Effects of Age and Parkinson's Disease on the Relationship between Vocal Fold Abductory Kinematics and Relative Fundamental Frequency

#,*,†,‡,§Jennifer M. Vojtech, and *,†,¶Cara E. Stepp, *†¶*Boston, and* ‡§*Natick, Massachusetts*

Summary: Purpose. This study reports on two experiments to examine vocal fold abduction and its relationship with relative fundamental frequency (RFF), considering two attributes that have been shown to elicit group differences in RFF: age (Experiment 1) and Parkinson's disease (PD; Experiment 2).

Methods. For both experiments, simultaneous acoustic and nasendoscopic recordings were collected as participants produced the utterance, /ifi/. RFF values were computed from the acoustic signal, whereas abduction duration and glottic angle at voicing offset were identified from the laryngoscopic images. In Experiment 1, 50 speakers with typical voices (18−83 years) were analyzed to examine (1A) the effects of speaker age on individual outcome measures (RFF, abduction duration, glottic angle) via Pearson's correlation coefficients, and (1B) the effects of abductory measures and age on RFF via an analysis of covariance. In Experiment 2, 20 speakers with PD and 20 matched controls were analyzed to examine (2A) the effects of group (with/without PD) on outcome measures via an analysis of variance, and (2B) the relationship of RFF with abduction duration, glottic angle, and age when considering group via an analysis of covariance.

Results. Age demonstrated a significant, negative relationship with glottic angle (1A) but was not a significant factor when examining the relationship of vocal fold abduction and RFF (1B). Speaker group (with/without PD) demonstrated a significant effect on measures of RFF and abduction duration (2A) but was not a significant factor when examining the relationship of vocal fold abduction and RFF (2B).

Conclusions. RFF is sensitive to changes in vocal fold abductory patterns during devoicing, irrespective of speaker age or PD status.

Key Words: Relative fundamental frequency−Vocal fold abduction Parkinson's disease−Voice−Age.

INTRODUCTION

Background

Laryngeal muscle tension and vocal fold abductory kinematics are thought to play key roles in the regulation of intervocalic offsets. Increased vocal fold tension[1,2] and abduction[3] have been observed during vowels that precede voiceless consonants, suggesting that these mechanisms work in combination to cease vocal fold vibration. Numerous mechanisms have been postulated to describe the physiological underpinnings of increased vocal fold tension during voicing offset, including increased cricothyroid activity,[4] thyroarytenoid activity,[5] and vocal fold stiffness from passive stretching that occurs as the result of changes in laryngeal height.[4,6] On the other hand, vocal fold abduction has been attributed to the posterior cricoarytenoid muscle, which acts as an antagonist to the cricothyroid and adductor muscles (ie, thyroarytenoid, interarytenoid, lateral cricoarytenoid)[7,8-10] to reduce the duration of vocal fold contact[11] and, thus, inhibit vocal fold vibration. Based on these mechanisms, the interplay of laryngeal muscle tension and vocal fold abduction may counteract each other during voicing offset to cease vocal fold vibration, wherein increased tension increases voice fundamental frequency (f_o) —or an increase in the frequency of vocal fold vibration— and vocal fold abduction lowers f_o.[13] These mechanisms may be captured by a non-invasive, objective measure called relative fundamental frequency (RFF).

RFF has been proposed as an acoustic estimate of the degree of baseline laryngeal muscle tension. As excessive and/or imbalanced laryngeal muscle forces have been implicated in a large proportion of voice disorders,[13] RFF may be useful for non-invasively assessing and tracking changes in laryngeal tension in the clinic. Reflecting short-term changes in instantaneous f_o during intervocalic offsets and onsets, RFF is calculated from the f_o values of the 10 voicing cycles immediately preceding and 10 voicing cycles immediately following the voiceless consonant in a vowel −voiceless consonant−vowel (VCV) production (Figure 1). These 20 instantaneous f_o values are then normalized to the steady-state f_o value of the corresponding vowel—as shown in *Eq. 1*—to enable comparisons of changes in f_o in units of

Accepted for publication March 8, 2022.

Disclosure: J.M. Vojtech receives financial compensation through employment by Delsys, Inc., a commercial company that manufactures, markets, and sells Electromyographic Sensors and other Physiological Measurement Systems. C.E. Stepp has received consulting fees from Altec, Inc. and 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.

Funding: This work was supported by the National Science Foundation under Grant No. 1247312 (J.M.V.) and the National Institutes of Health under Grant No. DC015570 (C.E.S.).

From the *Department of Biomedical Engineering, Boston University, Boston, Massachusetts; †Department of Speech, Language, and Hearing Sciences, Boston University, Boston, Massachusetts; ‡Delsys, Inc., Natick, Massachusetts; §Altec, Inc., Natick, Massachusetts; and the ¶Department of Otolaryngology − Head and Neck Surgery, Boston University School of Medicine, Boston, Massachusetts.

#Present Address: Delsys, Inc. and Altec, Inc., Natick, MA, 01760, USA,

Address correspondence and reprint requests to: Jennifer M. Vojtech, Boston University, 677 Beacon St., Boston, MA, 02215. E-mail: jvojtech@delsys.com

Journal of Voice, Vol. 38, No. 5, pp. 1008−1022

0892-1997

https://doi.org/10.1016/j.jvoice.2022.03.007

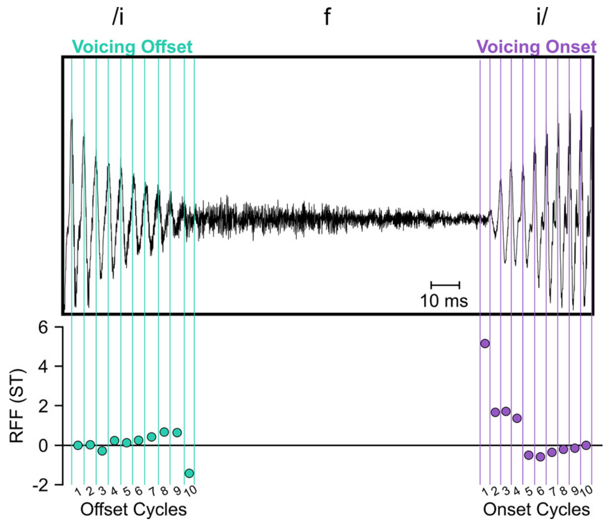

FIGURE 1. Acoustic waveform of the vowel−voiceless consonant−vowel production, /ifi/. Voicing cycles preceding the voiceless consonant, /f/, are marked as voicing offset (teal) whereas those following the /f/ are marked as voicing onset (purple) (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

semitones (ST) within and across speakers. This pattern of the 10 RFF values during voicing offset is thought to reflect the interplay of tension and abduction as these mechanisms act to slow down and eventually cease vocal fold vibration. Interestingly, the RFF values of young adults are characteristically stable or slightly decreasing during voicing offset,[12,14] whereas those of older adults are significantly lower in magnitude.[12,15]

$$RFF\ (ST) = 12 * \log_2\left(\frac{f_o}{f_o^{ref}}\right) \tag{1}$$

The dissimilarity in offset RFF trends between young and older adults suggests a potential difference in the mechanisms used for devoicing. Watson[12] proposed that older adults may not be able to produce transient increases in vocal fold tension to assist in devoicing, possibly as a byproduct of age-related vocal fold atrophy frequently observed in older adults (eg, breathiness, weakness, hoarseness, inability to sustain phonation, and/or atrophy of the vocalis muscle).[16] Under this hypothesis, older adults may rely on a prolonged abductory gesture rather than the combined efforts of tension and abduction. This would imply that vocal fold abduction begins earlier in the preceding vowel for older adults and may include an increased glottic angle at the time of voicing offset (ie, due to the vocal folds opening to cease vibration).[3] In spite of the theoretical backing to support this contribution of vocal fold abduction to devoicing and, in turn, to measures of RFF, no study to date has physiologically examined the role of abductory kinematics in intervocalic offsets.

The theoretical implications of a prolonged abductory gesture in older adults is also interesting to consider when comparing RFF values between adults with Parkinson's disease (PD) and age-matched controls.[14,15] PD is a progressive neurodegenerative disease that typically develops in middle to late life (with incidence rates rising rapidly after 60 years of age)[17] and affects the central and peripheral nervous systems.[18,19] Thought to be a product of neural processes and morphological changes to the muscles, increased muscle rigidity is one of the hallmark motor symptoms exhibited in PD.[20-24] Since the regulation of laryngeal muscle tension is a crucial component of phonation, it is no surprise that excessive intrinsic laryngeal muscle activity has been reported in adults with PD when compared to age-matched controls.[25,26] RFF is thought to reflect this disparity in baseline tension, wherein adults with PD exhibit even lower RFF values compared to controls.[14,15] Although it follows that the comparably lower offset RFF values in speakers with PD reflect increased levels of muscle rigidity that arise with PD, no study has provided evidence to support this notion. As vocal fold abduction is implicated alongside laryngeal muscle tension during voicing offset, it is possible that the RFF values observed in those with PD are the result of an even greater reliance on vocal fold abduction rather than from increased baseline levels of laryngeal muscle activity. Since the contribution of vocal fold abduction to RFF has not been physiologically examined, however, it is thus unclear how tension and abduction play a role in RFF values in older adults with PD.

Purpose of the current study

Although RFF shows promise as an acoustic estimate of the degree of baseline laryngeal muscle tension, the specific contribution of vocal fold abduction to RFF has not been physiologically assessed. The purpose of this study was to therefore characterize the relationship between vocal fold abduction and RFF at intervocalic offsets. The motivation for this work stems from prior speculations of a prolonged abductory gesture in older adults and concurrently increased levels of baseline laryngeal muscle tension in adults with PD. Because RFF is captured using an acoustic signal—which is useful for non-invasive clinical voice assessments, yet only *indirectly* reflects the glottal source—these hypotheses yet to be physiologically confirmed. As such, vocal fold vibratory and abductory kinematics during devoicing cannot be characterized in relation to RFF when only examining the acoustic signal. To carry out this investigation, measures of RFF at voicing offset cycle 10 were examined relative to duration of vocal fold abduction and glottic angle at voicing offset. These three measures were also investigated in the presence of 2 factors that have been previously hypothesized to impact the relationship between vocal fold abduction and RFF: speaker age and a diagnosis of idiopathic PD (ie, a disorder characterized by excessive laryngeal muscle tension).

Following previous work from Watson,[12] we first sought to characterize the relationship between speaker age and vocal fold abduction during intervocalic offsets, as well as the extent to which this relationship contributes to measures of RFF in vocally healthy speakers. It was hypothesized that vocal fold abduction (abduction duration, offset angle) would be significantly, positively related to speaker age, whereas RFF (at voicing offset cycle 10) would be significantly, negatively related to speaker age. Measures of vocal fold abduction were also hypothesized to exhibit a significant, negative relation with RFF at voicing offset (ie, greater abduction durations and larger glottic angles at voicing offset will be associated with lower RFF values at voicing offset cycle 10). These hypotheses were specifically tested in Experiment 1.

We then assessed the effects of idiopathic PD on vocal fold abduction, and how this relationship contributes to RFF at intervocalic offsets. It was hypothesized that measures of vocal fold abduction (abduction duration, offset angle) would not significantly differ between speakers with idiopathic PD and age-/sex-matched controls, but that speakers with idiopathic PD would exhibit statistically significantly lower RFF values at voicing offset. Furthermore, it was hypothesized that both vocal fold abduction and speaker group (control, PD) would be significantly predictive of RFF at voicing offset. These hypotheses were specifically tested in Experiment 2.

EXPERIMENT 1

Method

Participants

Fifty adults with typical voices (25 cisgender females, 25 cisgender males) aged 18−83 years (M = 43.5 years, SD = 21.8 years) were enrolled in the study. The Montreal Cognitive Assessment (MoCA) was administered to participants over the age of 50 to ensure all enrolled participants had the capacity to consent to study tasks via *a priori* cut-off of ≥21;[27] for the 24 participants over 50 years of age, the average MoCA score was 28.3 (SD = 1.3 years, *range* = 26−30). All participants were native English speakers, non-smokers, and had no history of speech, language, hearing, neurological, or voice problems. Participants with trained singing experience beyond grade school were excluded to minimize variability in phonatory behaviors that may occur when differentiating between singers and non-singers.[28] A certified voice-specializing speech-language pathologist screened all participants with typical voices for healthy vocal function via auditory-perceptual assessment and flexible nasendoscopic laryngeal imaging. A hearing screening of pulsed pure tones[29] was administered to each participant at frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz. Of this group, 38 adults passed the hearing screening under 25 dB HL at both ears (protocol based on American Speech-Language-Hearing Association[30]). The remaining 12 adults—all over 50 years of age—passed this hearing screening at frequencies of 125, 250, 500, and 1000 under 25 dB HL and at 2000 and 4000 Hz under 40 dB HL in at least one ear.[31] All participants provided informed, written consent in compliance with the Boston University Institutional Review Board. Detailed demographic information for the control speakers is reported in Table A.1. of *Appendix A*.

Recording procedures

Participants were trained to produce eight iterations of the VCV utterance, /ifi/, with equal stress (same pitch and duration) on both vowels. This VCV utterance was selected to visualize the vocal folds under endoscopy, wherein the voiceless consonant, /f/, has been shown to minimize within-speaker variations in the resulting acoustic signal.[32] Each participant was instructed to produce four /ifi/ utterances, take a breath, then produce the remaining four /ifi/ utterances.

Participants were then seated in a sound-attenuated booth and instrumented with a directional headset microphone (Shure SM35 XLR) placed 45° from the midline and 7 cm from the lips. Microphone signals were pre-amplified (Xenyx Behringer 802 Preamplifier) and digitized at 30 kHz (National Instruments 6312 USB). A flexible routine endoscope (Pentax, Model FNL-10RP3, 3.5-mm) was passed trans nasally over the soft palate and into the hypopharynx for laryngeal visualization using a steady xenon light source (300 W Kay-PENTAX Model 7162B). In the event that participant anatomy or comfort interfered with image acquisition using the routine endoscope, a flexible slim endoscope (Pentax, Model FNL-7RP3, 2.4-mm) was used instead. A nasal decongestant was offered to minimize participant discomfort as the endoscope was passed through the nasal cavity. Laryngeal images were captured at 1 kHz via a camera (FASTCAM Mini AX100l; Model 540K-C-16GB; 40-mm optical lens adapter) attached to the endoscope and were recorded using Photron Fastcam Viewer software (v.3.6.6). A frame rate of 1 kHz was chosen to sufficiently track the fundamental frequency of vocal fold vibration (85−255 Hz in adults[33]) and capture gross abductory gestures (104−227 ms in adults[34]) while permitting a reasonable spatial resolution (256 × 352 pixels with the 3.5-mm endoscope and 256 × 256 pixels with the 2.4-mm endoscope) and an appropriate video duration to successfully record all eight /ifi/ repetitions.

During the endoscopic procedure, participants were cued to produce the eight /ifi/ repetitions. This number of repetitions was selected based on the recording limitations of setup: the high-speed imaging and synchronized microphone recordings were restricted in duration to 7.940 seconds when the 3.5-mm endoscope was used and 8.734 seconds when the 2.4-mm endoscope was used. To account for trials in which productions at the end of the recording were incompletely captured or in cases where less than eight /ifi/ utterances were produced, each condition was repeated a minimum of two times. Additional trials were also recorded in the event that the vocal folds were not sufficiently captured (eg, due to obstruction from the epiglottis) or if the participant produced unequally stressed vowels. The length of the endoscopic procedure lasted approximately 5−10 minutes. A total of 740 productions were captured across the 50 speakers, with an average of 14.8 (*SD* = 3.2) /ifi/ productions captured per speaker.

DATA ANALYSIS

High-speed video processing

Technician training

A series of nine technicians were trained in glottic angle identification at a conventional framerate of 30 frames per second (fps) using methodology described by Diaz-Cadiz, McKenna, Vojtech and Stepp.[35] In brief, the technicians identified glottic angles extending from the anterior commissure along the medial vocal fold edge to the vocal process. These angles were then compared to angle markings made previously by a gold-standard technician. The technicians were required to meet a training standard of .80 via two-way mixed-effects intraclass correlation coefficients (ICC) for consistency of agreement. Using this methodology, the resulting average reliability for the nine technicians was *ICC(3,1)* = 0.89 (*SD* = 0.01, *range* = 0.88−0.91).

From here, the technicians were trained to use a semi-automated glottic angle extraction algorithm to identify glottic angles at 1000 fps. The algorithm was developed in MATLAB (version 9.3; The MathWorks, Natick, MA) to track the glottic angle over time within a VCV production, and is described by Diaz-Cadiz et al.[35] The algorithm carries out an automated glottic angle extraction process to process to identify the glottis, segment vocal fold edges, and estimate the glottic angle over time. The result of this three-step process is a glottic angle waveform for the VCV production, which is shown in a graphical user interface (GUI) alongside the time-aligned video frames and microphone signal (Figure 2). In the event that the technician did not agree with the results of the automated algorithm, the technician would manually mark glottic angles for the production at a down sampled rate of 50 Hz. The automated glottic angle extraction procedure would then run again, this time using the manual glottic angle data as a reference. Within the glottic angle tracking training, technicians were required to meet reliability standards of *ICC(3,1)* ≥ 0.80 compared to a gold-standard technician.[35] The resulting average reliability of the nine technicians was *ICC(3,1)* = 0.85 (*SD* = 0.04, *range* = 0.80−0.91). Following the training, the technicians analyzed the experimental data.

Video usability

Trained technicians manually inspected the video images comprising each /ifi/ production to determine whether the videos effectively captured the vocal folds during the transition into and out of the /f/. In the event that the vocal folds were obstructed (eg, by the epiglottis) or were not visible (eg, due to poor image contrast) during the recording, the production was regarded as "unusable" and removed from further analysis. Usable videos were then processed using a semi-automated glottic angle extraction algorithm[35] that tracks the glottic angle over time. If the vocal folds were not appropriately tracked, the technician could intervene by manually extracting vocal fold angles to inform the algorithm before running again. If errors still persisted following manual intervention, the technician marked the /ifi/ production as unusable.

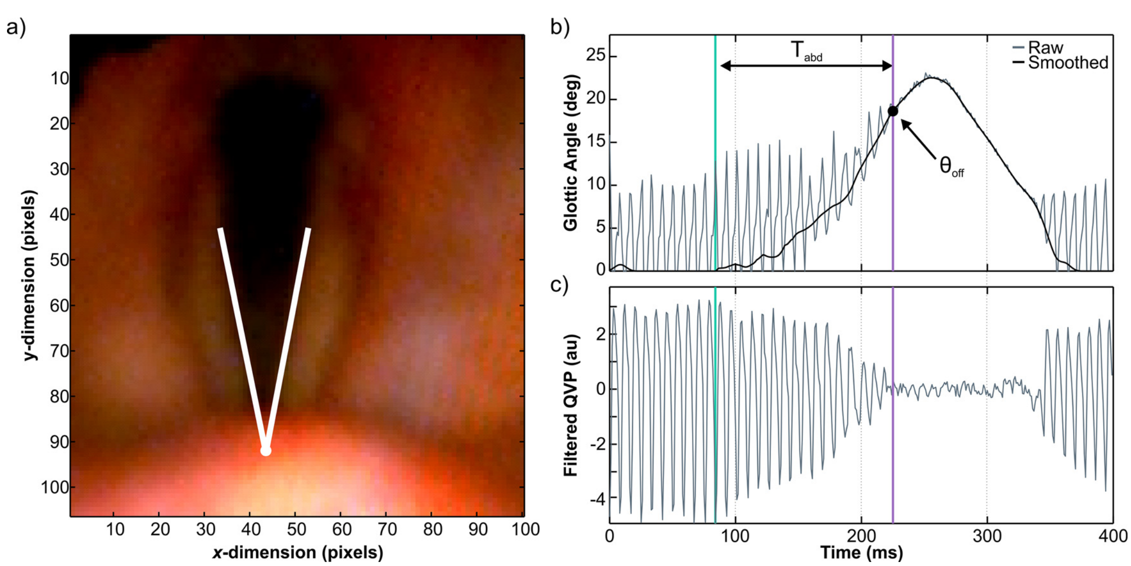

FIGURE 2. (A) View of the vocal folds under flexible nasendoscopy, with the glottic angle marked from the anterior commissure to the vocal processes, and (B) Upper panel: Raw glottic angle waveform (gray) with smoothed data overlay (black), and bottom panel: Filtered quick vibratory profile (QVP). Solid lines indicate the start of vocal fold abduction (green) and time of voicing offset (purple). Abduction duration (T_{abd}) and glottic angle at voicing offset (θ_{off}) are identified (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

The aforementioned technicians used the MATLAB-based algorithm to extract the glottic angle waveform for each /ifi/ production (N = 740). A single technician determined whether the /ifi/ production was usable and, if so, extracted the glottic angle waveform for the production. Of the 740 total /ifi/ productions, technicians accepted the automated algorithmic results in 68.4% of productions (N = 506), and accepted the algorithmic results after performing manual extraction techniques on 16.4% of productions (N = 121). The remaining 113 productions were removed from subsequent analyses due to demonstrated problems in video usability (9.7%; N = 72) or from errors in algorithmic estimation (5.5%; N = 41). Algorithmic reliability was not assessed since prior work indicates that the algorithm yields good reliability ICC ≥ 0.80.[35] The result of this analysis produced 627 usable /ifi/ productions for subsequent processing.

Vocal fold abduction

Vocal fold abduction was measured via abduction duration (milliseconds) and glottic angle at voicing offset (degrees). Technicians were presented with a MATLAB GUI that depicted time-aligned video frames, microphone signal, glottic angle waveform, and quick vibratory profile (QVP). In this analysis, the QVP was included as an alternative to the glottic angle waveform—extracted during the prior video usability processing steps—to assist technicians in discriminating the vibrating glottis in video frames of poor resolution, as well as to identify the window of time that contained the transition between the vowel, /i/, and voiceless consonant, /f/. The QVP was estimated using methodology from Ikuma, Kunduk, and McWhorter[36] to quantify the vibratory motion of the vocal folds via capturing changes in light intensity of the video frame. The video frame was first centered over the glottis; changes in light intensity were then calculated across the vertical and horizontal directions of the video frame.[35] The average of the minimum pixel intensity per row (for vertical profile) and column (for horizontal profile) were then calculated and summed to produce the QVP. From here, the QVP was high-pass filtered using a seventh order Butterworth filter using a cut-off frequency of 50 Hz to attenuate signal noise below a minimum f_o of 50 Hz.

Within the MATLAB GUI, three technicians used the glottic angle waveform and QVP to identify two timepoints: abduction onset and voicing offset. In this analysis, abduction onset was defined as the last full or maximum contact of the vocal folds during voicing offset, whereas voicing offset was measured as the termination of the last vibratory cycle before the voiceless consonant. If the arytenoid cartilages obstructed the view of the vocal folds during devoicing (eg, due to supraglottic constriction), abduction onset was considered as the time in which the arytenoid cartilages began to move away from one another. The technicians were instructed to corroborate the selected indices for abduction onset and voicing offset using the raw video frames in order to minimize errors that may occur if the glottic angle waveform failed to capture small glottal gaps during vibratory cycle phases and/or the QVP was confounded by lighting artifacts (eg, intensity saturation due to the epiglottis coming into view). The microphone signal was included within the GUI in the event that the glottic angle waveform and QVP both failed to properly track the vibrations of the vocal folds; in such cases, the technicians were able to use the GUI to indicate that the production needed to be rejected or reprocessed.

The technicians each reanalyzed 10% of participants in a separate sitting to ensure adequate intrarater reliability. When assessed via two-way mixed-effects ICCs for absolute agreement, intrarater reliability ranged from moderate to excellent (0.70−0.99) with an overall mean reliability of .98 (*95% CI* = 0.97−1.0). The three technicians also analyzed the high-speed video images of the same participant to assess interrater reliability. Interrater reliability was computed using two-way mixed-effects ICCs for consistency of agreement (single measures), producing an average reliability of .92 (*95% CI* = 0.83 −1.0) for abduction onset and voicing offset.

The duration of the abductory gesture (T_{abd}) was then measured by subtracting the timepoint corresponding to abduction onset from that of voicing offset. The abductory gesture was also characterized by extracting the glottic angle at t_{off} from the glottic angle waveform, called θ_{off}.

Acoustic signal processing

Semi-automated RFF estimated was performed on the microphone signals of the 627 usable /ifi/ productions via the aRFF-AP algorithm[37] in MATLAB (version 9.3). The algorithm required a technician to confirm the location of the midpoint of the voiceless consonant, /f/, in each /ifi/ production. Manual intervention was carried out in the event that the algorithm incorrectly located the /f/. RFF values were then automatically calculated from the vocal cycles closest to the /f/. RFF at voicing offset cycle 10 (RFF_{off_10}) was retained to examine the association of RFF with group and vocal fold abductory kinematics, as RFF_{off_10} has been reported to be reduced when speakers without voice disorders increase their vocal effort[38,39] and when speakers have voice disorders characterized by increased muscle tension (eg, vocal hyperfunction, laryngeal dystonia),[40,41] as well as being robust against varying abduction initiation times.[42] Voicing offset instances that were rejected during algorithmic processing (eg, due to voicing during the voiceless consonant, glottalization, or misarticulation) were removed from further analysis (185 /ifi/ productions).

Statistical analyses

The high-speed video and acoustic signal analyses resulted in one value for T_{abd}, θ_{off}, and RFF_{off_10} for each of the 50 speakers. Due to the large amount of rejected data, Pearson's product-moment correlation coefficients were first calculated to detect possible relationships between the number of missing or unusable /ifi/ productions and speaker age. Similarly, an independent t-test was used to examine the relationship between missing or unusable /ifi/ productions and speaker sex. Following, the impact of speaker age on vocal fold abduction was quantified by computing Pearson's product-moment correlation coefficients between speaker age and T_{abd}, θ_{off}, and RFF_{off_10}. The extent to which speaker age and vocal fold abduction contribute to measures of RFF in adults with typical voices was then assessed via an analysis of covariance (ANCOVA) model. To construct this model, speaker age, T_{abd}, and θ_{off} were included as covariates. Since older adult men have been shown to exhibit lower RFF values than older adult women,[16] speaker sex was included as a factor in the current ANCOVA to minimize confounding effects. RFF_{off_10} was implemented as the response variable. Significance was set *a priori* to $P < 0.05$ and partial eta squared (η_p^2) values were calculated as effect sizes.

RESULTS

Although no statistically significant relationship was observed between the number of missing or unusable /ifi/ productions and speaker age ($P = 0.386$), a significant relationship was observed for speaker sex ($P = 0.027$), wherein more /ifi/ productions were missing or unusable for female speakers ($M = 4.6$, $SD = 3.5$) than male speakers ($M = 2.5$, $SD = 2.9$). Figure 3 depicts the trends in RFF_{off_10}, T_{abd}, and θ_{off} across speaker age. Speaker age exhibited a significant, moderate relationship with θ_{off} ($r = -0.31$, $P = 0.027$; Figure 3C), wherein older speakers tended to have smaller glottic angles during intervocalic offsets. Age was not significantly correlated with RFF_{off_10} ($r = -0.09$, $P = 0.538$; Figure 3A) or T_{abd} ($r = -0.14$, $P = 0.346$; Figure 3B).

Table 1 summarizes the results of the ANCOVA examining relationship between age, sex, and measures of vocal fold abduction (ie, T_{abd}, θ_{off}) on RFF_{off_10}. The model accounted for 19.2% of the variance in the data for RFF_{off_10} (adjusted $R^2 = 12.0\%$). Whereas age, sex, and T_{abd} did not exhibit significant effects on RFF_{off_10}, θ_{off} demonstrated a significant, medium effect on RFF ($P = 0.019$, $\eta_p^2 = 0.12$).

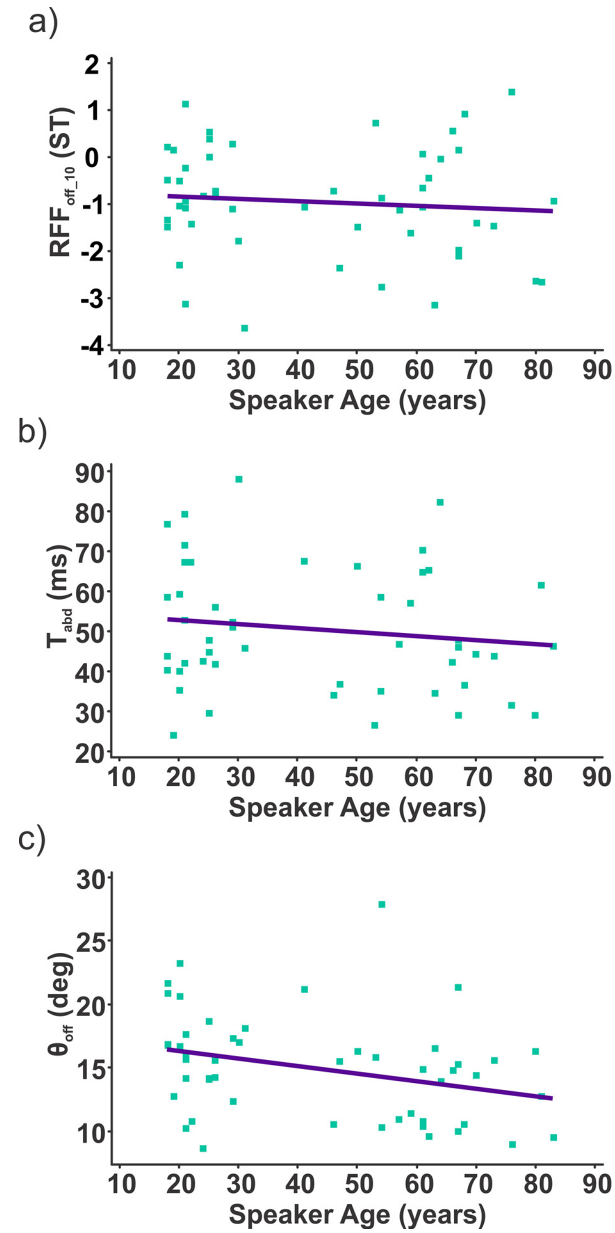

FIGURE 3. Scatter plots of speaker values (teal) for (A) relative fundamental frequency (RFF) at offset cycle 10, (B) abduction duration, and (C) glottic angle at voicing offset relative to speaker age. Lines of best fit are shown in purple (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

EXPERIMENT 2

Method

Participants

Twenty adults with idiopathic PD (6 cisgender females, 14 cisgender males) aged 50−75 years ($M = 65.0$ years, $SD = 7.4$ years) and 20 age- and sex-matched controls (six cisgender females, 14 cisgender males) aged 47−81 years ($M = 65.1$ years, $SD = 8.8$ years) were enrolled in the study. This sex distribution is consistent with the higher incidence of PD in men compared to women.[17,43] As in Experiment 1, all participants were native English speakers, non-smokers, and did not report trained singing experience beyond grade school. Of note, 14 of 20 of the control speakers constituted a subset of the participant sample examined in Experiment 1. The Montreal Cognitive Assessment (MoCA) was administered to each of the 40 participants to determine cognitive status and an *a priori* cut-off of ≥21 was set to ensure all included participants had the capacity to consent to the study tasks.[27] All participants provided informed, written consent in compliance with the Boston University Institutional Review Board.

Characteristics of speakers with PD

Within the current study, all adults with PD were diagnosed with idiopathic PD by a neurologist and were recorded

TABLE 1.
Results of the Analysis of Covariance Model Examining the Effects of Speaker Age, Speaker Sex, Abduction Duration, and Glottic Angle at Voicing Offset on RFF Offset Cycle 10

Effect	*df*	*F*	*P*	η_p^2	Effect Size Interpretation
Speaker Age	1	2.38	0.130	–	–
Speaker Sex	1	3.01	0.090	–	–
Abduction Duration (T_{abd})	1	0.02	0.895	–	–
Glottic Angle at Voicing Offset (θ_{off})	1	5.89	0.019	0.12	Medium

Note. Effect size interpretations based on criteria from Witte and Witte.
[1] Dashes (–) indicate non-significant findings ($P > 0.05$).

while on their usual carbidopa/levodopa medication schedule to preserve typical vocal function. Individuals who used deep brain stimulation devices (N = 4) were requested to turn their device off for the duration of data collection to minimize the potential impacts of deep brain stimulation on laryngeal function.

A speech-language pathologist specializing in voice disorders assessed the overall severity of dysphonia of each participant using the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V).[44] The average overall severity of dysphonia score for speakers with PD was 16.8 (SD = 10.6, *range* = 4.0−33.5). Additionally, the Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) was administered to each participant with PD to determine the extent of both motor and non-motor complications; each examination was administered and scored per protocol by a certified MDS-UPDRS administrator. The severity of motor complications was, on average, moderate (M = 43.7, SD = 17.7) and ranged from mild to severe (range = 13−91).[45] The average Hoehn-Yahr score was 2.0 (SD = 1.1) and ranged from 0 (no disability) to 4 (severe disability).[46,47] Table 2 shows demographic information for the 20 individuals with PD.

Characteristics of control speakers

Age- and sex-matched control speakers reported no history of speech, language, hearing, neurological, or voice problems. All control speakers were screened by a certified voice-specializing speech-language pathologist for healthy

TABLE 2.
Demographic Information of Participants with Parkinson's Disease (PD)

Participant	Sex	Age	CAPE-V OS	PD Characteristics		
				Years Post-Dx	*MDS-UPDRS-III*	*Hoehn-Yahr Scale*
PD1	M	60	30.1	7	54	2
PD2	F	62	5.6	9	49	3
PD3	F	70	14.7	6	77	4
PD4	M	50	7.1	0	17	0
PD5	M	55	18.4	21	49	3
PD6	M	62	10.0	3	50	2
PD7	F	74	30.6	24	59	2
PD8	M	73	33.6	9	19	1
PD9	M	67	6.4	4	63	3
PD10	M	67	19.4	2	38	2
PD11	M	62	27.9	13	47	2
PD12	M	59	4.0	2	23	2
PD13	M	73	6.8	3	23	1
PD14	M	68	5.0	6	38	2
PD15	F	73	33.3	8	52	2
PD16	M	75	22.1	1.5	68	2
PD17	F	65	15.4	10	48	3
PD18	M	68	8.5	1	52	3
PD19	M	65	28.3	1	35	0
PD20	F	51	9.7	5	13	0

Note. Dx, Diagnosis; CAPE-V OS, Consensus of Auditory-Perceptual Evaluation of Voice; Overall Severity of Dysphonia; MDS-UPDRS-III, Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale: Part III, Motor Examination.

vocal function via auditory-perceptual assessment (CAPE-V) and flexible nasendoscopic laryngeal imaging. The average score for overall severity of dysphonia of the control group was 10.6 (*SD* = 7.9, *range* = 1.7−34.2). Demographic information for the control speakers is reported in Table A.2. of *Appendix A*.

Hearing status

A hearing screening of pulsed pure tones[29] was administered to each participant at frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz (protocol based on American Speech-Language-Hearing Association[30]). Nineteen control speakers and 14 speakers with PD passed the hearing screening under at frequencies of 125, 250, 500, and 1000 under 25 dB HL and at 2000 and 4000 Hz under 40 dB HL in at least one ear.[31] The control speaker that did not meet these criteria exhibited a threshold of 40 dB HL at 125 and 250 Hz, as well as a threshold of 30 dB HL at 500 Hz. Of the six speakers with PD that did not meet these criteria, one speaker (PD14 in Table 2) demonstrated a threshold of 45 dB HL at 2000 Hz, two speakers (PD16, PD19) exhibited a threshold of 45 dB HL at 4000 Hz, two speakers (PD11, PD15) demonstrated a threshold of 50 dB HL at 4000 Hz, and one speaker (PD10) wore hearing aids during the course of the study and demonstrated thresholds of 45 dB HL at 125 Hz and 30 dB HL at 1000 Hz.

Recording procedures & data analysis

All data analyzed in the current experiment were collected during a nasendoscopic examination as described in Experiment 1. In total, there were 104 instances in which 16 full /ifi/ productions were not captured for a speaker; this process resulted in total of 536 /ifi/ productions ([40 participants × 16 /ifi/ productions] − 104 incomplete or missing /ifi/ productions) across the two speaker groups. In using the semi-automated MATLAB algorithm from Experiment 1[36] to extract the glottic angle waveform for each /ifi/ production, technicians accepted the automated algorithmic results in 61.2% of productions (*N* = 328) and accepted the algorithmic results after performing manual extraction techniques on 22.2% of productions (*N* = 119). Of the remaining productions, 10.4% were considered unusable (*N* = 56) and a further 6.2% were rejected due to errors in algorithmic estimation (*N* = 33). Finally, 124 voicing offset instances that were rejected during algorithmic processing of the microphone signal (eg, due to voicing during the voiceless consonant, glottalization, or misarticulation) were removed from further analysis. This process resulted in a total of 323 /ifi/ productions.

Using methodology described in Experiment 1, the analyses performed on the high-speed video images and microphone signals resulted in the following measures for each /i/-to-/f/ transition of the 323 /ifi/ productions: (1) T_{abd}, (2) θ_{off}, and (3) RFF_{off_10}. The resulting measures were then evaluated with a series of statistical models to determine the relationship between speaker group, RFF values, and vocal fold abductory kinematics. First, Pearson's product-moment correlation coefficients were calculated to detect possible relationships between the number of missing or unusable /ifi/ productions and speaker age. Independent *t*-tests were used to examine the relationship between missing or unusable /ifi/ productions and binary speaker variables of sex and group. Following, three separate analysis of variance (ANOVA) models were constructed to examine the effect of group (speakers with PD versus age- and sex-matched controls) on voicing offset measures of T_{abd}, θ_{off}, and RFF_{off_10}. In these models, each voicing offset measure was set as the response variable, group was set as a fixed factor, significance was set *a priori* to $P < 0.05$, and partial eta squared (η_p^2) values were calculated as effect sizes. An ANCOVA model was then constructed to examine the effects of covariates of speaker age, $T_{abd,}$ and θ_{off}, as well as the fixed factors of speaker group, sex, and relevant interactions of group × sex, group × T_{abd}, and group × θ_{off} on RFF_{off_10} (response variable). Significance was set *a priori* to $P < 0.05$ and η_p^2 values were calculated as effect sizes.

RESULTS

There were no statistically significant relationships between the number of missing or unusable /ifi/ productions and speaker variables of age ($P = 0.290$), sex ($P = 0.835$) or group ($P = 0.221$). Table 3 shows the results for the models examining the effects of group on each of the three voicing offset measures. Group was not a significant factor in the model for θ_{off} ($p = 0.476$, $\eta_p^2 = 0.01$), but was a significant factor in the models for both T_{abd} ($P = 0.034$, $\eta_p^2 = 0.11$) and RFF_{off_10} ($P = 0.021$, $\eta_p^2 = 0.13$).

TABLE 3.
Results of the Analysis of Variance Models Examining the Effects of Speaker Group on RFF at Offset Cycle 10, Abduction Duration, and Glottic Angle at Voicing Offset

Model	Effect	*df*	*F*	*P*	η_p^2	Effect Size Interpretation
RFF at Voicing Offset Cycle 10 (RFF_{off_10})	Group	1	5.80	0.021	0.13	Medium
Abduction Duration (T_{abd})	Group	1	4.84	0.034	0.11	Medium
Glottic Angle at Voicing Offset (θ_{off})	Group	1	0.52	0.476	–	–

Note. Effect size interpretations based on criteria from Witte and Witte.[48]
[1] Dashes (–) indicate non-significant findings ($P > 0.05$).

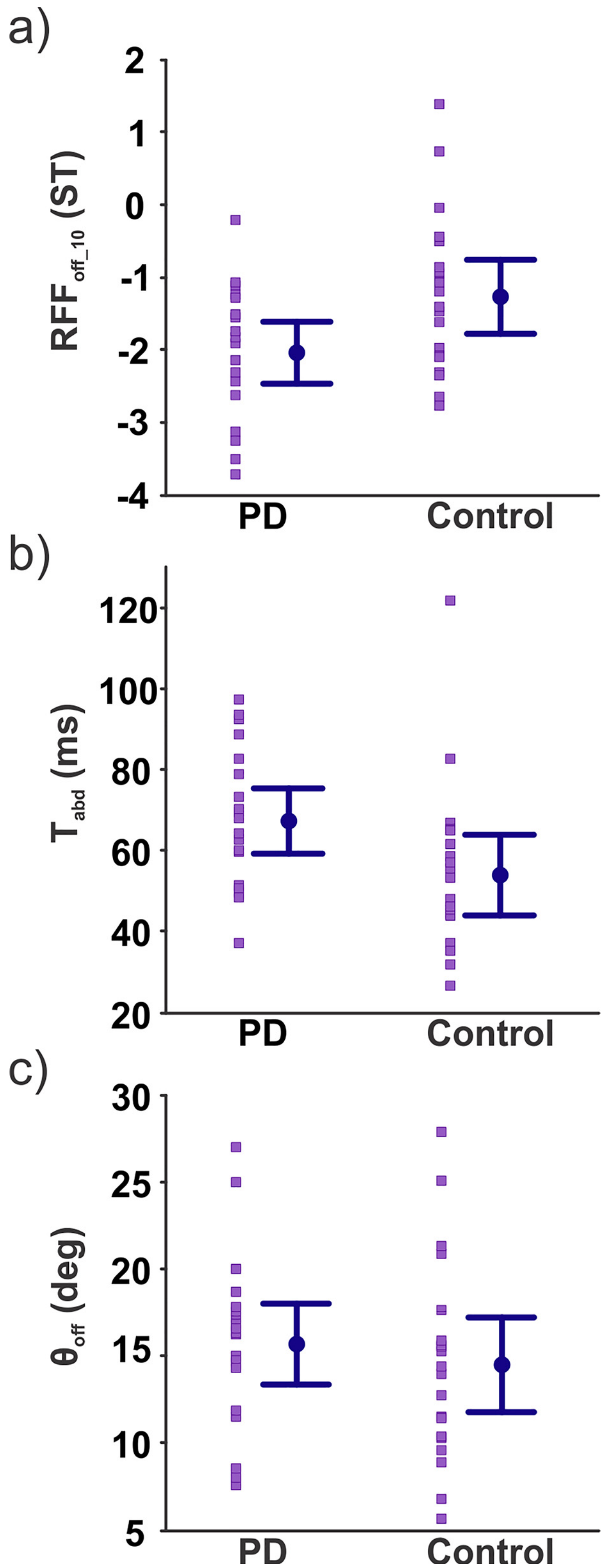

FIGURE 4. Individual speaker (purple squares) and mean (blue circles) values for (A) RFF at offset cycle 10, (B) glottic angle at voicing offset, and (C) abduction duration based on speaker group (speakers with PD: left; age- and sex-matched control speakers: right). Error bars show 95% confidence intervals (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Figure 4 shows individual values for these three variables based on speaker group. Speakers with PD exhibited larger, but less variable, values of T_{abd} (M = 67.4 ms, SD = 17.3 ms) compared to age- and sex-matched controls (M = 53.9 ms, SD = 21.2 ms). RFF values at voicing offset cycle 10 were substantially more negative, but also less variable for speakers with PD (M = -2.04 ST, SD = 0.92 ST) than controls (M = -1.27 ST, SD = 1.09 ST). Consistent with the lack of a group effect for θ_{off}, values for speakers with PD (M = 15.7 degrees, SD = 5.0 degrees were similar to those from age- and sex-matched controls (M = 14.5 degrees, SD = 5.8 degrees).

Table 4 summarizes the ANCOVA model examining the effects of θ_{off}, T_{abd}, age, group, sex, and relevant group interactions (group × sex, group × T_{abd}, group × θ_{off}) on RFF_{off_10}. The model accounted for 43.4% of the variance in the data for RFF_{off_10} (adjusted R^2 = 28.8%), with only θ_{off} (P = 0.003, η_p^2 = 0.25) and T_{abd} (P = 0.039, η_p^2 = 0.13) demonstrating significant effects on RFF.

DISCUSSION

The goal of this work was to investigate the relationship between speaker age, vocal fold abductory kinematics, and RFF. Two experiments were carried out to assess these relationships. The first experiment examined the effects of speaker age on both vocal fold abduction and RFF in individuals with typical voices. Simultaneous acoustic recordings and laryngeal images were captured via a microphone and flexible nasendoscope, respectively, as speakers with typical voices produced the VCV utterance, /ifi/. RFF was extracted from the acoustic signal, whereas abduction duration and glottic angle at voicing offset were computed from the laryngeal images to characterize vocal fold abduction. Vocal fold abduction measures and RFF at voicing offset cycle 10 were then analyzed with respect to speaker age within a cohort of speakers with typical voices. The aim of the second experiment was to examine how PD—which is characterized by excessive laryngeal muscle tension—affects vocal fold abduction and to what extent this relationship affects RFF. As such, abduction duration, glottic angle at voicing offset, and RFF at voicing offset cycle 10 were calculated as in the first experiment but were examined with respect to speaker group (speakers with PD versus age- and sex-matched control speakers).

Comparison of outcome metrics to the literature

In the current study, glottic angle at voicing offset was measured relative to the physiological cessation of vocal fold vibration. The mean glottic angle at voicing offset for the 50 speakers with typical voices in Experiment 1 was 14.9 degrees, and ranged from 8.6 to 27.8 degrees. In Experiment 2, mean glottic angle was 14.5 degrees (*range* = 5.6–27.8 degrees) for speakers with typical voices and 15.7 degrees (*range* = 7.6–27.0 degrees) for speakers with PD. Although numerous works have examined glottic angles during voicing offset, most values reported in the literature focus on

TABLE 4.
Results of the Analysis of Covariance Model Examining the Effects of Speaker Age, Abduction Duration, Glottic Angle at Voicing Offset, Group, and Relevant Interactions on RFF Offset Cycle 10

Effect	*df*	*F*	*P*	η_p^2	Effect Size Interpretation
Abduction Duration (T_{abd})	1	4.65	0.039	0.13	Medium
Glottic Angle at Voicing Offset (θ_{off})	1	10.25	0.003	0.25	Large
Speaker Age	1	0.682	0.415	–	–
Speaker Group	1	0.669	0.420	–	–
Speaker Sex	1	2.47	0.126	–	–
Speaker Group × Speaker Sex	1	0.12	0.733	–	–
Speaker Group × T_{abd}	1	0.14	0.714	–	–
Speaker Group × θ_{off}	1	1.96	0.172	–	–

Note. Effect size interpretations based on criteria from Witte and Witte.[48]
[1] Dashes (–) indicate non-significant findings ($P > 0.05$).

maximum abduction angles during non-speech tasks (eg, sniff or cough)[34,49,50] and are thus not comparable to those of the current study. The values obtained in this study are similar to those reported for speakers without voice disorders, which have been shown for speakers without voice disorders to range between 6 and 24 degrees during /ifi/ productions and between 5 and 18 degrees during /iti/ productions.[51]

As with glottic angle, abduction duration was measured relative to the termination of vocal fold vibration in the current work. Mean abduction duration in Experiment 1 was 50.6 ms for 50 speakers with typical voices. The speakers with typical voices in Experiment 2 exhibited a mean abduction duration of 53.9 ms, whereas speakers with PD demonstrated a mean duration of 67.4 ms. These values are greater than those reported in the literature. In determining the relationship between the acoustic signal and phonatory oscillatory events, Patel, Forrest and Hedges[52] reported a mean abduction duration of approximately 40 ms between the time of vocal fold cessation and the first incomplete vocal fold closure. However, the authors examined vocal behaviors during repetitions of the syllable /hi/ rather than /ifi/, as examined here. Differences in mean abduction duration may be a result of dissimilarities in the voicing mechanisms needed to produce a CV sequence (as in /hi/) compared to a VCV sequence (as in /ifi/), as well as possible mechanistic differences for producing /h/ compared to /f/. In addition, the authors used rigid laryngoscopic techniques to visualize the vocal folds, whereas the current study implemented flexible laryngoscopy. Since rigid laryngoscopy requires the tongue to be restricted during the endoscopic examination, it is possible that differences in laryngoscope type may have also led to divergent vocal behaviors—and, in turn, abduction durations—when producing the speech stimuli.

Measures of vocal fold abduction duration have also been obtained using electroglottography. For instance, Watson, Roark and Baken[53] simultaneously acquired acoustic and electroglottographic signals during sustained /ɑ/ productions. The authors estimated the physiological time of voicing offset via tracking the contact of the vocal folds through the electroglottographic recordings. Physiological times of voicing offset were compared to those obtained via the acoustic signal the authors used electroglottography to track the contact of the vocal folds. The average duration of the abductory gesture for 112 speakers without voice disorders (57 female, 55 male) was identified as 20.0 ms from the time of acoustic voicing offset. However, Watson, Roark and Baken[53] examined abduction duration in a linguistically unconstrained context as participants produced sustained /ɑ/ vowels as well as the words "hallways" and "always." The devoicing gesture does not require that the vocal mechanism be optimized in linguistically unconstrained contexts to meet the linguistic goals necessary to devoice and then rapidly reinitiate voice (as is required to produce the intervocalic /f/ of /ifi/ examined in the current study). Linguistically unconstrained voicing offsets are instead thought to facilitate the transition into tidal breathing.[53] It is therefore possible that the voicing mechanisms required to execute the devoicing gesture to produce /ifi/ require different abductory durations than that of a sustained /ɑ/ or the words "hallways" and "always."

In addition to differences in the stimuli acquired to assess voicing offset, the definition of abduction duration inherently differed between studies. Whereas the current study measured the start of abduction as the last maximum contact of the vocal folds as observed in the laryngoscopic images, Watson, Roark and Baken[53] identified this time from the cross-correlation of the amplitude of an electroglottograph signal that was band-pass filtered ±40% of the speaker's f_o. The current study considered the termination of the abduction gesture to be the cessation of vocal fold vibration, as determined from the laryngoscopic images. Watson, Roark and Baken[53] computed this time point using acoustic voicing offset; as it is not uncommon for voicing offset to occur earlier in the acoustic signal than the true cessation of vocal fold vibrations,[52] it is unsurprising that the absolute durations of the abductory gesture differ between studies.

The values obtained for RFF at voicing offset cycle 10 are comparable to those observed in the literature. In Experiment 1, the average RFF value at offset cycle 10 was -0.95 ST across the 50 speakers with typical voices. The speakers with typical voices in Experiment 2 exhibited a mean RFF of -1.23 ST, whereas speakers with PD demonstrated an average RFF of -2.04 ST. The mean values reported for the typical speakers are well within range of those reported in the literature for offset cycle 10.[12,14,54-56] Likewise, the observed mean RFF values at voicing offset cycle 10 in speakers with PD are similar to those reported in the literature: Goberman and Blomgren[14] reported an RFF value of -2.20 ST at offset cycle 10 for speakers with PD while on medication and Stepp[15] saw a mean offset RFF value of -1.90 ST for speakers with PD while on medication.

Effects of speaker age on vocal fold abduction and RFF

Prior work examining RFF proposed a relationship between vocal fold abduction and RFF, wherein abductory mechanisms during intervocalic offsets differ between young and older adults.[12] In particular, it has been hypothesized that older adults primarily rely on a prolonged vocal fold abductory gesture to devoice rather than laryngeal muscle tension as in young adults. Yet few studies have physiologically examined the supposed relationship between vocal fold abduction and RFF, or the impacts of speaker age on this relationship. Based on this prior work, it was hypothesized that vocal fold abduction (via metrics of abduction duration and glottic angle at voicing offset) would be positively related to speaker age whereas RFF at voicing offset cycle 10 would be negatively related to speaker age.

Although there were no significant correlations between age and either abduction duration or RFF, a significant negative relationship was observed between glottic angle and age. These results lend support to the notion that vocal fold abduction during intervocalic offsets differs with age, though not in the expected direction: larger glottic angles were expected with increasing speaker age due to the hypothesized “prolonged abductory gesture” in older adults, but current findings demonstrate the opposite. These results are not in line with speculations by Watson[12] that older speakers rely more than young speakers on vocal fold abduction (via longer abduction durations and/or glottic angles) to achieve devoicing. However, it must be considered that prior work such as that of Watson[12] examined voicing offsets relative to speaker age as a categorical factor comprising broad age groupings of “young” versus “older” adults. Speaker age was introduced in this study as a covariate to provide more informative interpretations of the underpinnings of voicing offset across the continuous spectrum of age. In particular, these results suggest that speakers may not be able to perform the quick changes in glottic angle necessary to assist in devoicing with increasing age, perhaps due to the effects of age-related vocal fold atrophy that present as vocal fold bowing, spindle-shaped glottal gap, prominent vocal processes, and thinning of the vocal fold mucosa.[16,57-62] Additional work is therefore needed to comprehensively assess the effects of vocal fold atrophy on vocal fold abductory kinematics.

Effects of Parkinson’s disease on vocal fold abduction and RFF

It was hypothesized that speakers with PD would exhibit significantly lower RFF values at voicing offset in the absence of significant differences in vocal fold abduction when compared to age- and sex-matched control speakers. These hypotheses were based on the findings of prior work demonstrating increased levels of laryngeal muscle activity in PD.[25,26] As hypothesized, speakers with PD exhibited lower average RFF values for voicing offset cycle 10 (M = -2.04 ST) compared to age- and sex-matched control speakers (M = -1.27 ST). However, speakers with PD also demonstrated significantly longer abduction durations (M = 67.4 ms) on average when compared to controls (M = 53.9 ms).

Although these findings do not support the hypothesis that older speakers with typical voices leverage a prolonged abductory gesture for devoicing, it is possible that speaker with PD require such a gesture to effectively terminate vocal fold vibration during intervocalic voicing offsets. One mechanism that could account for the observed trends is a prolonged abductory gesture to counteract increased baseline laryngeal muscle tension observed in PD.[25,26] It may not be possible for speakers with PD to further increase laryngeal activity from baseline to control the rapid changes in vocal fold vibration at voicing offset; instead, a prolonged vocal fold abductory gesture may be necessary to cease vocal fold vibration. An alternative mechanism that may contribute to longer abduction durations and lower RFF values in speakers with PD may be an overall slowness of movement in voice mechanisms. Specifically, bradykinesia (slower movements) and hypokinesia (slower, smaller movements) have been observed in PD, with many studies noting smaller and/or slower speech articulatory movements and impairments in the monitoring of speech movement timing at the segmental level.[63-69] A possible explanation for the observed prolonged abductory gesture may follow these observations in the articulatory domain, wherein voluntary movements of the voicing mechanism are slower (and for some speakers, smaller) in PD during linguistically constrained voicing offsets. Yet additional work is needed to address these speculations and, moreover, elucidate the mechanisms contributed to the observed prolonged abductory gesture in speakers with PD.

It was also hypothesized that both vocal fold abduction and speaker group (control, PD) would be significantly predictive of RFF at voicing offset. Abduction duration and glottic angle at voicing offset demonstrated medium- and large-sized significant effects, respectively, indicating that RFF is related to vocal fold abductory patterns during devoicing and—moreover—that changes in RFF during voicing offset may be captured in part via changes in abduction duration and glottic angle. These results support findings from Serry, Stepp and Peterson[42] in which computational models were used to demonstrate that RFF is sensitive to the duration of the abduction period. The authors suggested that this sensitivity, in combination with a dependency on abduction initiation time, leads to variations in RFF values observed clinically. Interestingly, neither speaker group nor its interactions (group × sex, group × T_{abd}, and group × θ_{off}) were significant predictors in the model for RFF. Taken together with findings from Serry, Stepp and Peterson,[42] our results suggest that—although vocal fold abductory kinematics may differ between some speakers with and without PD—the observed differences in RFF values at voicing offset may not be attributed to group differences in vocal fold abduction. These differences may instead be related to increased levels of baseline laryngeal muscle tension that have been observed in speakers with PD; however, it is also possible that another mechanism is at play (eg, laryngeal height). Thus, although these findings support a relationship between vocal fold abduction and RFF, additional work is needed to comprehensively investigate how laryngeal muscle tension and vocal fold abduction mechanisms differentially contribute to measures of RFF at intervocalic offsets.

Limitations and future directions

The experiments conducted in this study sought to determine the contribution of vocal fold abduction to RFF during intervocalic voicing offsets in speakers with typical voices and speakers with idiopathic PD. The identified relationships between vocal fold abduction and RFF may not be generalizable to speakers outside of these groups, particularly as speakers with idiopathic PD are only a subset of individuals with disorders that have has been associated with increased levels of intrinsic laryngeal muscle tension.[25,26] For instance, speakers with vocal hyperfunction may exhibit excessive or imbalanced laryngeal muscle forces.[70] However, the manifestation of vocal hyperfunction is broad, wherein hyper functional vocal behaviors may occur in the presence or absence of organic pathology (eg, vocal nodules), and may be the primary cause of a voice disorder or as a compensatory adaptation to glottal insufficiency. It is therefore unclear whether speakers who exhibit signs of vocal hyperfunction would demonstrate similar trends in vocal fold abductory kinematics as speakers with PD. Although this work attempts to provide insight about the relationship between vocal fold abduction and RFF, future work should aim to expand upon the patterns described here.

Within this vein, the current study was limited to examining the duration and magnitude of speaker abductory gestures during linguistically constrained voicing offsets. Future work should leverage a more comprehensive set of metrics to describe vocal fold abductory kinematics. For instance, timings corresponding to the start of the abductory gesture or the rate of change of the glottic angle could be examined to detect possible deficits in voice initiation or freezing associated with PD or voice disorders such as laryngeal dystonia. Such an investigation should examine additional phonetic contexts (eg, /iti/) to elucidate inter- and intra-group differences that may not be fully characterized when examining the VCV utterance, /ifi/.

A methodological limitation of this study is in using semi-automated algorithms to examine voicing offsets. In particular, two algorithms were used to extract metrics of vocal fold abductory kinematics from high-speed video images and estimates of RFF from the microphone signal; these algorithms were used to process data at a faster rate than would be possible manually. As a result, however, a large portion of data was rejected due to the conservative decision-making criteria of the algorithms (40.3% in Experiment 1, 39.7% in Experiment 2). Despite this loss of data, subsequent analyses were conducted on the full dataset of 50 speakers in Experiment 1 and 40 speakers in Experiment 2, wherein each speaker had an average of 8.8 and 8.8 usable /ifi/ productions for statistical analysis, respectively. This is notably larger than the recommended number of VCV productions for reliable RFF analysis.[42] Nonetheless, a significant relationship was observed between the number of missing or unusable /ifi/ productions and speaker sex within the first experiment. Although it is outside the scope of this study, future work should be carried out to identify and mitigate factors that may have led to more female /ifi/ productions being rejected during processing and/or less productions being collected during high-speed videoendoscopic recording (eg, discomfort during scoping).

Finally, further investigation is needed to characterize the relationship between vocal fold abduction and RFF in the presence of excessive and/or imbalanced laryngeal muscle forces, as tension was not precisely quantified in this work (eg, via laryngeal electromyography). Characterizing abductory kinematics in speakers with other voice disorders associated with excessive tension may also be a useful step toward isolating the differential contributions of tension and abduction in intervocalic voicing offsets.

CONCLUSION

The current study sought to examine the relationship between vocal fold abduction and RFF during intervocalic offsets. The results of this work support vocal fold abductory patterns as a potential mechanism of RFF, wherein changes in RFF during voicing offset may be captured in part via glottic angle. Despite previous speculations of a prolonged abductory gesture in older adults, however, vocal fold abductory patterns did not significantly affect RFF when considering speaker age. Additionally, vocal fold abductory patterns did not significantly affect RFF when examining speakers with and without idiopathic Parkinson's disease, a neurodegenerative disease for which excessive intrinsic laryngeal muscle forces have been observed. These findings suggest that differences in RFF values at intervocalic offsets are likely not due to differences in vocal fold abduction across groups but may instead be related to increased levels of baseline laryngeal muscle tension observed in PD or some other mechanism (eg, laryngeal height). Future work is needed to comprehensively investigate how laryngeal muscle tension and vocal fold abduction mechanisms differentially contribute to measures of RFF at intervocalic offsets.

Acknowledgments

The authors would like to acknowledge Victoria McKenna, Daniel Buckley, Dante Cilento, Austin Luong, Jacob Noordzij, Jr., Yeonggwang Park, and Manuel Díaz-Cadíz for their technical assistance with data acquisition and analysis.

REFERENCES

1. Boone DR, McFarlane SC, Von Berg SL, et al. The Voice and Voice Therapy. *Allyn & Bacon Communication Sciences and Disorders*. Pearson; 2014.
2. Löfqvist A, Baer T, McGarr NS, et al. The cricothyroid muscle in voicing control. *J Acoust Soc Am*. 1989;85:1314–1321.
3. Fukui N, Hirose H. Annual Report of the Institute of Phonetics University of Copenhagen. *Ann Bulletin Res Inst Logoped Phoniatr, University Copenhag*. 1983;17:61–71.
4. Stevens KN. Physics of laryngeal behavior and larynx modes. *Phonetica*. 1977;34:264–279.
5. Hirano M. Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatrica et Logopaedica*. 1974;26:89–94.
6. Sonninen A, Hurme P, Laukkanen AM. The external frame function in the control of pitch, register, and singing mode: Radiographic observations of a female singer. *J Voice*. 1999;13:319–340.
7. Choi HS, Berke GS, Ye M, et al. Function of the posterior cricoarytenoid muscle in phonation: In vivo laryngeal model. *Otolaryngol—Head Neck Surg*. 1993;109:1043–1051.
8. Faaborg-Andersen K. Electromyographic investigation of intrinsic laryngeal muscles in humans. *Acta Physiol Scand*. 1957;41:1–150.
9. Fujita M, Ludlow CL, Woodson GE, et al. A new surface electrode for recording from the posterior cricoarytenoid muscle. *The Laryngoscope*. 1989;99(3):316–320.
10. Hirano M. Vocal mechanisms in singing: Laryngological and phoniatric aspects. *J Voice*. 1988;2:51–69.
11. Rothenberg M, Mahshie JJ. Monitoring vocal fold abduction through vocal fold contact area. *J Speech Lang Hear Res*. 1988;31(3):338–351.
12. Watson BC. Fundamental frequency during phonetically governed devoicing in normal young and aged speakers. *J Acoust Soc America*. 1998;103:3642–3647.
13. Ramig LO, Verdolini K. Treatment efficacy: Voice disorders. *J Speech Lang Hear Res*. Feb 1998;41:S101–S116.
14. Goberman AE, Blomgren M. Fundamental frequency change during offset and onset of voicing in individuals with Parkinson disease. *J Voice*. 2008;22:178–191.
15. Stepp CE. Relative fundamental frequency during vocal onset and offset in older speakers with and without Parkinson's disease. *J Acoust Soc. Am*. 2013;133:1637–1643.
16. Takano S, Kimura M, Nito T, et al. Clinical analysis of presbylarynx —vocal fold atrophy in elderly individuals. *Auris Nasus Larynx*. 2010;37:461–464.
17. Van Den Eeden SK, Tanner CM, Bernstein AL, et al. Incidence of Parkinson's disease: Variation by age, gender, and race/ethnicity. *Am J Epidemiol*. 2003;157:1015–1022.
18. Schapira AHV, Chaudhuri KR, Jenner P. Non-motor features of Parkinson disease. *Nature Reviews Neuroscience*. Jul 2017;18:435–450.
19. Braak H, Del Tredici K, Rub U, et al. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol Aging*. 2003;24:197–211.
20. Dietz V, Quintern J, Berger W. Electrophysiological studies of gait in spasticity and rigidity: Evidence that altered mechanical properties of muscle contribute to hypertonia. *Brain*. 1981;104:431–449.
21. Edstrom L. Histochemical changes in upper motor lesions, Parkinsonism and disuse: Differential effect on white and red muscle fibres. *Experientia*. 1968;24:916–917.
22. Mu L, Sobotka S, Chen J, et al. Altered pharyngeal muscles in Parkinson disease. *J Neuropathol Exp Neurol*. 2012;71:520–530.
23. Rossi B, Siciliano G, Carboncini MC, et al. Muscle modifications in Parkinson's disease: Myoelectric manifestations. *Electroencephalogr Clin Neurophysiol*. 1996;101:211–218.
24. Watts RL, Wiegner AW, Young RR. Elastic properties of muscles measured at the elbow in man: II. Patients with Parkinsonian rigidity. *J Neurol Neurosurg Psychiatry*. 1986;49:1177–1181.
25. Zarzur AP, Duprat AC, Cataldo BO, et al. Laryngeal electromyography as a diagnostic tool for Parkinson's disease. *The Laryngoscope*. 2013;124:725–729.
26. Zarzur AP, Duprat AC, Shinzato G, et al. Laryngeal electromyography in adults with Parkinson's disease and voice complaints. *The Laryngoscope*. 2007;117:831–834.
27. Dalrymple-Alford JC, MacAskill MR, Nakas CT, et al. The MOCA. *Neurology*. 2010;75:1717.
28. Stepp CE, Heaton JT, Stadelman-Cohen TK, et al. Characteristics of phonatory function in singers and nonsingers with vocal fold nodules. *J Voice*. 2011;25:714–724.
29. Burk MH, Wiley TL. Continuous versus pulsed tones in audiometry. *Am J Audiol*. 2004;13:54–61.
30. American Speech-Language-Hearing Association. Guidelines for manual pure-tone threshold audiometry. Rockville, MD2005.
31. Schow RL. Considerations in selecting and validating an adult/elderly hearing screening protocol. *Ear Hear*. 1991;12:337–348.
32. Lien YS, Gattuccio CI, Stepp CE. Effects of phonetic context on relative fundamental frequency. *J Speech Lang Hear Res*. 2014;57:1259–1267.
33. Baken RJ, Orlikoff RF. *Clinical Measurement of Speech and Voice*. Singular Thomson Learning; 2000.
34. Dailey SH, Kobler JB, Hillman RE, et al. Endoscopic measurement of vocal fold movement during adduction and abduction. *Laryngoscope*. 2005;115:178–183.
35. Diaz-Cadiz M, McKenna VS, Vojtech JM, et al. Adductory vocal fold kinematic trajectories during conventional versus high-speed videoendoscopy. *J Speech Lang Hear Res*. 2019;62:1685–1706.
36. Ikuma T, Kunduk M, McWhorter AJ. Preprocessing techniques for high-speed videoendoscopy analysis. *J Voice*. 2013;27:500–505.
37. Vojtech JM, Segina RK, Buckley DP, et al. Refining algorithmic estimation of relative fundamental frequency: Accounting for sample characteristics and fundamental frequency estimation method. *J Acoust Soc Am*. 2019;146:3184.
38. Lien YS, Michener CM, Eadie TL, et al. Individual monitoring of vocal effort with relative fundamental frequency: Relationships with aerodynamics and listener perception. *J Speech, Lang Hear Res*. 2015;58:566–575.
39. McKenna VS, Heller Murray ES, Lien YS, et al. The relationship between relative fundamental frequency and a kinematic estimate of laryngeal stiffness in healthy adults. *J Speech, Lang Hear Res*. 2016;59:1283–1294.
40. Eadie TL, Stepp CE. Acoustic correlate of vocal effort in spasmodic dysphonia. *Ann Otol, Rhinol, Laryngol*. 2013;122:169–176.
41. Heller Murray ES, Lien YS, Van Stan JH, et al. Relative fundamental frequency distinguishes between phonotraumatic and non-phonotraumatic vocal hyperfunction. *J Speech, Lang Hear Res*. 2017;60:1507–1515.
42. Serry MA, Stepp CE, Peterson SD. Physics of phonation offset: Towards understanding relative fundamental frequency observations. *J Acoust Soc Am*. 2021;149:3654–3664.
43. Gillies GE, Pienaar IS, Vohra S, et al. Sex differences in Parkinson's disease. *Front Neuroendocrinol*. 2014;35:370–384.

44. Kempster GB, Gerratt BR, Verdolini Abbott K, et al. Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *Am J Speech-Lang Pathol*. 2009;18:124–132.
45. Martínez-Martín P, Rodríguez-Blázquez C, Mario A, et al. Parkinson's disease severity levels and MDS-Unified Parkinson's Disease Rating Scale. . Parkinson Related Disord. 2015;21:50–54.
46. Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society task force report on the hoehn and yahr staging scale: Status and recommendations the Movement Disorder Society task force on rating scales for Parkinson's disease. *Movement Disord*. 2004;19:1020–1028.
47. Hoehn MM, Yahr MD. Parkinsonism: Onset, progression and mortality. *Neurology*. 1967;17:427–442.
48. Witte RS, Witte JS. *Statistics*. Wiley; 2010.
49. Iwahashi T, Ogawa M, Hosokawa K, et al. A detailed motion analysis of the angular velocity between the vocal folds during throat clearing using high-speed digital imaging. *J Voice*. 2016;30:770.e1-8.
50. Poletto CJ, Verdun LP, Strominger R, et al. Correspondence between laryngeal vocal fold movement and muscle activity during speech and nonspeech gestures. *J App Physiol*. 2004;97:858–866.
51. Park Y, Wang F, Díaz-Cádiz ME, et al. Vocal fold kinematics and relative fundamental frequency as a function of obstruent type and speaker age. *J Acoust Soc Am*. 2021;149:2189–2199.
52. Patel RR, Forrest K, Hedges D. Relationship between acoustic voice onset and offset and selected instances of oscillatory onset and offset in young healthy men and women. *J Voice*. 2017;31:389.e9-389.e17.
53. Watson BC, Roark RM, Baken RJ. Vocal release time: A quantification of vocal offset. *J Voice*. 2012;26:682–687.
54. Robb MP, Smith AB. Fundamental frequency onset and offset behavior: A comparative study of children and adults. *J Speech Lang Hear Res*. 2002;45:446–456.
55. Stepp CE, Hillman RE, Heaton JT. The impact of vocal hyperfunction on relative fundamental frequency during voicing offset and onset. *J Speech Lang Hear Res*. 2010;53:1220–1226.
56. Stepp CE, Sawin DE, Eadie TL. The relationship between perception of vocal effort and relative fundamental frequency during voicing offset and onset. *J Speech Lang HearRes*. 2012;55:1887–1896.
57. Omori K, Slavit DH, Matos C, et al. Vocal fold atrophy: Quantitative glottic measurement and vocal function. *Ann Otol, Rhinol, Laryngol*. 1997;106:544–551.
58. Bloch I, Behrman A. Quantitative analysis of videostroboscopic images in presbylarynges. *Laryngoscope*. 2001;111:2022–2027.
59. Isshiki N, Shoji K, Kojima H, et al. Vocal fold atrophy and its surgical treatment. *Ann Otol, Rhinol Laryngol*. 1996;105:182–188.
60. Angerstein W. Vocal changes and laryngeal modifications in the elderly (presbyphonia and presbylarynx). *Laryngo-Rhino-Otologie*. 2018;97:772–776.
61. Pontes P, Brasolotto A, Behlau M. Glottic characteristics and voice complaint in the elderly. *J Voice*. 2005;19:84–94.
62. Rodeño MT, Sánchez-Fernández JM, Rivera-Pomar JM. Histochemical and morphometrical ageing changes in human vocal cord muscles. *Acta Otolaryngol*. 1993;113:445–449.
63. Ackermann H, Gröne BF, Hoch G, et al. Speech freezing in parkinson's disease: a kinematic analysis of orofacial movements by means of electromagnetic articulography. *Folia Phoniatrica et Logopaedica*. 1993;45:84–89.
64. Ackermann H, Konczak J, Hertrich I. The temporal control of repetitive articulatory movements in parkinson's disease. *Brain Lang*. 1997;56:312–319.
65. Connor NP, Abbs JH, Cole KJ, et al. Parkinsonian deficits in serial multiarticulate movements for speech. *Brain*. 1989;112:997–1009.
66. Forrest K, Weismer G. Dynamic aspects of lower lip movement in parkinsonian and neurologically normal geriatric speakers' production of stress. *J Speech Lang Hear Res*. 1995;38:260–272.
67. Forrest K, Weismer G, Turner GS. Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults. *J Acoust Soc Am*. 1989;85:2608–2622.
68. Yunusova Y, Weismer G. Kent ray d, rusche nicole m. breath-group intelligibility in dysarthria. *J Speech, Lang Hear Res*. 2005;48:1294–1310.
69. Kearney E, Giles R, Haworth B, et al. Sentence-level movements in parkinson's disease: loud, clear, and slow speech. *J Speech, Lang Hear Res*. 2017;60:3426–3440.
70. Hillman RE, Holmberg EB, Perkell JS, et al. Objective assessment of vocal hyperfunction: An experimental framework and initial results. *J Speech Hear Res*. 1989;32:373–392.

APPENDIX A

Table A.1 and A.2

TABLE A.1.
Demographic Information of Participants Examined in Experiment 1.

Participants 1–25				Participants 26–50			
ID	Sex	Age	CAPE-V OS	ID	Sex	Age	CAPE-V OS
C1	F	18	8.8	C26	M	50	5.6
C2	M	21	3.3	C27	F	59	9.8
C3	F	21	7.8	C28	F	30	0.6
C4	M	18	5.6	C29	M	18	6.8
C5	F	29	3.3	C30	F	21	23.5
C6	F	20	4.1	C31	M	25	6.4
C7	F	18	3.3	C32	F	25	3.4
C8	F	63	10.0	C33	M	41	3.8
C9	M	53	6.3	C34	M	22	5.4
C10	M	67	18.1	C35	M	26	6.1
C11	M	46	15.7	C36	M	61	3.0
C12	F	26	4.4	C37	M	62	4.8
C13	F	19	6.3	C38	M	21	2.7
C14	M	54	6.6	C39	M	67	15.4
C15	F	66	7.4	C40	F	67	10.7
C16	M	20	2.1	C41	F	54	6.0
C17	F	80	22.5	C42	M	76	34.2
C18	M	29	3.0	C43	F	61	6.1
C19	M	31	1.9	C44	F	68	113
C20	F	25	6.6	C45	F	61	8.4
C21	F	73	4.6	C46	F	83	25.6
C22	M	81	24.2	C47	M	24	2.4
C23	F	20	3.4	C48	F	57	6.3
C24	M	47	7.8	C49	F	70	5.6
C25	M	21	5.8	C50	M	64	9.7

Note. CAPE-V OS = Consensus of Auditory-Perceptual Evaluation of Voice, Overall Severity of Dysphonia

TABLE A.2.
Demographic Information of Participants without Parkinson's Disease Examined in Experiment 2

Participant	Sex	Age	CAPE-V OS
C1	M	67	1.7
C2	M	53	6.3
C3	M	67	18.1
C4	M	54	6.6
C5	F	73	12.5
C6	F	73	4.6
C7	M	81	24.2
C8	M	47	7.8
C9	F	59	9.8
C10	M	67	7.1
C11	M	61	3.0
C12	M	62	4.8
C13	M	67	15.4
C14	F	54	6.0
C15	M	76	34.2
C16	F	62	17.7
C17	M	66	10.7
C18	F	70	5.6
C19	M	67	6.7
C20	M	64	9.7

Note. CAPE-V OS, Consensus of Auditory-Perceptual Evaluation of Voice, Overall Severity of Dysphonia