Feminize Me!

How Anatomy Influences the Sound of a Voice

Marik Roos, Veronika Weber
Department of Musicology, University of Vienna

DAGA 2021 Wien
15th – 18th of August 2021
Vienna, Austria

Background

Each vocal tract is unique and so is the sound of a voice. The perception not only of personality traits and emotional states in speech, but also of the speaker's gender can be predicted more accurately by timbre rather than fundamental frequency (Roos et al., 2016). Also, voices that are either perceived as male or female show significant differences in timbre features such as formant structures and Mel Frequency Cepstral Coefficients, especially mfcc2, 4, 5, 7, 9 and 13 (Reuter et al., 2016).

The aim of this study is to gain a better understanding of the underlying connections between the vocal anatomy and the sound it produces by analyzing the timbre changes as results of different methods of surgical alteration of the vocal tract.

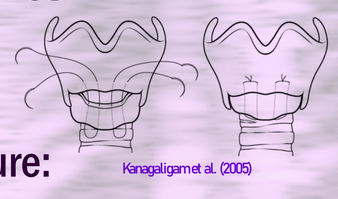
COI disclosure: This study is not to be understood as recommendation for any specific method and has neither been commissioned nor funded.

Methods

N=258 recordings of patients who underwent so called „Voice Feminization Surgery", in particular Cricothyroidale Approximation (CTA), Feminization Laryngoplasty (FemLar), FemLar plus Thyrohyoid Elevation (FLT), Vocal Folds Shortening and Retrodisplacement of Anterior Commissure (VFSRAC) and Vocal Fold Webbing (VFW) were analyzed using „Feature Extractor" by Isabella Czedik-Eysenberg with MIRToolBox and Mining Suite plus the Python library „Parselmouth" (Praat). Paired as well as independent samples *t*-tests and regression along with factor analyses were conducted between singular vowels and timbre feature mean values of each voice to estimate significant differences between pre- and post-op timbres as well as different methods of altering the vocal tract. Decreased tonal or band energy were not interpreted due to possible effects of the individual healing process on a voice's physical strength and roughness features.

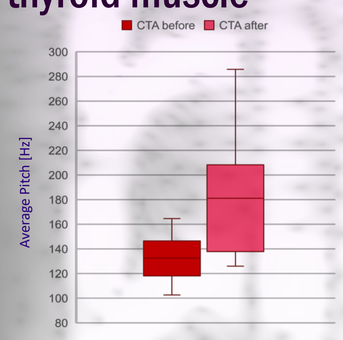
Results

CTA


Kanagaligam et al. (2005)

Procedure:

- Thyroid and cricoid cartilage stitched together
- Permanent contraction, thus release of crico-thyroid muscle



Significant effects:

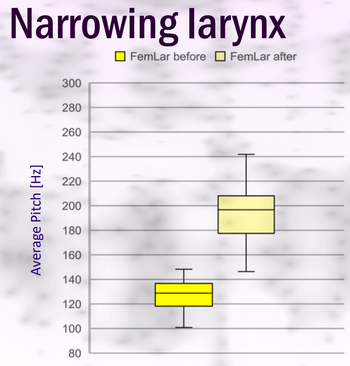
- Pitch increase (t=2.906, p=.009, d=1.300, df=18)
- Formant changes: –
- Spectral centroid: –

FemLar


Kunachak, et al. (2000)

Procedure:

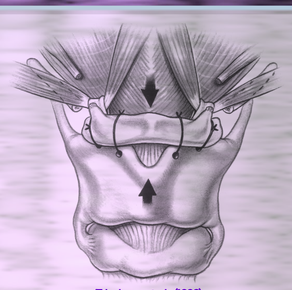
- Removing front of cricoid cartilage and part of vocal folds
- Shortening vocal folds
- Narrowing larynx



Significant effects:

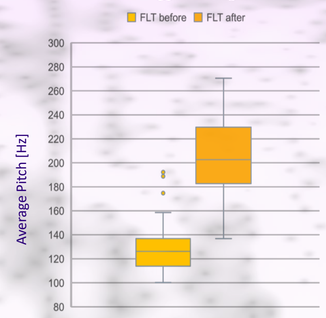
- Pitch increase (t=9.601, p<.001, d=3.506, df=28)
- Formant changes: –
- Spectral centroid: –

FLT


Friedman et al. (1998)

Procedure:

- Thyroid cartilage and hyoid bone stitched together
- Shortening larynx



Significant effects:

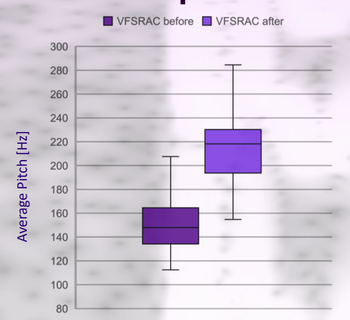
- Pitch increase (t=14.422, p<.001, d=2.856, df=100)
- All formants lower (d=.468)
- Increased spectral centroid (t=3.092, p=.003, d=.433)

VFSRAC


Kim (2017)

Procedure:

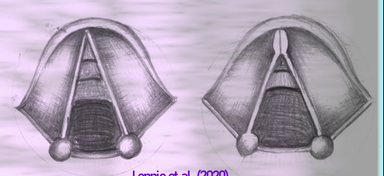
- Vocal folds stitched together (~1/3 of length)
- Extra stability by additional stitches to posterior larynx



Significant effects:

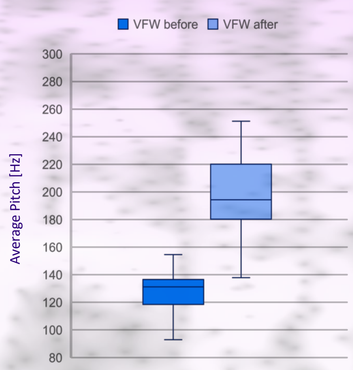
- Pitch increase (t=10.672, p<.001, d=2.551, df=68)
- Formant changes: –
- Increased spectral centroid (t=3.715, p<.001, d=.628)

VFW


Lennie et al. (2020)

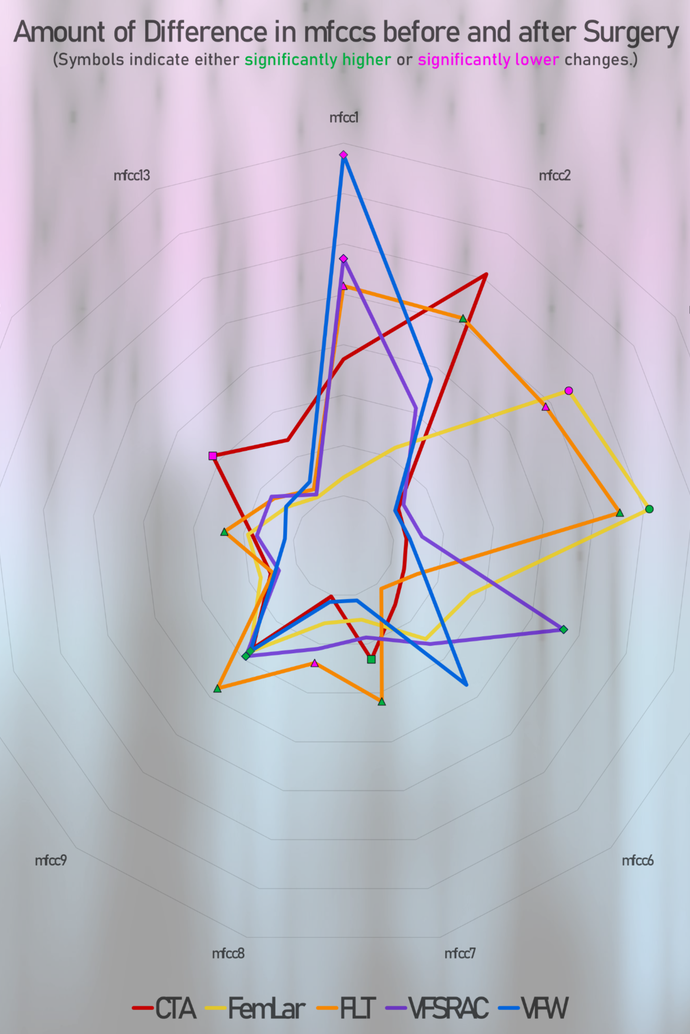
Procedure:

- Vocal folds stitched together (~1/3 of length)



Significant effects:

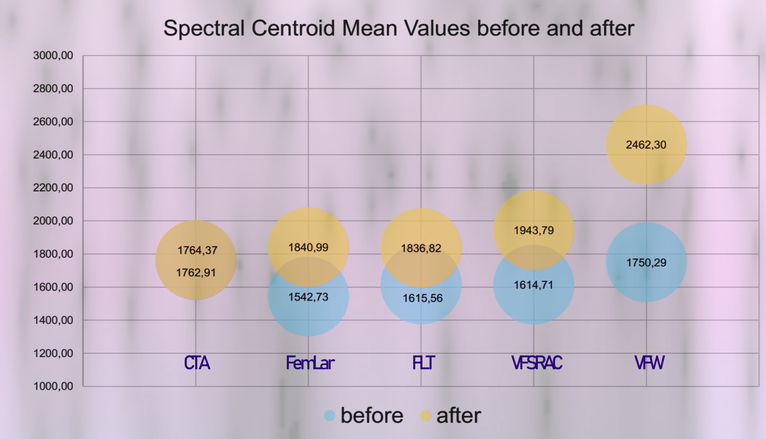
- Increase of pitch (t=9.616, p<.001, d=3.205, df=34)
- Lower 1st formant (t=2.521, p=.022, d=.594)
- Increased spectral centroid (t=3.033, p=.008, d=.715)

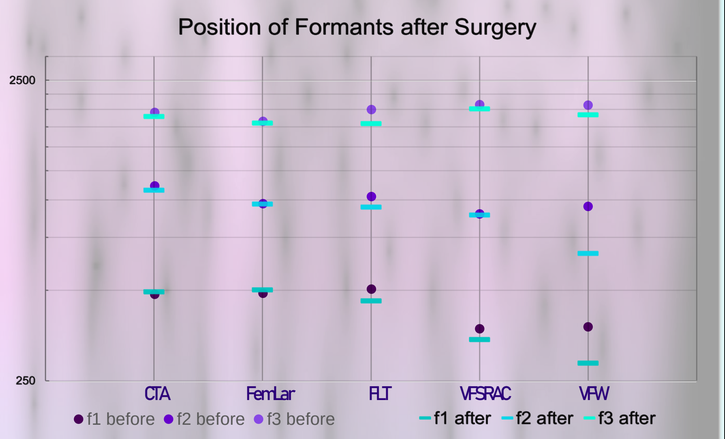


MFCC Changes with Medium or Large Effect Size

	CTA [df=9]	FemLar [df=14]	FLT [df=50]	VFSRAC [df=34]	VFW [df=17]
mfcc1			d=.421 p=.004	d=.399 p=.024	d=.626 p=.017
mfcc2			d=.548 p<.001		
mfcc3		d=.687 p=.019	d=.791 p<.001		
mfcc4		d=1.291 p<.001	d=.843 p<.001		
mfcc5				d=.712 p<.001	
mfcc6				d=.333 n. s.	d=.468 n. s.
mfcc7	d=.842 p=.011		d=.550 p<.001		
mfcc8			d=.378 p=.009		
mfcc9	d=.548 n. s.	d=.456 n. s.	d=.942 p<.001	d=.536 p=.003	d=.546 p=.033
mfcc10					
mfcc11			d=.500 p<.001		
mfcc12	d=.948 p=.015				
mfcc13	d=.602 n. s.				

higher after surgery & *lower after surgery*





Discussion

The data indicate that the shortening of the vocal tract (CTA, FLT) predicts less increase of spectral centroid and higher mfcc7. This could also explain the significantly lower formant structures for FLT, but not for VFW, where a similar effect could be observed in spite of a greatly increased SC mean.

The increase of mfcc5 for VFSRAC could be explained by a more stable suspension of the vocal folds. The shortening of vocal folds by suturing may influence mfcc6 while shortening by reducing the size of the voice box seems to predict changes in mfccs3&4. The reduction of (any) resonating tissue could be responsible for the increase of mfcc9. Further research is necessary to evaluate these findings.

References

Friedman, M., Lim, J., & Tanyeri, H. (1998). Thyrohyoid Suspension for correction of obstructive sleep apnea. *Operative Techniques in Otolaryngology - Head and Neck Surgeries, 9*(3), S. 138-141. <> Kanagalingam, J., Georgalas, C., Wood, G., Ahluwalia, S., Sandhu, G., & Cheesman, A. (2005). Cricothyroid Approximation and Subluxation in 21 Male-to-Female Transsexuals. *The Laryngoscope*, 115, S. 611-618. <> Kim, H. (2017). A New Conceptual Approach for Voice Feminization: 12 Years of Experience. *The Laryngoscope*, 127, S. 1102-1108. <> Kunachak, S., Prakunhungsit, S., & Sujjalak, K. (2000). Thyroid cartilage and vocal fold reduction: a new phonosurgical method for male-to-female transsexuals. *Ann Otol Rhinol Laryngol*, 109(11), S. 1082-1086. <> Lartillot, O., & Toiviainen, P. (2007). A Matlab Toolbox for Musical Feature Extraction from Audio. *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-07)*, S. 237-244. <> Lennie, Y., Leareng, K., & Evered, L. (2020). Perioperative considerations for transgender women undergoing routine surgery: a narrative review. *British Journal of Anaesthesia*, 124(6), S. 702-711. <> Reuter, C., Wöllner, C., Czedik-Eysenberg, I., & Hohagen, J. (2016). *Tönend bewegte Normen: Akustische und kinetische Morphs bei Sängerinnen und Sängern*. Poster presented at "Akustik und Musikalische Hörwahrnehmung": 32. Jahrestagung 2016 der Deutschen Gesellschaft für Musikpsychologie (DGM), Wien, Österreich. <> Roos, M., Reuter, C. (2016). *Ich höre, wer Du bist! Über die Wahrnehmung geschlechtsspezifischer Persönlichkeitsmerkmale in der menschlichen Stimme*. Poster presented at "Akustik und Musikalische Hörwahrnehmung": 32. Jahrestagung der Deutschen Gesellschaft für Musikpsychologie (DGM), Wien, Österreich.