Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Short-Term Acoustic Effects of Speech Therapy in Transgender Women: A Randomized Controlled Trial

Clara Leyns,[a] Julie Daelman,[a] Anke Adriaansen,[a] Peter Tomassen,[b] Dominique Morsomme,[c] Guy T'Sjoen,[d,e] and Evelien D'haeseleer[a,f]

[a]*Center for Speech and Language Sciences (CESLAS), Department of Rehabilitation Sciences, Ghent University, Belgium.*
[b]*Department of Head and Neck Surgery, Ghent University Hospital, Belgium.*
[c]*Department of Speech and Language Therapy, University of Liège, Belgium.*
[d]*Department of Endocrinology, Ghent University Hospital, Belgium.*
[e]*Center for Sexology and Gender, Ghent University Hospital, Belgium.*
[f]*Department of Otorhinolaryngology, Ghent University Hospital, Belgium.*

Abstract

Purpose: This study measured and compared the acoustic short-term effects of pitch elevation training (PET) and articulation-resonance training (ART) and the combination of both programs, in transgender women.

Method: A randomized controlled study with cross-over design was used. Thirty transgender women were included and received 14 weeks of speech training. All participants started with 4 weeks of sham training; after which they were randomly assigned to one of two groups: One group continued with PET (5 weeks), followed by ART (5 weeks); the second group received both trainings in opposite order. Participants were recorded 4 times, in between the training blocks: pre, post 1 (after sham), post 2 (after training 1), and post 3 (after training 2). Speech samples included a sustained vowel, continuous speech during reading, and spontaneous speech and were analyzed using Praat software. Fundamental frequency (f_o), intensity, voice range profile, vowel formant frequencies ($F_{1\text{-}2\text{-}3\text{-}4\text{-}5}$ of /a/-/i/-/u/), formant contrasts, vowel space, and vocal quality (Acoustic Voice Quality Index) were determined.

Results and Conclusions: Fundamental frequencies increased after both the PET and ART program, with a higher increase after PET. The combination of both interventions showed a mean increase of the f_o of 49 Hz during a sustained vowel, 49 Hz during reading, and 29 Hz during spontaneous speech. However, the lower limit (percentile 5) of the f_o during spontaneous speech did not change. Higher values were detected for $F_{1\text{-}2}$ of /a/, F_3 of /u/, and vowel space after PET and ART separately. $F_{1\text{-}2\text{-}3}$ of /a/, $F_{1\text{-}3\text{-}4}$ of /u/, vowel space, and formant contrasts increased after the combination of PET and ART; hence, the combination induced more increases in formant frequencies. Intensity and voice quality measurements did not change. No order effect was detected; that is, starting with PET or ART did not change the outcome.

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Some transgender individuals experience that their voice, speech, and communication may not be congruent with their personal gender presentation and identity, which can negatively impact integration in society and psychosocial functioning (Adler et al., 2018; Colton & Casper, 1996; Davies et al., 2015; Kennedy & Thibeault, 2020; Mills et al., 2017). Some transgender individuals try to self-modificate their voice or can achieve a satisfying voice through gender-affirming hormone treatment, for example, in case of testosterone treatment. If they are not able to do so, voice and communication training by a speech-language pathologist (SLP) is the intervention of choice to develop a more feminine or masculine communication (Adler et al., 2018; E. Coleman et al., 2012; Davies et al., 2015). More often, transgender individuals who were presumed male at birth (transgender women) seek help from these health care providers as the hormone therapy does not affect their voice after puberty (Gooren, 2005; Gray & Courey, 2019; Hancock & Garabedian, 2013; Quinn & Swain, 2018).

Looking at the voice and communication training for transgender women, feminization can include several aspects of the speech such as adjusting the speaking fundamental frequency (f_o), f_o range, intonation patterns, loudness, vocal quality, and resonance (Dacakis, 2002; Leyns, Papeleu, et al., 2021). A systematic review by Leung et al. (2018) showed that the aspects that are salient in listener perceptions of speaker gender are primarily f_o of the voice and secondly resonance characteristics. The results of this systematic review with meta-analysis suggested that the f_o of the voice contributes for 41.6% of the variance in gender perception. This percentage suggests that listeners' perceptions may not change from male to female or masculine to feminine by altering pitch alone.

Besides the speaking f_o, resonance has been reported as the second most widely studied vocal domain concerning listener perceptions of speaker gender. Hardy et al. (2020) found that vowel formant frequencies were identified as significant predictors of masculinity-femininity ratings. A recent study by Leung et al. (2021) highlighted a substantially larger contribution of vowel formant frequencies to listener perceptions of speaker gender and vocal femininity-masculinity relative to speaking f_o than has previously been reported. Resonance depends on the length and shape of the vocal tract, which can be altered to change the formant frequencies (De Bodt et al., 2015; Meister et al., 2017).

Mean formant frequency values for cisgender males are 20% lower than those of cisgender females (R. O. Coleman, 1983). A study by Günzburger (1995) mentioned that the differences in formant frequencies are too large to be generated by merely anatomical characteristics, that is, smaller female physical dimensions of the cavities of the head and neck (R. O. Coleman, 1971). When cisgender men and women are asked to imitate masculinity, formant frequencies ($F_{1\text{-}2\text{-}3\text{-}4}$) dropped, inducing a smaller vowel space (Cartei et al., 2012). The opposite happened when imitating female voices. Gallena et al. (2018) investigated gender perception of the voice after increasing both the f_o and formant frequencies. If the f_o is in the gender-ambiguous zone (150185 Hz; Mordaunt 2006), the voice of transgender women is nevertheless often perceived as that of cisgender men when the formant frequencies are still in the male area.

It is possible to determine several therapy goals based on the systematic review of Leung et al. (2018), pointing out the main ingredients of listening perception of speaker gender. However, it is

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

not yet clear whether voice and communication training focusing on these goals is successful; that is, transgender women sound more feminine after the intervention and are satisfied with the outcome. The review by Leyns, Papeleu, et al. (2021) researched the outcome of speech feminization therapy for transgender women and found 14 studies. Most therapy programs focused on pitch and/or resonance. However, due to the heterogeneity and vaguely described speech training programs, the outcomes of these studies are hard to interpret. Most studies included weekly 60-min sessions, and the number of sessions ranged from one to 90.

Concerning the outcome of therapy programs reported in the review by Leyns, Papeleu, et al. (2021), positive results have been described. f_o of sustained vowels increased between 4 Hz and 110 Hz, 14-71 Hz during reading, and 159 Hz during spontaneous speech (Gelfer & Van Dong, 2013; Hancock & Garabedian, 2013; Hancock & Helenius, 2012; Kaye et al., 1993; Meszáros et al., 2005; Mount & Salmon, 1988; Quinn & Swain, 2018). It is important to acknowledge that in general, most f_o post measurements of the studies that were focused on raising the pitch are still in the gender-ambiguous zone (150-185 Hz; Mordaunt 2006). Not only raising the speaking pitch to a value higher than 180 Hz is necessary in order to be perceived female during gender perception. The findings in the systematic review by Leung et al. (2018) suggested that speaking in the range of 140 Hz as a lower limit and 300 Hz as an upper limit would also contribute to listener perceptions that the speaker is female. However, not a lot of studies described frequency range characteristics.

Looking at the training outcome results of the formant frequencies, increases have been observed for the first three formant frequencies, but not for all vowels and not each formant in each study (Carew et al., 2007; Gelfer & Tice, 2013; Hancock & Helenius, 2012; Mount & Salmon, 1988). Consequently, it might be possible that resonance outcomes can be altered to support a more feminine perception of the voice. A recent study by Leyns, Corthals, et al. (2021) discovered that there were subtle changes in formant frequencies after a single session of 30 min performing lip spreading and cork exercises, but not in every formant frequency or vowel. However, the authors concluded that further research was needed with a more extensive therapy program.

Preliminary results reporting speech therapy with transgender women are promising and suggest that voice and communication training could result in vocal changes (Carew et al., 2007; Dacakis, 2000; Gelfer & Tice, 2013; Gelfer & Van Dong, 2013; Meszáros et al., 2005; Söderpalm et al., 2004; Van Borsel et al., 2000). The quality and risk of bias of these effectiveness studies were measured in the review by Leyns, Papeleu, et al. (2021) with the QUALSYST Tool (Kmet et al., 2004). The tool revealed total summary scores ranging from 18% to 82%, with a mean score of 49%. The poor quality of these studies makes it hard to interpret the acoustic and perceptual effects of the interventions, caused by methodological issues such as retrospective study designs (Dacakis, 2000; Hancock & Garabedian, 2013; Söderpalm et al., 2004), small samples sizes (Bralley et al., 1978; Hancock & Helenius, 2012; Kalra, 1978; Kaye et al., 1993; Mount & Salmon, 1988; Quinn & Swain, 2018), heterogeneous study populations (Söderpalm et al., 2004), vaguely described therapy contents (Dacakis, 2000; Söderpalm et al., 2004), one-dimensional approaches of voice assessment (Dacakis, 2000; Gelfer & Van Dong, 2013; Mészáros et al., 2005; Söderpalm et al., 2004), and risk for experimenter bias (Carew et al., 2007; Dacakis, 2000; Gelfer & Tice, 2013; Gelfer & Van Dong, 2013;

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Hancock & Helenius, 2012; Mészáros et al., 2005; Söderpalm et al., 2004). There is a need for well-designed prospective randomized controlled trials, including bigger sample sizes, well-described therapy contents, multidimensional voice assessments, and blinded investigators. Additionally, few studies mentioned the view of transgender individuals themselves on their voice or their needs in speech therapy. It would be interesting to include those perspectives as well in future research.

The aim of this study was to measure and compare the acoustic short-term effects of pitch elevation training (PET) and articulation-resonance training (ART) in transgender women using a randomized sham-controlled trial. This aim can be divided into two purposes: firstly, the impact of each program separately, that is, sham training, PET, and ART and, secondly, the impact of the combination of all training programs. It was hypothesized that the PET would cause an increase in the f_o measures during a sustained vowel, reading, and spontaneous speech (Gelfer & Van Dong, 2013; Hancock & Garabedian, 2013; Hancock & Helenius, 2012; Kaye et al., 1993; Meszáros et al., 2005; Quinn & Swain, 2018). ART would increase the formant frequencies and, therefore, the vowel space as well (Carew et al., 2007).

METHOD

This research project was completed according to the Declaration of Helsinki and approved by the Ethics Committee of the Ghent University Hospital with the following registration number: B670201941335. This trial has been registered on ClinicalTrials.gov, a resource provided by the U.S. National Library of Medicine. Its unique identifier is NCT04708600. The CONsolidated Standards Of Reporting Trials Non-Pharmacologic Treatment Interventions checklist was used to report the intervention specifics (Boutron et al., 2017) and was added in Appendix A. A written informed consent was signed by each participant.

PARTICIPANTS

Thirty-five transgender women were initially included in the study. Five participants dropped out in the course of the project, due to the weekly commute to the clinic (n = 2), a change of mind about therapy expectations (e.g., impossible to practice at home and preference for phonosurgery instead, *n* = 2), and an unexpected move to another country (n = 1). Thus, a total of 30 transgender women were included. They were recruited through the gender team of the Ghent University Hospital (Belgium). All participants had not yet received any speech training to feminize the voice. Inclusion criteria were an established diagnosis of gender dysphoria and female gender identity confirmed by the interdisciplinary gender team at the Ghent University Hospital (Belgium), age between 18 and 70 years, selfreported normal hearing, Dutch speaker, with genderaffirming hormonal treatment (both estrogens and antiandrogens, or after orchidectomy), a female gender role, and seeking voice feminization care. Exclusion criteria were as follows: a history of neurological disorders, previous phonosurgery or voice and communication training to feminize the voice, organic pathology of the vocal folds (observed by videolaryngostroboscopic examination of the vocal folds), or smoking. The videolaryngostroboscopic examination included producing the sustained vowel /i/ at habitual pitch and loudness followed by a low-to-high

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

glissando and was performed by an ear, nose, and throat (ENT) doctor specialized in transgender voice care (P. T.). Participants who smoked in the past, but quit at least 1 month prior to the start of the training, were not excluded. The age of the participants ranged from 18 to 57 years, with a mean of 30.93 years (*SD* = 10.167).

STUDY DESIGN

This study used a cross-over design (see Figure 1), and participants were randomly assigned to a group and received 14 weeks (=15 hr) of speech training. All participants started with 4 weeks of sham training (1 hr 15 min per session), after which they were randomly assigned to one of two groups. Simple randomization was used where participants were assigned to one of the two groups based on the chronological recruitment (C. L.), alternating Groups 1 and 2. One group continued with 5 weeks of PET (1 hr per session), followed by 5 weeks of ART (1 hr per session), and the second group received both trainings in the opposite order. Participants were recorded 4 times during the study, in between the training blocks: pre, post 1 (after sham), post 2 (after training 1), and post 3 (after training 2). Recruitment started in September 2019 and ended in June 2021. This study aimed to (a) measure the impact of each program separately, that is, sham training (pre - post 1), pitch elevation (pre - post 2), and ART (pre - post 2), and (b) measure the impact of the combination of all training programs (pre - post 3).

SPEECH TRAINING

Figure 1. *Study design. ART = articulation-resonance training; PET = pitch elevation training*

Randomized controlled trial – short-term parallel design								
	n							
group 1	15	pre	sham	post 1	PET	post 2	ART	post 3
group 2	15	pre	sham	post 1	ART	post 2	PET	post 3
							Cross-over design	

All participants received the speech training in a sound- treated room at Ghent University Hospital. The interventions were carried out by a certified SLP (C. L.) with experience in the field of transgender voice. The clinician knew beforehand which group the participant would be assigned to. Sham training lasted for 4 weeks (5 hr), one session of 75 min per week, and included discussing vocal hygiene, anatomy, voice characteristics, nonverbal communication, relaxation, and breathing exercises. Both the PET and ART lasted for 5 weeks (5 hr), one session of 60 min per week. A detailed description of the sessions of PET and ART can be found in Appendix B. All participants received a print of all exercises performed during the sessions (bundled per therapy program, i.e., ART and PET) with further explanation in order to have some background information for them to read during home practice. Additionally, they received a homework chart (see Appendix C) where they could indicate whether they practiced or not. They were encouraged to exercise twice a day, 10 min each.

SPEECH ASSESSMENT

The speech samples of the transgender women were recorded 4 times during the study: a premeasurement, post 1 after sham training, post 2 after intervention 1, and post 3 after intervention 2. The participants were recorded in a speech lab at Ghent University Hospital with a Samson C01U Pro USB Studio Condenser Microphone, digitized at a sampling rate of 44.1 kHz and analyzed with the Praat software program for acoustic analysis (Institute of Phonetic Sciences; Boersma, 2002). The mouth-to- microphone distance was 15 cm during every recording. The calibration procedure of Maryn and Zarowski (2015) was used to calibrate the microphone. This consisted of comparing the dB intensity levels of the microphone and a sonometer after recording white noise, resulting in a calibration factor that can be used for acoustic analyses. The signal-to-noise ratio (SNR) was measured for every assessment and reported in dB. The speech assessments were conducted by experimenters (J. D., A. A.) who were certified SLPs with experience in the field of transgender voice. They were blinded to group allocation (i.e., PET or ART) and study process. Study data were collected and managed using the REDCap (Research Electronic Data Capture) tools hosted at Ghent University. REDCap (Harris et al., 2009, 2019) is a secure, web-based software platform designed to support data capture for research studies.

F_0

f_o was calculated of a sustained vowel /a/, continuous speech consisting of a phonetically balanced text, "Papa en Marloes" (Van de Weijer & Slis, 1991), and spontaneous speech answering the questions, "What is your favorite movie/series?" "What is your favorite book?" and "What do you like to do in your free time?" Mean, median, and percentiles 5-95 were measured with Praat software, using the Analyse periodicity > To Pitch > Query "mean" and "quantile" functions.

INTENSITY (SOUND PRESSURE LEVEL)

The median intensity of each speech sample was calculated in both the sustained vowel /a/, phonetically balanced text "Papa en Marloes" (Van de Weijer & Slis, 1991), three sentences each including four /a/ vowels, four /i/ vowels and four /u/ vowels, and spontaneous speech. Praat software was used as well to calculate the median intensity, "To Intensity > Query 'quantile'."

VOICE RANGE PROFILE

The voice range profile (VRP) was determined by the Computerized Speech Lab (CSL; Model 4500, KayPENTAX), using a Shure SM-48 microphone located at a distance of 15 cm from the mouth and angled at 45°. The procedure outlined by Heylen et al. (1998) was used. This assessment included determination of the highest and the lowest f_o (F-high, F-low) and intensity (I-high, I-low), and the frequency (F-range) and intensity range (I-range). Participants were instructed to produce the vowel /a/ for at least 2 s using a habitual pitch and loudness, a minimal pitch, a minimal intensity, a

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

maximal pitch, and a maximal intensity, respectively. Each production was modeled by the experimenters, and the participants received visual and verbal encouragement.

FORMANT FREQUENCIES

Vowels /a/, /i/, and /u/ were manually extracted by the investigator (C. L.) from three sentences containing four vowels of /a/, /i/, and /u/ ("Maarten luistert vaak naar de radio," "Sofie zegt dat niemand vier kopjes koffie per dag drinkt," and "Ik moet vandaag de koe geen voer geven, zei de boer"). The extraction was conducted with the program Praat using the "View and Edit" and "Annotate" pane. In order to capture vowel samples that were sufficiently long enough to perform acoustic analyses, the samples were concatenated into one sound chain per vowel. Vowels /a/, /i/, and /u/ were selected as they stand for extreme tongue and lip positions during the production of speech. Formant frequencies ($F_{1\text{-}2\text{-}3\text{-}4\text{-}5}$), vowel space (Hz^2), and formant contrasts (F_2 contrast /i/-/u/, F_1 contrast /a/-/i/, F_1 contrast /a/-/u/, in Hz) were calculated from these vowel chains. Praat software was used to automatically calculate these parameters, using formant frequency and vowel space scripts (Corthals, 2019, 2020).

ACOUSTIC VOICE QUALITY INDEX

The Acoustic Voice Quality Index (AVQI) is an objective, multiparameter approach to quantify dysphonia severity on the basis of both sustained vowels and continuous speech (Maryn et al., 2010). The AVQI consists of a weighted combination of six time-domain (i.e., shimmer local [SL], shimmer local decibels [SLdB], and harmonics-to-noise ratio [HNR]), frequency-domain (i.e., general slope of the spectrum [slope] and tilt of the regression line through the spectrum [tilt]), and quefrency-domain (i.e., smoothed cepstral peak prominence [CPPs]) measures (Maryn et al., 2010). The index is constructed as 2.571 (3.295 - 0.111 CPPs - 0.073 HNR - 0.213 SL + 2.789 SLdB - 0.032 slope + 0.077 tilt) and ranges from 0 to 10. A higher index indicates a worse vocal quality. The threshold score separating normophonic from dysphonic persons in Dutch is 2.95 (Maryn et al., 2010).

STATISTICAL ANALYSIS

SPSS 25.0 (SPSS Corp.) was used for the statistical analysis of the data. Analyses were conducted at $\alpha = .05$. A linear mixed model was used to compare the acoustic voice measurements between the groups and between measurements at pre, post 1, post 2, and post 3. Time, Group, and Time x Group interactions were specified as fixed factors. A random intercept for subjects was included, and within-group effects of time were determined using pairwise comparisons. Also, restricted maximum likelihood estimations and scaled identity covariance structures were used during the analyses. In addition, effect sizes were calculated for the Time x Group interactions using Cohen's *ds* (small [d = 0.2], medium [d = 0.5], and large [d = 0.8]), dividing the estimated mean by the standard deviation of a linear null model on the baseline data (Cohen, 1988; Feingold, 2013). The intrarater reliability of the manual vowel extraction has been examined with two-way mixed intraclass correlation coefficients (ICCs) and type absolute agreement (single measures) and interpreted with the Altman (1990) classification (ICC < .20: poor, .21-.40: fair, .41-.60: moderate,

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

.61-.80: good, .81-1.00: very good). Sample size calculation is based on a pilot study of our research group (Leyns, Corthals, et al., 2021) in 13 transgender women receiving articulation training. A sample of 12 transgender women per group will be required to detect a difference in vowel space of 65 Hz^2 using a paired-samples *t* test assuming a standard deviation of 72 Hz, a power of 80%, and a significance level of .05. These sample size calculations are performed with IBM SPSS Statistics 28. Taking into account a mean drop-out rate of 15%, the sample sizes are increased to 14 (= 12/(1-0.15)) transgender women per group.

RESULTS

ACOUSTIC ANALYSIS

For every speech assessment, the SNR was measured. Samples with an SNR below 20 were excluded *(n* = 0). The mean SNR was 37.5 dB *(SD* = 5.32 dB). The ICCs to check the intrarater reliability of the investigator who manually selected the vowels for formant frequency analyses were very good (F1 /a/: 0.989, 95% CI [0.895, 0.999]; vowel space: 0.998, 95% CI [0.987, 1.000]).

Tables 1-4 show the parameters with significant measures, and Tables D1-D4 with nonsignificant measures can be found in Appendix D. Tables 1 and 3 show the results of each separate program, that is, pitch elevation and ART, comparing pre and post 2 measurements (Purpose 1). Sham training, comparing pre and post 1 measurements, did not show any significant differences. There were no significant differences at the premeasurement between the two groups (PET and ART) for all parameters.

Tables 2 and 4 display the results of the whole training program, comparing pre and post 3 measurements (Purpose 2). The possible order effect was also described in these tables.

The baseline at the premeasurement was compared between groups. The median intensity during spontaneous speech was the only parameter that was significantly higher *(p* = .011) in the group starting with PET compared to the group starting with ART.

F_0

When looking at the separate training programs (see Table 1), significant differences have been found for several parameters. The f_o of the sustained vowel /a:/ (mean, median), reading (mean, median, percentile 5 [pc5]), and spontaneous speech (mean, median) was found significantly higher at post 2 compared to premeasurements. Percentile 95 (pc95) during reading, however, only significantly increased for PET, and pc5 during spontaneous speech did not significantly increase for none of the programs. PET caused an increase of 47, 46, and 27 Hz of the median f_o of the sustained vowel, reading, and spontaneous speech, respectively. ART on the other hand induced an increase of 24, 25, and 12 Hz, respectively. For most parameters, the increase of PET was significantly higher than the increase caused by ART. Effect sizes (Cohen's *d)* were calculated for all significant differences, which were all large for PET, and medium to large for ART.

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

The combination of both programs (see Table 2) caused a significant increase in f_o for all parameters, except for pc5, comparing the pre and post 3 results. The median sustained vowel, reading, and spontaneous speech was raised with 49, 49, and 29 Hz, respectively. All significant differences had large effect sizes. There was no order effect detected; that is, people starting with PET or ART ended with the same outcome. On Figure 2, the evolution of both groups over time can be observed.

INTENSITY (SOUND PRESSURE LEVEL)

Concerning the outcome of the separate programs (see Table 1), no differences were found; however, the combination of both (see Table 2) induced a significant higher intensity (+3.9 dB) of the sustained vowel (medium effect size). An order effect was detected for the median intensity during spontaneous speech; that is, people starting with PET or ART ended with a different outcome.

VRP

The VRP showed a significantly higher F-high after PET, with a mean increase of 112 Hz (medium effect size). Comparing the programs, the ART showed a significantly higher difference in F-low compared to PET (see Table 1).

The combination (see Table 2) did not result in any significant differences in the VRP, and no order effect was observed.

FORMANT FREQUENCIES

The separate programs PET and ART (see Table 3) caused an increase of both F_1 and F_2 of /a/ and F_3 of /u/. Additionally, the vowel space significantly increased with 59 and 65 Hz^2 after PET and ART, respectively. The F_1 /a/-/i/ and /a/-/u/ contrast was significantly higher at post 2 for ART and F_2 /i/-/u/ for PET. These significant differences had medium effect sizes.

Concerning the combination of both programs (see Table 4), F_1, F_2, and F_3 of /a/; F_1, F_3, and F_4 of /u/; the vowel space; and all three formant contrasts increased significantly. Also, for formant frequencies, no order effect was observed in this study. Small to large effect sizes were detected. On Figure 3, the outcome of the vowel space can be observed for both groups.

VOCAL QUALITY

Regarding the vocal quality, captured by the AVQI, no significant differences could be detected for either the separate programs or the combination of both (see Tables 1 and 2). Starting with PET or ART did not result in a different outcome, that is, the order effect.

The premeasurement demonstrated a mean AVQI of 3.1, which is above the threshold score of 2.95 (in Dutch) and indicates a dysphonic voice. However, after the whole program, this value dropped to a mean of 2.7, which illustrates a normal vocal quality.

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

DISCUSSION

This study aimed to measure the acoustic short-term effects of pitch elevation and ART in transgender women using a randomized sham-controlled trial. Firstly, the effects of each program were described, that is, sham training (pre - post 1), pitch elevation (pre - post 2), and ART (pre - post 2). Secondly, the combination of all programs was investigated (pre - post 3).

In total, 30 transgender women completed the full program. Sham training did not show any significant differences for any of the acoustic parameters. This corresponds with the objective of this sham training and the hypothesis that the sham training does not change voice characteristics. Unlike drug trials and some medical interventions, voice therapy trials cannot easily blind participants to the treatment they receive or trigger placebo effects (Bos-Clark & Carding, 2011). This finding is valuable in a research context, as it also means that the voice parameters that are known to have an effect on gender perception will not spontaneously change through contact and a therapist relationship. In other words, we may cautiously conclude that there is no placebo effect in speech feminization training, though a study with a bigger sample size should confirm this statement. To what extent the participants naturally expect an effect from these exercises remains an open question, compared with the Hawthorne Effect (Fernald et al., 2012), although this sham training had active ingredients (e.g., breathing exercises, nonverbal communication training). Consequently, researchers might be able to include these topics in sham training for voice modification for transgender persons in a research context.

Concerning the effects on the f_o parameters, significant differences were found, for both the separate programs as the combination of both. All f_o parameters changed during PET, except for pc5 of spontaneous speech. The same happened during ART, although pc95 of reading did not change either. It is surprising that the f_o changed during both PET and ART, as the ART did not focus on raising the pitch. The higher f_o values may have been induced by extra laryngeal height, trained during larynx elevation in ART, since this usually implies greater vocal fold tension. This result relates to the finding of a study by Carew et al. (2007) and Leyns, Corthals, et al. (2021). They observed an increased f_o as well, even though their participants did not receive any PET before the start of their study. They reported this as a side effect of participants modeling the voice of a female clinician during speech modification sessions. As a substantial amount of the therapy administered during tasks involved repetition of a clinician model, it is possible that these imitative tasks resulted in participants increasing their f_o to more closely resemble that of the clinician, as a part of a convergence process. Carew et al. (2007) mentioned that for some transgender women, targeting an increase in pitch may not be necessary, because an increase of 30 Hz may occur incidentally while targeting other aspects of voice, such as oral resonance. Looking at the difference between ART-PET, the increase in f_o after PET was higher for most parameters. Figure 2 shows a steep increase in f_o for PET at the post 2 measurement, ending at post 3 with the same outcome as ART. This was confirmed in the statistics when looking at the combination of both programs. These results showed that there was no order effect between the programs and all parameters increased at post 3, except for pc5 of spontaneous speech. The mean increase was 49, 49, and 29 Hz for the sustained vowel /a/, reading, and spontaneous speech,

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

respectively. In the systematic review by Leyns, Papeleu, et al. (2021), an increase of the f_o of a sustained vowel was reported between 4 and 100 Hz (Hancock & Garabedian, 2013; Hancock & Helenius, 2012; Mount & Salmon, 1988). f_o during reading was between 14 and 71 Hz (Carew et al., 2007; Gelfer & Tice, 2013; Gelfer & Van Dong, 2013; Hancock & Garabedian, 2013; Hancock & Helenius, 2012; Kaye et al., 1993; Meszáros et al., 2005; Quinn & Swain, 2018; Söderpalm et al., 2004) and spontaneous speech between 1 and 59 Hz (Bralley et al., 1978; Dacakis, 2000; Gelfer & Tice, 2013; Gelfer & Van Dong, 2013; Hancock & Garabedian, 2013; Hancock & Helenius, 2012; Kalra, 1978; Quinn & Swain, 2018). The median f_o during the sustained vowel, reading, and spontaneous speech at post 3 in this study was 176 (*SD* = 34.0 Hz), 167 (*SD* = 25.6 Hz), and 145 Hz (*SD* = 20.1 Hz), respectively, which are still in the gender-ambiguous zone, that is, 150-185 Hz (Mordaunt, 2006). Eight out of 30 participants did not reach the gender-ambiguous zone during a sustained vowel at post 3. Eight and even 20 participants did not reach that zone during reading and spontaneous speech, respectively. The findings in the systematic review by Leung et al. (2018) showed that not only raising the speaking pitch to a value higher than 180 Hz is necessary in order to be perceived female during gender perception. They suggested that speaking in the range of 140 Hz as a lower limit and 300 Hz as an upper limit would also contribute to listener perceptions that the speaker is female. The mean lower limit (pc5) and upper limit (pc95) in this study is 117 and 242 Hz, respectively, during reading, which might implicate that the increase in f_o would not be enough to be perceived as female by listeners.

The PET and ART showed increased fundamental frequencies, both separately and the combination of the two programs. A 20-Hz difference was observed between the increase in spontaneous speech and the sustained vowel and reading, indicating that the spontaneous speech was harder to increase compared to the other speech tasks. Iwarsson (2015) described nine factors relevant to behavioral learning and changing voice behavior habits, such as cue-altering, attention exercises, and repetition. It would be interesting to spend more time on the generalization process during voice modification training, including all nine factors. Furthermore, the 5th percentile of the f_o during spontaneous speech did not significantly differ, for each program separately, nor the combination of programs. In the future, more time should be spent on generalization of the elevated pitch, as the lower limit of f_o is important for gender perception.

The intensity measurements did not significantly change after one of the programs; however, after the combination of both, the intensity of the sustained vowel significantly increased with 3.9 dB. Participants were accustomed to the speech assessment by the time they were at the post 3 measurement. They might have felt more comfortable to produce a sustained vowel and therefore produced a louder vowel than they did at the premeasurement. Additionally, forward resonance and articulation were trained during the ART; that is, more forward projection might have been used while producing the vowel /a/, although this effect was not observed in reading nor spontaneous speech. During spontaneous speech, an order effect was detected; that is, starting with PET or ART resulted in a different outcome. This can be caused by the significant baseline difference at the premeasurement. This is the first study, to our knowledge, to include intensity measurements during both a sustained vowel, reading, and spontaneous speech.

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

The VRP was examined by means of the CSL (Model 4500, KayPENTAX). A significantly higher F-high at post 2 was observed after the PET. Elevating the pitch by using various glissando patterns (see Appendix B for details) during the training might have caused this increase. Hancock and Helenius (2012) and Hancock and Garabedian (2013) saw an increase of the F-high of 52 and 114 Hz, respectively. Meszáros et al. (2005) examined the pitch and intensity range as well, observing a diminished pitch range caused by the elevation of the lower limit, and an increased intensity range, although no statistical data were reported. There was a significant difference between PET-ART for the F-low; that is, the F-low increases with 8 Hz during ART but decreases 2 Hz during PET. None of the parameters changed significantly after the combination of both programs. The observed effect after one of both programs is therefore temporary. The combination of both programs does not influence the voice capacity in terms of frequency and intensity range.

Looking at the formant frequencies of this study, derived from extracted vowels /a/, /i/, and /u/, some significant differences were observed. F_1 and F_2 of vowel /a/ and F_3 of vowel /u/ and vowel space significantly increased after both the PET and ART separately. These changes could be expected from the ART, as the F_1, F_2, and F_3 are correlated with a bigger jaw drop, forward tongue position, and lip spreading, and these aspects were trained during the ART. However, these formants also increased during PET. It could be possible that the participant tried to model the clinician as well, for example, used more lip spreading and clear speech during the pitch elevation exercises, although this was not specifically trained during these sessions. The combination of PET and ART caused a significantly higher $F_{1\text{-}2\text{-}3}$ of vowel /a/, $F_{1\text{-}3\text{-}4}$ of vowel / u/, vowel space, and all formant contrasts. As previously mentioned, higher $F_{1\text{-}2\text{-}3}$ values could be expected, as these physiological correlates were trained during the sessions. Hypothetically, the significantly higher F_4 can be related to a possible shortening of the laryngeal cavity or laryngeal height (Takemoto et al., 2006). Other studies found higher F_1 values during /a/ and /o/ (Carew et al., 2007), /i/ (Gelfer & Tice, 2013), all vowels (Hancock & Helenius, 2012), higher F_2 values during /a/ (Carew et al., 2007), all vowels (Hancock & Helenius, 2012), and higher F_3 for all three vowels (Carew et al., 2007). The results of this study are not completely in line with previous research, as no significant differences were found for vowel /i/. This study extracted the vowels from running speech, rather than performing acoustic analyses on isolated vowels. When extracting vowels from a phonetically balanced text that is being read aloud, they are more representative for actual daily communication. Previous research included either extracted vowels (Carew et al., 2007; Gelfer & Tice, 2013; Gelfer & Van Dong, 2013) or isolated vowels (Hancock & Helenius, 2012; Mount & Salmon, 1988). After PET and ART, some formant frequencies changed, but not every formant nor for every vowel. During the ART, the participants learned one technique per week (see Appendix B for details), that is, 1 hr of lip spreading, 1 hr of forward tongue position, and so forth. This might not have been enough for the participants to really master all of the techniques.

Feminization of the voice during voice and communication training includes both adjustment of several voice characteristics (Dacakis, 2002; Leyns, Papeleu, et al., 2021) and prevention of vocal hyperfunction (Palmer et al., 2012). Therefore, it is crucial to measure and examine vocal quality. Vocal quality measurements were described by the AVQI (Maryn et al., 2010). During the premeasurement, dysphonic values were found, which could be caused by nervousness as it was

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

the start of the study and the fact that they are not accustomed to speaking in a microphone. However, their level of stress was not measured in this study. For both the separate programs, as the combination of both, no significant differences were detected. It should be mentioned, although not significant, that the AVQI improved slightly; that is, the AVQI decreased to normophonic values (cutoff score is 2.95). This is the first study that implemented the AVQI in an effectiveness study for voice feminization. It can be concluded that there is no decline in vocal quality after this program and therefore both PET and ART could be considered a safe approach to feminize speech.

It seems that the combination of ART and PET is important to reach more differences in vocal characteristics. No order effect was observed; hence, we might conclude that it does not matter whether you start with pitch training or articulation-resonance therapy. This is an interesting finding for clinical application. Effect sizes for the significantly changed parameters were medium to large, indicating that these effects are clinically meaningful. This is the first effectiveness study to include effect sizes. Therefore, it is hard to compare these results with previous studies. Due to the combination of programs and the combination of several articulation techniques, it is unclear which component contributes to which vocal outcome. Future research that investigates each separate component is necessary to determine this. In this study, the effect of the combination of techniques is measured.

This study is the first randomized controlled trial in this research field. A previous review on the effects of speech therapy for transgender women concluded that there is a need for studies with randomized controlled trial designs, blinded investigators, bigger sample sizes, and well-described therapy programs (Leyns, Papeleu, et al., 2021). The study was designed to meet these requirements by using a sham training and, therefore, control participants, investigators blinded to group allocation and study process, a sample size of 30 participants, and a detailed description of the therapy content. However, the clinician who provided intervention was not blinded to group allocation during recruitment, as simple randomization, that is, alternating Groups 1 and 2, was used. This could have caused bias during the recruitment process. Future research should also include long-term acoustic follow-up measurements and perceptual effects of both listeners and participants.

For clinicians, SLPs, researchers, ENT doctors, or clients, it is crucial to know the effects of speech therapy. It is also important to understand that every transgender person has an individual need and differences in training effects can be observed. Some might want to train different targets, some undergo phonosurgery, and some choose to have no intervention (Nolan et al., 2019). However, clinical trials are necessary to determine which therapy program or specific exercise contributes to which acoustic effect. Future research should focus on the perceptual effects both by listeners and participants, and which factors can predict progress in speech training.

Conclusions

This study demonstrated that the separate programs of PET and ART can increase fundamental frequencies and formant frequencies of transgender women. The sham training did not cause any significant differences. The mean increase was 49, 49, and 29 Hz for the f_o during sustained vowel

/a/, reading, and spontaneous speech, respectively, after the whole program. Pc5 of the f_o during spontaneous speech did not increase. $F_{1\text{-}2\text{-}3}$ of /a/, $F_{1\text{-}3\text{-}4}$ of /u/, vowel space, and formant contrasts increased after the combination of PET and ART, which is more than the observed effects after each separate program. In order to achieve more differences in formant frequencies, it is needed to include both programs. Long-term follow-up and perceptual effects should be included in future research.

Data Availability Statement

The data sets generated during and/or analyzed during this study are not publicly available due to ethical reasons but are available from the corresponding author on reasonable request.

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Table 1. *f_o, intensity, and vocal quality measurements of each separate program*

	PET					ART					PET vs. ART
Variable	Pre *(M, SD)*	Post 2 (M, SD)	Mean difference	*p* value	Cohen's *d*	Pre (M, SD)	Post 2 (M, SD)	Mean difference	p value	Cohen's d	*p* value
f_o (Hz)											
Sustained vowel (M)	122.4, 31.4	170, 38.56	47.6	< .001	1.25	129.7, 29.03	154.1, 28.37	24.4	< .001	0.64	.008
Sustained vowel *(Mdn)*	123.3, 31.09	170.3, 38.13	47.0	< .001	1.22	129.7, 29.04	154.1, 28.35	24.4	< .001	0.63	.010
Reading *(M)*	127.4, 17.6	172.3, 25.47	44.9	< .001	1.45	124.1, 21.81	150.9, 20.25	26.8	< .001	0.86	.012
Reading *(Mdn)*	118.2, 17.14	164.6, 29.44	46.4	< .001	1.50	117.6, 19.76	142.8, 19.78	25.2	< .001	0.82	.005
Reading (pc5)	90.1, 16.41	112, 14.89	21.9	< .001	1.04	94.6, 19.83	108.1, 17.66	13.5	.002	0.64	.122
Reading (pc95)	168.5, 32.19	240.1, 42.17	71.6	< .001	1.23	170, 79.6	205.8, 42.21	35.8	.201		.069
Spontaneous (M)	119.9, 15.52	148.3, 21.13	28.4	< .001	1.27	123.5, 20.29	134.5, 20.27	11.0	.002	0.49	.001
Spontaneous *(Mdn)*	114.4, 14.58	141.5, 21.29	27.0	< .001	1.21	117.3, 19.32	128.9, 20.11	11.6	.001	0.52	.002
Spontaneous (pc95)	155.7, 21.92	199.4, 29.93	43.7	< .001	1.29	155.7, 28.6	173.5, 27.05	17.7	.008	0.52	.002
Voice range profile (dB and Hz)											
F-low	86.9, 10.93	85.1, 14.26	1.9	1.000		91.5, 20.27	99.9, 26.19	8.4	.115		.017
F-high	529.6, 187.39	641.2, 182.27	111.6	.026	0.54	663.4, 296.52	651.4, 208.92	12.0	1.000		.058

Note. Bold data are significant p values (level α = .05 is used). *f_o* = fundamental frequency; PET = pitch elevation training; ART = articulation-resonance training; pc5 = percentile 5; pc95 = percentile 95; F-low = lowest fundamental frequency; F-high = highest fundamental frequency

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Table 2. *f_o, intensity, and vocal quality measurements of combination of ART and PET.*

Variable	Pre (M, *SD)*	Post 3 (M, SD)	Mean difference	*p* value	Cohen's *d*	Order effect *p* value
fo (Hz)						
Sustained vowel (M)	126, 29.94	175.2, 34.05	49.2	< .001	1.29	.909
Sustained vowel *(Mdn)*	126.5, 29.74	175.7, 34.04	49.3	< .001	1.28	.935
Reading *(M)*	125.7, 19.54	175, 25.44	49.2	< .001	1.59	.316
Reading *(Mdn)*	117.9, 18.18	166.8, 25.62	48.9	< .001	1.58	.391
Reading (pc5)	92.4, 18.03	116.7, 21.71	24.3	< .001	1.15	.549
Reading (pc95)	169.2, 59.66	242.3, 42.71	73.1	< .001	1.25	.435
Spontaneous (M)	121.7, 17.85	150.2, 19.73	28.5	< .001	1.27	.720
Spontaneous *(Mdn)*	115.9, 16.88	144.6, 20.09	28.7	< .001	1.29	.730
Spontaneous (pc95)	155.7, 25.04	203.5, 27.18	47.8	< .001	1.41	.479
Intensity (dB)						
Sustained vowel *(Mdn)*	71.5, 7.12	75.5, 5.77	3.9	.008	0.60	.205
Spontaneous *(Mdn)*	64, 5.7	66.5, 5.54	2.6	.079		.029

Note. Bold data are significant *p* values (level α = .05 is used). f_o = fundamental frequency; ART = articulation-resonance training; PET = pitch elevation training; pc5 = percentile 5; pc95 = percentile 95

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Table 3. *Formant frequency measurements of each separate program*

	PET					ART					PET vs. ART
Variable	Pre (M, SD)	Post 2 *(M, SD)*	Mean difference	*p* value	Cohen's *d*	Pre (*M, SD*)	Post 2 (M, SD)	Mean difference	*p* value	Cohen's *d*	*p* value
Median formant frequencies											
/a/											
F_1	705.1,91.5	772.4, 110.18	67.3	.003	0.64	709.7, 89.56	784.8, 104.73	75.2	.012	0.71	.786
F_2	1453.3, 148.33	1529.2, 158.01	75.9	.012	0.54	1481.4, 103.91	1571.9, 110.13	90.6	.003	0.65	.663
/u/											
F_3	2368.6, 153.51	2465.2, 140.8	96.6	.001	0.52	2393.2, 161.72	2523.4, 220.22	130.2	.003	0.70	.423
Vowel space (Hz^2)											
F_1-F_2	251.1,92.79	309.9, 108.72	58.9	.039	0.63	235.6, 73.62	300.4, 98.06	64.8	.002	0.69	.823
Formant contrasts (Hz)											
F_2/i/-/u/	1269.2,158.38	1388.5, 182.6	119.4	.003	0.63	1274.1, 208.77	1345.9, 218.09	71.8	.273		.315
F_1 /a/-/i/	404.5, 117.31	447.3, 117.52	42.8	.620		372.7, 89.54	456.5, 129.36	83.7	.003	0.73	.231
F_1 /a/-/u/	380.2, 94.45	426.1, 108.56	45.9	.448		365.9, 88.84	436.6, 120.19	70.7	.021	0.66	.466

Note. Bold data are significant *p* values (level α = .05 is used). PET = pitch elevation training; ART = articulation-resonance training

Table 4. *Formant frequency measurements of combination of ART and PET*

Variable	Pre (M, SD)	Post 3 (M, SD)	Mean difference	*p* value	Cohen's *d*	Order effect *p* value
Median formant frequencies						
/a/						
F_1	707.4, 88.99	796.6, 114.71	89.2	< .001	0.84	.894
F_2	1467.3, 126.64	1580.6, 145.32	113.3	< .001	0.81	.500
F_3	2450.7, 162.28	2542, 218.78	91.4	.047	0.50	.334
/u/						
F_1	308.5, 29.53	325.1, 37.75	16.7	.009	0.52	.944
F_3	2380.9, 155.43	2519.5, 190.12	138.7	< .001	0.74	.374
F_4	3510.1, 238.61	3635, 271.13	124.9	.008	0.50	.439
Vowel space (Hz^2)						
F_1-F_2	243.4, 82.67	307.5, 105.74	64.2	< .001	0.68	.430
Formant contrasts (Hz)						
F_2/i/-/u/	1271.6, 182.09	1343.9, 202.95	72.3	.017	0.38	.662
F_1/a/-/i/	388.6, 103.8	458.5, 124.06	69.8	.001	0.61	.512
F_1 /a/-/u/	373, 90.39	437.3, 122.02	64.3	.002	0.60	.466

Note. Bold data are significant *p* values (level α = .05 is used). ART = articulation-resonance training; PET = pitch elevation training

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Figure 2. *Evolution of the median f_o during a sustained vowel, reading, and spontaneous speech (Hz) for both groups. f_o = fundamental frequency.*

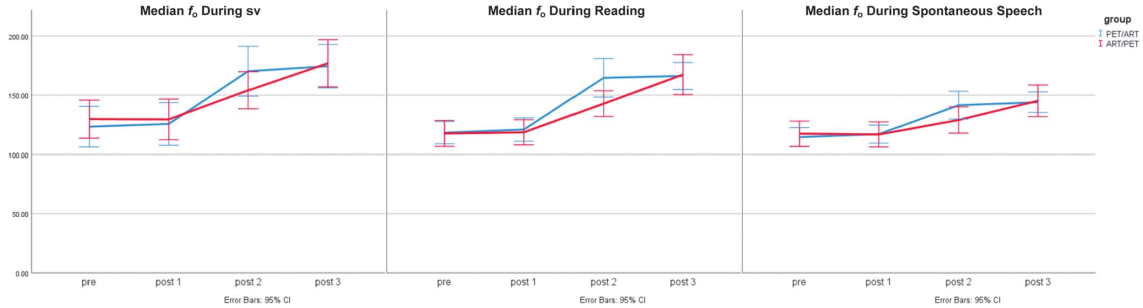
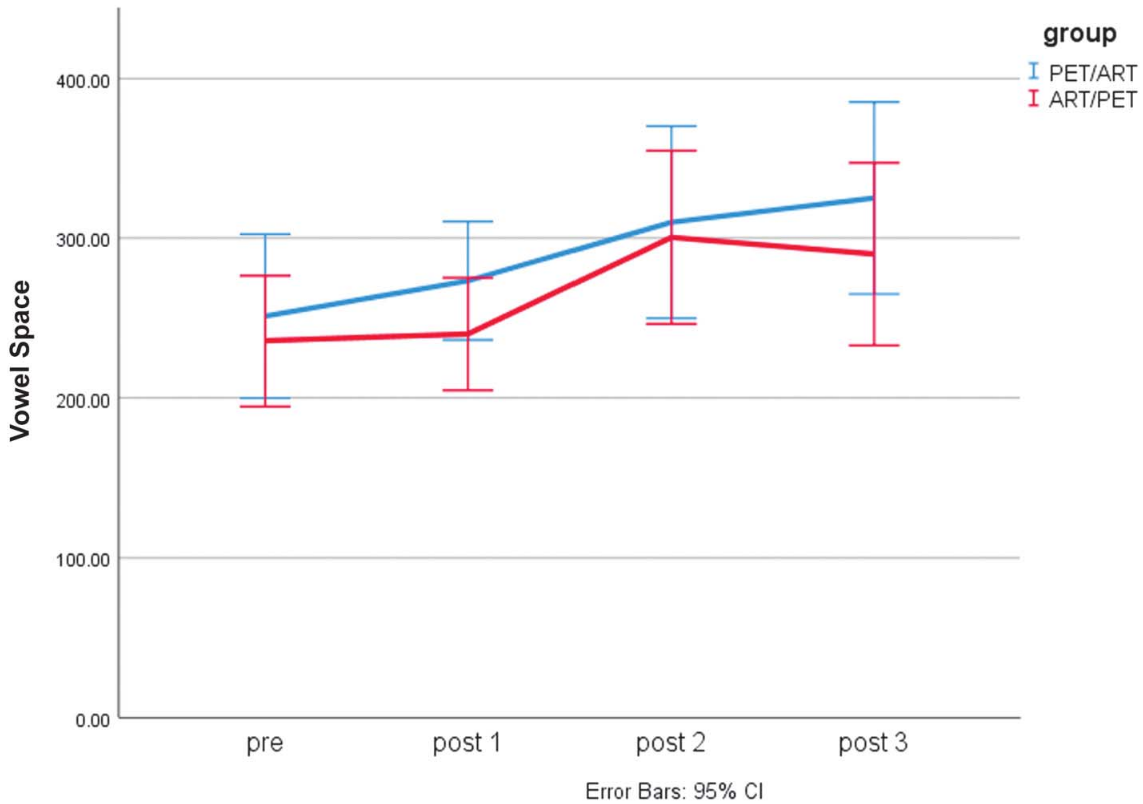


Figure 3. *Evolution of the vowel space (Hz^2) for both groups. PET = pitch elevation training; ART = articulation-resonance training; CI = confidence interval.*



Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

References

Adler, R. K., Hirsch, S., & Pickering, J. (2018). Voice and communication therapy for the transgender/gender diverse client: A comprehensive clinical guide. Plural.

Altman, D. G. (1990). Practical statistics for medical research (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/ 9780429258589

Boersma, P. P. G. (2002). Praat, a system for doing phonetics by computer. Glot International, 5.

Bos-Clark, M., & Carding, P. (2011). Effectiveness of voice therapy in functional dysphonia: Where are we now? Current Opinion in Otolaryngology & Head and Neck Surgery, 19(3), 160-164. https://doi.org/10.1097/MOO.0b013e3283448f85

Boutron, I., Altman, D. G., Moher, D., Schulz, K. F., Ravaud, P., & CONSORT NPT Group. (2017). CONSORT statement for randomized trials of nonpharmacologic treatments: A 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. Annals of Internal Medicine, 167(1), 40-47. https://doi.org/10.7326/M17-0046

Bralley, R. C., Bull, G. L., Gore, C. H., & Edgerton, M. T. (1978). Evaluation of vocal pitch in male transsexuals. Journal of Communication Disorders, 11(5), 443-449. https://doi.org/ 10.1016/0021-9924(78)90037-0

Carew, L., Dacakis, G., & Oates, J. (2007). The effectiveness of oral resonance therapy on the perception of femininity of voice in male-to-female transsexuals. Journal of Voice, 21(5), 591-603. https://doi.Org/10.1016/j.jvoice.2006.05.005

Cartei, V., Cowles, H. W., & Reby, D. (2012). Spontaneous voice gender imitation abilities in adult speakers. PLOS ONE, 7(2), Article e31353. https://doi.org/10.1371/journal.pone. 0031353

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Erlbaum.

Coleman, E., Bockting, W., Botzer, M., Cohen-Kettenis, P., DeCuypere, G., Feldman, J., Fraser, L., Green, J., Knudson, G., Meyer, W. J., Monstrey, S., Adler, R. K., Brown, G. R., Devor, A. H., Ehrbar, R., Ettner, R., Eyler, E., Garofalo, R., Karasic, D. H., ... Zucker, K. (2012). Standards of care for the health of transsexual, transgender, and gender-nonconforming people, version 7. International Journal of Transgenderism, 13(4), 165-232. https://doi.org/10.1080/ 15532739.2011.700873

Coleman, R. O. (1971). Male and female voice quality and its relationship to vowel formant frequencies. Journal of Speech and Hearing Research, 14(3), 565-577. https://doi.org/10.1044/ jshr.1403.565

Coleman, R. O. (1983). Acoustic correlates of speaker sex identification: Implications for the transsexual voice. Journal of Sex Research, 19(3), 293-295. https://doi.org/10.1080/00224498309551189

Colton, R., & Casper, J. (1996). Understanding voice problems: A physiological perspective for diagnosis and treatment. Lippincott Williams & Wilkins.

Corthals, P. (2019). Vowel space. Ghent University.

Corthals, P. (2020). Info Formanten. Ghent University.

Dacakis, G. (2000). Long-term maintenance of fundamental frequency increases in male-to-female transsexuals. Journal of Voice: Official Journal of the Voice Foundation, 14(4), 549556. https://doi.org/10.1016/s0892-1997(00)80010-7

Dacakis, G. (2002). The role of voice therapy in male-to-female transsexuals. Current Opinion in Otolaryngology & Head and Neck Surgery, 10(3), 173-177. https://doi.org/10.1097/00020840- 200206000-00003

Davies, S., Papp, V. G., & Antoni, C. (2015). Voice and communication change for gender nonconforming individuals: Giving voice to the person inside. International Journal of Transgenderism, 16(3), 117-159. https://doi.org/10.1080/15532739.2015.1075931

De Bodt, M., Heylen, L., Mertens, F., Vanderwegen, J., & Van de Heyning, P. (2015). Stemstoornissen. Handboek voor de klinische praktijk. Zesde, herziene uitgave: 2015 (Vol. 5) [Voice disorders. Guidebook for clinical practice]. Maklu.

Feingold, A. (2013). A regression framework for effect size assessments in longitudinal modeling of group differences. Review of General Psychology, 17(1), 111-121. https://doi.org/10.1037/ a0030048

Fernald, D. H., Coombs, L., DeAlleaume, L., West, D., & Parnes, B. (2012). An assessment of the Hawthorne effect in practicebased research. The Journal of the American Board of Family Medicine, 25(1), 83-86. https://doi.org/10.3122/jabfm.2012.01. 110019

Gallena, S. J. K., Stickels, B., & Stickels, E. (2018). Gender perception after raising vowel fundamental and formant frequencies: Considerations for oral resonance research. Journal of Voice, 32(5), 592-601. https://doi.org/10.1016Zj.jvoice.2017. 06.023

Gelfer, M. P., & Tice, R. M. (2013). Perceptual and acoustic outcomes of voice therapy for male-to-female transgender individuals immediately after therapy and 15 months later. Journal of Voice, 27(3), 335-347. https://doi.org/10.1016/j.jvoice. 2012.07.009

Gelfer, M. P., & Van Dong, B. R. (2013). A preliminary study on the use of vocal function exercises to improve voice in male-to-female transgender clients. Journal of Voice, 27(3), 321-334. https://doi.org/10.1016/j.jvoice.2012.07.008

Gooren, L. (2005). Hormone treatment of the adult transsexual patient. Hormone Research in Paediatrics, 64(Suppl. 2), 31-36. https://doi.org/10.1159/000087751

Gray, M. L., & Courey, M. S. (2019). Transgender voice and communication. Otolaryngologic Clinics of North America, 52(4), 713-722. https://doi.org/10.1016Tj.otc.2019.03.007

Günzburger, D. (1995). Acoustic and perceptual implications of the transsexual voice. Archives of Sexual Behavior, 24(3), 339348. https://doi.org/10.1007/BF01541604

Hancock, A. B., & Garabedian, L. M. (2013). Transgender voice and communication treatment: A retrospective chart review of 25 cases. International Journal of Language & Communication Disorders, 48(1), 54-65. https://doi.org/10.1111/j.1460-6984. 2012.00185.x

Hancock, A. B., & Helenius, L. (2012). Adolescent male-to- female transgender voice and communication therapy. Journal of Communication Disorders, 45(5), 313-324. https://doi.org/ 10.1016/j.jcomdis.2012.06.008

Hardy, T. L. D., Boliek, C. A., Aalto, D., Lewicke, J., Wells, K., & Rieger, J. M. (2020). Contributions of voice and nonverbal communication to perceived masculinity-femininity for cisgender and transgender communicators. Journal of Speech, Language, and Hearing Research, 63(4), 931-947. https://doi.org/10.1044/2019_jslhr-19-00387

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S. N., & REDCap Consortium. (2019). The REDCap consortium: Building an international community of software platform partners. Journal of Biomedical Informatics, 95, 103208. https://doi.org/10.1016/j.jbi.2019.103208 '

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of Biomedical Informatics, 42(2), 377-381. https://doi.org/10.1016/j.jbi.2008.08.010

Heylen, L., Wuyts, F. L., Mertens, F., Bodt, M. D., Pattyn, J., Croux, C., & de Heyning, P. H. V. (1998). Evaluation of the vocal performance of children using a voice range profile index. Journal of Speech, Language, and Hearing Research, 41(2), 232-238. https://doi.org/10.1044/jslhr.4102.232

Iwarsson, J. (2015). Facilitating behavioral learning and habit change in voice therapy—Theoretic premises and practical strategies. Logopedics Phoniatrics Vocology, 40(4), 179-186. https://doi.org/10.3109/14015439.2014.936498

Kalra, M. (1978). Voice therapy in the case of a transsexual. British Journal of Sexual Medicine, 5(40), 47-50.

Kaye, J., Bortz, M. A., & Tuomi, S. K. (1993). Evaluation of the effectiveness of voice therapy with a male-to-female transsexual subject. Scandinavian Journal of Logopedics and Phoniatrics, 18(2-3), 105-109. https://doi.org/10.3109/14015439309101356

Kennedy, E., & Thibeault, S. L. (2020). Voice-gender incongruence and voice health information-seeking behaviors in the transgender community. American Journal of Speech-Language Pathology, 29(3), 1563-1573. https://doi.org/10.1044/2020_ajslp-19-00188

Kmet, L. M., Cook, L. S., & Lee, R. C. (2004). Standard quality assessment criteria for evaluating primary research papers from a variety of fields. https://doi.org/10.7939/R37M04F16

Leung, Y., Oates, J., & Chan, S. P. (2018). Voice, articulation, and prosody contribute to listener perceptions of speaker

gender: A systematic review and meta-analysis. Journal of Speech, Language, and Hearing Research, 61(2), 266-297. https://doi.org/10.1044/2017_jslhr-s-17-0067

Leung, Y., Oates, J., Chan, S.-P., & Papp, V. (2021). Associations between speaking fundamental frequency, vowel formant frequencies, and listener perceptions of speaker gender and vocal femininity-masculinity. Journal of Speech, Language, and Hearing Research, 64(7), 2600-2622. https://doi.org/10.1044/2021_JSLHR-20-00747

Leyns, C., Corthals, P., Cosyns, M., Papeleu, T., Van Borsel, J., Morsomme, D., T'Sjoen, G., & D'haeseleer, E. (2021). Acoustic and perceptual effects of articulation exercises in transgender women. Journal of Voice. https://doi.org/10.10167j.jvoice. 2021.06.03 3

Leyns, C., Papeleu, T., Tomassen, P., T'Sjoen, G., & D'haeseleer, E. (2021). Effects of speech therapy for transgender women: A systematic review. International Journal of Transgender Health, 1-21. https://doi.org/10.1080/26895269.2021.1915224

Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. Journal of Voice, 24(5), 540-555. https://doi.org/10.1016Zj.jvoice.2008.12.014

Maryn, Y., & Zarowski, A. (2015). Calibration of clinical audio recording and analysis systems for sound intensity measurement. American Journal of Speech-Language Pathology, 24(4), 608-618. https://doi.org/10.1044/2015_AJSLP-14-0082

Meister, J., Kuhn, H., Shehata-Dieler, W., Hagen, R., & Kleinsasser, N. (2017). Perceptual analysis of the male-to- female transgender voice after glottoplasty-the telephone test. The Laryngoscope, 127(4), 875-881. https://doi.org/10.1002/ lary.26110

Mészáros, K., Csokonai Vitéz, L., Szabolcs, I., G6th, M., Kovacs, L., Gôrômbei, Z., & Hacki, T. (2005). Efficacy of conservative voice treatment in male-to-female transsexuals. Folia Phonia- trica et Logopaedica, 57(2), 111-118. https://doi.org/10.1159/ 000083572

Mills, M., Stoneham, G., & Georgiadou, I. (2017). Expanding the evidence: Developments and innovations in clinical practice, training and competency within voice and communication therapy for trans and gender diverse people. International Journal of Transgenderism, 18(3), 328-342. https://doi.org/10. 1080/15532739.2017.1329049

Mordaunt, M. (2006). Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide.

Mount, K. H., & Salmon, S. J. (1988). Changing the vocal characteristics of a postoperative transsexual patient: A longitudinal study. Journal of Communication Disorders, 21(3), 229238. https://doi.org/10.1016/0021-9924(88)90031-7

Nolan, I. T., Morrison, S. D., Arowojolu, O., Crowe, C. S., Massie, J. P., Adler, R. K., Chaiet, S. R., & Francis, D. O. (2019). The role of voice therapy and phonosurgery in transgender vocal feminization. The Journal of Craniofacial Surgery, 30(5), 13681375. https://doi.org/10.1097/scs.0000000000005132

Palmer, D., Dietsch, A., & Searl, J. (2012). Endoscopic and stroboscopic presentation of the larynx in male-to-female transsexual persons. Journal of Voice, 26(1), 117-126. https:// doi.org/10.1016/j.jvoice.2010.10.014

Quinn, S., & Swain, N. (2018). Efficacy of intensive voice femini- sation therapy in a transgender young offender. Journal of Communication Disorders, 72, 1-15. https://doi.org/10.1016fj. jcomdis.2018.02.001

Sôderpalm, E., Larsson, A., & Almquist, S.-Â. (2004). Evaluation of a consecutive group of transsexual individuals referred for vocal intervention in the west of Sweden. Logopedics Phoniatrics Vocology, 29(1), 18-30. https://doi.org/10. 1080/14015430310021618

Takemoto, A. S., Adachi, S., Kitamura, T., Mokhtari, P., & Honda, K. (2006). Acoustic roles of the laryngeal cavity in vocal tract resonance. The Journal of the Acoustical Society of America, 120(4), 2228-2238. https://doi.org/10.1121/1.2261270

Van Borsel, J., De Cuypere, G., Rubens, R., & Destaerke, B. (2000). Voice problems in female-to-male transsexuals. International Journal of Language & Communication Disorders, 35(3), 427-442. https://doi.org/10.1080/136828200410672

Van de Weijer, J., & Slis, I. (1991). Nasaliteitsmeting met de nas- ometer [Nasality measurements with the nasometer]. Logopedie en Foniatrie, 63, 97-101

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Appendix A *(p 1 of 3) 2017 CONSORT Checklist of Information to Include When Reporting a Randomized Trial Assessing Nonpharmacologic Treatments (NPTs)*

Section/topic item	**Checklist item no.**	**CONSORT item**	**Extension for NPT trials**	**Reported (page number)**
Title and abstract				
	1a	Identification as a randomized trial in the title		145
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT or abstracts)	*Refer to CONSORT extension for abstracts for NPT trials*	145
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale		145–147
	2b	Specific objectives or hypotheses		147
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	When applicable, how care providers were allocated to each trial group	147
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		N/A
Participants	4a	Eligibility criteria for participants	When applicable, eligibility criteria for centers and for *care providers*	147
	4b	Settings and locations where the data were collected		147–148
Interventions†	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Precise details of both the experimental treatment and comparator	147–148 + Appendixes B–C
	5a		Description of the different components of the interventions and, when applicable, description of the procedure for tailoring the interventions to individual participants.	147–148 + Appendices B–C
	5b		Details *of whether and* how the interventions were standardized.	/
	5c		*Details of whether and how adherence of care providers to the protocol was assessed or enhanced*	/
	5d		*Details of whether and how adherence of participants to interventions was assessed or enhanced*	147–148 + Appendix C
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed		148–149
	6b	Any changes to trial outcomes after the trial commenced, with reasons		N/A
Sample size	7a	*How sample size was determined*	When applicable, details of whether and how the clustering by care providers or centers was addressed	149
	7b	When applicable, explanation of any interim analyses and stopping guidelines		N/A

(Table continues)

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Appendix A *(p 2 of 3) 2017 CONSORT Checklist of Information to Include When Reporting a Randomized Trial Assessing Nonpharmacologic Treatments (NPTs)*

Section/topic item	Checklist item no.	CONSORT item	Extension for NPT trials	Reported (page number)
Randomization:				
- Sequence generation	8a	Method used to generate the random allocation sequence		147
	8b	Type of randomization; details of any restriction (such as blocking and block size)		147
- Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned		N/A
- Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions		147
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	If done, who was blinded after assignment to interventions (e.g., participants, care providers, *those administering co-interventions,* those assessing outcomes) and how	148
	11b	If relevant, description of the similarity of interventions		N/A
	11c		*If blinding was not possible, description of any attempts to limit bias*	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	When applicable, details of whether and how the clustering by care providers or centers was addressed	149
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		149

(Table continues)

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Appendix A *(p 3 of 3) 2017 CONSORT Checklist of Information to Include When Reporting a Randomized Trial Assessing Nonpharmacologic Treatments (NPTs)*

Section/topic item	Checklist item no.	CONSORT item	Extension for NPT trials	Reported (page number)
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome	The number of care providers or centers performing the intervention in each group and the number of patients treated by each care provider or in each center	147
	13b	For each group, losses and exclusions after randomization, together with reasons		147
	13c		*For each group, the delay between randomization and the initiation of the intervention*	/
	new		Details of the experimental treatment and comparator as they were implemented	Appendix B
Recruitment	14a	Dates defining the periods of recruitment and follow-up		/
	14b	Why the trial ended or was stopped		/
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centers (volume) in each group.	147–148
Numbers analyzed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups		N/A
Outcomes and estimation	17a	*For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)*		Tables 1-2-3-4 and additional tables in Appendix D
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		N/A
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory		Tables 1-2-3-4 and additional tables in Appendix D
Harms	19	All important harms or unintended effects in each group (for specific guidance, see CONSORT for harms)		N/A
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centers in each group	151–156 (throughout Discussion section)
Generalizability	21	Generalizability (external validity, applicability) of the trial findings	Generalizability (external validity) of the trial findings according to the intervention, comparators, patients, and care providers and centers involved in the trial	156
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		151–156 (throughout Discussion section)
Other information				
Registration	23	Registration number and name of trial registry		147
Protocol	24	Where the full trial protocol can be accessed, if available		N/A
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		N/A

Note. Additions or modifications to the 2010 CONSORT checklist have been made. Modifications of the extension appear in italics. CONSORT = Consolidated Standards of Reporting Trials; N/A = not applicable.
[*]Items 5, 5a, 5b, 5c, and 5d are consistent with the Template for Intervention Description and Replication checklist)

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Appendix B *(p. 1 of 4)*

Therapy Program

1. **Pitch elevation training**

   a. Session 1

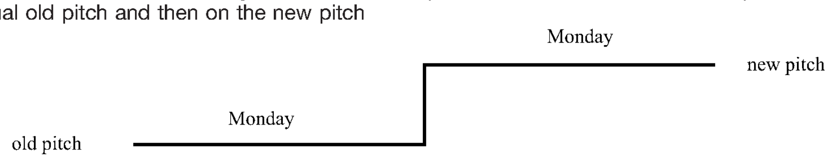
      i. Auditory discrimination with a piano
      ii. Glissando patterns (using biofeedback real-time pitch of Computerized Speech Lab (Kay Elemetrics)

         1. From habitual 'old' pitch to the 'new', higher pitch in isolated nasal consonants (approximately till 160 Hz); from habitual old pitch to highest pitch, etc.

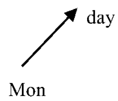
      iii. Adding consonant–vowel–consonant combinations
      iv. Explanation of biofeedback tool to use at home: smartphone app Voice Pitch Analyzer

   b. Session 2

      i. Repetition of glissando patterns
      ii. Automatic sequences (counting from 1 to 10, days of the week, months of the year, etc.) starting on habitual old pitch and then on the new pitch



      iii. Automatic sequences with gliding from old to new pitch within the word



   c. Session 3

      i. Short warm-up with glissando patterns
      ii. Speaking with the new pitch, making sure there are a lot of upward intonation patterns

         1. Short expressions (e.g., 'Go away', 'Mum and dad', 'up and down', 'be careful', etc.)
         2. Building up sentences
            Can you
            Can you put
            Can you put those files
            Can you put those files in the storage room later on?
         3. Short sentences
         4. Poems
         5. Texts

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Appendix B *(p. 2 of 4)*

Therapy Program

d. Session 4

   i. Introduction of water resistance therapy, using a resonance tube (2 cm under water)

      1. Bubbling without phonation
      2. Phonation with old pitch in the tube
      3. Phonation with new pitch in the tube
      4. Glissando patterns in the tube
      5. Short sentences with new pitch in the tube, then without tube
      6. Poems with new pitch in the tube, then without tube
      7. Texts with new pitch in the tube, then without tube
      8. Spontaneous speech, answering in the tube, then without tube

         a. Short answers (1 sentence)
         b. Longer answers (2-3 sentences)
         c. Conversation

e. Session 5

   i. Repetition of water resistance therapy
   ii. Straw phonation

      1. Blowing without phonation
      2. Phonation on new pitch
      3. Glissando patterns
      4. Spontaneous speech, answering in the straw, then without straw

2. **Articulation-resonance training**

a. Session 1

   i. Lip spreading

      1. Alternating with making an /u/ and /i/ movement of the lips (= discrimination with lip protrusion and lip spreading).
         Using a mirror to look at the lip movements.
      2. Alternating with making an /u/ and /i/ sound.
      3. Alternating with making an /e/ and /y/ sound.
      4. Consonant + /i/ combinations, consonant + /e/ combinations, feeling the easy lip spreading
      5. Trying to reach lip spreading when doing consonant + /u/ and /y/ combinations
      6. Monosyllable words with /i/, /e/, /u/ and /y/
      7. Multisyllable words with /i/, /e/, /u/ and /y/
      8. Sentences with /i/
      9. Sentences with /e/
      10. Sentences with all combinations
      11. Text
      12. Spontaneous speech

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Appendix B *(p. 3 of 4)*

Therapy Program

b. Session 2

i. Repetition of lip spreading
ii. Forward tongue position

1. Awareness of the tongue muscle: nonspeech oral motor exercises
   Using a mirror to look at the tongue movements
2. Moving the tongue from front to back when producing vowels
3. Pronouncing /i/ (high vowel) and feeling the forward tongue position with a high back of the tongue
4. Starting from /i/ sound and gliding to other vowels, trying to reach forward tongue and high back of the tongue
5. Words with /i/ (high vowel)
6. Words with /y/ (high vowel)
7. Words with /a/ (low vowel)
8. Sentences with /a/
9. Words with /o/ (low vowel)
10. Sentences with /o/
11. Texts
12. Spontaneous speech

c. Session 3

i. Repetition of forward tongue position
ii. Larynx elevation through twang

1. Awareness exercise: yawning (downward movement of the larynx) and swallowing (upward movement of the larynx)
2. Listening to twang sound such as crying baby, goat sounds, etc.
3. Adding twang to vowel /a/
4. Decreasing twang to vowel /a/
5. Consonant + /a/ + consonant + /a/ + consonant + /a/ with twang
6. Words with /a/
7. Sentences with /a/
8. Texts
9. Spontaneous speech

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Appendix B *(p. 4 of 4)*

Therapy Program

d. Session 4

   i. Repetition of larynx elevation
   ii. Forward resonance

      1. Discrimination between chest resonance and head resonance, saying /o/ vowel
      2. Putting a finger on left and right nostril and saying "hmmm," feeling forward airflow
      3. Nasal consonant /m/ + vowel
      4. Words with initial /m/
      5. Extra exercise to feel forward resonance

         a. Stand in front of a wall about 50 cm away.
         b. Place your head against the wall, comfortably.
         c. Your arms hang loose by your side.
         d. Place your tongue on your hard palate and start with a "nnnn" sound. Make a few glissandos to high and low frequencies.
         e. Repeat the previous step but now place the back part of your tongue on your soft palate. Make a "ng" sound and a few glissandos. By placing the head against the wall you feel the resonances better in your head.

   iii. Clear speech

      1. Combinations of consonants and vowels, pronouncing slow and then very fast, trying to pronounce clearly and precisely

         a. Tippetiptiptip tappetaptaptap toppetoptoptop
         b. Tanatanta tenetente tinitinti tonotonto
         c. Prieke prokke prakke pro prieke prokke prakke pro
         d. . . .

      2. Word combinations, 3x slow and 3x fast
      3. Cork exercise: using a cork with a diameter of 23 mm and length of 45 mm.

         a. Placing the upper front of the cork (approximately 2–3 mm) between their front teeth and reading words out loud with large and precise articulation movements. After a block of long nouns (6–9 syllables), they removed the cork and used the same large articulation movements to pronounce the same block of words.
         b. Tongue twisters with and without cork
         c. Text: reading sentences with and without cork

      4. Spontaneous speech

e. Session 5

   i. Repetition of all articulation-resonance techniques, spending most time on forward resonance and clear speech
   ii. Generalization of all articulation-resonance techniques

      1. Texts
      2. Spontaneous speech

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

Appendix C *Homework chart.*

Oefenschema *(Homework chart)*

Teken een smiley bij de dagen wanneer je geoefend hebt. ***(draw a smiley/circle/cross on the days that you practiced)***

Ideaal: 2x per dag telkens 10 min. (ideal: 2x a day for 10 minutes each)

	Maandag *(Monday)*		Dinsdag *(Tuesday)*		Woensdag *(Wednesday)*		Donderdag *(Thursday)*		Vrijdag *(Friday)*		Zaterdag *(Saturday)*		Zondag *(Sunday)*	
Week 5														
Week 6														
Week 7														
Week 8														
Week 9														
	Maandag *(Monday)*		Dinsdag *(Tuesday)*		Woensdag *(Wednesday)*		Donderdag *(Thursday)*		Vrijdag *(Friday)*		Zaterdag *(Saturday)*		Zondag *(Sunday)*	
Week 10														
Week 11														
Week 12														
Week 13														
Week 14														

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

***Appendix D**. Tables With Nonsignificant Values (p. 1 of 2)*

***Table D1.** f_o, intensity, and vocal quality measurements of each separate program.*

	PET				ART				PET vs. ART
Variable	Pre (M, SD)	Post 2 (M, SD)	Mean difference	*P* value	Pre (M, SD)	Post 2 (M, SD)	Mean difference	*P* value	*P* value
f_o (Hz)									
Spontaneous (pc5)	91.4, 18.15	104, 23.81	12.6	.427	92.6, 22.43	95, 23.23	2.4	1.000	.207
Intensity (dB)									
Sustained vowel (Mdn)	73.1, 7.64	75.8, 8.37	2.7	.586	69.9, 6.43	73.8, 5.12	3.8	.218	.64
Reading *(Mdn)*	67.5, 4.95	69.0, 6.48	1.5	1.000	64, 5.82	67.4, 4.61	3.4	.188	.368
Spontaneous *(Mdn)*	66.6, 5.87	67.9, 6.41	1.3	1.000	61.3, 4.25	64.4, 7.2	3.1	.261	.37
Voice range profile (dB and Hz)									
I-low	57.5, 3.31	56.8, 5.17	0.7	1.000	58.3, 2.5	58.5, 3.27	0.1	1.000	.532
I-high	102.5, 9.09	102.7, 7.17	0.2	1.000	103.8, 6.62	102.3, 9.75	1.5	1.000	.439
Vocal quality									
AVQI	3.2, 1.31	2.7, 0.83	0.5	.842	3, 1.03	3.1, 1.02	0.1	1.000	.149

Note. f_o = fundamental frequency; PET = pitch elevation training; ART = articulation-resonance training; pc5 = percentile 5; I-low = lowest intensity; I-high = highest intensity; AVQI = Acoustic Voice Quality Index.

***Table D2.** f_o, intensity, and vocal quality measurements of combination of ART and PET.*

Variable	Pre (M, SD)	Post 3 (M, SD)	Mean difference	*p* value	Order effect *p* value
f_o (Hz)					
Spontaneous (pc5)	92, 20.06	101.6, 28.21	9.6	.116	.934
Intensity (dB)					
Reading *(Mdn)*	65.7, 5.59	67.3, 5.07	1.6	.751	.112
Voice range profile (dB and Hz)					
I-low	57.9, 2.91	58.3, 4.10	0.4	1.000	.486
I-high	103.1, 7.85	105, 8.25	1.9	.477	.878
F-low	89.2, 16.17	93.6, 24.51	4.4	.244	.223
F-high	596.5, 253.04	681.2, 192.94	84.7	.06	.417
Vocal quality					
AVQI	3.1, 1.16	2.7, 0.9	0.4	.347	.728

Note. f_o = fundamental frequency; ART = articulation-resonance training; PET = pitch elevation training; pc5 = percentile 5; I-low = lowest intensity; I-high = highest intensity; F-low = lowest fundamental frequency; F-high = highest fundamental frequency; AVQI = Acoustic Voice Quality Index.

Published in : American Journal of Speech-Language Pathology (2023), vol. 32, pp. 145–168
DOI: 10.1044/2022_AJSLP-22-00135
Status : Postprint (Author's version)

***Appendix D**. Tables With Nonsignificant Values (p. 2 of 2)*

***Table D3.** Formant frequency measurements of each separate program.*

Variable	PET				ART				PET vs. ART
	Pre (M, SD)	Post 2 *(M, SD)*	Mean difference	*P* value	Pre (M, SD)	Post 2 (M, SD)	Mean difference	*P* value	*P* value
Median formant frequencies									
/a/									
F_3	2469.1, 169.98	2522.1, 223.85	53.0	1.000	2432.2, 157.9	2509.1, 138.98	76.9	.776	.722
F_4	3640.4, 192.22	3667.5, 367.66	27.1	1.000	3618.2, 239.81	3564.9, 318.58	53.4	1.000	.475
F_5	4560, 352.89	4624.4, 427.97	64.5	1.000	4464.8, 387.8	4488.8, 401.01	24.0	1.000	.784
/i/									
F_1	290.3, 40.69	300.1, 29.74	9.8	1.000	296.4, 32.5	292.3, 36.81	4.1	1.000	.297
F_2	2153.6, 152.04	2179.9, 267.16	26.3	1.000	2165.1, 200.77	2244, 147.76	78.9	1.000	.512
F_3	2811.7, 197.8	2887.3, 199.56	75.6	.249	2879.7, 200.24	2987.7, 262.36	108.0	.625	.664
F_4	3619.2, 277.59	3714.8, 307.78	95.6	.733	3682.8, 302.41	3731.6, 345.07	48.8	1.000	.642
F_5	4283.9, 329.94	4455.6, 458.5	171.7	.855	4507.7, 369.4	4561.4, 430.94	53.7	1.000	.46
/u/									
F_1	311.2, 23.3	321.4, 24.32	10.1	.704	305.7, 35.31	310.8, 39.64	5.1	1.000	.619
F_2	930.8, 93	907.8, 100.72	23.0	1.000	942.8, 87.76	972.5, 146.41	29.7	1.000	.144
F_4	3492.7, 216.14	3609.2, 235.57	116.5	.054	3527.5, 265.66	3671, 223.07	143.4	.153	.72
F_5	4308.7, 251.93	4400.2, 220.11	91.5	1.000	4381.6, 219.94	4446.5, 343.83	64.9	1.000	.797

Note. PET = pitch elevation training; ART = articulation-resonance training

***Table D4.** Formant frequency measurements of combination of ART and PET.*

Variable	Pre (M, SD)	Post 3 (M, SD)	Mean difference	*p* value	Order effect p value
Median formant frequencies					
/a/					
F_4	3629.3, 213.84	3622.3, 325.46	7.0	1.000	.147
F_5	4512.4, 367.51	4565.7, 431.7	53.3	1.000	.325
/i/					
F_1	293.3, 36.32	305.5, 41.52	12.2	.422	.968
F_2	2159.4, 175.08	2201.9, 245.67	42.5	1.000	.992
F_3	2845.7, 198.6	2922.9, 208.97	77.2	.244	.261
F_4	3651, 287.04	3767.2, 277.36	116.2	.137	.632
F_5	4395.8, 362.46	4521, 319.85	125.2	.713	.072
/u/					
F_2	936.8, 89.05	964.4, 139.7	27.6	.759	.247
F_5	4345.2, 235.31	4438, 330.38	92.8	.451	.511

Note. ART = articulation-resonance training; PET = pitch elevation training