Voice Feminization: Voice Therapy vs. Surgical Intervention: A Systematic Review

Carson Leon-Gambetta  
*University of Vermont*, cleongam@uvm.edu

Ruth Huttner  
*University of Vermont*, Ruth.Huttner@uvm.edu

Amanda Matyas  
*University of Vermont*, amatyas@uvm.edu

Follow this and additional works at: https://scholarworks.uvm.edu/csdms

Part of the Speech Pathology and Audiology Commons

**Recommended Citation**

https://scholarworks.uvm.edu/csdms/5

This Project is brought to you for free and open access by the College of Nursing and Health Sciences at ScholarWorks @ UVM. It has been accepted for inclusion in Communication Sciences and Disorders: Systematic Review Publications by an authorized administrator of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.
Voice Feminization: Voice Therapy vs. Surgical Intervention: A Systematic Review

Authors: Ruth Huttnner, Carson Leon-Gambetta, and Amanda Matyas

Department of Communication Sciences and Disorders

College of Nursing and Health Sciences

University of Vermont
Abstract

Purpose: Transgender individuals often seek to alter their vocal characteristics. For Male to Female (MtF) transgender individuals, attaining a feminine voice may be particularly challenging. The objective of this systematic review is to determine whether MtF transgender individuals who receive voice feminization therapy alone or Wendler’s Glottoplasty (WG) surgical intervention with subsequent voice therapy yield greater outcomes in frequency and self-perception.

Method: A systematic review of the literature was conducted by using PubMed and Ovid to search terms pertaining to voice feminization. The articles were reviewed and appraised by the authors for inclusionary criteria, exclusionary criteria, and quality. Inclusionary criteria included: 1) adult MtF Transgender individuals, 2) pre and post measures of fundamental frequency (fo), 3) post puberty age, 4) measure of perception of femininity, and 5) patients who underwent WG (articles pertaining to surgical intervention only).

Results: A total of 82 articles were identified and 12 met inclusionary criteria for this systematic review. Overall, the quality of the studies was moderate. Outcome measures included frequency range and frequency gain. The authors were unable to compare measurements of self-perception and perception of femininity due to the variability in assessments.

Conclusions: Patients who opted for surgical intervention experienced a greater increase in fo but a decrease in frequency range. Conversely, patients who participated in voice feminization therapy alone were found to exhibit smaller gains in fo and an increase in frequency range. This implies that both voice feminization therapy and surgical intervention are effective methods for increasing the frequency of voice. Limitations of this research include the subjective nature of
self-perception measures, variability in measurements of perception of femininity, and overall limited research regarding this population.
Background

The Lesbian, Gay, Bisexual, Transgender, Queer, and other identities (LGBTQ+) community face unique challenges in healthcare. Transgender people identify as a gender that is different than the one they were assigned at birth (based on primary sexual characteristics). The term ‘transgender’ includes both transgender individuals and transsexual individuals (i.e., individuals who undergo surgery to align their primary sexual characteristics with their gender identity).

Presently, .0039% of adults in the world identify as transgender (Meerwijk & Sevelius, 2017). As a result, there is need for specialized health care for this small community. For the purpose of this systematic review, we will use the umbrella term ‘Trans’ to refer to both transgender and transsexual individuals. A multitude of risk factors impact both the physical and mental health of persons who identify as Trans. Notably, Trans individuals are four times more likely to live in poverty, twice as likely to be unemployed and homeless, and four times more likely to be HIV-positive (Meerwijk & Sevelius, 2017) than individuals who do not identify as Trans. Additionally, 41% of Trans individuals report having attempted suicide at least once in their lifetime (Meerwijk & Sevelius, 2017). These findings demonstrate that resources and appropriate intervention are of utmost importance for the Trans population to increase quality of life.

While some Trans individuals seek sexual reassignment surgery, many Trans individuals also explore options to alter their secondary sexual characteristics to align with their gender identity. Secondary sexual characteristics include body hair, hip size, laryngeal prominence (i.e., Adam’s apple), and voice. For male to female (MtF) Trans individuals (also referred to as ‘Trans women’), attaining a feminine voice presents a unique challenge. Of particular note, the
introduction of estrogen does not impact the frequency of voice for MtF Trans individuals undergoing hormone therapy (Mastronikolis, Remacle, Biagini, Kiagiadaki, & Lawson, 2013). As a result, many MtF Trans women seek to feminize their voices using alternative means. Conversely, female to male transgender individuals receive high levels of testosterone during hormone therapy, which deepens the frequency of their voices (Irwig, 2017). Currently, MtF Trans individuals have two main tracts for achieving a more feminine voice: Voice feminization Therapy (VT) and Surgical Voice Feminization (SVF).

**Voice Therapy**

Despite potential benefits, there are inherent risks associated with both VT and SVF. While there are no standard procedures for VT, common treatment goals include, but are not limited to, increasing fundamental frequency, utilizing more feminine intonation, and attaining forward resonance. VT often includes vocal exercises to increase pitch and using audio recordings to increase self-monitoring of intonation. Clear benefits of VT include the non-invasive nature and lower cost compared to surgery. Individuals who attempt to raise their vocal pitch without the guidance of a licensed speech language pathologist and/or adhering to proper vocal hygiene risk developing vocal fold nodules.

**Surgical Voice Feminization**

While there are many SVF options, current research indicates that Wendler’s Glottoplasty (WG) is superior to other surgical interventions available. WG results in increased frequency range and speaking fundamental frequency ($f_0$) when compared to the popular cricothyroid approximation method (Kelly et al., 2018). WG sutures the anterior third of the vocal folds, effectively decreasing their length and increasing pitch (Mastronikolis et al., 2013). However, risks associated with anesthesia and high cost of surgery are potential considerations.
Currently, there is limited research directly comparing VT to SVF. Trans women would benefit from a comprehensive comparison of the interventions available to address voice feminization to make informed decisions regarding their healthcare. Therefore, it is critical to determine how frequency, voice quality, gender perception to unfamiliar listeners, and self-perception are impacted in MtF Trans individuals undergoing VT compared to SVF. This systematic review seeks to consolidate research to compare VT to SVF.

**Methods**

*Search Strategy & Article Selection*

The primary electronic database used for this systematic review was Ovid PubMed-1946. The authors initially utilized mapping features of PubMed to obtain Medical Subject Headings (MeSH) terms pertaining to the research question. These MeSH headings were then combined using Boolean operators (see Figure 1) on Ovid MEDLINE to acquire a preliminary total of 82 articles. Using Ovid’s advanced searching, the authors first combined all MeSH terms for ‘Trans’ and then combined those with the MeSH terms for ‘vocal surgery’. This yielded results for the examination of outcomes of SVF in Trans individuals. Separately, the MeSH terms for ‘Trans’ were also combined with a search of MeSH terms for ‘voice therapy’. This yielded research results concerning Trans individuals and VT. The authors utilized other research databases (e.g. PsychINFO, EBSCOhost, CommDisDome) by searching the same MESH headings and Boolean operators. However, no novel literature was found in these secondary searches.

When selecting articles for this systematic review, the authors considered various inclusionary and exclusionary criteria (see Figure 2). The population examined were MtF Trans individuals who had undergone either VT or SVF in an effort to increase the pitch and the degree
to which their voice is perceived as female. A total of 82 articles were found and 12 were
deemed appropriate based on inclusionary and exclusionary criteria.

Then, the authors carried out a quality assessment on the 12 articles to be included in this
systematic review. Final articles were selected based on overall quality and integrity of the
study. The validity section and quality assessment sections below fully describe the selection
process.

Figure 1: Boolean Operators

<table>
<thead>
<tr>
<th>Database Interface and Version</th>
<th>Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovid MEDLINE 1946-November 8, 2018</td>
<td>(exp Speech Therapy/ or exp Transsexualism OR exp Speech Therapy/ or exp Transsexualism OR exp Speech Therapy/ and exp Transsexualism/ and exp Transgender Persons OR exp Speech Therapy/ and exp Transgender Persons/ OR Transgender Persons OR Transsexualism OR Feminization/ OR Gender Dysphoria OR transexual*.mp. OR transgender*.mp. OR transexual*.mp.) AND (glottoplasty.mp. OR Laryngoplasty OR exp Glottis/ OR exp GLOTTIS/su [Surgery]) OR [(exp Speech Therapy/ or exp Transsexualism OR exp Speech Therapy/ or exp Transsexualism OR exp Speech Therapy/ and exp Transsexualism/ and exp Transgender Persons OR exp Speech Therapy/ and exp Transgender Persons/ OR Transgender Persons OR Transsexualism OR Feminization/ OR Gender Dysphoria OR transexual*.mp. OR transgender*.mp. OR transexual*.mp.) AND (exp Pitch Perception/ OR Speech Perception/ OR Speech Therapy/ OR exp Voice Training/)]</td>
</tr>
</tbody>
</table>

Inclusion/Exclusion Criteria

Each article was independently screened by two reviewers for relevance of title and
quality of information contained in the abstract. Articles deemed relevant by both reviewers were
read in their entirety to determine if they met the inclusionary criteria detailed in Figure 2.

Articles that did not meet these criteria were excluded from this review. Exclusionary criteria
included intersex individuals. The anatomical/physiological features and voice goals may be
different for intersex individuals as compared to MtF Trans individuals. For this reason, this population was not included in this systematic review.

**Figure 2: Inclusion Criteria**

<table>
<thead>
<tr>
<th>Inclusionary Criteria</th>
<th>Rationale (if applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voice Feminization Therapy Articles</strong></td>
<td></td>
</tr>
<tr>
<td>MtF transgender individuals</td>
<td>Perception of voice is impacted by fundamental frequency. Although other factors contribute to the femininity of one’s voice (e.g. intonation, resonance, quality), fundamental frequency is essential in achieving a feminine voice.</td>
</tr>
<tr>
<td>Objective measurement of fundamental frequency pre/post therapy</td>
<td></td>
</tr>
<tr>
<td>Post puberty age range</td>
<td>There are significant changes to hormone levels and muscle development during puberty. Due to the unpredictable nature of how these changes will impact vocal frequency and quality, only those post puberty at the time of transition will be included.</td>
</tr>
<tr>
<td>Self-perception measurement</td>
<td>Oftentimes, individuals who have increased their fundamental frequency to within the typical cisgender female voice range will still be perceived as masculine by others and themselves. As such, a measure of self-perception is crucial to determining successful VT.</td>
</tr>
<tr>
<td><strong>Surgical Voice Feminization Articles</strong></td>
<td></td>
</tr>
<tr>
<td>MtF transgender individuals</td>
<td>Perception of voice is impacted by fundamental frequency. Although other factors contribute to the femininity of one’s voice (e.g. intonation, resonance, quality), fundamental frequency is essential in achieving a feminine voice.</td>
</tr>
<tr>
<td>Objective measurement of fundamental frequency pre/post therapy</td>
<td></td>
</tr>
<tr>
<td>Post puberty age range</td>
<td>There are significant changes to hormone levels and muscle development during puberty. Due to the unpredictable nature of how these changes will impact vocal frequency and quality, only those post-puberty at the time of transition will be included.</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Self-perception measurement</td>
<td>Oftentimes, individuals who have increased their fundamental frequency to within the typical cisgender female voice range will still be perceived as masculine by others and themselves. As such, a measure of self-perception is crucial to determining successful voice modification surgery.</td>
</tr>
<tr>
<td>Patients who underwent Wendler’s Glottoplasty</td>
<td></td>
</tr>
</tbody>
</table>
Quality Assessment

Studies were given a descriptive term of quality based on percentage of study characteristics each one contained (Low: 0-25% of SC; Moderate: 25%-75% of SC; High 75-100% of SC). It should be noted that all study characteristics were valued equally based on judgment of the authors. To determine the quality of each article the form in Figure 3 was used.
Data Extraction

Data including number of participants (n), mean age, age range, study design, speaking \( f_0 \) gained, statistical significance of \( f_0 \) gained, frequency range, frequency range gain/loss, mean number of sessions of voice therapy, and duration of therapy was extracted as appropriate from each article and entered into an Excel spreadsheet. It should be noted that quantitative and qualitative data were examined separately.

Quantitative data (e.g., speaking \( f_0 \) gained, frequency range) was combined and divided by the number of participants for each outcome measure (VT and SVF separately) to determine the average outcome of each measure. The results are displayed in Figures 5 and 6. Due to the variety of qualitative measures used to examine perception and voice quality, results were compared generally across studies. These findings can be found in the discussion section.

Results

Twelve studies met the inclusion criteria for this systematic review: six articles examined WG and six articles examined VT. All participants were MtF Trans individuals who ranged from 16 to 64 years of age. Articles reported outcomes using various assessments (e.g., Voice Handicap Index, Hirano GBRAS scale, Transsexual Voice Questionnaire, measures of satisfaction by the participants, and measures of femininity by unfamiliar listeners), as well as the outcome measures identified in the inclusionary criteria. The PRISMA flow chart (Figure 4) outlines the progression of study selection.
Quality Assessment

The authors reviewed a total of 12 studies relevant to the research question. These studies were primarily rated ‘moderate’ in their level of quality. The majority of studies were retrospective and prospective treatment studies. Of the 12 studies, the mean number of participants was 13 and five articles included study control groups. All studies included in this review included baseline measures. Five out of six studies included an assessment to account for loss to follow-up when appropriate (i.e., VT articles). Out of the 12 articles, nine indicated inclusionary criteria and four described exclusionary criteria. Harms and benefits of VT or GP were discussed in the majority of articles reviewed (10/12). Due to the nature of retrospective treatment studies, blinding, reliability, and validity were not fully discussed in each article. The authors noted that measurement bias, researcher bias, selection bias, performance bias, impact bias, cognitive bias, and negative selection bias.
<table>
<thead>
<tr>
<th>First author (year)</th>
<th>Type of Study</th>
<th>Number of participants</th>
<th>Baseline Comparability</th>
<th>Desired Study Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Meszaros (2005)</strong></td>
<td>Prospective Treatment Study</td>
<td>+3</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td><strong>Mastonikolis (2013)</strong></td>
<td>Retrospective Treatment Study</td>
<td>+29</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td><strong>Meister (2016)</strong></td>
<td>Outcomes Research Study</td>
<td>+21</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td><strong>Carew (2007)</strong></td>
<td>Intervention Study</td>
<td>+10</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td><strong>Hancock (2012)</strong></td>
<td>Retrospective Chart Review</td>
<td>+25</td>
<td>++</td>
<td>N.A.</td>
</tr>
<tr>
<td><strong>Gelfer (2012)</strong></td>
<td>Prospective Treatment Study</td>
<td>+5</td>
<td>++</td>
<td>N.A.</td>
</tr>
<tr>
<td><strong>Mora (2018)</strong></td>
<td>Retrospective Treatment Study</td>
<td>+23</td>
<td>++</td>
<td>N.A.</td>
</tr>
<tr>
<td><strong>Kelly (2018)</strong></td>
<td>Retrospective Treatment Study</td>
<td>+13</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td><strong>Remacle (2011)</strong></td>
<td>Retrospective Treatment Study</td>
<td>+15</td>
<td>++</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

Legend:
- **Participants relevant to research question**
- **Experimental group**
- **Control group (as appropriate)**
- **Presented/Achieved**
- **Loss to follow-up assessed**
- **Inclusion criteria specified**
- **Exclusion criteria specified**
- **Interventions & competitors described/defined**
- **Harms and benefits of clinical outcomes considered (as appropriate)**
- **Blinding Present (as appropriate)**
- **Reliability presented**
- **Reliability achieved**
Effects of Intervention

Individuals receiving WG demonstrated a mean frequency increase of 53.75 Hz compared to VT (mean frequency increase of 34.06 Hz). Conversely, individuals receiving WG demonstrated a mean loss of 68.91 Hz in frequency range compared to 109.07 Hz gained in frequency range for VT. The authors were unable to objectively compare measurements of self-perception of voice due to the variability in assessments. Despite the lack of standardized measures to compare self-perception, overall, studies found self-perception of voice to improve following WG and VT intervention. Specifically, five out of six studies examining VT found some improvement in satisfaction of voice or self-perception of femininity and four out of four studies examining WG accompanied by VT found an improvement in satisfaction of voice or self-perception of femininity. Two studies on VT did not report on these measures.

Figure 5. Study Characteristics Table (WG)
<table>
<thead>
<tr>
<th>Author (year)</th>
<th>n</th>
<th>Age range (mean)</th>
<th>Design</th>
<th>Frequency gain (Hz)</th>
<th>Statistical significance</th>
<th>Frequency range pre/post surgery (Hz)</th>
<th>Average frequency range loss (Hz)</th>
<th>Level of quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group B: 12</td>
<td>Group A: 16-39 (28.6)</td>
<td>Group B: 45-59 (51.9)</td>
<td>Retrospective treatment study</td>
<td>Group B: 43</td>
<td>Group B: 345.8/227.6</td>
<td>Group B: -118.2</td>
<td></td>
</tr>
<tr>
<td>Casado (2016)</td>
<td>10</td>
<td>30-52 (39.9)</td>
<td>Retrospective Study</td>
<td>106</td>
<td>P=0.005</td>
<td>NA</td>
<td>NA</td>
<td>High</td>
</tr>
<tr>
<td>Kelly (2018)</td>
<td>13*</td>
<td>NA**</td>
<td>Retrospective Study</td>
<td>31.3</td>
<td>P=0.001</td>
<td>N/A reported in semitones</td>
<td>NA</td>
<td>High</td>
</tr>
<tr>
<td>Meister (2016)</td>
<td>21</td>
<td>24-57 (42)</td>
<td>Outcomes Research Study</td>
<td>43</td>
<td>P=0.001</td>
<td>317/310</td>
<td>-7</td>
<td>Moderate</td>
</tr>
<tr>
<td>Remacle (2011)</td>
<td>15</td>
<td>22-57 (42.5)</td>
<td>Retrospective Treatment Study</td>
<td>44</td>
<td>P=0.006</td>
<td>438/241</td>
<td>-197</td>
<td>Low</td>
</tr>
<tr>
<td>Mora (2018)</td>
<td>23*</td>
<td>19-50 (35)</td>
<td>Retrospective study</td>
<td>43</td>
<td>P&gt;0.05</td>
<td>NA</td>
<td>NA</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Figure 6. Study Characteristics (VT)

Discussion
This systematic review found that individuals who receive WG exhibit higher frequency gains than those who opt for VT. However, patients who underwent WG demonstrated a more restricted frequency range post-surgery compared to VT patients who exhibited an increased frequency range. These results may have varying implications for different Trans individuals. For example, a professional voice user (e.g. singer, actress) may desire a greater frequency range and may be limited by the results of WG. Alternatively, other MtF Trans individuals with fewer vocal demands may be less concerned with a decreased frequency range. This is an important factor for patients and healthcare providers to consider when choosing an intervention for voice feminization.

Additionally, self-perception of voice is a significant indicator of the success of intervention. It should be noted that, based on the articles reviewed, it cannot be concluded that patients who underwent WG will be perceived as more feminine than those who opted for VT. Because there is currently no standardized measure to assess self-perception of voice, it is difficult to draw objective conclusions regarding the femininity outcomes of one intervention over the other. However, positive trends were generally noted across studies. This suggests that VT and WG may be an effective method to improve self-perception in MtF Trans individuals.

Future research/limitations

While this systematic review highlighted some important findings for MtF transgender individuals, the studies reviewed were not without limitations. Limitations of these studies included lack of randomized control trials, potential biases, lack of standardized measurements of self-perception, and lack of standardized measurements of femininity of voice. The majority of articles reviewed were either retrospective studies or prospective treatment studies. Future researchers may consider conducting randomized control trial studies, as well as obtaining larger
sample sizes for increased generalizability. Randomized control trials would increase fidelity of research. A number of potential biases exist in this literature. Many articles selected patients who underwent intervention at the same location. As a result, these articles may demonstrate convenience sampling and therefore lack generalizability. Additionally, due to the nature of intervention, researchers may seek positive outcomes for participants, thus risking research bias. Finally, patients who are not satisfied with results are less likely to return for follow-up (i.e. negative selection bias). Future researchers should consider developing and utilizing a standard measure of self-perception of femininity as well as directly compare WG and VT femininity outcomes utilizing blinded unfamiliar listeners. Due to limited research, the authors of this systematic review were unable to examine all the components that influence perception of femininity. This includes resonance, articulation, breathiness, intonation, and word choice (Mora, Cobeta, Becerra, & Lucio, 2018). Due to the aforementioned limitations, this systematic review is not comprehensive in its scope. However, it does provide a template for the direction of future research.

Conclusions

Individuals seeking to alter their vocal characteristics, including undergoing voice feminization surgery, should consult with a speech language pathologist and other related health professionals to make an informed health care decision. Research suggests that both WG and VT are beneficial in altering vocal characteristics. For this reason, clinicians should work with their clients to consider the client’s long-term goals when suggesting treatment intervention. In addition to goals, clinicians should consider potential risks of VT (e.g., developing vocal nodules) and of WG (e.g., risks associated with anesthesia and high cost of surgery). Overall,
research pertaining to the Trans community is limited. As such, additional research is needed to support this population and their voice goals.

References


