Voice and Speech Changes in Transmasculine Individuals Following Circumlaryngeal Massage and Laryngeal Reposturing

Kimberly L. Dahl,² Felicia A. François,² Daniel P. Buckley,²,³ and Cara E. Stepp²,³,⁴

²Department of Speech, Language & Hearing Sciences, Boston University, MA ³Department of Otolaryngology—Head & Neck Surgery, Boston University School of Medicine, MA ⁴Department of Biomedical Engineering, Boston University, MA

ARTICLE INFO

Article History:
Received August 9, 2021
Revision received January 3, 2022
Accepted January 24, 2022

Editor-in-Chief: Katherine C. Hustad
Editor: Rita R. Patel

https://doi.org/10.1044/2022_AJSLP-21-00245

ABSTRACT

Purpose: The purpose of this study was to measure the short-term effects of circumlaryngeal massage and laryngeal reposturing on acoustic and perceptual characteristics of voice in transmasculine individuals.

Method: Fifteen transmasculine individuals underwent one session of sequential circumlaryngeal massage and laryngeal reposturing with a speech-language pathologist. Voice recordings were collected at three time points—baseline, postmassage, and postreposturing. Fundamental frequency (f₀), formant frequencies, and relative fundamental frequency (RFF; an acoustic correlate of laryngeal tension) were measured. Estimates of vocal tract length (VTL) were derived from formant frequencies. Twelve listeners rated the perceived masculinity of participants’ voices at each time point. Repeated-measures analyses of variance measured the effect of time point on f₀, estimated VTL, RFF, and perceived voice masculinity. Significant effects were evaluated with post hoc Tukey’s tests.

Results: Between baseline and end of the session, f₀ decreased, VTL increased, and participant voices were perceived as more masculine, all with statistically significant differences. RFF did not differ significantly at any time point. Outcomes were highly variable at the individual level.

Conclusion: Circumlaryngeal massage and laryngeal reposturing have short-term effects on select acoustic (f₀, estimated VTL) and perceptual characteristics (listener-assigned voice masculinity) of voice in transmasculine individuals.

Supplemental Material: https://doi.org/10.23641/asha.19529299

Individuals whose gender differs from the sex they were assigned at birth make up a growing segment of the population (Meerwijk & Sevelius, 2017). These gender-diverse individuals face health inequities and an insufficient evidence base to support gender-affirming care (James et al., 2016). The evidence base is particularly lacking in voice care for transmasculine individuals. The term transmasculine encompasses different genders, including transgender men, some nonbinary individuals, and many others whose gender differs from an assigned female sex. The paucity of evidence for gender-affirming voice care does not reflect the importance of voice to this population. Transmasculine individuals ranked voice as the most important trait among those they sought to change with gender-affirming care (Hodges-Simeon et al., 2021). Likewise, a large majority of transmasculine individuals (83%) have reported incongruence between their gender and voice (Kennedy & Thibeault, 2020). Providing effective voice care to this population is therefore of critical importance.

Hormone replacement therapy (HRT) with exogenous testosterone is a common gender-affirming treatment for transmasculine individuals (Mueller et al., 2007; Wierckx et al., 2014). HRT induces changes in voice and speech in some transmasculine individuals, including...
decreased mean fundamental frequency ($f_o$), decreased $f_o$ range, decreased vowel formant frequencies, and changes in voice quality (Azul et al., 2018; Cler et al., 2020; Hancock et al., 2017; Papp, 2011). However, due to barriers to health care access (James et al., 2016; Stroumsa et al., 2020) and individual preferences regarding the wide-ranging effects of testosterone (Schneiders, 2014), not all transmasculine people undergo HRT. Among those who do, some find its effects insufficient to address their voice concerns (Azul et al., 2021). Up to 29% of transmasculine individuals report continuing to be misgendered (i.e., being referred to in a way that is incongruent with their gender identity)—especially on the phone, where voice is the most salient cue of gender—even months and years after onset of HRT (Azul et al., 2018; Van Borsel et al., 2000). Nearly one third (31%) of transmasculine participants who had undergone HRT for at least a year expressed interest in voice training for further voice masculinization (Van Borsel et al., 2000).

Relying on HRT to meet the needs of those transmasculine people who seek to masculinize their voice leaves a substantial portion of this population unserved. Other evidence-based voice care options are needed. Behavioral approaches to voice masculinization may provide an alternative or supplement to HRT. However, there remains little evidence of the effectiveness of behavioral approaches in addressing voice concerns in this population. Such concerns often relate to gendered voice characteristics (e.g., $f_o$), excessive laryngeal muscle tension, and the misattribution of gender by others (Azul et al., 2021). Approaches expected to affect $f_o$ and laryngeal muscle tension are circumlaryngeal massage and laryngeal reposturing (Aronson, 1990; Dagli et al., 2008; Mathieson et al., 2009; Roy, 2008; Roy & Bless, 1998; Roy & Ferguson, 2001). These techniques are thought to reduce laryngeal tension and facilitate a lower resting position of the larynx, thus lengthening the vocal tract in individuals with muscle tension-related voice disorders. The corresponding acoustic effects of these physiological changes include decreased $f_o$ (Dagli et al., 2008; Roy et al., 2017) and decreased formant frequencies (Dagli et al., 2008; Roy & Ferguson, 2001).

Both $f_o$ and formant frequencies relate to perception of gender, with lower $f_o$ and formant frequencies corresponding to a greater likelihood of being perceived as male or masculine (Avery & Liss, 1996; Dahl & Mahler, 2020; Gelfer & Bennett, 2013; Gelfer & Schofield, 2000; Hardy et al., 2020; Hillenbrand & Clark, 2009; Munson, 2007; Skuk & Schweinberger, 2014), although this relationship has not been specifically demonstrated in transmasculine speakers. Because circumlaryngeal massage and laryngeal reposturing achieve these same acoustic effects, these approaches may facilitate desired voice and speech changes for individuals who would like to be perceived as male or masculine. However, these effects have primarily been observed in individuals known to have excessive laryngeal tension due to muscle tension dysphonia (Mathieson et al., 2009; Roy & Ferguson, 2001; Roy & Leeper, 1993) or mutational falsetto (Roy et al., 2017). Expecting similar effects in transmasculine individuals may therefore depend upon an assumption of elevated laryngeal tension in these speakers as well.

Self-reports of excessive laryngeal tension have indeed been documented in this population (Azul et al., 2021); however, no objective assessments of elevated laryngeal tension in transmasculine speakers have been published. For example, no study of transmasculine speakers has included laryngoscopic findings indicative of elevated tension, electromyography, or relative fundamental frequency (RFF). RFF is an acoustic correlate of laryngeal tension that decreases when speakers with typical voices deliberately produce voice with increased tension or strain (Lien et al., 2015; McKenna & Stepp, 2018). Although there is no clinical threshold for RFF that distinguishes between typical and elevated levels of tension, RFF is lower in individuals with excessive laryngeal tension compared with speakers with typical voices (Heller Murray et al., 2017; Stepp et al., 2010) and increases after voice therapy for vocal hyperfunction (Roy et al., 2016; Stepp et al., 2011). An increase in RFF after voice training with transmasculine individuals might therefore suggest a decrease in laryngeal tension.

Two single-participant case studies have explored the use of circumlaryngeal massage and laryngeal reposturing with transmasculine individuals (Buckley et al., 2020; Myers & Bell, 2020). Buckley et al. (2020) focused exclusively on these techniques. Myers and Bell (2020) incorporated them into an approach focusing more heavily on vocal function exercises. In both cases, the voice training induced changes in voice acoustics—reduced $f_o$ (Buckley et al., 2020; Myers & Bell, 2020) and reduced formant frequencies (Buckley et al., 2020)—and perceptions of voice gender (Myers & Bell, 2020) and masculinity (Buckley et al., 2020) for each transmasculine participant. However, given the differences in voice training approaches and the single-participant design of these studies, the observed effects of circumlaryngeal massage and laryngeal reposturing cannot be extended to the transmasculine population in general. Furthermore, neither study objectively measured laryngeal tension before nor after intervention; thus, it is unclear if the effects of these approaches are the result of relief from elevated muscle tension, as assumed in other populations, or if other mechanisms are involved.

The purpose of this study was to measure the short-term effects of circumlaryngeal massage and laryngeal reposturing on acoustic and perceptual characteristics of voice and speech in transmasculine individuals. Outcome
measures included $f_o$, estimated vocal tract length (VTL; derived from formant frequencies), RFF, and listener perceptions of voice masculinity. We hypothesized that $f_o$ would decrease, estimated VTL would increase, RFF would increase, and participants’ voices would be perceived as more masculine after completing one session of circumlaryngeal massage and laryngeal reposturing.

Method

Participants

Participants were 15 transmasculine adults between the ages of 19 and 53 years ($M = 27.3$ years, $SD = 8.2$ years). All participants self-identified as transmasculine on a lab screening form. They also reported their specific genders, which included man, trans, nonbinary, gender-fluid, genderqueer, and demi boy. Participants were native speakers of American English. Most participants (13/15) passed a hearing screening at 25 dB HL at octaves from 125 to 4000 Hz. One participant had an elevated threshold (40 dB HL) at 500 Hz in one ear. Another participant had moderate-to-severe hearing loss, with elevated thresholds (30–70 dB HL) across most frequencies in both ears. One participant reported a remote diagnosis of muscle tension dysphonia; no other participant reported any history of speech, voice, language, or hearing disorders.

Most participants (13/15; 81%) were undergoing HRT with exogenous testosterone at the time of the study and had done so for an average of 2.4 years ($SD = 1.9$ years, range: 5 months to 6 years). The timing and extent of HRT-induced voice changes varies (Cosyns et al., 2014; Hancock et al., 2017; Irwig et al., 2017; Nygren et al., 2016), but all 13 of these participants reported experiencing voice change with HRT, with changes in pitch and voice quality being most common. Nine participants (60%) reported that their ideal voice differed from their current voice; this difference was documented in their responses to items on the Transgender Self-Evaluation Questionnaire (TSEQ; Davies, 2006) in which participants evaluated their current and ideal vocal gender (i.e., how “male” or “female” their voice sounds). Because both previously experienced HRT-induced voice changes and self-perception of voice could influence voice training outcomes, the participants were later subdivided by these characteristics for supplementary, qualitative analysis.

Nine participants (60%) had previously engaged in behavioral voice change. All but one of these participants took a self-guided approach, ranging from “casual exploration” to watching YouTube videos on pitch lowering. One participant reported working on “voice deepening” with a singing voice coach. No participant had worked with a speech-language pathologist (SLP) on voice modification. All participants completed written consent in compliance with the Boston University Institutional Review Board.

Voice Training Protocol

Each participant underwent a sequence of two laryngeal manipulation procedures in a single study session: (a) circumlaryngeal massage and (b) laryngeal reposturing. All voice training was led by a single SLP who specialized in voice and had experience providing voice care to gender-diverse individuals.

Circumlaryngeal Massage

Circumlaryngeal massage was administered first and lasted 15 min. The participant sat in a procedure chair without speaking for the entirety of the massage. The SLP administered a standardized massage protocol adapted from published descriptions (Aronson, 1990; Buckley et al., 2020; Mathieson et al., 2009). The massage consisted of downward sternocleidomastoid pull-downs (1.5 min), thyrohyoid space anterior–posterior massage (3.5 min), circular anterior–posterior massage of the suprahypoid muscles (2.5 min), dynamic lateral stretches of the thyroid cartilage (1.5 min), static lateral stretches of the thyroid cartilage (1 min/side), thyroid cartilage pull-downs (2 min), and hyoid pull-downs (2 min).

The aim of the circumlaryngeal massage was to reduce extrinsic laryngeal muscle tension and allow for a lower resting position of the larynx. Two subsequent effects of a lower laryngeal position were expected, on the basis of previous findings—a decrease in $f_o$ (Buckley et al., 2020; Roy et al., 1997) and a lengthening of the vocal tract, as reflected in a decrease in formant frequencies (Buckley et al., 2020; Roy & Ferguson, 2001). With decreases in $f_o$ and formant frequencies, an increase in perceived voice masculinity was also expected, on the basis of previous findings in cisgender speakers (Avery & Liss, 1996; Gelfer & Bennett, 2013; Hillenbrand & Clark, 2009; Munson, 2007). Finally, a reduction in laryngeal muscle tension was expected to result in an increase in RFF, given the relationship between laryngeal tension and RFF (Park & Stepp, 2019; Roy et al., 2016; Stepp et al., 2010).

Laryngeal Reposturing

The second training component consisted of 15 min of active laryngeal reposturing during a hierarchy of speech tasks. First, the SLP trialed thyroid pull-down and hyoid push-back maneuvers on the participant as the participant produced their voice. The SLP used these trials to determine which laryngeal reposturing technique yielded a lower laryngeal position without inducing strain or other
The participant then continued practicing laryngeal reposturing without manual manipulation through more complex speech tasks, including rote speech (e.g., counting to 10, reciting months of the year), oral reading, and conversation. Although the rate of progress through these tasks differed by participant, all participants did reach conversational practice during this training. Negative practice was used intermittently to facilitate awareness of the larynx at different elevations and the corresponding perceptual characteristics of voice.

The aim of laryngeal reposturing was again to facilitate a lower laryngeal position with the subsequent effects of a lower $f_0$ and longer vocal tract. These acoustic changes were expected to result in a further increase in perceived voice masculinity. No additional increase in RFF was anticipated, as a release of baseline tension was expected after circumlaryngeal massage. Given the participants’ active engagement in this second component and the focus on a lower laryngeal position during actual speech tasks, greater changes in $f_0$, VTL, and perceived voice masculinity were expected over those seen after the passive circumlaryngeal massage.

**Acoustic Recordings**

Acoustic recordings were collected at three time points—baseline, postmassage, and postreposturing. Recordings were completed in a sound-treated booth. Participants wore an omnidirectional headset microphone (MX153; Shure, Niles, IL) placed 7 cm from the participant’s lips at a 45° angle (Patel et al., 2018). Signals were recorded in Sonar Artist (Cakewalk) and sampled at 44.1 kHz.

During each recording, the participant completed a set of four speech tasks. Participants first sustained the vowels /a, i, u, æ, ʌ/ for 3 s in elongated productions of the words “pot, seed, coop, cat, hut,” with three repetitions of each. The intention was to conduct formant frequency analysis on sustained /ʌ/, but additional vowels were included in the recording to ensure sufficient productions of a single vowel for the collection of stable formant values. Participants then repeated nine productions of voiced–voiceless–voiced phoneme sequences (i.e., /afæ/) for calculation of RFF. Finally, participants read the first paragraph of the Rainbow Passage (Fairbanks, 1960) and provided a 30- to 60-s sample of spontaneous speech. Two types of connected speech tasks were included to support the collection of perceptual measures, as described in the Listeners and Listening Procedures section below. At the first two time points—baseline and postmassage—participants were instructed to complete the speech tasks using their typical speaking voice. At the final time point—postreposturing—participants were instructed to complete the speech tasks using the manner of voice production practiced during the laryngeal reposturing activity, without manual manipulation of the larynx.

**Listeners and Listening Procedure**

Listeners were 12 adults ($M = 22.3$ years, $SD = 3.8$ years, range: 18–30 years) with no history of speech, voice, language, or hearing disorders. Listeners included cisgender women ($n = 4$), cisgender men ($n = 4$), transgender men ($n = 2$), a nonbinary person, and an agender person. All were native speakers of American English.

All listeners completed the listening procedure remotely with the online behavioral research platform Gorilla Experiment Builder (gorilla.sc). Listeners completed a single study session from their home or other quiet environment using a personal laptop and wired earbuds or headphones. A study staff member supervised all study sessions via videoconference.

The listening procedure began with a volume adjustment and headphone screening task (Milne et al., 2020) drawn from the Gorilla open materials repository. Listeners first played a 4-s clip of white noise while setting their computer volume to a comfortable level. They then completed an antiphase tone discrimination task in which they identified the quietest of three 200-Hz sinusoidal tones of different loudness levels. Listeners passed the headphone screening by making three correct identifications of the quietest tone.

Listeners then completed the primary listening task, which was divided into two sections according to the type of speech stimuli presented—reading passage (two sentences) and spontaneous speech (5–8 s). Reading passage...
samples were included to collect voice masculinity perceptions while controlling for the potentially confounding variable of content. However, because the characteristics of speech and voice during oral reading differ from spontaneous productions (Daly & Zue, 1992; Laan, 1997), a reading sample may not be truly representative of a participant’s voice as used in daily life. Thus, samples of spontaneous speech were also included to capture more ecologically valid perceptions of voice masculinity.

The order of the reading and spontaneous speech sections was counterbalanced across listeners. In each section, listeners rated the masculinity of the voices in the samples according to a paired comparisons method (Thurstone, 1994), which allowed for the detection of small, within-speaker changes in perceived voice masculinity. Listeners heard two recordings separated by a beep and selected the clip in which the speaker’s voice sounded more masculine. Masculinity was not defined for the listeners, who were expected to draw on their own varied experiences with gender and masculinity to make their selections.

All permutations of speech sample pairings for each speaker (six pairings/speaker) were presented in random order (i.e., baseline/postmassage, baseline/postreposturing, and postmassage/postreposturing, in both orders). Listeners were instructed to let both samples play to completion at least once before making their selection. They were allowed to play each pair of recordings a second time to account for potential disruptions in the listener’s environment that could cause inattention during the first presentation. The listening task took an average of 1 hr 20 min to complete. Participants were offered a 10-min break between the two sections of the listening task.

Data Analysis

Fundamental Frequency

All voice recordings were analyzed in Praat (Boersma & Weenink, 2015) by a trained technician. The technician extracted \( f_o \) traces from each reading passage and spontaneous speech sample. Prior to data extraction, the technician visually inspected the samples and adjusted the pitch settings as needed to ensure accurate \( f_o \) tracking in Praat. Mean \( f_o \) for both reading and spontaneous speech was calculated for each participant to account for possible differences in this measure according to speech task (Daly & Zue, 1992). For illustration purposes, each participant’s \( f_o \) was converted to semitones (ST) at each time point, with the baseline \( f_o \) as the reference value. Semitone conversions were calculated using Equation 1 where \( f_1 \) was the baseline \( f_o \), and \( f_2 \) was the \( f_o \) of a given time point.

\[
\text{ST} = 39.86 \times \log_{10} \frac{f_2}{f_1}.
\]

VTL

Vowel formant frequencies were also measured in Praat by a trained technician. Formant frequencies were measured from the middle portion of the elongated vowel /æ/ or /ʌ/. The latter vowel was used for one participant who did not produce a stable /æ/. These vowels were selected because their vocal tract configurations are most similar to a tube with a consistent cross-sectional area (Stevens, 1998). The length of the vocal tract can thus be estimated based on the formant frequencies of these vowels (Wakita, 1977). Specifically, VTL estimates were derived, in centimeters, from the mean of the third and fourth formant frequencies. These estimates were calculated with Equation 2, where \( n \) is the formant number, \( c \) the speed of sound in air (34,300 cm/s), and \( F_n \) the formant frequency in Hz (Wakita, 1977).

\[
\text{VTL} = \frac{(2n-1)c}{4F_n^2}.
\]

For illustration purposes, changes in VTL for each participant were calculated as the percent change between baseline VTL and each subsequent time point.

RFF

RFF is an acoustic correlate of laryngeal tension (Park & Stepp, 2019; Roy et al., 2016; Stepp et al., 2010), thought to decrease as tension increases. It is measured during the laryngeal gesture that occurs in transitions between voiced and voiceless phonemes. RFF was calculated by the first author using an automated algorithm in MATLAB (MathWorks; Vojtech et al., 2019). The algorithm returned RFF values for the 10 voicing cycles preceding and following the voiceless consonant in productions of /af/ (up to nine productions) at each time point.

Voice Masculinity Ratings

Ratings of voice masculinity were analyzed according to Thurstone’s law of comparative judgment with case V simplification (Thurstone, 1994). This process transforms listener ratings into a scale value that places each voice sample on a continuum of voice masculinity. Scale values were calculated based on the proportion of times in which a voice sample was chosen as more masculine than other samples (Kaiser & Serlin, 1978; Krus & Krus, 1977). A higher scale value corresponded to a voice that listeners perceived as more masculine.

Reliability

Reliability of \( f_o \) and formant frequency estimates were calculated as Pearson correlation coefficients, on the basis of repeated analysis of 13% of samples. For
in Intrerrater reliability, a second technician reanalyzed a subsample. Intrerrater reliability, a second technician reanalyzed a subsample.

Intrerrater reliability of listeners’ perceptual ratings was calculated as percent agreement between repeated presentations of voice sample pairings. Intrerrater reliability of listener ratings was calculated as percent agreement between listeners for all ratings.

**Statistical Analysis**

Seven repeated-measures analyses of variance (ANOVAs) were constructed to measure the main effect of time point (baseline, postmassage, and postreposturing) on the outcomes of $f_o$ (Hz), VTL (cm), RFF (ST), and voice masculinity rating (scale value). The main effect of voicing cycle and its interaction with time point were also added to the model for RFF. Separate ANOVAs were constructed for $f_o$ and voice masculinity ratings based on data from the two speech tasks—reading and spontaneous speech. Effect sizes for each ANOVA factor were calculated as squared partial curvilinear correlations ($\eta^2$). Effect sizes of $\sim 0.01$ were classified as small, $\sim 0.09$ medium, and $> 0.25$ large (Witte & Witte, 2009). Statistically significant effects were evaluated with post hoc Tukey’s tests. Cohen’s $d$ effect sizes quantified the magnitude of significant differences and were designated according to recommendations specific to speech, language, and hearing sciences (Gaeta & Brydges, 2020) as small (0.25), medium (0.55), or large (> 0.93). Significance was set a priori at $\alpha = .05$. Statistical analysis was conducted in Minitab (Version 19, Minitab Inc.).

**Results**

**Acoustic Measures**

**Fundamental Frequency**

Participants’ baseline $f_o$ ranged from 85 to 186 Hz during reading ($M = 125$ Hz, $SD = 30$ Hz) and 86 to 164 Hz during spontaneous speech ($M = 120$ Hz, $SD = 27$ Hz). There was a significant effect of time point on $f_o$ derived from the reading passage, $F(2, 28) = 6.18, p < .05, \eta^2_p = .31$, and from spontaneous speech, $F(2, 28) = 5.88, p < .05, \eta^2_p = .30$, both with large effect sizes. Post hoc testing showed that, at the group level, mean $f_o$ decreased by 1.51 ST from baseline to postlaryngeal massage for reading-based $f_o$ estimates ($d = 0.46; p_{adj} < .05$). Mean $f_o$ decreased between the postmassage and reposturing time points by 0.27 ST for reading-based $f_o$ estimates ($d = 0.36; p_{adj} < .05$) and by 0.73 ST for spontaneous speech estimates ($d = 0.52; p_{adj} < .05$). There was no statistically significant change in $f_o$ between baseline and postmassage for either connected speech task. At the individual level, the effects of circumlaryngeal massage and laryngeal reposturing on $f_o$ were highly variable. These findings are illustrated in Figure 1 and statistical results provided in Table 1. All $f_o$ outcomes on an individual speaker basis are presented in Supplemental Material S1.

**VTL**

Participants’ mean estimated VTL at baseline was 15.7 cm ($SD = 0.8$ cm; range: 14.6–17.5 cm). There was a significant effect of time point on estimated VTL, with large effect size, $F(2, 28) = 6.25, p < .05, \eta^2_p = .31$. Estimated vocal tract increased, on average, 0.42 cm between baseline and postreposturing ($d = 0.14, p_{adj} < .05$) and 0.34 cm from postmassage to postreposturing ($d = 0.11, p_{adj} < .05$). VTL did not change significantly between baseline and postmassage. Again, individual results were highly variable, as shown in Figure 2 and Supplemental Material S1. Statistical results are presented in Table 1.

**RFF**

At baseline, participants’ RFF offset values were relatively stable with a slight decrease before the voiceless consonant. RFF at voicing onset was relatively high, followed by a quick decrease to 0 ST. This pattern held true across all time points, with no significant differences in RFF values throughout the session. Both the pattern and values of RFF at offset and onset were consistent with those reported for speakers without excessive laryngeal tension (Heller Murray et al., 2017; McKenna & Stepp, 2018; Robb & Smith, 2002). RFF at each time point is shown in Figure 3. Statistical results are listed in Table 1.

**Perception of Voice Masculinity**

There was a significant effect of time point on listener perceptions of voice masculinity for both reading, $F(2, 28) = 98.99, p < .05, \eta^2_p = .88$, and spontaneous speech samples, $F(2, 28) = 16.77, p < .05, \eta^2_p = .55$, both with large effect sizes. Listeners perceived participants’ voices to be more masculine, on average, postreposturing than at baseline for both reading ($d = 10.13, p_{adj} < .05$) and spontaneous speech samples ($d = 8.63, p_{adj} < .05$), with large effect sizes. Perceived voice masculinity also increased significantly between postmassage and reposturing for both reading ($d = 9.94, p_{adj} < .05$) and spontaneous speech samples ($d = 11.49, p_{adj} < .05$), with large effect sizes. Perceived voice masculinity did not change from baseline to postmassage. Scale values of perceived voice masculinity are shown in Figure 4 and statistical results listed in Table 1. All perceptual outcomes by speaker are presented in Supplemental Material S1.
Reliability

Reliability of $f_o$ and formant analyses was high. Intrarater reliability was $r = .97$ for $f_o$ and $r = .99$ for formant frequencies. Interrater reliability was $r = .98$ for $f_o$ and $r = .99$ for formant frequencies. Intrarater agreement for perception of voice masculinity averaged 0.71 ($SD = 0.09$, range: 0.56–0.87). Interrater agreement for perception of voice masculinity averaged 0.67 ($SD = 0.05$, range: 0.55–0.75). This level of agreement is consistent with past research showing relatively high variability in the perception of gender and gender-related concepts (Azul et al., 2018; Buckley et al., 2020; Cler et al., 2020), to which the listener brings a vast set of experiences with genders and gender expressions—both their own and others’.

Discussion

This study measured the short-term effects of circumlaryngeal massage and laryngeal reposturing on voice and speech in transmasculine individuals. We found that combining these approaches led to several significant changes in acoustic and perceptual characteristics of voice and speech. These changes were observed without any indication of co-occurring changes in laryngeal tension that would suggest elevated levels of tension at baseline or at any point during voice training.

Acoustic Changes in Voice and Speech

Circumlaryngeal massage followed by laryngeal reposturing induced a significant $f_o$ decrease of 1.3–1.5 ST during reading and spontaneous speech, respectively, in a sample of transmasculine individuals. This is consistent with $f_o$ outcomes in individuals with mutational falsetto (Dagli et al., 2008; Roy et al., 2017), although the magnitude of average $f_o$ decrease in this previously studied group—individuals with known excessive laryngeal tension at baseline—was greater at 4.8 ST (Roy et al., 2017). Our result is also consistent with $f_o$ changes in Buckley et al.’s (2020) case study using an identical laryngeal manipulation approach with a single transmasculine participant. This finding suggests that circumlaryngeal massage and laryngeal reposturing may play a role in alleviating concerns about pitch that are often expressed by transmasculine individuals (Azul et al., 2018, 2021).

The $f_o$ outcomes in our sample were highly variable, however, ranging from a decrease of 5.4 ST to an increase of 2.4 ST in spontaneous speech samples (−5.1 to 0.4 ST during reading). Although a majority of participants (11/15; 73%) did experience a decrease in $f_o$, it may be that this approach, in which participants are not explicitly...
instructed to lower their pitch, is ineffective at decreasing $f_0$ for certain individuals. Alternatively, some individuals may require multiple sessions of laryngeal manipulation to achieve substantial decreases in $f_0$.

Formant-derived estimates of VTL also increased significantly after circumlaryngeal massage followed by laryngeal reposturing. Again, this is consistent with previous literature documenting similar changes in individuals with muscle tension dysphonia (Roy & Ferguson, 2001) and a single transmasculine individual who underwent these same laryngeal manipulation techniques (Buckley et al., 2020).

The lengthening of the vocal tract after circumlaryngeal massage and laryngeal reposturing is theorized to result from the release of excessive muscle tension, which allows for a lower resting position of the larynx (Aronson, 1990). Although we indirectly measured the position of the larynx using formant-derived VTL estimates, our $f_0$ results provide further support for a lowering of the larynx among our participants. As the larynx lowers, the natural curvature of the cervical spine tilts the cricoid cartilage anteriorly, thus reducing vocal fold tension and causing a decrease in $f_0$ (Honda et al., 1999). The co-occurrence of increased VTL estimates and decreased $f_0$ in our findings suggest that circumlaryngeal massage and laryngeal reposturing did result in laryngeal lowering.

This laryngeal lowering, however, does not appear to be the result of released muscle tension, which is the mechanism by which these techniques are assumed to induce voice changes in individuals with tension-related voice disorders (Roy & Bless, 1998). We hypothesized that RFF, an acoustic correlate of laryngeal tension (Roy et al., 2016; Stepp et al., 2010), would increase over the course of the study session, indicating a reduction in laryngeal tension. That was not the case; RFF remained unchanged throughout the session, with values at each time point similar to those previously reported for adults with typical voices (Heller Murray et al., 2017; McKenna & Stepp, 2018; Robb & Smith, 2002). Thus, these results suggest that, as a group, the participants had typical levels of laryngeal tension, consistent with the appropriate levels necessary to sustain healthy voice function.

These RFF results suggest that the acoustic changes observed may not have depended on the release of excessive or imbalanced extrinsic laryngeal tension in this sample of transmasculine individuals. Rather, circumlaryngeal massage may have simply given participants, who had a typical level of extrinsic laryngeal tension at baseline, greater laryngeal mobility. Laryngeal reposturing then incorporated both somatosensory and auditory feedback, which was effective in facilitating a lower laryngeal position during speech. This suggestion could be confirmed in future studies by comparing outcomes of the present approach with one in which participants are simply instructed to speak with a lower laryngeal position without facilitative manual techniques. Future work may also consider whether factors such as previous voice training experience affect an individual’s volitional control over their laryngeal position.

### Perceptual Changes in Voice

In addition to acoustic changes, circumlaryngeal massage followed by laryngeal reposturing induced perceptual

---

**Table 1.** Results of repeated-measures analyses of variance for $f_0$, vocal tract length, RFF, and perceived voice masculinity.

<table>
<thead>
<tr>
<th>Effect</th>
<th>$df$</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2_p$</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_0$ (reading)</td>
<td>2</td>
<td>6.18</td>
<td>.006*</td>
<td>.31</td>
<td>Large</td>
</tr>
<tr>
<td>Time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_0$ (spontaneous speech)</td>
<td>2</td>
<td>5.88</td>
<td>.007*</td>
<td>.30</td>
<td>Large</td>
</tr>
<tr>
<td>Time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocal tract length</td>
<td>2</td>
<td>6.25</td>
<td>.006*</td>
<td>.31</td>
<td>Large</td>
</tr>
<tr>
<td>Time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFF offset</td>
<td>2</td>
<td>2.92</td>
<td>.055</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Cycle</td>
<td>9</td>
<td>5.80</td>
<td>&lt; .001*</td>
<td>.14</td>
<td>Medium</td>
</tr>
<tr>
<td>Time Point × Cycle</td>
<td>18</td>
<td>.47</td>
<td>.971</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>RFF onset</td>
<td>2</td>
<td>0</td>
<td>.998</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Cycle</td>
<td>9</td>
<td>230.64</td>
<td>&lt; .001*</td>
<td>5.38</td>
<td>Large</td>
</tr>
<tr>
<td>Time Point × Cycle</td>
<td>18</td>
<td>0.09</td>
<td>1.000</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Voice masculinity (reading)</td>
<td>2</td>
<td>98.99</td>
<td>&lt; .001*</td>
<td>.88</td>
<td>Large</td>
</tr>
<tr>
<td>Time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice masculinity (spontaneous speech)</td>
<td>2</td>
<td>16.77</td>
<td>&lt; .001*</td>
<td>.55</td>
<td>Large</td>
</tr>
<tr>
<td>Time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** $f_0$ = fundamental frequency; RFF = relative fundamental frequency; $df$ = degrees of freedom; NS = not significant; — = not applicable for nonsignificant findings.

*Significant at $p < .05$. 

---

---

According to the image, the text discusses the effectiveness of laryngeal manipulation for lowering pitch, with some individuals requiring multiple sessions for significant decreases. Formant-derived estimates of vocal tract length increased significantly after circumlaryngeal massage and laryngeal reposturing, consistent with previous literature. The study also investigated the role of released muscle tension in these changes; however, RFF results did not support this mechanism for the group as a whole. The laryngeal lowering observed is theorized to result from the release of muscle tension, but further studies are needed to confirm this mechanism. Perceptual changes in voice were also noted, with significant differences found for measures like $f_0$, vocal tract length, and RFF.
voice changes in our sample. Specifically, listeners perceived participants’ voices to be significantly more masculine postreposturing. This has important implications for addressing transmasculine individuals’ reported concerns about being misgendered (Azul et al., 2021), especially in contexts like phone calls in which voice is the most salient cue of the speaker’s gender. Although voice masculinity is not equivalent to gender attribution, these two aspects of listener perceptions are strongly related (Leung et al., 2021). Thus an increase in the perceived voice masculinity of our participants may also be expected to influence listener’s attribution of gender.

At the group level, this increase in perceived voice masculinity coincided with a decrease in \( f_o \) and increase in VTL. This study was not sufficiently powered to measure correlations between these acoustic and perceptual voice characteristics. However, this group-level finding suggests that the relationships between \( f_o \) and formant frequencies previously identified in cisgender and transfeminine speakers (Avery & Liss, 1996; Dahl & Mahler, 2020; Gelfer & Bennett, 2013; Gelfer & Schofield, 2000; Hardy et al., 2020; Hillenbrand & Clark, 2009; Munson, 2007; Skuk & Schweinberger, 2014) may also apply to perceptions of voice masculinity in transmasculine speakers. Specifically, decreased \( f_o \) and increased VTL may correspond to the perception of a more masculine voice in transmasculine individuals. These relationships should be explicitly measured in future studies of transmasculine speakers.

**Variability in Individual Outcomes**

The effects of circumlaryngeal massage and laryngeal reposturing were highly variable at the individual level. This study was not sufficiently powered to quantitatively evaluate the wide range of results. However, we did qualitatively assess differences both within and across speakers.

Within-speaker variability was reflected across speech tasks. Many speakers showed similar changes in \( f_o \) (8/15; 53%) and perceived voice masculinity (12/15; 80%) for both reading and spontaneous speech. Two speakers, however, experienced a decrease in reading \( f_o \), but an increase in spontaneous speech \( f_o \). Two speakers (including one of the two just mentioned) showed similarly conflicting results for perceived voice masculinity; listeners rated their voices as more masculine when based on a reading excerpt but less masculine when based on spontaneous speech. In other speakers, a change was observed at the end of the
session in only one speech task ($f_0$: 5/10, 33%; perceived voice masculinity: 3/15; 20%). It may therefore be important to evaluate voice training progress using a variety of speech tasks for transmasculine individuals who seek to change their $f_0$ or listener-perceived voice masculinity.

Within-speaker differences across outcome measures were also noted. Six speakers (40%) demonstrated the expected changes across $f_0$, VTL, and perceived voice masculinity. Seven others (47%) experienced the expected changes in two of these measures. In most cases (9/15; 60%), the predicted negative relationship between $f_0$ and perceived voice masculinity held true, that is, the voices of those speakers whose $f_0$ decreased were also perceived to sound more masculine. However, this was not always the case. Three speakers demonstrated a positive relationship between $f_0$ and perceived voice masculinity, such that even though their $f_0$ increased by the end of the session, their voices were perceived as more masculine, or vice versa for one speaker. This likely reflects the role that voice and speech characteristics other than $f_0$ play in perceptions of gender and masculinity (see Leung et al., 2018, for a review).

Similarly, the predicted negative relationship between $f_0$ and VTL was seen in a majority of speakers (8/15; 53%). In three speakers, however, a decrease in $f_0$ corresponded to a decrease in VTL. Lowering the larynx can cause a decrease in $f_0$ (Honda et al., 1999), as noted above. Laryngeal height was specifically targeted in this study, and so the expectation was that most participants would experience both a decrease in $f_0$ and an increase in VTL. However, speakers may also adjust their $f_0$ by other means, including altering intrinsic laryngeal muscle tension (Hirano et al., 1969) or subglottal pressure (Titze, 1989). Those speakers who experienced this seemingly contradictory response of $f_0$ and VTL may therefore have employed a strategy other than laryngeal lowering to change their $f_0$.

There was less individual variability in RFF outcomes; nearly all participants saw no substantial change in this measure and presented with RFF values consistent with a typical level of laryngeal tension at all time points (Heller Murray et al., 2017; McKenna & Stepp, 2018; Robb & Smith, 2002). Nevertheless, RFF increased at offset (Cycle 10) and onset (Cycle 1) by a magnitude consistent with differences between speakers with and without excessive laryngeal tension (Stepp et al., 2010) for two speakers in our sample. VTL also increased by the end of the session for one of these speakers but decreased for the other. These qualitative findings further suggest that the overall changes observed were not dependent on a release of even typical levels of laryngeal tension in this sample.

The individual variability described above prompted us to further evaluate what role individual characteristics may have played in the observed outcomes. One characteristic considered was length of exposure to HRT. HRT induces voice changes in some transmasculine individuals,
but the timing and extent of these effects is variable (Azul et al., 2021; Hancock et al., 2017). We divided our sample into those participants with 1 year or less of HRT ($n = 5$; $M = 5$ months, $SD = 5.2$ months, range: 0–12 months) and those with over a year of HRT ($n = 10$; $M = 35.2$ months, $SD = 23$ months, range: 14–72 months). Qualitatively, the group with shorter HRT exposure demonstrated a greater change in $f_o$ (−2.0 ST), a smaller increase in VTL (1.6%), and a greater increase in perceived voice masculinity (2.7 units) between baseline and postreposting than those with longer HRT exposure (−0.5 ST, 3.2%, and 2.1 units, respectively; see Figure 5). The differences in $f_o$ and perceived voice masculinity reported here are from spontaneous speech samples; the same patterns were observed in reading samples. Outcomes for the two participants who were not undergoing HRT were similar to the rest of the shorter HRT exposure group, except that one of these participants demonstrated an atypically large change in their reading passage $f_o$ (−3.4 ST).

A second characteristic of interest was participants’ self-perception of voice, as indicated by their responses to the TSEQ’s Likert scale evaluations of current and ideal vocal gender (Davies, 2006). We interpreted a mismatch between current and ideal vocal genders as an indicator of voice dissatisfaction, which may influence an individual’s motivation to achieve voice change. We divided the group into participants who reported that their current voice differed in vocal gender from their ideal voice ($n = 9$) and those with no such difference ($n = 6$). There was some overlap between this grouping and the HRT grouping; 44% of those with a mismatch between current and ideal voices were also part of the short HRT exposure group. Qualitatively, the participants whose current and ideal vocal gender did not match demonstrated greater changes in $f_o$ (−1.5 ST), VTL (2.9%), and perceived voice masculinity (2.8 units) than those who reported no such difference in vocal gender (−0.2 ST, 2.4%, and 1.5 units, respectively; see Figure 6). Again, the $f_o$ and perceived voice masculinity values reported here are derived from spontaneous speech, with identical patterns observed in reading samples.

These findings must be interpreted with caution given the subdivided sample sizes, overlap between groupings, and baseline group differences. For example, two groups—those with shorter HRT exposure and those with a mismatch of current and ideal vocal gender—had a higher mean $f_o$ at baseline than their respective comparison groups; a floor effect among the groups with lower $f_o$ at baseline might therefore explain differences in outcomes. However, this exploratory, qualitative analysis suggests that these characteristics—HRT duration and self-evaluation of voice—warrant further examination in future work to determine what role they play in predicting individual outcomes. This study demonstrates that significant voice changes occur with behavioral voice training even after a year of HRT and even in individuals satisfied with their vocal gender. However, the subgroup comparison suggests that the extent of the changes may be mediated by length
of HRT exposure and an individual’s own level of voice satisfaction.

Clinical Implications

The evidence to support voice care for transmasculine individuals is sparse. To our knowledge, this study is the first to evaluate voice training outcomes in transmasculine individuals beyond a single-participant case study design. This study offers a first step in building evidence-based approaches to effectively support transmasculine individuals and their providers in addressing voice concerns. It specifically demonstrates that circumlaryngeal massage and laryngeal reposturing may facilitate voice masculinization in transmasculine individuals, including those who have already undergone testosterone-induced voice changes. It also demonstrates that acoustic and perceptual outcomes may vary across speech tasks for this population. Providers should therefore derive measures such as $f_0$ and perceived voice masculinity from both reading and spontaneous speech tasks to best monitor the effectiveness of any voice modification approach.

Limitations and Future Directions

This study measured short-term changes in voice and speech that were achieved in a single session of voice training. We therefore make no conclusions about the maintenance of these voice changes beyond the study session, while noting that Buckley et al. (2020) found that their participant returned to baseline for all acoustic and perceptual measures within 3 days of an identical voice training protocol. Nor can we determine what outcomes might be achieved with multiple sessions of these approaches. Future work should investigate longer term effects of these voice training techniques for transmasculine individuals and the effects of multiple sessions.

The laryngeal manipulation techniques applied in this study were administered sequentially, with circumlaryngeal massage always preceding laryngeal reposturing. This sequence was theoretically motivated, allowing for the release of baseline laryngeal tension through massage to best facilitate the active reposturing of the larynx. However, this fixed order did not allow us to differentiate the relative contributions of each approach to the acoustic and perceptual changes observed. Although no statistically significant changes were identified after circumlaryngeal massage, this component may play an important role in facilitating the outcomes achieved with laryngeal reposturing.

Finally, our selection of acoustic and perceptual measures was motivated by transmasculine individuals’ reported voice concerns (Azul et al., 2021), documented outcomes of circumlaryngeal massage and laryngeal reposturing (Dagli et al., 2008; Roy & Ferguson, 2001; Roy et al., 2017), and...
relationships between acoustics and listener perceptions of gender and masculinity (Avery & Liss, 1996; Gelfer & Bennett, 2013; Hillenbrand & Clark, 2009; Munson, 2007). Self-perceptions of voice were collected only at baseline. Thus, we could not directly determine how circulatoryngeal massage and laryngeal reposturing may affect important patient-reported outcome measures such as voice satisfaction and voice–gender congruence. Given the importance of voice to transmasculine individuals (Hodges-Simeon et al., 2021), the relationships between acoustic characteristics and voice satisfaction (Deuster et al., 2016; Nygren et al., 2016), and the prevalence of self-reported voice-gender incongruence in this population (Kennedy & Thibeault, 2020), future work should evaluate how patient-reported outcome measures such as voice satisfaction (Deuster et al., 2016; Nygren et al., 2021), the relationships between acoustic characteristics and voice function in gender diverse people assigned female at birth. Journal of Voice, 35(4), 662.e15–662.e34. https://doi.org/10.1016/j.jvoice.2020.01.001


