) vermilion border of the upper lip at midline (UL). An additional sensor was attached to the vermilion border of the lower lip at midline (LL; see Figure 3 for schematic of all sensor locations). These locations were selected because prior work has shown that sensor placements at the TB, TF, UL, and LL regions provide useful data on the articulation of words and sentences (Wang et al., 2016). These four sensors were attached with dental adhesive (high-viscosity PeriAcryl, GluStich, Inc., Delta, BC, Canada). All participants began experimental speech tasks at least 10 min after sensors were placed on

Figure 3. Schematic of electromagnetic articulography sensor placement. Seven sensors were placed on the head as labeled in the circle at the top of the schematic: H1 = head correction sensor 1 (placed over the left mastoid); H2 = head correction sensor 2 (placed over the right mastoid); H3 = head correction sensor 3 (placed between the upper central incisors); TF = tongue front sensor, placed on the tongue blade (5 mm from tongue apex); TB = tongue body sensor (10–20 mm from TF); UL = upper lip sensor (placed on the upper lip vermilion border at midline); LL = lower lip sensor (placed on the lower lip vermilion border at midline).



the lips and tongue to allow for habituation to sensors prior to experimental data collection (Dromey et al., 2018).

Respiratory Data Analysis

Custom MATLAB (MathWorks, 2016) scripts were used to time align the respiratory kinematic, articulatory kinematic, and acoustic data and to separate each speech production trial by condition (baseline, mild vocal effort, and maximal vocal effort). Lung volume changes over time were approximated using respiratory inductance plethysmography, by calculating a weighted sum of the output signals from the rib cage and abdomen coils (Konno & Mead, 1967). On the basis of prior work in typical young adults (Banzett et al., 1995; McKenna & Huber 2019), a factor of 2:1 (rib cage to abdomen coil signals) has shown reliable estimates of measured contributions of the rib cage and abdomen to overall lung volume. Thus, a set factor of 2:1 was used in this work to determine a weighted sum for analysis.

Prior to analyzing respiratory measures during speech production, each speaker's resting expiratory level (REL) and MD were calculated. The REL was defined as the average termination level across the tidal breathing recording. MD was defined as the largest breathing excursion (maximal inhalation level to maximal exhalation level) during the MD recording. A custom MATLAB (MathWorks, 2016) script was used to visualize and determine the troughs and peaks from tidal breathing to calculate REL (dotted line; see panel A on Figure 4). The experimenter manually selected each trough and peak from a display of the displacement over time, and the script determined the exact minimum or maximum value closest to the location of the click. The same script was used to determine trough and peaks for the MD (vertical line with arrows; see panel A on Figure 4).

A custom MATLAB (MathWorks, 2016) script was used to calculate LVSI and LVST, referenced to REL (Heller Murray et al., 2018). Using this method, positive values indicate a lung volume above the REL and negative values indicate a lung volume below the REL. Because vital capacity (maximum breathing excursion in liters) was not determined in this work, MD was used to express respiratory measures on a common scale across speakers (Rochet-Capellan & Fuchs, 2013). The following equations were used to determine LVSI and LVST relative to REL as % MD:

$$LVSI = \frac{(Inspiration\ at\ speech\ initiation - REL)}{MD} \times 100; \tag{1}$$

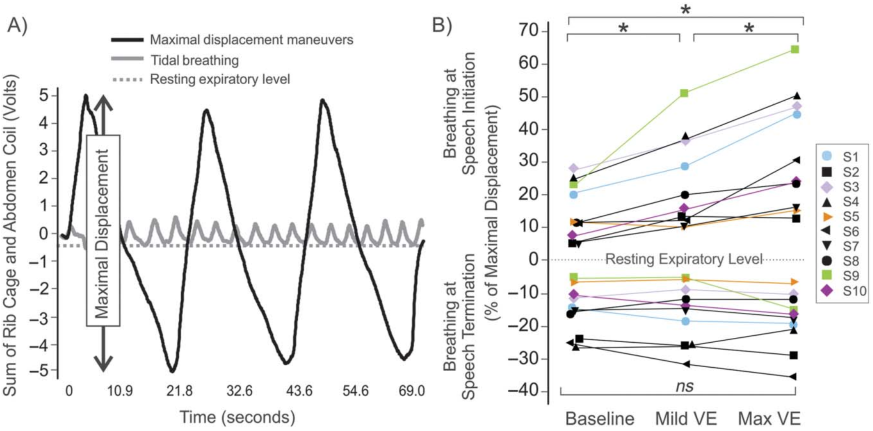
$$LVST = \frac{(Expiration\ at\ speech\ termination - REL)}{MD} \times 100. \tag{2}$$

Articulatory Data Analysis

EMA data were exported from NDI WaveFront v1.1.1 software. Custom MATLAB (MathWorks, 2016) scripts were used to low-pass filter the data using a third-order Butterworth filter. A 5-Hz cutoff was used for the three reference sensors, and a 20-Hz cutoff was used for all other sensors (per Tiede et al., 2010). EMA data were corrected for head movement using the three reference sensors (H1, H2, and H3) and referenced to each speaker's own oral cavity using the maxillary occlusal plane measurement. The resulting data for each speaker had an origin corresponding to the midline point between the back molars, aligned with the diastema of the upper central incisors.

For the experimental (speech) data, each trial and condition were extracted after time aligning the respiratory and articulatory measures with the acoustic microphone signals. Sensor trajectories in the *y* (vertical) and *z* (anteroposterior) dimensions were determined for the TB, TF, UL, and LL sensors. Prior work indicates that movement in the *x* (lateral) dimension is small and does not contain meaningful information during speech in individuals with typical patterns of articulation (Westbury, 1994). For this reason, the *x* dimension was omitted from analyses of the experimental data (as in prior studies in speakers with typical voices; Cler et al., 2017; Wang et al., 2016). For each of the tongue sensors (TB and TF), the AKVS was calculated using the methods described in Whitfield et al. (2018). The AKVS has shown sensitivity to kinematic, sentence-level changes in supraglottal articulations that accompany changes in voice production (i.e.,

Figure 4. Panel A: Tidal breathing (gray line) and three maximal displacement (MD) maneuvers (black line) are shown for an example participant (S9). The *y*-axis is the summed signal acquired using the respiratory inductive plethysmograph system recordings of the rib cage and abdomen coils, with a set 2:1 correction factor applied. The resting expiratory level (REL), corresponding to the average of the expiratory troughs in the tidal breathing measurement, is denoted by the dotted grey line. The magnitude of the MD is labeled with arrows and corresponds to the difference between the amplitude of the highest peak and the lowest trough produced across the three maneuvers. Panel B: Breathing at speech initiation and termination are displayed as the percent of each female (light-colored shapes) and male (black shapes) speaker's MD for the baseline, mild vocal effort (VE), and maximum VE conditions. The zero value on the *y*-axis (shown in the dotted line) corresponds to each speaker's REL. *Indicates statistically significant differences between speaking conditions, and *ns* denotes no statistically significant differences.



increased intensity; Whitfield et al., 2018) in speakers with typical speech. This measure was chosen to quantify the overall extent of articulatory excursions (articulatory kinematics) during running speech in this work. The AKVS corresponds to the square root of the generalized variance of the *y* and *z* trajectories of the sensor. The generalized variance is defined as the product of the variance in *y* trajectory, variance in the *z* trajectory, and the proportion of unshared variance between the two dimension trajectories. The AKVS was calculated for the TB and TF sensors for each trial and averaged within condition, resulting in six AKVS values for each speaker (two for each condition).

To examine changes in LA, which can accompany changes in voice production (Dromey et al., 2007), the displacement between the UL and LL sensors was calculated in the *y* dimension for each trial and condition. LA was averaged across trials for each condition, resulting in three values for each speaker (one for each condition).

Descriptive Analysis: Auditory-Perceptual Ratings and Acoustic Measures of Voice

Descriptive analysis was used to explore whether there were auditory-perceptual and acoustic changes to voice when speakers were told to increase their vocal effort. To examine auditory-perceptual measures, a voice-specializing speech-language pathologist, who was blinded to the experimental paradigm, study hypotheses, and participant identity, completed auditory-perceptual ratings using the Consensus Auditory-perceptual Evaluation of Voice (Kempster et al., 2009). All stimuli were peak-amplitude normalized and presented using a custom-written MATLAB (MathWorks, 2016) script. Three of the five repetitions of each vocal effort condition were selected randomly for each speaker and used as auditory stimuli for auditory-perceptual ratings. In addition, 14% of stimuli were repeated for intrajudge reliability calculations, and a Pearson's correlation coefficient was used to determine relations between ratings.

Acoustic measures of voice quality and speech intensity were determined from the acoustic microphone signal. To examine acoustic measures reflecting voice quality during running speech (Awan et al., 2010; Shrivastav & Sapienza, 2003), the smoothed cepstral peak prominence (CPP) and the ratio of high-to-low spectral energy (H/L ratio) were calculated using Praat for all stimuli after silences/pauses had been removed using a custom MATLAB script. The same acoustic stimuli were used to calculate speech intensity across speaking conditions. First, Praat was used to determine the mean amplitude for each reading passage using the "Get Intensity" feature. To determine the intensity of the signal in dB SPL, a ratio was generated from the calibration process: the sound-level meter value from the calibration (the dB SPL value) divided by the mean amplitude of the electrolarynx recording in Praat. The resulting ratio was multiplied by the mean amplitude for each passage to yield dB SPL values for each trial.

No statistical tests were conducted on the auditory-perceptual and voice quality acoustic measures because they were not associated with the study hypotheses. Statistical tests were conducted on the speech intensity measure because it was associated with the study hypotheses.

Statistical Analysis

Because of the small sample size in this study, nonparametric statistical analyses were conducted. Nonparametric Friedman tests of differences among repeated measures were conducted for each outcome measure, with vocal effort modulations as the condition (baseline, mild vocal effort, and maximal vocal effort). The six outcome measures were LVSI and speech termination (LVSI and LVST), AKVS for the TB and TF sensors, LA, and speech intensity. Factor effect sizes were quantified as Kendall's *W* values derived from the chi-square statistic. Kendall's *W* values of less than 0.3 were considered small effect sizes, values between 0.3 and 0.5 were considered medium effect sizes, and values greater 0.5 were considered large effect sizes (Kendall & Smith, 1939). A *p* value of .05 or less was considered to be statistically significant. For Friedman tests that demonstrated statistically significant effect of condition, post hoc Wilcoxon signed-rank tests were conducted to assess differences between each of the three conditions.

Results

Respiratory Measures

The Friedman test revealed a statistically significant effect of condition for LVSI ($df = 2$, $\chi^2 = 16.20$, $p < .001$), with a large effect size (Kendall's $W = 0.81$). The Wilcoxon singed-rank test post hoc analyses revealed statistically significant changes in LVSI between the baseline and mild vocal effort conditions, $W(10) = 53$, $p = .011$, the mild vocal effort and maximum vocal effort conditions, $W(10) = 54$, $p = .006$, and the baseline and maximum vocal effort conditions, $W(10) = 55$, $p = .006$. In contrast, condition was not a statistically significant factor for LVST ($df = 2$, $\chi^2 = 3.80$, $p = .150$). However, at a group level, speakers demonstrated increases in LVSI and a trend for decreases in LVST when vocal effort increased (see panel B; Figure 4). The group averages showed that the LVSI was 16.3% (*SD* = 8.1%) at baseline, 24.8% (*SD* = 13.7%) during mild vocal effort, and 33.7% (*SD* = 17.0%) during maximum vocal effort. For LVST, the average lung volumes were −15.5% (*SD* = 7.4%) at baseline, −16.2% (*SD* = 8.8%) during mild vocal effort, and −18.2% (*SD* = 8.4%) during maximum vocal effort.

Articulatory Measures

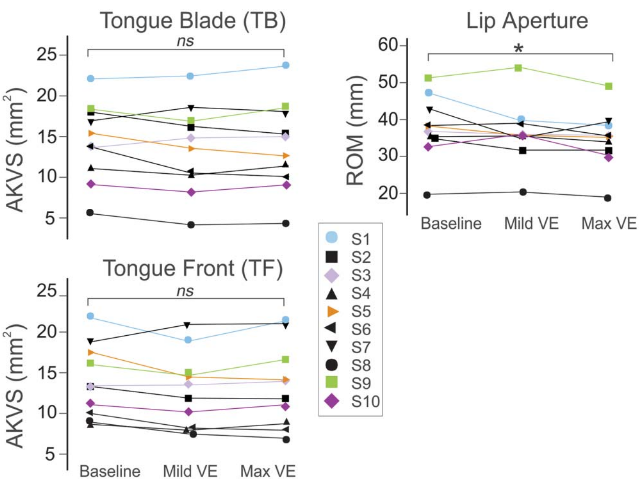
The Friedman test revealed no statistically significant effect of condition for the AKVS for the TB sensor ($df = 2$, $\chi^2 = 1.40$, $p = .497$) or the TF sensor ($df = 2$, $\chi^2 = 2.60$, $p = .273$). For the TB sensor, the average AKVS was 14.4 mm^2 (*SD* = 5.0 mm^2) at baseline, 13.6 mm^2 (*SD* = 5.3 mm^2) during mild vocal effort, and 13.8 mm^2 (5.4 mm^2) during maximum vocal effort. The average AKVS for the TF sensor was 13.9 mm^2 (*SD* = 4.4 mm^2) at baseline, 12.8 mm^2 (*SD* = 4.5 mm^2) during mild vocal effort, and 13.3 mm^2 (*SD* = 5.0 mm^2) during maximum vocal effort (see Figure 5).

For LA, the Friedman test showed a statistically significant effect of condition ($df = 2$, $\chi^2 = 12.20$, $p = .002$), with a large effect size (Kendall's $W = 0.61$). The Wilcoxon signed-rank tests revealed a statistically significant difference between the baseline and maximum vocal effort condition, $W(10) = 0$, $p = .006$. The Wilcoxon statistic was equal to zero, which indicated that all speakers demonstrated smaller LA in the maximum vocal effort condition compared with the baseline condition. No statistical differences in LA were found between the measures taken during baseline and mild vocal effort, $W(10) = 19$, $p = .415$, or mild vocal effort and maximum vocal effort conditions, $W(10) = 8$, $p = .053$. The average LA was 38.0 mm (*SD* = 8.4 mm) at baseline, 36.7 mm (*SD* = 8.2 mm) during mild vocal effort, and 35.2 mm (*SD* = 7.5 mm) during maximum vocal effort (see Figure 5).

Descriptive Analysis: Auditory-Perceptual Ratings and Acoustic Measures of Voice

The auditory-perceptual ratings by the speech-language pathologist validated that voice quality did change with

Figure 5. The articulatory kinematic measures are shown for female (light-colored shapes) and male (black shapes) speakers as a function of speaking condition: baseline, mild vocal effort (VE), and maximum VE. The articulatory kinematic vowel space (AKVS) of the tongue body (TB) sensor is shown in the upper panel, the AKVS of the tongue front (TF) sensor is shown in the middle panel, and lip aperture is shown on the bottom panel. *Indicates statistically significant differences between speaking conditions, and *ns* denotes no statistically significant differences.



increases in vocal effort at the individual and group level (see Supplemental Material S1). For the four percepts used in the study (overall severity of dysphonia, roughness, breathiness, and strain), the intrajudge reliability determined via a Pearson's correlation coefficient showed a strong association between repeated ratings (ranging from $r = .94$ to $r = .95$ across percepts). On average, ratings of overall severity of dysphonia, roughness, breathiness, and strain all increased with increases in vocal effort. When comparing the baseline to the maximal vocal effort condition, average ratings increased from 6.8 (*SD* = 6.6) to 28. 6 (*SD* =19.8) for overall severity, 3.7 (*SD* = 4.0) to 16.3 (*SD* = 12.9) for roughness, 0.5 (*SD* = 1.7) to 9.8 (*SD* = 12.8) for breathiness, and 4.3 (*SD* = 7.1) to 25.8 (*SD* = 19.6) for strain. Auditory-perceptual ratings were plotted against experimental variables of respiratory and articulatory kinematics, but no trends were seen.

At a group level, the acoustic measures showed small changes with increases in vocal effort. Between the baseline and maximal vocal effort conditions, the average smoothed CPP decreased from 8.5 dB (*SD* = 1.1 dB) to 8.1 dB (*SD* = 1.2 dB) and the ratio of high-to-low spectral energy (H/L ratio) increased from −24.0 (*SD* = 2.8) to −22.6 (*SD* = 3.3). Acoustic measures were plotted against experimental variables of respiratory and articulatory kinematics, but no trends were observed.

The Friedman test did not show a statistically significant effect of condition on speech intensity ($df = 2$, $\chi^2 = 2.60$, $p = .27$). The average speech intensity was 71.7 dB SPL (*SD* = 3.22 dB) in the baseline condition, 71.4 dB SPL (*SD* = 2.40 dB) in the mild vocal effort condition, and 71.7 dB SPL (*SD* = 2.45 dB) in the maximum vocal effort condition.

Discussion

This study investigated how increasing vocal effort while maintaining constant speech intensity affected respiratory and articulatory kinematics in speakers with typical voices. On the basis of earlier work, we hypothesized that measures of LVSI and LVST would both increase with increasing vocal effort. We also hypothesized that AKVS would reduce with increased vocal effort, which would be reflected as decreases in measures of AKVS and LA. A statistically significant effect of condition was found for LVSI and LA, but not for LVST or measures of AKVS.

Respiratory Kinematic Findings

This work investigated how respiratory kinematics were impacted when speech intensity was maintained at typical levels and shows agreement with our hypothesis. Results indicated that there was a statistically significant effect of speaking condition on LVSI and that LVSI increased as vocal effort increased across all conditions. These findings agree with prior observations (McKenna et al., 2017), of larger total LVEs, characterized by a trend for increased LVSI and LVST, during increased vocal effort. This suggests that, regardless of whether speech intensity is maintained relatively constant or slightly increased, LVSI increases when speakers with typical voices increase their vocal effort. In another study, Sundarrajan et al. (2017) found lower calibrated values of LVSI and greater LVST when speakers with typical voices reported increased vocal effort and decreased speech intensity after a vocal loading task. Thus, decreased intensity may differentially affect respiratory kinematics (compared with steady or increasing intensity) during increased vocal effort in speakers with typical voices (i.e., a lower LVSI rather than a higher LVSI).

The impact of speech intensity on respiratory patterns during increased vocal effort may also differ between speakers with typical voices and speakers with voice disorders. Lower LVST has been reported in populations who report increased effort when speaking, both with and without laryngeal pathology, compared with speakers with no voice complaints (Hixon & Putnam, 1983; Iwarsson & Sundberg, 1999; Lowell et al., 2004, 2008). However, the impact of vocal intensity on LVSI may interact with the presence of laryngeal pathology because a lesion on the vocal folds would result in greater air leakage during speech. In speakers with vocal nodules, lower LVST has been observed at both typical and loud speech intensities (Iwarsson & Sundberg, 1999; Sapienza et al., 1997), whereas lower LVSI was found during loud speech (Iwarsson & Sundberg, 1999) and greater LVSI was found during typical speech (Sapienza et al., 1997). Instead, high-voice users who report increased speaking effort, without known concurrent laryngeal pathology, show lower LVST and lower LVSI for both typical and loud speech (Lowell et al., 2008). The latter would be the closer comparison to typical speakers with increased vocal effort (in the absence of laryngeal pathology), yet the current findings show increased LVSI with increased vocal effort when speech intensity is held constant. In speakers with typical voices, McKenna et al. (2017) report increases in speech intensity (5 dB, on average) with increased vocal effort, whereas Sundarrajan et al. (2017) found that speakers decreased their speech intensity after a vocal loading task that increased vocal effort. This difference in findings may stem from the type of task: A vocal loading task is likely to be more representative of chronic vocal effort as seen in voice disorders (rather than a temporary change via instruction), which could differentially impact speech intensity. Hence, the interaction between respiratory kinematics and speech intensity also needs to be examined in speakers with and without chronic vocal effort to clarify this discrepancy.

The type of speech task is also likely to have impacted the respiratory measures reported here. This study employed a reading passage with long sentences that was designed to induce larger ranges in LVSI and LVST. Specifically, longer sentences may have induced speakers to use more breath to finish a sentence prior to their next breath (lowering LVST). For example, the average baseline condition LVST (−15.5% of maximum displacement relative to REL) was lower than typical speech LVST reported in prior work using calibrated respiratory measures in liters (1.68% of vital capacity relative to REL; Lowell et al., 2008). The normalized LVST value represents the lung volume termination as a percent of the maximal breathing range and relative to REL. Thus, a negative value indicates that speakers are exhaling beyond their typical expiratory level (as observed in this work), and a value close to zero indicates that speakers are exhaling close to their typical expiratory level during tidal breathing (REL). For this reason, the low values reported for the baseline condition LVST may result from speech task itself.

Articulatory Kinematic Findings

To our knowledge, this is the first investigation of articulatory kinematics in speakers with typical voices modulating their vocal effort. On the basis of articulatory–acoustic evidence from prior work in speaker populations with elevated vocal effort (Dromey et al., 2008; Roy & Ferguson, 2001), we hypothesized that measures of articulatory excursion would demonstrate reductions as vocal effort increased. The current results demonstrated a statistically significant difference in LA consistent with this hypothesis: LA reduced for all speakers in the maximum vocal effort condition compared with baseline. However, the AKVS for sensors placed on the TB and blade (TF) did not show statistically significant differences between speaking conditions.

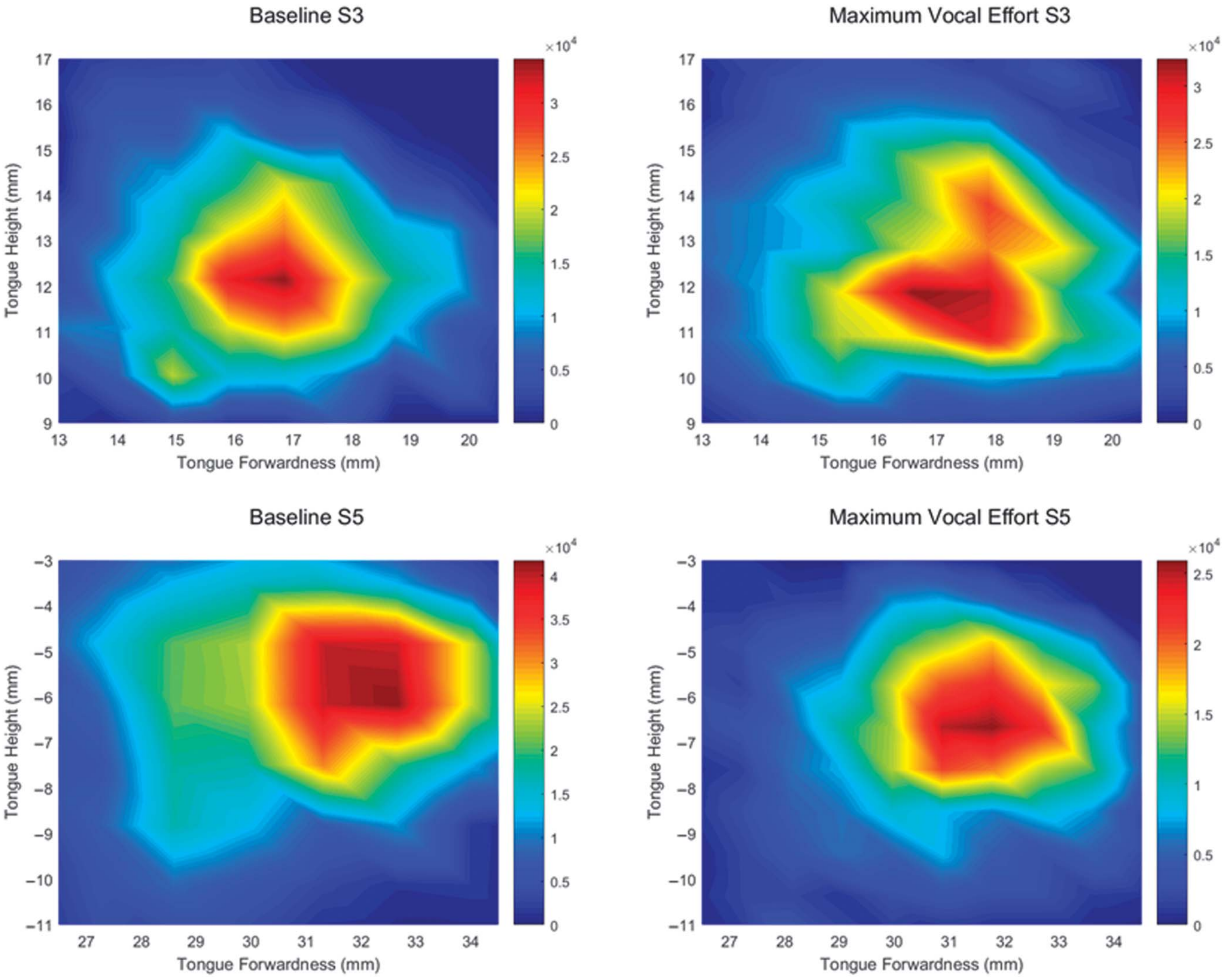
This is also the first work to examine LA with increasing vocal effort in speakers with typical voices; nevertheless, prior work has demonstrated associations between coordination of upper and lower lip kinematics and laryngeal spasms in speakers with spasmodic dysphonia (Dromey et al., 2007), who report increased vocal effort (Shoffel-Havakuk et al., 2019). Articulatory–acoustic evidence of decreased vowel space area (characterized by decreases in the first vowel formant, F_1) in speakers with HVDs also supports that lip kinematics may be implicated in voice disorders with vocal effort (Roy et al., 2009). Variations in formant frequencies of vowels can result from a number of sources along the articulatory tract (e.g., laryngeal height, tongue position, and jaw movement), which include LA (Fant, 1970). Specifically, greater LA is thought to widen the oral cavity, which would result in an F_1 increase (Harrington et al., 2000). Thus, reduced LA observed in this study could be related to a reduction in F_1 that would align with the findings of Roy et al. (2009).

Although the observed reductions in LA with increased vocal effort in this study imply interactions between vocal effort and articulatory kinematics, the articulatory kinematic changes by condition could have resulted from various mechanisms. Typical speakers show respiratory and laryngeal modifications with increases in vocal effort that also include increases in speech intensity (McKenna et al., 2017); thus, instructing speakers to maintain level intensity during the current task could have resulted in compensatory articulatory movements (reducing LA would reduce output speech intensity). Speakers with voice disorders who report elevated vocal effort do not show increases in speech intensity, presumably because the laryngeal system is functioning inefficiently during speech, so these speakers may not show the same type of changes to lip kinematics. This could be specifically explored in future studies.

Overall, the AKVS measures demonstrated minimal changes across speaking conditions, but individual speakers showed trends for increases or decreases with greater vocal effort (example trajectories in Figure 6). Prior work has found tongue displacements increase with increased speech intensity (Mefferd & Green, 2010; Whitfield et al., 2018); however, there was no statistical effect of speaking condition on speech intensity, which suggests that these trends in AKVS were not driven by speech intensity. For example, the speaker trajectory showing increased AKVS with vocal effort (S3; see Figure 6) had an average speech intensity of 70.54 dB SPL at baseline and 70.41 dB SPL during maximum vocal effort. AKVS is also affected by the phonetic context of the sentences produced, so the AKVS measures here are not directly comparable to those reported by Whitfield et al. (2018). Yet, the average AKVS values reported here (12.75–14.38 mm^2) are within the range reported by Whitfield et al. (2018) across two phonemically different sentences (10.90–19.40 mm^2), which provides support for the generalizability of this measure across different tasks. Individual variability in articulatory strategies in speakers with typical voices has been reported in previous work examining articulatory kinematics between normal and clear speech (Perkell et al., 2002).

The differences seen in articulatory strategies here may result from speaker-specific recruitment of physiological mechanisms to achieve their targets for experiential anchors of vocal effort. For example, some speakers may have shown reductions in AKVS by focusing on increasing laryngeal muscle tension and adductory force as seen in HVDs (e.g., resulting in greater tension in the tongue base, which would restrict tongue movement). Other speakers in the study may have primarily focused on experiential anchors related to increased abduction of the vocal folds (resulting in a breathy voice quality) without increased intrinsic muscle tension. In addition to the possibility of

Figure 6. Data from two example trajectories (time-varying displacement) for the tongue front (TF) sensors are depicted as density plots in the anteroposterior (tongue forwardness) and vertical dimensions (tongue height). When increasing vocal effort from baseline (left) to the maximum vocal effort condition (right), speaker S5 demonstrated decreases in articulatory-working space (upper panel) and speaker S7 demonstrated increases in articulatory-working space (lower panel). These changes are reflected in the articulatory–acoustic vowel space (AKVS) measurements for these speakers in Figure 5.



using different mechanisms within the laryngeal system, speakers may have also employed different mechanisms across speech subsystems (placing more emphasis on a laryngeal compared with a respiratory strategy to increase vocal effort or vice versa). The variability in speaker strategies may also result from differences in laryngeal height during increased vocal effort: Increasing laryngeal height with increased vocal effort (as seen in HVDs) would shorten the vocal tract and allow tongue movements to be less restricted. However, a recent study reported no changes to articulatory–acoustic estimates of vocal tract length during increased vocal effort in speakers with typical voices (Groll et al., 2020). Taken together with earlier work, the current findings suggest that speakers may have variable patterns in AKVS due to a combination of physiological modifications and this should be explored further in future work.

Prior work has also demonstrated a positive relationship between increases in vocal effort and speech intensity in speakers with typical voices (McKenna et al., 2017; Rosenthal et al., 2014). Thus, some speakers may have utilized experiential anchors of vocal effort that were associated with increased speech intensity, even if their speech intensity was not increasing. On the basis of prior work, it is likely that the motor plans related to greater speech intensity would entail increases in AKVS (Whitfield et al., 2018). However, the AKVS of the tongue sensors demonstrated both increases and decreases across speakers with increasing vocal effort, so this pattern was not observed at a group level in this work.

Auditory-Perceptual and Acoustic Findings

Auditory-perceptual ratings of voice and acoustic measures of voice quality were used in this study to confirm that there were changes in voice across vocal effort conditions. The auditory-perceptual data revealed that speakers exhibited changes across all percepts (see Supplemental Material S1), but the acoustic measures demonstrated minimal changes across vocal effort conditions. One explanation for the lack of changes seen in acoustic measures may be that speakers used a combination of increasing strain and breathiness as vocal effort increased. Measures of CPP and smoothed CPPs have shown both increases and decreases with increased vocal strain (Anand et al., 2019; Lowell, Kelley, Awan, et al., 2012) and decreases with increased breathy voice quality (Hillenbrand & Houde, 1996). A similar situation may have occurred with the ratio of high-to-low spectral energy (H/L ratio): Breathy voice quality would result in less harmonic energy at higher frequencies, but a more strained or pressed voice quality would result in more harmonic energy at higher frequencies. Hence, if speakers employed a combination of these strategies, including both strain and breathiness, the CPPs and H/L ratio measures, averaged across speakers, may not have been able to reflect the nature of voice changes with increased vocal effort.

Finally, there was no statistical effect of speaking condition on speech intensity, which supports that the speakers in this work were able to follow the instruction to maintain constant speech intensity when increasing their vocal effort. This is important to clarify because prior work has demonstrated statistical effects of minimal changes in speech intensity with vocal effort when speakers were not given specific instructions about intensity (increases of 0.97 dB with mild vocal effort and 2.76 dB with maximum vocal effort; McKenna & Stepp, 2018).

Limitations

In addition to the limitations discussed, no measures of laryngeal subsystem physiology were examined in this study. This limits the interpretation of speech subsystem involvements during increased vocal effort because the functional contributions of the laryngeal system cannot be clarified. Ideally, speech kinematics would be examined in conjunction with measures of laryngeal physiology in order to best describe the impact of vocal effort on speech production. However, it was not feasible to simultaneously examine measures of laryngeal physiology (e.g., videoendoscopy) in conjunction with EMA and a multisentence running speech task used in this study. Prior work has examined measures related to the laryngeal system during increases in vocal effort in speakers with typical voices (Lien et al., 2015; McKenna et al., 2017), so this information can serve as a basis to speculate about the involvement of similar mechanisms in this work. To resolve such issues, future studies should examine physiological, acoustical, aerodynamic, and respiratory kinematic measures, in addition to articulatory kinematic measures, when increasing vocal effort.

Another limitation of this work is that the respiratory measures were not calibrated with direct lung volume measurements in liters and were not compared between speakers. With calibrated measurements, lung volume can be expressed in liters with respect to an individual's vital capacity (a speaker's full range of inhalation and exhalation in liters). With this information, the relative contributions of the rib cage and abdomen to overall breathing excursions can be determined in liters (see McKenna & Huber, 2019 for review). Because only within-speaker changes were being examined, the respiratory signal in this work was expressed as a function of each speaker's MD: the difference between the largest possible inhalation and exhalation. MD has previously been used to describe within-speaker respiratory patterns (Rochet-Capellan & Fuchs, 2013; Werner et al., 2021). However, this measure is only an approximation of lung volume changes and cannot be used to directly compare this study to prior work using lung volume measures calibrated in liters.

These study findings also have a limited degree of generalizability to other speaker populations. As discussed earlier, investigations into vocal effort in speakers with typical voices and speakers with HVDs do not always show agreement (Dromey & Ramig, 1998; Iwarsson & Sundberg, 1999). Chronically increased vocal effort (as seen in HVDs) and experimentally increased vocal effort (elicited with instruction during a study) may have different underlying mechanisms; therefore, this work cannot directly be extended to speakers with voice disorders. Furthermore, vocal effort is also impacted environmental factors, such as background noise and proximity to the listener (Bermúdez de Alvear et al., 2011; Hunter et al., 2020; Pelegrín-García et al., 2011), which were not examined here. Lastly, this work used a small sample size. Additional studies would be needed to extend these preliminary findings to a larger sample of speakers both with and without voice disorders.

Conclusions

Respiratory and articulatory kinematics were examined during increases in vocal effort, without changes in speech intensity, in speakers with typical voices. On average, speakers exhibited larger LVSI and reduced LA during increases in vocal effort. To our knowledge, this is the first study to demonstrate evidence that articulatory kinematics are affected by laryngeal-level changes during modulations of vocal effort in speakers with typical voices. This work provides preliminary support that modulations in vocal effort interact with the articulatory subsystem. This hypothesis should be further examined in future studies that include both speakers with typical voices and speakers who report chronic elevation of vocal effort in order to elucidate the mechanisms underlying vocal effort across speech subsystems.

Acknowledgments

This work was supported by Grants R01 DC015570 (awarded to Cara E. Stepp), T32 DC013017 (awarded to Cara E. Stepp and Christopher A. Moore), and F31 DC019032 (awarded to Defne Abur) from the National Institute of Deafness and Other Communication Disorders. It was also supported by a Graduate Fellow Award from the Rafik B. Hariri Institute for Computing and Computational Science and Engineering (awarded to Defne Abur). The authors would like to thank Monique C. Tardif and Nicole M. Enos for help with data collection; Mark Tiede for kindly sharing scripts for filtering and applying movement corrections to the articulatory kinematic data; Gabriel J. Cler and Elizabeth S. Heller Murray for help with instrumentation and data preprocessing; and Katherine Brown, Felicia Francois, and Katelyn Carone for help with participant recruitment.

References

Altman, K. W., Atkinson, C., & Lazarus, C. (2005). Current and emerging concepts in muscle tension dysphonia: A 30-month review. *Journal of Voice, 19*(2), 261–267. https://doi.org/10.1016/j.jvoice.2004.03.007

American Speech-Language-Hearing Association. (1997). *Guidelines for audiologic screening*.

Anand, S., Kopf, L. M., Shrivastav, R., & Eddins, D. A. (2019). Objective indices of perceived vocal strain. *Journal of Voice, 33*(6), 838–845. https://doi.org/10.1016/j.jvoice.2018.06.005

Aronson, A. E. (1990). *Clinical voice disorders: An interdisciplinary approach*. Thieme.

Awan, S. N. (2011). The effect of smoking on the dysphonia severity index in females. *Folia Phoniatrica et Logopaedica, 63*(2), 65–71. https://doi.org/10.1159/000316142

Awan, S. N., Roy, N., Jetté, M. E., Meltzner, G. S., & Hillman, R. E. (2010). Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics, 24*(9), 742–758. https://doi.org/10.3109/02699206.2010.492446

Bach, K. K., Belafsky, P. C., Wasylik, K., Postma, G. N., & Koufman, J. A. (2005). Validity and reliability of the glottal function index. *Archives of Otolaryngology–Head & Neck Surgery, 131*(11), 961–964. https://doi.org/10.1001/archotol.131.11.961

Banister, E. W. (1979). The perception of effort: An inductive approach. *European Journal of Applied Physiology and Occupational Physiology, 41*(2), 141–150. https://doi.org/10.1007/BF00421661

Banzett, R. B., Mahan, S. T., Garner, D. M., Brughera, A., & Loring, S. H. (1995). A simple and reliable method to calibrate respiratory magnetometers and Respitrace. *Journal of Applied Physiology, 79*(6), 2169–2176. https://doi.org/10.1152/jappl.1995.79.6.2169

Bermúdez de Alvear, R. M., Baron, F. J., & Martinez-Arquero, A. G. (2011). School teachers' vocal use, risk factors, and voice disorder prevalence: Guidelines to detect teachers with current voice problems. *Folia Phoniatrica et Logopaedica, 63*(4), 209–215. https://doi.org/10.1159/000316310

Berry, D. A., Verdolini, K., Montequin, D. W., Hess, M. M., Chan, R. W., & Titze, I. R. (2001). A quantitative output-cost ratio in voice production. *Journal of Speech, Language, and Hearing Research, 44*(1), 29–37. https://doi.org/10.1044/1092-4388(2001/003)

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise, 14*(5), 377–381. https://doi.org/10.1249/00005768-198205000-00012

Cler, G. J., Lee, J. C., Mittelman, T., Stepp, C. E., & Bohland, J. W. (2017). Kinematic analysis of speech sound sequencing errors induced by delayed auditory feedback. *Journal of Speech, Language, and Hearing Research, 60*(6S), 1695–1711. https://doi.org/10.1044/2017_JSLHR-S-16-0234

Dastolfo, C., Gartner-Schmidt, J., Yu, L., Carnes, O., & Gillespie, A. I. (2016). Aerodynamic outcomes of four common voice disorders: Moving toward disorder-specific assessment. *Journal of Voice, 30*(3), 301–307. https://doi.org/10.1016/j.jvoice.2015.03.017

Dromey, C., Hunter, E., & Nissen, S. L. (2018). Speech adaptation to kinematic recording sensors: Perceptual and acoustic findings. *Journal of Speech, Language, and Hearing Research, 61*(3), 593–603. https://doi.org/10.1044/2017_JSLHR-S-17-0169

Dromey, C., Nissen, S. L., Roy, N., & Merrill, R. M. (2008). Articulatory changes following treatment of muscle tension dysphonia: Preliminary acoustic evidence. *Journal of Speech, Language, and Hearing Research, 51*(1), 196–208. https://doi.org/10.1044/1092-4388(2008/015)

Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate. *Journal of Speech, Language, and Hearing Research, 41*(5), 1003–1018. https://doi.org/10.1044/jslhr.4105.1003

Dromey, C., Reese, A., & Howey, S. (2007). Lip kinematics in spasmodic dysphonia before and after treatment with botulinum toxin. *Journal of Medical Speech-Language Pathology, 15,* 263–277.

Espinoza, V. M., Zanartu, M., Van Stan, J. H., Mehta, D. D., & Hillman, R. E. (2017). Glottal aerodynamic measures in women with phonotraumatic and nonphonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research, 60*(8), 2159–2169. https://doi.org/10.1044/2017_JSLHR-S-16-0337

Fant, G. (1970). *Acoustic theory of speech production*. Walter de Gruyter.

Fink, B. R. (1978). *Laryngeal biomechanics*. Harvard University Press.

Finnegan, E. M., Luschei, E. S., & Hoffman, H. T. (2000). Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech. *Journal of Speech,* *Language, and Hearing Research, 43*(4), 934–950. https://doi.org/10.1044/jslhr.4304.934

Friedman, A. D., Hillman, R. E., Landau-Zemer, T., Burns, J. A., & Zeitels, S. M. (2013). Voice outcomes for photoangiolytic KTP laser treatment of early glottic cancer. *Annals of Otology, Rhinology & Laryngology, 122*(3), 151–158. https://doi.org/10.1177/000348941312200302

Fryd, A. S., Van Stan, J. H., Hillman, R. E., & Mehta, D. D. (2016). Estimating subglottal pressure from neck-surface acceleration during normal voice production. *Journal of Speech, Language, and Hearing Research, 59*(6), 1335–1345. https://doi.org/10.1044/2016_JSLHR-S-15-0430

Grillo, E. U., & Verdolini, K. (2008). Evidence for distinguishing pressed, normal, resonant, and breathy voice qualities by laryngeal resistance and vocal efficiency in vocally trained subjects. *Journal of Voice, 22*(5), 546–552. https://doi.org/10.1016/j.jvoice.2006.12.008

Groll, M. D., McKenna, V. S., Hablani, S., & Stepp, C. E. (2020). Formant-estimated vocal tract length and extrinsic laryngeal muscle activation during modulation of vocal effort in healthy speakers. *Journal of Speech, Language, and Hearing Research, 63*(5), 1395–1403. https://doi.org/10.1044/2020_JSLHR-19-00234

Harrington, J., Fletcher, J., & Beckman, M. E. (2000). Manner and place conflicts in the articulation of accent in Australian English. In M. B. Broe (Ed.), *Papers in laboratory phonology V* (pp. 40–55). Cambridge University Press.

Heller Murray, E. S., Michener, C. M., Enflo, L., Cler, G. J., & Stepp, C. E. (2018). The impact of glottal configuration on speech breathing. *Journal of Voice, 32*(4), 420–427. https://doi.org/10.1016/j.jvoice.2017.07.001

Higgins, M. B., Carney, A. E., & Schulte, L. (1994). Physiological assessment of speech and voice production of adults with hearing loss. *Journal of Speech and Hearing Research, 37*(3), 510–521. https://doi.org/10.1044/jshr.3703.510

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research, 39*(2), 311–321. https://doi.org/10.1044/jshr.3902.311

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

Hixon, T., & Putnam, A. (1983). *Voice disorders in relation to respiratory kinematics*. Paper presented at the Seminars in Speech and Language.

Holmberg, E. B., Doyle, P., Perkell, J. S., Hammarberg, B., & Hillman, R. E. (2003). Aerodynamic and acoustic voice measurements of patients with vocal nodules: Variation in baseline and changes across voice therapy. *Journal of Voice, 17*(3), 269–282. https://doi.org/10.1067/S0892-1997(03)00076-6

Huber, J. E., & Chandrasekaran, B. (2006). Effects of increasing sound pressure level on lip and jaw movement parameters and consistency in young adults. *Journal of Speech, Language, and Hearing Research, 49*(6), 1368–1379. https://doi.org/10.1044/1092-4388(2006/098)

Huber, J. E., Chandrasekaran, B., & Wolstencroft, J. J. (2005). Changes to respiratory mechanisms during speech as a result of different cues to increase loudness. *Journal of Applied Physiology, 98*(6), 2177–2184. https://doi.org/10.1152/japplphysiol.01239.2004

Hunter, E. J., Cantor-Cutiva, L. C., van Leer, E., van Mersbergen, M., Nanjundeswaran, C. D., Bottalico, P., Sandage, M. J., & Whitling, S. (2020). Toward a consensus description of vocal effort, vocal load, vocal loading, and vocal fatigue. *Journal of Speech, Language, and Hearing Research, 63*(2), 509–532. https://doi.org/10.1044/2019_JSLHR-19-00057

Iwarsson, J., & Sundberg, J. (1999). Breathing behaviors during speech in healthy females and patients with vocal fold nodules. *Logopedics Phoniatrics Vocology, 24*(4), 154–169. https://doi.org/10.1080/140154399435002

Jiang, J. J., & Titze, I. R. (1994). Measurement of vocal fold intraglottal pressure and impact stress. *Journal of Voice, 8*(2), 132–144. https://doi.org/10.1016/S0892-1997(05)80305-4

Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology, 18*(2), 124–132. https://doi.org/10.1044/1058-0360(2008/08-0017)

Kendall, M. G., & Smith, B. B. (1939). The problem of *m* rankings. *The Annals of Mathematical Statistics, 10*(3), 275–287. https://doi.org/10.1214/aoms/1177732186

Konno, K., & Mead, J. (1967). Measurement of the separate volume changes of rib cage and abdomen during breathing. *Journal of Applied Physiology, 22*(3), 407–422. https://doi.org/10.1152/jappl.1967.22.3.407

Lien, Y. A., Michener, C. M., Eadie, T. L., & Stepp, C. E. (2015). Individual monitoring of vocal effort with relative fundamental frequency: Relationships with aerodynamics and listener perception. *Journal of Speech, Language, and Hearing Research, 58*(3), 566–575. https://doi.org/10.1044/2015_JSLHR-S-14-0194

Lowell, S. Y., Barkmeier-Kraemer, J. M., Hoit, J. D., & Story, B. H. (2008). Respiratory and laryngeal function during spontaneous speaking in teachers with voice disorders. *Journal of Speech, Language, and Hearing Research, 51*(2), 333–349. https://doi.org/10.1044/1092-4388(2008/025)

Lowell, S. Y., Hoit, J. D., & Barkmeier, J. M. (2004). *Speech breathing behavior in muscle tension dysphonia: Preliminary findings*. Paper presented at the American Speech-Language-Hearing Association Convention.

Lowell, S. Y., Kelley, R. T., Awan, S. N., Colton, R. H., & Chan, N. H. (2012). Spectral- and cepstral-based acoustic features of dysphonic, strained voice quality. *Annals of Otology, Rhinology & Laryngology, 121*(8), 539–548. https://doi.org/10.1177/000348941212100808

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. (2009). Treatment for spasmodic dysphonia: Limitations of current approaches. *Current Opinion in Otolaryngology, Head and Neck Surgery, 17*(3), 160–165. https://doi.org/10.1097/MOO.0b013e32832aef6f

Ludlow, C. L. (2011). Spasmodic dysphonia: A laryngeal control disorder specific to speech. *Journal of Neuroscience, 31*(3), 793–797. https://doi.org/10.1523/JNEUROSCI.2758-10.2011

Macklin, C. C. (1925). X-ray studies on bronchial movements. *American Journal of Anatomy, 35*(2), 303–329. https://doi.org/10.1002/aja.1000350208

MathWorks. (2016). *MATLAB 2016 (Version b)*. https://www.mathworks.com/help/matlab/release-notes-R2016b.html

McCall, G. N., Colton, R. H., & Rabuzzi, D. D. (1973). Preliminary EMG investigation of certain intrinsic and extrinsic laryngeal muscles in patients with spasmodic dysphonia. *Journal of the Acoustical Society of America, 53*(1), 345. https://doi.org/10.1121/1.1982437

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., & Huber, J. E. (2019). The accuracy of respiratory calibration methods for estimating lung volume during speech breathing: A comparison of four methods across three adult cohorts. *Journal of Speech, Language, and Hearing Research, 62*(8), 2632–2644. https://doi.org/10.1044/2019_JSLHR-S-18-0478

McKenna, V. S., Llico, A. F., Mehta, D. D., Perkell, J. S., & Stepp, C. E. (2017). Magnitude of neck-surface Vibration as an estimate of subglottal pressure during modulations of vocal effort and intensity in healthy speakers. *Journal of Speech, Language, and Hearing Research, 60*(12), 3404–3416. https://doi.org/10.1044/2017_JSLHR-S-17-0180

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

Mefferd, A. (2015). Articulatory-to-acoustic relations in talkers with dysarthria: A first analysis. *Journal of Speech, Language, and Hearing Research, 58*(3), 576–589. https://doi.org/10.1044/2015_JSLHR-S-14-0188

Mefferd, A., & Dietrich, M. S. (2020). Tongue-and jaw-specific articulatory changes and their acoustic consequences in talkers with dysarthria due to amyotrophic lateral sclerosis: Effects of loud, clear, and slow speech. *Journal of Speech, Language, and Hearing Research, 63*(8), 2625–2636. https://doi.org/10.1044/2020_JSLHR-19-00309

Mefferd, A., & Green, J. R. (2010). Articulatory-to-acoustic relations in response to speaking rate and loudness manipulations. *Journal of Speech, Language, and Hearing Research, 53*(5), 1206–1219. https://doi.org/10.1044/1092-4388(2010/09-0083)

Mendes, A. P., Brown, W. S., Sapienza, C., & Rothman, H. B. (2006). Effects of vocal training on respiratory kinematics during singing tasks. *Folia Phoniatrica et Logopaedica, 58*(5), 363–377. https://doi.org/10.1159/000094570

Merrill, R. M., Roy, N., & Lowe, J. (2013). Voice-related symptoms and their effects on quality of life. *Annals of Otology, Rhinology & Laryngology, 122*(6), 404–411. https://doi.org/10.1177/000348941312200610

Morrison, M., Rammage, L. A., Belisle, G. M., Pullan, C. B., & Nichol, H. (1983). Muscular tension dysphonia. *The Journal of Otolaryngology, 12*(5), 302–306.

Pelegrín-García, D., Smits, B., Brunskog, J., & Jeong, C.-H. (2011). Vocal effort with changing talker-to-listener distance in different acoustic environments. *The Journal of the Acoustical Society of America, 129*(4), 1981–1990. https://doi.org/10.1121/1.3552881

Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., & Jackson, M. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *The Journal of the Acoustical Society of America, 92*(6), 3078–3096. https://doi.org/10.1121/1.404204

Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H. (2002). Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. I. A preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America, 112*(4), 1627–1641. https://doi.org/10.1121/1.1506369

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

Robb, M. P., & Chen, Y. (2008). Vowel space in Mandarin-accented English. *Asia Pacific Journal of Speech, Language and Hearing, 11*(3), 175–188. https://doi.org/10.1179/136132808805297179

Rochet-Capellan, A., & Fuchs, S. (2013). Changes in breathing while listening to read speech: The effect of reader and speech mode. *Frontiers in Psychology, 4,* 906. https://doi.org/10.3389/fpsyg.2013.00906

Rosenthal, A. L., Lowell, S. Y., & Colton, R. H. (2014). Aerodynamic and acoustic features of vocal effort. *Journal of Voice, 28*(2), 144–153. https://doi.org/10.1016/j.jvoice.2013.09.007

Rosner, B. S., & Pickering, J. B. (1994). *Vowel perception and production*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198521389.001.0001

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), 169–175.

Roy, N., Merrill, R. M., Gray, S. D., & Smith, E. M. (2005). Voice disorders in the general population: Prevalence, risk factors, and occupational impact. *The Laryngoscope, 115*(11), 1988–1995. https://doi.org/10.1097/01.mlg.0000179174.32345.41

Roy, N., Nissen, S. L., Dromey, C., & Sapir, S. (2009). Articulatory changes in muscle tension dysphonia: Evidence of vowel space expansion following manual circumlaryngeal therapy. *Journal of Communication Disorders, 42*(2), 124–135. https://doi.org/10.1016/j.jcomdis.2008.10.001

Sampaio, M. C., dos Reis, E. J. F. B., Carvalho, F. M., Porto, L. A., & Araújo, T. M. (2012). Vocal effort and voice handicap among teachers. *Journal of Voice, 26*(6), 820.e15–820.e18. https://doi.org/10.1016/j.jvoice.2012.06.003

Sapienza, C. M., & Stathopoulos, E. T. (1994). Respiratory and laryngeal measures of children and women with bilateral vocal fold nodules. *Journal of Speech and Hearing Research, 37*(6), 1229–1243. https://doi.org/10.1044/jshr.3706.1229

Sapienza, C. M., Stathopoulos, E. T., & Brown, W., Jr. (1997). Speech breathing during reading in women with vocal nodules. *Journal of Voice, 11*(2), 195–201. https://doi.org/10.1016/S0892-1997(97)80078-1

Schönle, P. W., Gräbe, K., Wenig, P., Höhne, J., Schrader, J., & Conrad, B. (1987). Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract. *Brain and Language, 31*(1), 26–35. https://doi.org/10.1016/0093-934X(87)90058-7

Shoffel-Havakuk, H., Marks, K. L., Morton, M., Johns, M. M., III, & Hapner, E. R. (2019). Validation of the OMNI vocal effort scale in the treatment of adductor spasmodic dysphonia. *The Laryngoscope, 129*(2), 448–453. https://doi.org/10.1002/lary.27430

Shrivastav, R., & Sapienza, C. M. (2003). Objective measures of breathy voice quality obtained using an auditory model. *The Journal of the Acoustical Society of America, 114*(4), 2217–2224. https://doi.org/10.1121/1.1605414

Smith, E., Gray, S. D., Dove, H., Kirchner, L., & Heras, H. (1997). Frequency and effects of teachers' voice problems. *Journal of Voice, 11*(1), 81–87. https://doi.org/10.1016/S0892-1997(97)80027-6

Smith, E., Taylor, M., Mendoza, M., Barkmeier, J., Lemke, J., & Hoffman, H. (1998). Spasmodic dysphonia and vocal fold paralysis: Outcomes of voice problems on work-related functioning. *Journal of Voice, 12*(2), 223–232. https://doi.org/10.1016/S0892-1997(98)80042-8

Södersten, M., Ternström, S., & Bohman, M. (2005). Loud speech in realistic environmental noise: Phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *Journal of Voice, 19*(1), 29–46. https://doi.org/10.1016/j.jvoice.2004.05.002

Stager, S. V., Bielamowicz, S. A., Regnell, J. R., Gupta, A., & Barkmeier, J. M. (2000). Supraglottic activity. *Journal of Speech, Language, and Hearing Research, 43*(1), 229–238. https://doi.org/10.1044/jslhr.4301.229

Stathopoulos, E. T., & Sapienza, C. M. (1997). Developmental changes in laryngeal and respiratory function with variations in sound pressure level. *Journal of Speech, Language, and Hearing Research, 40*(3), 595–614. https://doi.org/10.1044/jslhr.4003.595

Stepp, C. E., Hillman, R. E., & Heaton, J. T. (2010). A virtual trajectory model predicts differences in vocal fold kinematics in individuals with vocal hyperfunction. *The Journal of the Acoustical Society of America, 127*(5), 3166–3176. https://doi.org/10.1121/1.3365257

Sundarrajan, A., Huber, J. E., & Sivasankar, M. P. (2017). Respiratory and laryngeal changes with vocal loading in younger and older individuals. *Journal of Speech, Language, and Hearing Research, 60*(9), 2551–2556. https://doi.org/10.1044/2017_JSLHR-S-17-0106

Sundberg, J., Leanderson, R., & von Euler, C. (1989). Activity relationship between diaphragm and cricothyroid muscles. *Journal of Voice, 3*(3), 225–232. https://doi.org/10.1016/S0892-1997(89)80004-9

Švec, J. G., & Granqvist, S. (2018). Tutorial and guidelines on measurement of sound pressure level in voice and speech. *Journal of Speech, Language, and Hearing Research, 61*(3), 441–461. https://doi.org/10.1044/2017_JSLHR-S-17-0095

Tiede, M., Bundgaard-Nielsen, R., Kroos, C., Gibert, G., Attina, V., Kasisopa, B., Vatikiotis-Bateson, E., & Best, C. (2010). Speech articulator movements recorded from facing talkers using two electromagnetic articulometer systems simultaneously. *Proceedings of Meetings on Acoustics*, *11,* Article 060007. https://doi.org/10.1121/1.4772388

van Leer, E., & van Mersbergen, M. (2017). Using the Borg CR10 physical exertion scale to measure patient-perceived vocal effort pre and post treatment. *Journal of Voice, 31*(3), 389.e19–389.e25. https://doi.org/10.1016/j.jvoice.2016.09.023

Van Mersbergen, M., Patrick, C., & Glaze, L. (2008). Functional dysphonia during mental imagery: Testing the trait theory of voice disorders. *Journal of Speech, Language, and Hearing Research, 51*(6), 1405–1423. https://doi.org/10.1044/1092-4388(2008/06-0216)

Vatikiotis-Bateson, E., & Kelso, J. S. (1993). Rhythm type and articulatory dynamics in English, French and Japanese. *Journal of Phonetics, 21*(3), 231–265. https://doi.org/10.1016/S0095-4470(19)31338-5

Wang, J., Samal, A., Rong, P., & Green, J. R. (2016). An optimal set of flesh points on tongue and lips for speech-movement classification. *Journal of Speech, Language, and Hearing Research, 59*(1), 15–26. https://doi.org/10.1044/2015_JSLHR-S-14-0112

Werner, R., Fuchs, S., Trouvain, J., & Möbius, B. (2021). *Inhalations in speech: Acoustic and physiological characteristics*. Paper presented at the Interspeech, pp. 3186–3190. https://doi.org/10.21437/Interspeech.2021-1262

Westbury, J. R. (1994). On coordinate systems and the representation of articulatory movements. *The Journal of the Acoustical Society of America, 95*(4), 2271–2273. https://doi.org/10.1121/1.408638

Whitfield, J. A., Dromey, C., & Palmer, P. (2018). Examining acoustic and kinematic measures of articulatory working space: Effects of speech intensity. *Journal of Speech, Language, and Hearing Research, 61*(5), 1104–1117. https://doi.org/10.1044/2018_JSLHR-S-17-0388

Zenker, W., & Zenker, A. (1960). Über die Regelung der Stimmlippenspannung durch von aussen eingreifende Mechanismen. *Folia Phoniatrica et Logopaedica, 12*(1), 1–36. https://doi.org/10.1159/000262846

Zietels, S. M., Burns, J. A., Lopez-Guerra, G., Anderson, R., & Hillman, R. E. (2008). Photoangiolytic laser treatment of early glottic cancer: A new management strategy. *Annals of Otology, Rhinology & Laryngology, 117*(Suppl. 7), 3–24. https://doi.org/10.1177/00034894081170S701

Zuskin, E., Mustajbegovic, J., Schachter, E. N., Kern, J., Vitale, K., Pucarin-Cvetkovic, J., Chiarelli, A., Milosevic, M., & Jelinic, J. D. (2009). Respiratory function in wind instrument players. *La Medicina del Lavoro, 100*(2), 133–141.