JASA ARTICLE

Vocal fold kinematics and relative fundamental frequency as a function of obstruent type and speaker age

Yeonggwang Park,[1,a)] Feng Wang,[2] Manuel Díaz-Cádiz,[1] Jennifer M. Vojtech,[2,b)] Matti D. Groll,[2,b)] and Cara E. Stepp[1,c)]

[1]Department of Speech, Language, and Hearing Sciences, Boston University, 635 Commonwealth Avenue, Boston, Massachusetts 02215, USA

[2]Department of Biomedical Engineering, Boston University, 44 Cummington Mall, Boston, Massachusetts 02215, USA

ABSTRACT:
The acoustic measure, relative fundamental frequency (RFF), has been proposed as an objective metric for assessing vocal hyperfunction; however, its underlying physiological mechanisms have not yet been fully characterized. This study aimed to characterize the relationship between RFF and vocal fold kinematics. Simultaneous acoustic and high-speed videoendoscopic (HSV) recordings were collected as younger and older speakers repeated the utterances /ifi/ and /iti/. RFF values at voicing offsets and onsets surrounding the obstruents were estimated from acoustic recordings, whereas glottal angles, durations of voicing offset and onset, and a kinematic estimate of laryngeal stiffness (KS) were obtained from HSV images. No differences were found between younger and older speakers for any measure. RFF did not differ between the two obstruents at voicing offset; however, fricatives necessitated larger glottal angles and longer durations to devoice. RFF values were lower and glottal angles were greater for stops relative to fricatives at voicing onset. KS values were greater in stops relative to fricatives. The less adducted vocal folds with greater KS and lower RFF at voicing onset for stops relative to fricatives in this study were in accordance with prior speculations that decreased vocal fold contact area and increased laryngeal stiffness may decrease RFF.
 https://doi.org/10.1121/10.0003961

(Received 16 May 2020; revised 2 March 2021; accepted 11 March 2021; published online 1 April 2021)

[Editor: Zhaoyan Zhang] Pages: 2189–2199

I. INTRODUCTION

Relative fundamental frequency (RFF) is an acoustic measure that quantifies short-term changes in fundamental frequency (f_o) during the transition into and out of a voiceless consonant (e.g., in a vowel–voiceless consonant–vowel, or VCV, utterance). RFF has been operationally defined as comparing the short-term f_o values of the ten voicing cycles immediately before and after the voiceless consonant to a steady-state value (Stepp *et al.*, 2010a; Watson, 1998). In this way, RFF examines changes in f_o as a speaker terminates (vowel into voiceless consonant, "voicing offset") and re-initiates (voiceless consonant into vowel, "voicing onset") phonation. Because RFF captures changes in short-term f_o, resulting RFF values reflect changes in the vibratory rate of the vocal folds during voicing offsets and onsets.

RFF shows promise as a non-invasive, objective measure for assessing vocal hyperfunction (Roy *et al.*, 2016; Stepp *et al.*, 2010b; Stepp *et al.*, 2011). Vocal hyperfunction is a common feature of voice disorders and is described as excessive or imbalanced laryngeal muscle forces due to daily vocal overuse and/or misuse (Hillman *et al.*, 1989). Vocal hyperfunction may also present as increased vocal effort (i.e., perceived exertion of a vocalist to a perceived communication scenario; Hunter *et al.*, 2020) and (para)laryngeal stiffness while speaking (Morrison, 1997). Previous studies have shown that individuals with vocal hyperfunction produce lower average RFF values than those of typical speakers (Roy *et al.*, 2016; Stepp *et al.*, 2010b; Stepp *et al.*, 2011), implicating a possible relationship between laryngeal muscle tension and RFF. Indeed, prior work indicates a relationship between RFF and listener perceptions of vocal effort (Lien *et al.*, 2015; McKenna and Stepp, 2018; Stepp *et al.*, 2012), and between RFF and the degree of laryngeal stiffness when estimated *via* vocal fold kinematics (McKenna *et al.*, 2016). Taken together, these studies support the potential of RFF to reflect increased vocal effort and laryngeal muscle tension in individuals with vocal hyperfunction.

A. Physiological Mechanisms of RFF

Despite the potential for RFF as an objective tool for assessing vocal hyperfunction, the underlying physiological mechanisms are still unclear. Studies in typical speakers have shown that RFF values are generally increased in voicing cycles closest to the voiceless consonant and then

a)Electronic mail: yppark@bu.edu, ORCID: 0000-0003-0736-5371.
b)Also at: Department of Speech, Language, and Hearing Sciences, Boston University, 635 Commonwealth Avenue, Boston, MA 02215, USA.
c)Also at: Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, 72 East Concord Street, Boston, MA 02118, USA.

decrease toward zero during voicing onset in typical speakers (House and Fairbanks, 1953; Stepp *et al.*, 2010b; Watson, 1998). In other words, voice f_o is generally increased from steady-state values at the start of voicing onset, but normalizes to steady-state over time; this increase in f_o in voicing cycles near the voiceless obstruent may be the result of increased longitudinal tension in the vocal folds from activation of the cricothyroid (CT) muscle. Specifically, Stevens (1977) postulated that increased longitudinal vocal fold tension might aid in the devoicing needed to produce intervocalic voiceless consonants. It is possible that this tension is carried over to the first few cycles of voicing after voiceless consonants (Stevens, 1977), which would increase the f_o values—and thus RFF values—of these cycles. Indeed, contributions from the CT muscle in devoicing have been observed *via* laryngeal electromyographic experiments, in which CT muscle activity increased both before and during the production of voiceless consonants in VCV utterances (Löfqvist *et al.*, 1989). In contrast, CT muscle activity did not increase during this period when the same speakers produced *voiced* consonants (Löfqvist *et al.*, 1989).

The magnitude of this temporarily increased CT muscle activity surrounding voiceless consonant production is hypothesized to be reduced in vocal hyperfunction and to result in the decreased RFF values compared to typical speakers. The increased baseline laryngeal (and paralaryngeal) muscle tension, which are thought to be involved in vocal hyperfunction (Hillman *et al.*, 1989; Roy, 2008), have been suggested to restrict the ability of an individual with vocal hyperfunction to vary CT muscle activity for voiceless consonant production compared to typical voices (Stepp *et al.*, 2010b; Stepp *et al.*, 2011). Thus, the restricted CT muscle activity in vocal hyperfunction may decrease the magnitude of f_o increases surrounding intervocalic obstruent production.

In conjunction with increased CT activity, the vocal folds abduct to assist in devoicing during voicing offset. During vocal fold abduction, the vocal folds pull apart from each other, and the decreases in the vocal fold contact area during abduction have been hypothesized to lead to decreases in RFF during voicing offset (Watson, 1998).[1] Watson (1998) postulated that decreased vocal fold contact area might result in decreased recoil forces as well as less abrupt closure of the vocal folds. The decreased recoil forces and less abrupt vocal fold closure may lead to increases in contacting and decontacting phases of the vibratory cycle, which ultimately increase the voicing cycle period, and thus decrease f_o (Watson, 1998). Abduction, along with increased CT activity for devoicing, is thought to produce the relatively stable or slightly decreasing offset RFF values observed in young adults with healthy voices (Stepp *et al.*, 2010b; Watson, 1998). Although the hypothesis describing the effect of decreased vocal fold contact area on RFF (Watson, 1998) is theoretically reasonable, it has not yet been supported with empirical evidence.

B. Relationship between RFF and vocal fold kinematics

In this study, we sought to evaluate the hypothesis from Watson (1998) by examining differences in RFF and vocal fold kinematics between two voiceless obstruents: a fricative and a stop. We used voiceless obstruents only since RFF can be obtained from VCV contexts with a voiceless consonant where there is a clear voicing offset and onset surrounding the consonant. Prior work examining voiceless obstruents within VCV productions shows that fricatives tend to result in higher RFF values at voicing onset relative to stops (Lien *et al.*, 2014; Roy *et al.*, 2016). The effects of intervocalic obstruents on RFF values at voicing offset are less clear: Lien *et al.* (2014) found fricatives to lead to higher RFF values than stops at voicing offset, whereas Roy *et al.* (2016) found stops to result in higher RFF values at voicing offset. We postulate that these conflicting findings may be due to differences in vocal fold kinematics during intervocalic obstruent production. Specifically, the degree of abduction and adduction are thought to differ between voiceless, aspirated fricatives and stops (Löfqvist *et al.*, 1995; McGarr and Löfqvist, 1988). However, there has been no direct evidence for differences *via* vocal fold kinematic features in voicing transitions into and out of the obstruent. In the following section, we summarize previous findings on possible differences in vocal fold kinematics between fricatives and stops that motivated our research questions and hypotheses.

1. Glottal angles and duration of voicing transition

Two kinematic features are thought to differ between voiceless fricatives and stops are the glottal angle and the duration of voicing transition (i.e., duration of voicing offset or onset). From videoendoscopic images of the vocal fold movement, glottal angles at voicing offset and onset can be estimated as the glottal angles at the time points at which vocal fold vibration ceases (offset) or reinitiates (onset) during intervocalic obstruent production. The duration of voicing offset can be estimated as the duration between the start of abduction and the cessation of the vocal fold vibration, whereas the duration of voicing onset can be estimated as the duration between the reinitiation of the vocal fold vibration and the end of adduction. Discrepancies in these features are likely the result of differences in the mechanisms necessary to devoice and reinitiate voicing when producing an intervocalic fricative rather than a stop.

Voiceless fricatives are produced *via* a partial constriction—unlike voiceless stops, which require a full oral constriction—and may require additional strategies to devoice. Continuous airflow in the absence of a complete oral constriction may keep the vocal folds vibrating longer, resulting in a greater duration of voicing offset. One strategy that may be used during intervocalic fricative production is an increased reliance on vocal fold abduction to pull the vocal folds apart and terminate vocal fold vibration: increased reliance on vocal fold abduction may lead to a larger abductory angle, and as a result, a smaller vocal fold contact area during voicing offset. Indeed, voiceless fricatives have been shown to have a greater degree of abduction than voiceless stops (McGarr and Löfqvist, 1988; Yoshioka *et al.*, 1982). In addition, voiceless fricatives may also require an earlier start of abduction than voiceless stops do because frication noise, which would require glottal opening, starts immediately after the preceding vowel, whereas voiceless stops start with a period of silence after the preceding vowel. Overall, these results suggest that the transition from vowel to voiceless fricative is marked by larger glottal angles and a longer voicing offset duration than the transition from vowel to voiceless stop. The increase in voicing offset may also decrease RFF since more vibratory cycles would be affected by the decreased vocal fold contact area; thus, longer voicing offset before fricatives than stops may be related to lower offset RFF for fricatives relative to stops.

In contrast to these features during voicing offset, it is thought that voiceless aspirated stops require a larger adductory angle and a longer voicing onset duration than voiceless fricatives. Specifically, previous work has shown that vowels are produced with larger glottal flow rates and open quotients when following a voiceless aspirated stop rather than a voiceless fricative (Löfqvist *et al.*, 1995). These findings suggest that voicing onset occurred when the vocal folds were at a larger glottal angle, which may be required due to aspiration before voicing onset in American stops. Because the vocal folds are at a larger glottal angle at the start of phonation, it is possible that voicing onset necessitates a longer duration when following a voiceless stop rather than a fricative. As a result, it is thought that glottal angle and duration of voicing onset are larger when transitioning out of a voiceless stop rather than a voiceless fricative; thus, RFF would be lower at voicing onset after a stop than after a fricative.

2. Vocal fold stiffness

In addition to glottal angle and duration of voicing transition, the kinematic estimate of laryngeal stiffness (KS) may also be used to elucidate differences in vocal fold kinematics between voiceless fricatives and stops. KS has been estimated as the ratio of the maximum adductory velocity during adduction to the magnitude of changes in the glottal angle during the adductory gesture (McKenna *et al.*, 2019). Previous studies used KS to indirectly estimate vocal fold stiffness during adductory gestures (Dailey *et al.*, 2005; Stepp *et al.*, 2010a), and KS has been shown to positively correlate with stiffness values of the thyroarytenoid (TA), lateral cricoarytenoid (LCA), and posterior cricoarytenoid (PCA) muscles in a simple virtual trajectory model of vocal fold kinematics (Stepp *et al.*, 2010a). KS has also been shown to negatively correlate with voicing offset RFF values when typical speakers were instructed to modulate their vocal effort (McKenna *et al.*, 2016). These findings suggest that KS may be a useful tool for examining differences in laryngeal stiffness during the transition into and out of voiceless fricatives and stops.

Prior work suggests that voiceless stops may be produced with stiffer laryngeal adductor muscles, and thus produce greater KS values, than voiceless fricatives. Collier *et al.* (1979) used hooked-wire electromyography to investigate the TA and LCA muscles during intervocalic fricative and stop productions. The authors determined that the intervocalic production of an /f/ resulted in a higher degree of reduction in the TA and LCA muscle activity, before and during the consonant than an intervocalic /t/. Based on this finding, we suspect that intervocalic voiceless stop production may result in greater KS values (stiffer adductors) from having less relaxed TA and LCA activity than voiceless fricative production. Greater KS values for stops relative to fricatives may be related to lower RFF at both voicing offset and onset for stops relative to fricatives since the TA muscle activity at voicing transitions may decrease RFF values (although KS value does not represent the TA muscle activity alone but together with the LCA and PCA muscles; see Stepp *et al.*, 2010b). The TA muscle is likely to diminish the f_o-raising effect of the CT muscle since the TA muscles have shown to have an antagonistic effect on CT muscle in f_o control (Chhetri *et al.*, 2012). Thus, a smaller reduction in TA muscle activity for stops relative to fricatives may result in both greater KS values and lower RFF for stops relative to fricatives.

3. Effects of age on RFF and vocal fold kinematics

Age-related differences in voice production occur as a result of age-related morphological and muscular changes in the vocal folds (Honjo and Isshiki, 1980; Rodeno *et al.*, 1993; Sato and Hirano, 1997; Sato *et al.*, 2002; Watson, 1998). Older adults often exhibit stiffer vocal folds (Sato and Hirano, 1997; Sato *et al.*, 2002) and may produce voice with less adducted vocal folds due to vocal fold atrophy or bowing (Honjo and Isshiki, 1980). Moreover, older adults have been reported to produce lower RFF values than younger adults, perhaps because of a greater reliance on the abductory gesture to assist in devoicing, as suggested by Watson (1998).

C. Purpose of the current study

The current study aimed to use simultaneous acoustic and high-speed videoendoscopic (HSV) recordings to determine the relationship between vocal fold kinematics and RFF of voiceless aspirated obstruents. Acoustic signals were captured using a microphone for use in manually calculating RFF. HSV images were used to examine vocal fold movement, from which glottal angles at voicing offset (Θ_{off}) and onset (Θ_{on}) as well as durations of voicing offset (d_{off}) and onset (d_{on}) were extracted. These images were also used to compute KS.

Our hypotheses are described in detail below, as well as schematized *via* glottal angle waveforms in Fig. 1. With respect to vocal fold kinematics, we hypothesized that,

https://doi.org/10.1121/10.0003961

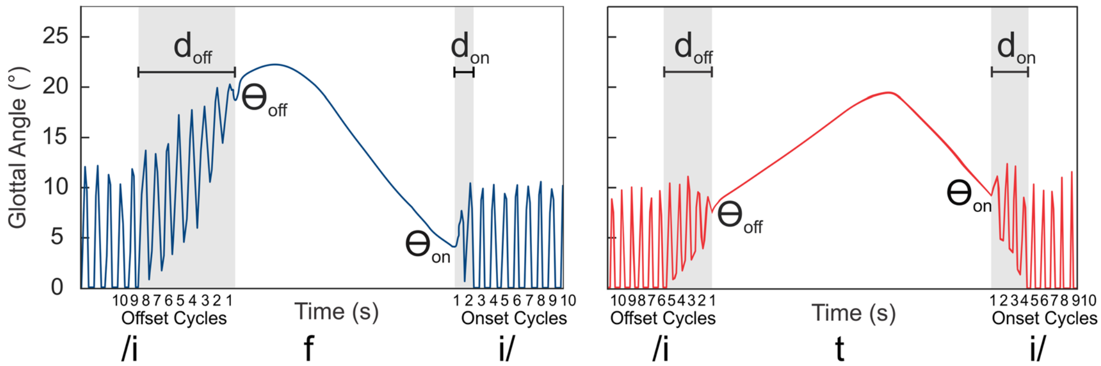

FIG. 1. (Color online) Schematics of hypothesized glottal angle waveforms extracted during /ifi/ and /iti/ VCV productions (abbreviations: Θ_{off} and Θ_{on}, glottal angles at voicing offset and onset; d_{off} and d_{on}, durations of voicing offset and onset).

relative to fricatives: (1) Θ_{off} and d_{off} would be lower in voiceless aspirated stops due to less reliance on and later timing of abduction for devoicing; (2) Θ_{on} and d_{on} would be greater in voiceless aspirated stops due to the production of aspiration before voicing onset; and (3) KS would be greater in voiceless aspirated stops due to higher TA muscle activity. With respect to changes in vocal fold kinematics with age, we hypothesized that, compared to younger adults, older adults would exhibit: (1) greater Θ_{off} and d_{off} due to an increased reliance on the abductory gesture to devoice; (2) greater Θ_{on} and d_{on} due to age-related morphological and muscular changes in the vocal folds; and (3) greater KS due to increased laryngeal stiffness. We further hypothesized that RFF at voicing offset (offset RFF) would not differ between the two obstruents due to the antagonistic effects of smaller glottal angles (increasing RFF) and stiffer laryngeal adductors (decreasing RFF) in voiceless stops, whereas RFF at voicing onset (onset RFF) would be lower in voiceless stops than in voiceless fricatives due to greater glottal angles and stiffer laryngeal adductors acting together to decrease RFF during voicing onset. We also hypothesized that, compared to younger adults, older adults would produce lower offset and onset RFF values due to greater glottal angles and increased laryngeal stiffness at both voicing offset and onset.

II. METHOD

A. Participants

Twenty typical speakers were recruited to participate in the study, including ten younger adults (five females, five males; mean age = 22.7 years, range = 19–26 years) and ten older adults (five females, five males; mean age = 62.1 years, range = 53–76 years). Participants were native English speakers, were non-smokers, and had no history of speech, language, hearing, or voice disorders. All participants were screened for normal vocal function by a certified speech-language pathologist. Speakers provided written consent prior to participation in compliance with the Boston University Institutional Review Board.

B. Recording procedures

Participants were trained to produce iterations of the utterances /ifi/ and /iti/. Each /ifi/ and /iti/ set consisted of eight consecutive productions with a pause in the middle (e.g., /ifi ifi ifi ifi, pause, /ifi ifi ifi ifi/). For this study, /ifi/ was selected to examine the effects of an intervocalic, voiceless fricative (i.e., /f/) on vocal fold kinematics and resulting RFF values. Similarly, /iti/ was selected to examine the effects of an intervocalic, voiceless stop (i.e., /t/). Participants were instructed to produce these stimuli with equal stress (same pitch and duration) on both vowels using a comfortable voice (Park and Stepp, 2019).

Prior to recording, a directional headset microphone (SM35 XLR, Shure) was placed 7 cm away from the corner of the mouth at a 45-degree angle from the midline. A speech-language pathologist trained to perform transnasal endoscopy examination visualized participants' larynges using a flexible nasendoscope. A pediatric endoscope (Pentax, model FNL-7RP3, 2.4 mm) was used for 14 of the participants and an adult nasendoscope (Pentax, model FNL-10RP3, 3.5 mm) was used on the remainder based on the clearance of nasal passages and their tolerance to the endoscope. All examinations were performed in a sound-treated room at Boston University. A FASTCAM Mini AX100 camera (model 540 K-C-16GB), operated at a 256×256 pixel resolution and at a frame rate of 1000 frames/s, was attached to the endoscope by a 40-mm optical lens adaptor along with a steady xenon light source (300 W KayPentax model 7162B) for high-speed visualization. The video images were obtained through Photron Fastcam Viewer software (version 3.6.6). Although the endoscopy procedure could cause discomfort, a numbing agent was not provided, due to its effect on laryngeal sensory feedback. However, the participants were presented with the option of using a nasal decongestant. After a clear view of the larynx was obtained, the participants produced one set of eight /ifi/s followed by one set of eight /iti/s. If the participant produced unequally stressed vowels or vocal fry or if the speech-language pathologist failed to acquire clear larynx images, participants were instructed to repeat the set. The acoustic signals were recorded, amplified with Xenyx Behringer 802

Preamplifier, and digitalized at 30 kHz using a data acquisition board (DAQ; National Instruments 6312 USB) to time-synchronize acoustic recordings with the HSV images.

C. Data analysis

1. HSV image analysis

A series of kinematic measures was extracted semi-automatically from HSV images using a graphical user interface described in Diaz-Cadiz *et al.* (2019). In brief, the algorithm uses differences in pixel intensities between the glottis and the vocal folds to estimate the glottal angle during laryngeal articulatory and/or vibratory movements as a function of time. Figure 2 presents an example of a glottal angle waveform from an /ifi/ utterance. Within the graphical user interface, a single technician examined the glottal angle waveform and HSV images in order to locate a series of pertinent time points necessary to estimate vocal fold kinematics, as described below. The technician corroborated the raw HSV images in order to minimize errors that may occur in time point identification if the glottal angle waveform failed to capture small glottal gaps or movements when the differences in pixel intensities between the glottis and the vocal folds were unclear. In these cases, the time points were determined from the raw HSV images.

Two time points were extracted through manual examination of the glottal angle waveform and HSV images during devoicing: the start of abduction (t_{abd}) and the time of voicing offset (t_{off}). The technician marked t_{abd} as the last complete contact of the vocal folds during the voicing offset. If the vocal folds had never reached full closure during the preceding vowel, t_{abd} was marked as the time at which the last maximum closure of the vocal folds occurred before the glottal angle started to increase for abduction (18.3%). If arytenoid cartilages blocked the view of the vocal folds, t_{abd} was marked as the time at which the two arytenoid cartilages started to move apart from one another (14.1%). The location of t_{off} was considered as the termination of the last vibratory cycle prior to the obstruent. Two time points were then identified during the reinitiation of voicing: the time of voicing onset (t_{on}) and the termination of adduction (t_{add}). The location of t_{on} was marked as the time point in which the vocal folds first started to vibrate after the obstruent. The technician marked t_{add} at the time point in which the first full or maximum (22.1%) vocal fold closure during voicing onset was achieved, or, if the view of the vocal folds was blocked, when the arytenoid cartilages stopped moving toward each other (18.6%).

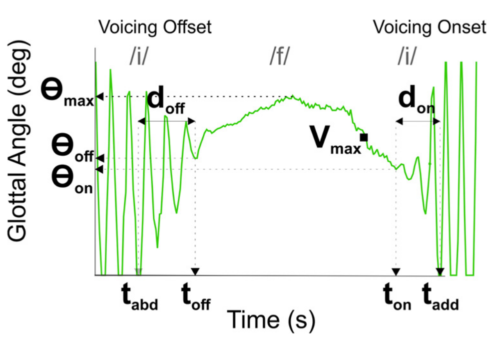

FIG. 2. (Color online) An actual glottal angle waveform during a participant's production of /ifi/ surrounding the obstruent (abbreviations: Θ_{max}, maximum glottal angle; V_{max}, maximum adductory velocity; Θ_{off} and Θ_{on}, glottal angles at voicing offset and onset; d_{off} and d_{on}, durations of voicing offset and onset; t_{abd}, start of abduction; t_{off}, voicing offset; t_{on}, voicing onset; t_{add}, end of adduction).

After locating these time points using the glottal angle waveform, four estimates of vocal fold kinematics were obtained:

(1) Θ_{off}: Glottal angle at voicing offset, extracted at t_{off}
(2) d_{off}: Duration of voicing offset, calculated as $t_{off} - t_{abd}$
(3) Θ_{on}: Glottal angle at voicing onset, extracted at t_{on}
(4) d_{on}: Duration of voicing offset, calculated as $t_{add} - t_{on}$

Finally, a KS was calculated as the ratio of the maximum adductory velocity during the adductory gesture prior to voicing onset (V_{max}) to the magnitude of changes in the glottal angle during the adductory gesture, which was the maximum glottal angle (Θ_{max}) in most cases (or in case of incomplete vocal fold closure after voicing onset, the maximum glottal angle minus the minimum glottal angle after voicing onset). V_{max} was estimated from the sigmoidal fit of the adductory gesture that spanned from Θ_{max} to Θ_{on}.

In sum, five vocal fold kinematic measures were obtained from this analysis: Θ_{off}, Θ_{on}, d_{off}, d_{on}, and KS. These measures were averaged within participants for all /ifi/ and /iti/ productions to enable comparisons across participants in terms of obstruent type (i.e., /f/ or /t/) and age group (i.e., younger or older adult).

2. Manual RFF estimation

A single trained technician[2] carried out manual RFF estimation using Praat software (version 6.0.21). The technician first examined each /ifi/ and /iti/ waveform to determine whether there was evidence of voicing during the associated obstruent. If there was clear evidence of voicing, such as glottal pulses present during the obstruent or a short duration of aspiration (voice onset time < 40 ms), the RFF production was immediately rejected (2.7%). Otherwise, the technician proceeded with RFF computation, as follows: (1) the boundary between voiced and voiceless speech segments was identified for both voicing offset and onset, (2) the voicing cycles closest to the voice offset boundary (offset cycle 10) and voice onset boundary (onset cycle 1) were located, (3) the nine voicing cycles before voice offset cycle 10 and after voice onset cycle 1 were selected *via* examining the general waveform shape of the vibratory cycles, (4) the instantaneous f_o of each voicing cycle was calculated as the inverse of cycle period, which was estimated with the autocorrelation method in Praat, and (5) RFF (semitones; ST) was calculated using Eq. (1),

$$\mathrm{RFF\,(ST)} = 12 \cdot \log_2\left(\frac{f_o}{f_o^{ref}}\right). \tag{1}$$

In Eq. (1), f_o^{ref} refers to the f_o of the cycles located furthest from the obstruent, assumed to be part of steady-state voicing. Specifically, the f_o of offset cycle 1 was used as f_o^{ref} for offset cycles and the f_o of onset cycle 10 was used as f_o^{ref} for onset cycles (Stepp *et al.*, 2010a).

RFF values were then examined to determine whether offset or onset instances were valid following criteria set by Vojtech and Heller Murray (2019). If the RFF value closest to the reference cycle (offset cycle 2, onset cycle 9) was larger than 0.8 ST, which suggested that steady-state voicing was not achieved, this specific instance was rejected (1.2%). Any RFF instances with evidence of vocal fry (low frequency, irregular glottal pulses in the spectrogram) were also rejected, since accurate estimation of f_o during vocal fry is not possible (2.9%). We also excluded RFF instances in which HSV images could not be analyzed due to obstruction of the view of the vocal folds or poor resolution of the images (1.9%). In total, 51 of 581 (8.8%) offset and/or onset instances were rejected. RFF values from offset cycle 10 (offset 10 RFF) and onset cycle 1 (onset 1 RFF) values were averaged within participants for /ifi/ and /iti/. These values were considered for further examination, as they have been shown to reflect the largest differences between individuals with healthy and hyperfunctional voices (Stepp *et al.*, 2011).

We confirmed that most of the RFF stimuli were produced with similar f_o values during both vowels by calculating the ST differences between f_os of offset cycle 1 and onset cycle 10, which were furthest from the consonants. A ST is equivalent to a half step in musical notation, which is typically perceived as the smallest musical interval (Zarate *et al.*, 2012). The ST differences, on average, were near zero (mean = 0.36 ST, standard deviation, SD = 0.41 ST). Although baseline f_o and loudness have been reported to have minor effects on RFF (Park and Stepp, 2019), we also checked the differences in f_o and dB sound pressure level (SPL) between /ifi/ and /iti/ productions and found minor differences (f_o; mean = 0.66 ST, SD = 0.89 ST; dB SPL: mean = 1.23 dB SPL, SD = 0.89 dB SPL).

3. Reliability

Prior to carrying out statistical analyses, the reliability of the extracted vocal fold kinematic measures and RFF values was assessed. To assess intrarater reliability, the primary technician reanalyzed 20% of HSV and RFF samples in a separate sitting. Interrater reliability was assessed by comparing HSV time markings and RFF measures between the primary technician and an additional technician who carried out HSV image analysis and manual RFF estimation on 20% of samples. Intrarater and interrater reliability were each calculated *via* two-way mixed-effects intraclass correlation coefficients (ICC) for absolute agreement (single measures).

Reliability measures are presented in Table I. Overall, the reliability of HSV-based time markings were excellent (Portney and Watkins, 2000), with intrarater reliability reaching a mean ICC = 0.98 (SD = 0.02, range = 0.96–0.99) and

TABLE I. Intra- and inter-reliability values (intraclass correlation coefficients; single measures, absolute agreement, a 2-way mixed-effects model) of all measures (95% confidence interval, CI, in parentheses).

Measures	Intrarater Reliability	Interrater Reliability
RFF	0.87	0.76
	(0.81–0.91)	(0.66–0.83)
t_{abd}	0.96	0.98
	(0.89–0.98)	(0.96–0.99)
t_{off}	0.99	0.99
	(0.99–0.99)	(0.90–0.99)
t_{on}	0.99	0.99
	(0.99–0.99)	(0.99–0.99)
t_{add}	0.99	0.99
	(0.99–0.99)	(0.98–0.99)

interrater reliability averaging at ICC = 0.99 (SD = 0.01, range = 0.98–0.99). Intrarater reliability of RFF markings was good (Portney and Watkins, 2000), with a mean intrarater reliability of ICC = 0.87 and interrater reliability of ICC = 0.76.

4. Statistical analysis

Wilcoxon signed-rank tests were performed on RFF (offset 10, onset 1) and vocal fold kinematic measures (Θ_{off}, Θ_{on}, d_{off}, $d_{on,}$ and KS) to evaluate hypothesized differences in these measures between /ifi/ and /iti/ within participants. To assess within-participant relationships between RFF and kinematic variables we used repeated-measures correlations (Bakdash and Marusich, 2017), as previous studies have reported that correlations between RFF and other measures are stronger within participants than across participants (Lien *et al.*, 2015; McKenna *et al.*, 2016). Mann-Whitney U tests were then performed to compare group differences in RFF and vocal fold kinematics between younger and older adults. Nonparametric tests (Wilcoxon signed-rank tests and Mann-Whitney U tests) were used because of the small sample size. Resulting p values were adjusted using the Bonferroni correction to account for 20 tests (seven Wilcoxon signed-rank tests, six repeated measures correlations, seven Mann-Whitney U tests; i.e., p = 0.05/20 = 0.0025).

III. RESULTS

Figure 3 presents the mean RFF and vocal fold kinematic measures for /ifi/ and /iti/ productions across all participants. Offset 10 RFF values were not statistically significantly different between /ifi/ and /iti/, whereas onset 1 RFF values were statistically significantly greater in /ifi/ productions (p = 0.002). The vocal fold kinematic measures, Θ_{off} and d_{off}, were statistically significantly greater in /ifi/ relative to /iti/ ($p < 0.001$), whereas Θ_{on} and KS were statistically greater in /iti/ than in /ifi/ (p = 0.002). Although these tests showed statistically significant differences at the group level, there was considerable inter-participant variability. Some participants showed small differences in these

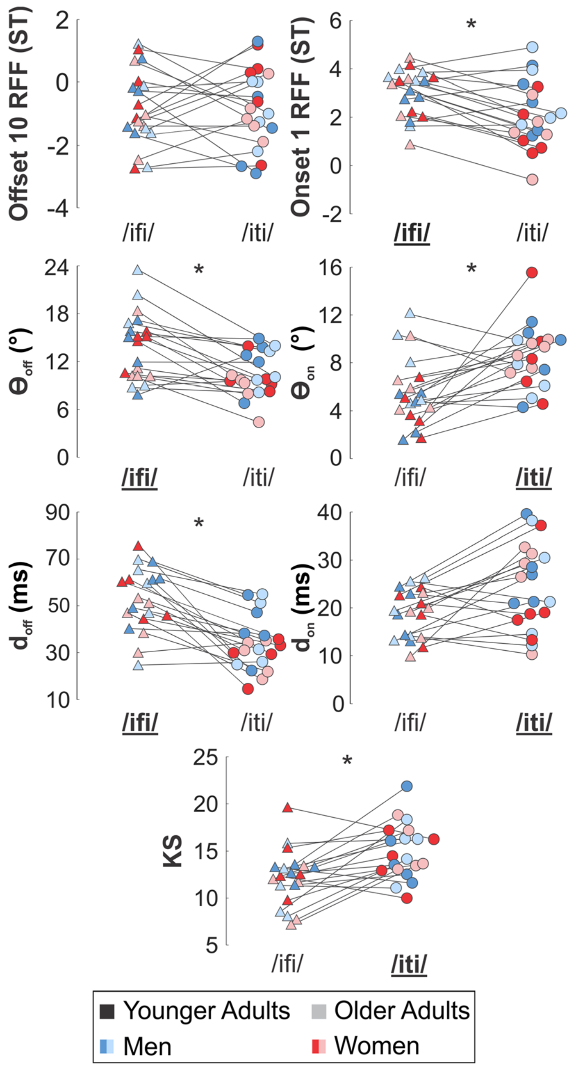

FIG. 3. (Color online) RFF and the vocal fold kinematic measures from /ifi/ and /iti/. Lines connect data from each participant. Within each measure, the bolded and underlined /ifi/ or /iti/ on the x axis represents which utterance was hypothesized to have greater values for the given measure. Asterisks indicate statistical significance at $p = 0.0025$ (abbreviations: Θ_{off} and Θ_{on}, glottal angles at voicing offset and onset; d_{off} and d_{on}, durations of voicing offset and onset; KS, kinematic estimate of laryngeal stiffness).

measures between fricatives or stops and some even showed patterns opposite of the group. Average d_{on} values were greater in /iti/ than in /ifi/, but this difference did not reach statistical significance ($p = 0.007$).

Repeated measures correlations were performed to assess possible within-participant associations between RFF and vocal fold kinematic measures. Offset 10 RFF was not statistically significantly correlated with Θ_{off} ($r = -0.20$, $p = 0.19$), d_{off} ($r = -0.24$, $p = 0.15$), or KS ($r = 0.21$, $p = 0.18$). On the other hand, onset 1 RFF was moderately correlated with Θ_{on} ($r = -0.46$, $p = 0.02$), d_{on} ($r = -0.42$, $p = 0.03$), and KS ($r = -0.36$, $p = 0.05$); however, these relationships were not statistically significant.

Mann-Whitney U tests to compare younger and older speakers did not show statistically significant differences for any of the measures (Fig. 3).

IV. DISCUSSION

The present study sought to examine the relationship between vocal fold kinematics and RFF during voicing offset and onset. To carry out this investigation, we chose two voiceless obstruents, /f/ and /t/, to characterize the effects of intervocalic fricative and stop productions on resulting measures. We further examined the effect of speaker age on vocal fold kinematics and RFF.

A. RFF and vocal fold kinematics during intervocalic fricatives and stops

1. Voicing offset

As hypothesized, offset 10 RFF values were not statistically significantly different between /ifi/ and /iti/. These results may be due to the effects of less stiff laryngeal adductors (hypothesized to increase RFF) canceling out the effects of more abducted vocal folds (hypothesized to decrease RFF) at voicing offset prior to fricative production. We suspect that our findings of lower kinematic laryngeal stiffness during /ifi/ may be due to the higher degree of relaxation of the TA muscles observed prior to and during the production of a fricative compared to the degree of relaxation of the TA muscles observed prior to and during the production of a stop consonant (Collier *et al.*, 1979). Less TA activity may increase offset RFF prior to fricative production due to the antagonistic relationship between the TA and CT muscles in f_o control (Chhetri *et al.*, 2012). KS showed a statistically significant, negative correlation with offset 10 RFF values in McKenna *et al.* (2016); however, we did not observe a statistically significant association between KS and offset 10 RFF. This may be the result of differences in methodology. McKenna *et al.* (2016) included recordings of speakers modulating their vocal effort level and thus had a wider range of KS (7.3–31.9) compared to the range of KS in our study (7.2–21.8). Additionally, McKenna *et al.* (2016) used only /ifi/ for the recording stimulus, whereas we examined two different obstruents in RFF stimuli that could vary in other dimensions (e.g., glottal angle and duration of voicing offset) that may affect RFF.

Greater glottal angles and longer durations of voicing offset were generally observed prior to fricative production relative to stop production. This finding is consistent with previous observations of a greater degree of abduction during fricatives than during stops (McGarr and Löfqvist, 1988; Yoshioka *et al.*, 1982). McGarr and Löfqvist (1988)—who examined transillumination images of vocal fold movements during VCV stimuli for both voiceless fricatives and stops in a single healthy speaker—demonstrated that fricatives resulted in higher maximum glottal areas and abduction velocities when compared to stops. Moreover, a hooked-wire electromyographic experiment showed that fricatives led to greater PCA muscle activity than stops (Yoshioka *et al.*, 1982). Although this study provides additional evidence for increased reliance on abduction prior to fricative production (i.e., as the PCA is an abductor muscle), it was also only conducted in a single participant. The results of the current study provide additional evidence that fricative production involves a greater degree and a longer period of vocal fold abduction than stop production.

Greater glottal angles and durations of voicing offset resulting from greater abduction in fricative production were thought to lower offset 10 RFF based on the hypothesis from Watson (1998). However, we also hypothesized that there would be a co-occurring effect of lower KS (hypothesized to increase RFF), thereby canceling out possible RFF-decreasing effects of greater glottal angle and duration of voicing offset. Thus, we hypothesized that offset 10 RFF would not be statistically different between the obstruents. Unsurprisingly, we found no statistically significant difference in offset 10 RFF values between the obstruents and no statistically significant correlations between offset 10 RFF values and glottal angle or duration of voicing offset.

2. Voicing onset

In the current study, onset 1 RFF values were generally lower and glottal angles were greater at voicing onset for intervocalic stops relative to fricatives. Although some participants showed different trends between the obstruents due to possible intra- and inter-speaker variability in these measures, as previously observed in RFF (Heller Murray *et al.*, 2017), the statistical tests indicated significantly lower onset 1 RFF and greater glottal angles at voicing onset for stops relative to fricatives. Our results are consistent with prior work showing that onset 1 RFF is lower (Lien *et al.*, 2014; Roy *et al.*, 2016) and minimum glottal flow rates at voicing onset (which may indicate how open the glottis is and thus glottal angles at voicing onset) are higher (Löfqvist *et al.*, 1995) for stops relative to fricatives. The larger glottal angle may be required to produce the burst and aspiration before the intervocalic aspirated stop, /t/. Furthermore, these findings are consistent with the hypothesis from Watson (1998), which posited a negative relationship between vocal fold contact area and RFF. A moderate, negative correlation was also identified between onset 1 RFF and Θ_{on} ($r = -0.46$, $p = 0.02$); however, this relationship was not statistically significant. This may have been due to the limited range of Θ_{on} ($0°$–$15.5°$) being produced by individuals with healthy voices. Future studies should therefore examine the relationship between glottal angle and RFF at voicing onset in a wider range of vocal function, including individuals with voice disorders and/or individuals self-modulating their vocal effort.

Based on these results, we suspect that the degree of vocal fold closure at voicing offset and onset surrounding an obstruent may be an important factor that contributes to RFF. Specifically, Watson (1998) suggested that decreased vocal fold contact area in a vibratory cycle would decrease the magnitudes of both recoil forces (thus, opening the vocal folds) and abrupt closure, which would increase the period and thus decrease the frequency of the cycle. Results from prior computational modeling also support this hypothesis: when the vocal folds were positioned at a greater glottal angle prior to phonation (i.e., decreased vocal fold contact area), the experimenters observed decreases in the closed quotient of the vibratory cycle and the resultant f_o (Zhang, 2016). A decrease in the closed quotient suggests a decrease in vocal fold contact area as the vocal folds may remain closed for a shorter time in the vibratory cycle from slower opening and closing of the vocal folds. These hypotheses may be indirectly supported by the greater glottal angles at voicing onset and the lower onset 1 RFF values observed in the current study. Additionally, because the degree of vocal fold closure is decreased during voicing offset and onset, the duration and timing of abduction and adduction may also affect RFF. Future studies could further investigate the effects of duration and timing of abduction and adduction on RFF with speakers with both typical and disordered voices who may show wide ranges of duration and timing of abduction and adduction.

Lower onset 1 RFF values following intervocalic stops may also be the result of a higher KS observed during stop production relative to fricative production. Lien *et al.* (2014) reasoned that onset 1 RFF was higher for fricatives in their study because the TA muscles are more relaxed during fricative production than during stop productions, as shown in Collier *et al.* (1979). Our result showing lower stiffness during /ifi/ than /iti/ also supports the notion that the TA muscles may be more relaxed during fricative production than during stop production. Although the exact reason behind the different degree of TA muscle relaxation in fricative and stop productions is not yet known, earlier timing of abduction for frication production in fricatives relative to stops may contribute to this difference because earlier abduction would require adductor muscles, including the TA muscles, to relax earlier. Less relaxed TA activity during stop production could decrease onset RFF since the CT (which aids in devoicing) and TA muscles exhibit antagonistic effects on f_o when activated together (Chhetri *et al.*, 2012). Although we observed both higher kinematic laryngeal stiffness and lower onset 1 RFF during /iti/ productions, we did not find a statistically significant relationship between KS and onset 1 RFF. This finding is similar to the results from McKenna *et al.* (2016), in which only 40% of the participants showed moderate, negative relationships between KS and onset 1 RFF when the speakers modulated their vocal effort.

B. RFF and vocal fold kinematics in younger and older adults

We recruited both younger and older adults to evaluate the associations between RFF and the kinematic measures in a wide range of voices, as well as to investigate the differences between younger and older adult voices due to aging. In contrast to our hypotheses, we did not find statistically significant differences between age groups in any of our measures. Due to the observations of Watson (1998), we hypothesized that RFF would be lower in older adults. We further hypothesized that glottal angles, voicing transition durations, and KS would be higher in older adults due to increased difficulty with voicing offset and onset, as well as increased vocal fold stiffness due to age-related changes to the laryngeal system (Honjo and Isshiki, 1980; Rodeno *et al.*, 1993; Sato and Hirano, 1997; Sato *et al.*, 2002; Watson, 1998). The anatomical changes in aging laryngeal systems have also been mirrored by the changes in voice quality and vocal function as a result of age in previous studies (Dehqan *et al.*, 2012; Xue and Deliyski, 2001). A possible reason that we did not find differences as a function of speaker age may be because our older group was younger than the older groups recruited in previous studies. Older adults in these previous studies had an average age of roughly 75 years; however, the age of our older adult participants ranged from 53 to 76 years, with an average of only 62 years. Additionally, our older female participants especially may not have had enough age-related anatomical changes in the larynx as there is some evidence suggesting that aging in the larynx may appear later in females than in males (Pufe *et al.*, 2004). Due to this discrepancy, the anatomical characteristics of our older adult group may be different from those of these prior studies. Futures studies should evaluate older adults aged over 75 years to assess possible differences in RFF and vocal fold kinematic measures.

C. Implications for vocal hyperfunction

Although all of the participants studied had healthy voices, our findings may be used to inform future hypotheses about potential laryngeal mechanisms underlying differences in RFF in individuals with vocal hyperfunction. RFF has been shown to be lower in individuals with vocal hyperfunction than in individuals with typical voices in several studies (Heller Murray *et al.*, 2017; Roy *et al.*, 2016; Stepp *et al.*, 2010a; Stepp *et al.*, 2011). Yet the original physiological hypothesis about lower RFF in individuals with vocal hyperfunction was based on reduced activity of the CT muscle before and during the voiceless consonant as a result of increased baseline laryngeal tension (Stepp *et al.*, 2010a; Stepp *et al.*, 2011). Heller Murray *et al.* (2017) recently hypothesized that increased transverse laryngeal tension (tightly approximated vocal folds) may also result in an increased degree of abduction to terminate vocal fold vibration, which was thought to decrease offset RFF by Watson (1998). In line with their hypotheses, we also suspect that individuals with vocal hyperfunction may require additional mechanisms to achieve devoicing if the activity of the CT muscle—which is hypothesized to assist in devoicing—is reduced due to increased baseline laryngeal tension. An example of this mechanism could be an increased degree of abduction, which was also observed in intervocalic voiceless fricative production (Collier *et al.*, 1979; Yoshioka *et al.*, 1982). Thus, individuals with vocal hyperfunction may have increased glottal angles and durations of voicing offset, which may decrease their offset RFF. This speculation is consistent with the hypothesis proposed by Watson (1998). A possible difference in the degree of adduction between typical speakers and speakers with vocal hyperfunction may also be a factor in explaining their differences in RFF. Previous studies also suggest that increases in glottal angles and durations of voicing onset may be due to glottal insufficiency in individuals with vocal hyperfunction (Galindo *et al.*, 2017; Woo, 2017). Future HSV studies should directly investigate these hypotheses on voicing offset and onset features of individuals with vocal hyperfunction when compared to healthy voices.

D. Limitations

One limitation of this study is the set order of /ifi/ and /iti/ productions during HSV. Recording order could affect vocal function if there was a difference in the effect of flexible endoscopy on the participant over time (e.g., due to discomfort). Although psychological stress has been associated with changes in vocal function (Dietrich and Verdolini Abbott, 2012; Helou *et al.*, 2013), the effect of flexible endoscopy without anesthetic on voice over time has not been studied. Considering the relatively short duration of endoscopy and possible variations in how participants tolerate the endoscope, we suspect that changing the order of recording would not substantially affect our results.

The frame rate used for videoendoscopic recordings is another possible limitation of this study. Our videos were recorded at a frame rate of 1000 frames/s, which is well above the typical speaking f_o of adults, and technicians manually selected t_{abd} and t_{add} based on subjective evaluations of HSV images and glottal angle waveforms. If the indicators for the actual initiation or termination of voicing offset or onset (e.g., the last complete glottal closure before the start of abduction) were not captured due to the frame rate, adjacent cycles would have been selected, potentially leading to less precise measurements for d_{off} and d_{on}. However, the difference we observed in d_{off} between /ifi/ and /iti/ was an average of 19 ms, which is around two to four times greater than typical fundamental periods of speaking voices (5–10 ms). Thus, we suspect that the frame rate is unlikely to have affected the results of the study.

Different places of articulation between /ifi/ and /iti/ may have also affected our results. Different places of articulation may result in differences in vocal tract length and supraglottal pressure applied to the vocal folds when the obstructions in the oral cavity occur during consonant production. Although the impact of the place of articulation on our results is likely to be small since the places of articulations of the fricative and stop are very close (i.e., /f/ and /t/), differences in supraglottal pressure between productions of two obstruents in our study are unknown and the possible differences may have affected our results.

Differences in speaking f_os between sexes may also affect RFF results. Females who typically have higher f_o values are likely to have their reference f_os for RFF estimation (offset cycle 1 and onset cycle 10) closer to the consonant in duration than males since the periods of the cycles are shorter. A previous study also observed differences in RFF between sexes (Stepp, 2013). There may be also differences in f_o changes due to aging between sexes (Honjo and Isshiki, 1980). However, a sex effect in the current study should be minimized due to sex-matching the composition of the groups, except for a possible interaction between age and sex due to potential differences in aging progression between sexes. Future studies could further investigate the effect of sex on RFF.

Although not investigated in this study, individual differences in voice onset time, the duration from the release of stops to voicing onset, may have also affected our results. Voice onset time is different from the duration of voicing onset investigated in this study. Voice onset time measures the duration of the unvoiced portion after the release of stops and right before voicing onset, whereas the duration of voicing onset measured the duration of the voiced portion from voicing onset to the end of adduction. Previous studies have shown that individual speakers differ in their voice onset time (Morris *et al.*, 2008) and that, on average, females produce longer voice onset times (Robb *et al.*, 2005). Because voice onset time measures the event right before voicing onset in stop production, different voice onset times may affect voicing onset measures obtained from voiceless stop consonants.

V. CONCLUSIONS

In the current study, we examined the differences in RFF and vocal fold kinematics between intervocalic voiceless fricatives and stops. Our results suggest that lower onset RFF values, smaller glottal angles and durations of voicing offset, larger glottal angles at voicing onset, and greater kinematic estimates of laryngeal stiffness are observed in the production of intervocalic aspirated stops compared to fricatives. The lower onset RFF may be related to less adducted vocal folds and stiffer laryngeal adductors required during stop production. This work indirectly supports prior speculation of a positive relationship between the degree of vocal fold contact area and RFF. Future studies should use computational modeling to examine this hypothesis directly.

ACKNOWLEDGMENTS

This work was supported by Grant Nos. DC015570 (CES) and DC013017 (CAM & CES) from the National Institute on Deafness and Other Communication Disorders, 1247312 (JMV) from the National Science Foundation, and student research award (FW) from the Undergraduate Research Opportunities Program at Boston University. Thanks to Daniel Buckley, Kimberly Dahl, and Surbhi Hablani for assistance with data recording, to Austin Luong, Jacob Noordzij Jr., and Samantha Shank for help with data analysis, and to Joseph Perkell, Daryush Mehta, Robert Hillman, and Matias Zañartu for their advice on interpretation of the results. C.E.S. has received consulting fees from Altec, Inc./Delsys, Inc., companies focused on developing and commercializing technologies related to human movement. The interests of C.E.S. were reviewed and are managed by Boston University in accordance with their conflict of interest policies. All other authors have declared that no competing interests existed at the time of publication.

[1]Abduction may also result in changes in vocal fold length (Sawashima *et al.*, 1981), which may contribute to RFF. However, we did not investigate vocal fold length in this study because our methodology using flexible endoscopy is not sufficient to measure it accurately.

[2]The dataset used to train individuals in manual RFF estimation is a separate dataset from that described here and may be downloaded from https://sites.bu.edu/stepplab/research/rff/.

Bakdash, J. Z., and Marusich, L. R. (**2017**). "Repeated measures correlation," Front. Psychol. **8**, 456.

Chhetri, D. K., Neubauer, J., and Berry, D. A. (**2012**). "Neuromuscular control of fundamental frequency and glottal posture at phonation onset," J. Acoust. Soc. Am. **131**(2), 1401–1412.

Collier, R., Lisker, L., Hirose, H., and Ushijima, T. (**1979**). "Voicing in intervocalic stops and fricatives in Dutch," J. Phon. **7**, 357–373.

Dailey, S. H., Kobler, J. B., Hillman, R. E., Tangrom, K., Thananart, E., Mauri, M., and Zeitels, S. M. (**2005**). "Endoscopic measurement of vocal fold movement during adduction and abduction," Laryngoscope **115**(1), 178–183.

Dehqan, A., Scherer, R. C., Dashti, G., Ansari-Moghaddam, A., and Fanaie, S. (**2012**). "The effects of aging on acoustic parameters of voice," Folia Phon. Logopaed. **64**(6), 265–270.

Diaz-Cadiz, M., McKenna, V. S., Vojtech, J. M., and Stepp, C. E. (**2019**). "Adductory vocal fold kinematic trajectories during conventional versus high-speed videoendoscopy," J. Speech Lang. Hear. Res. **62**(6), 1685–1706.

Dietrich, M., and Verdolini Abbott, K. (**2012**). "Vocal function in introverts and extraverts during a psychological stress reactivity protocol," J. Speech Lang. Hear. Res. **55**(3), 973–987.

Galindo, G. E., Peterson, S. D., Erath, B. D., Castro, C., Hillman, R. E., and Zanartu, M. (**2017**). "Modeling the pathophysiology of phonotraumatic vocal hyperfunction with a triangular glottal model of the vocal folds," J. Speech Lang. Hear. Res. **60**(9), 2452–2471.

Heller Murray, E. S., Lien, Y. S., Van Stan, J. H., Mehta, D. D., Hillman, R. E., Pieter Noordzij, J., and Stepp, C. E. (**2017**). "Relative fundamental frequency distinguishes between phonotraumatic and non-phonotraumatic vocal hyperfunction," J. Speech Lang. Hear. Res. **60**(6), 1507–1515.

Helou, L. B., Wang, W., Ashmore, R. C., Rosen, C. A., and Abbott, K. V. (**2013**). "Intrinsic laryngeal muscle activity in response to autonomic nervous system activation," Laryngoscope **123**(11), 2756–2765.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., and Vaughan, C. (**1989**). "Objective assessment of vocal hyperfunction: An experimental framework and initial results," J. Speech Hear. Res. **32**(2), 373–392.

Honjo, I., and Isshiki, N. (**1980**). "Laryngoscopic and voice characteristics of aged persons," Arch. Otolaryngol. **106**(3), 149–150.

House, A. S., and Fairbanks, G. (**1953**). "The influence of consonant environment upon the secondary acoustical characteristics of vowels," J. Acoust. Soc. Am. **25**, 105–113.

Hunter, E. J., Cantor-Cutiva, L. C., van Leer, E., van Mersbergen, M., Nanjundeswaran, C. D., Bottalico, P., Sandage, M. J., and Whitling, S. (**2020**). "Toward a consensus description of vocal effort, vocal load, vocal loading, and vocal fatigue," J. Speech Lang. Hear. Res. **63**(2), 509–532.

Lien, Y. S., Gattuccio, C. I., and Stepp, C. E. (**2014**). "Effects of phonetic context on relative fundamental frequency," J. Speech Lang. Hear. Res. **57**(4), 1259–1267.

Lien, Y. S., Michener, C. M., Eadie, T. L., and Stepp, C. E. (**2015**). "Individual monitoring of vocal effort with relative fundamental frequency: Relationships with aerodynamics and listener perception," J. Speech Lang. Hear. Res. **58**(3), 566–575.

Löfqvist, A., Baer, T., McGarr, N. S., and Story, R. S. (**1989**). "The cricothyroid muscle in voicing control," J. Acoust. Soc. Am. **85**(3), 1314–1321.

https://doi.org/10.1121/10.0003961

Löfqvist, A., Koenig, L. L., and McGowan, R. S. (**1995**). "Vocal-tract aerodynamics in /aCa/ utterances: Measurements," Speech Commun. **16**(1), 49–66.

McGarr, N. S., and Löfqvist, A. (**1988**). "Laryngeal kinematics in voiceless obstruents produced by hearing-impaired speakers," J. Speech Hear. Res. **31**(2), 234–239.

McKenna, V. S., Diaz-Cadiz, M. E., Shembel, A. C., Enos, N. M., and Stepp, C. E. (**2019**). "The relationship between physiological mechanisms and the self-perception of vocal effort," J. Speech Lang. Hear. Res. **62**(4), 815–834.

McKenna, V. S., Heller Murray, E. S., Lien, Y. S., and Stepp, C. E. (**2016**). "The relationship between relative fundamental frequency and a kinematic estimate of laryngeal stiffness in healthy adults," J. Speech Lang. Hear. Res. **59**(6), 1283–1294.

McKenna, V. S., and Stepp, C. E. (**2018**). "The relationship between acoustical and perceptual measures of vocal effort," J. Acoust. Soc. Am. **144**(3), 1643–1658.

Morris, R. J., McCrea, C. R., and Herring, K. D. (**2008**). "Voice onset time differences between adult males and females: Isolated syllables," J. Phon. **36**(2), 308–317.

Morrison, M. D. (**1997**). "Pattern recognition in muscle misuse voice disorders: How I do it," J. Voice **11**(1), 108–114.

Park, Y., and Stepp, C. E. (**2019**). "The effects of stress type, vowel identity, baseline f_0, and loudness on the relative fundamental frequency of individuals with healthy voices," J. Voice **33**(5), 603–610.

Portney, L. G., and Watkins, M. P. (**2000**). *Foundations of Clinical Research: Applications to Practice* (Prentice-Hall, Englewood Cliffs, NJ).

Pufe, T., Mentlein, R., Tsokos, M., Steven, P., Varoga, D., Goldring, M. B., Tillmann, B. N., and Paulsen, F. P. (**2004**). "VEGF expression in adult permanent thyroid cartilage: Implications for lack of cartilage ossification," Bone **35**(2), 543–552.

Robb, M., Gilbert, H., and Lerman, J. (**2005**). "Influence of gender and environmental setting on voice onset time," Folia Phon. Logopaed. **57**(3), 125–133.

Rodeno, M. T., Sanchez-Fernandez, J. M., and Rivera-Pomar, J. M. (**1993**). "Histochemical and morphometrical ageing changes in human vocal cord muscles," Acta Otolaryngol. **113**(3), 445–449.

Roy, N. (**2008**). "Assessment and treatment of musculoskeletal tension in hyperfunctional voice disorders," Int. J. Speech Lang. Pathol. **10**(4), 195–209.

Roy, N., Fetrow, R. A., Merrill, R. M., and Dromey, C. (**2016**). "Exploring the clinical utility of relative fundamental frequency as an objective measure of vocal hyperfunction," J. Speech Lang. Hear. Res. **59**(5), 1002–1017.

Sato, K., and Hirano, M. (**1997**). "Age-related changes of elastic fibers in the superficial layer of the lamina propria of vocal folds," Ann. Otol. Rhinol. Laryngol. **106**(1), 44–48.

Sato, K., Hirano, M., and Nakashima, T. (**2002**). "Age-related changes of collagenous fibers in the human vocal fold mucosa," Ann. Otol. Rhinol. Laryngol. **111**(1), 15–20.

Sawashima, M., Hirose, H., Hibi, S., Yoshioka, H., Kawase, N., and Yamada, M. (**1981**). "Measurements of the vocal fold length by use of stereoendoscope—A preliminary study," Annu. Bull. Res. Inst. Logoped. Phoniatr. **15**, 9–16.

Stepp, C. E. (**2013**). "Relative fundamental frequency during vocal onset and offset in older speakers with and without Parkinson's disease," J. Acoust. Soc. Am. **133**(3), 1637–1643.

Stepp, C. E., Hillman, R. E., and Heaton, J. T. (**2010a**). "A virtual trajectory model predicts differences in vocal fold kinematics in individuals with vocal hyperfunction," J. Acoust. Soc. Am. **127**(5), 3166–3176.

Stepp, C. E., Hillman, R. E., and Heaton, J. T. (**2010b**). "The impact of vocal hyperfunction on relative fundamental frequency during voicing offset and onset," J. Speech Lang. Hear. Res. **53**(5), 1220–1226.

Stepp, C. E., Merchant, G. R., Heaton, J. T., and Hillman, R. E. (**2011**). "Effects of voice therapy on relative fundamental frequency during voicing offset and onset in patients with vocal hyperfunction," J. Speech Lang. Hear. Res. **54**(5), 1260–1266.

Stepp, C. E., Sawin, D. E., and Eadie, T. L. (**2012**). "The relationship between perception of vocal effort and relative fundamental frequency during voicing offset and onset," J. Speech Lang. Hear. Res. **55**(6), 1887–1896.

Stevens, K. N. (**1977**). "Physics of laryngeal behavior and larynx modes," Phonetica **34**(4), 264–279.

Vojtech, J. M., and Heller Murray, E. (**2019**). *Tutorial for Manual Relative Fundamental Frequency (RFF) Estimation Using Praat* (Boston University, Boston, MA).

Watson, B. C. (**1998**). "Fundamental frequency during phonetically governed devoicing in normal young and aged speakers," J. Acoust. Soc. Am. **103**(6), 3642–3647.

Woo, P. (**2017**). "High-speed imaging of vocal fold vibration onset delay: Normal versus abnormal," J. Voice **31**(3), 307–312.

Xue, S. A., and Deliyski, D. (**2001**). "Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications," Educat. Gerontol. **27**(2), 159–168.

Yoshioka, H., Löfqvist, A., and Collier, R. (**1982**). "Laryngeal adjustments in Dutch voiceless obstruent production," Annu. Bull. Res. Inst. Logped. Phoniatr. **16**, 27–35.

Zarate, J., Ritson, C. R., and Poeppel, D. (**2012**). "Pitch-interval discrimination and musical expertise: Is the semitone a perceptual boundary?," J. Acoust. Soc. Am. **132**(2), 984–993.

Zhang, Z. (**2016**). "Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model," J. Acoust Soc. Am. **139**(4), 1493–1507.