Proceedings Paper:

Hughes, C., Brown, G. orcid.org/0000-0001-8565-5476, Ma, N. et al. (1 more author) (2024) Acoustic effects of facial feminisation surgery on speech and singing: A case study. In: Proceedings of Interspeech 2024. Interspeech 2024, 01-05 Sep 2024, Kos island, Greece. International Speech Communication Association (ISCA) , pp. 3065-3069.

https://doi.org/10.21437/Interspeech.2024-1132

Interspeech 2024
1-5 September 2024, Kos, Greece

Acoustic Effects of Facial Feminisation Surgery on Speech and Singing: A Case Study

Cliodhna Hughes[1], *Guy Brown*[1], *Ning Ma*[1], *Nicola Dibben*[1]

[1]University of Sheffield, United Kingdom

chughes9@sheffield.ac.uk, g.j.brown@sheffield.ac.uk, n.ma@sheffield.ac.uk, n.j.dibben@sheffield.ac.uk

Abstract

Transfeminine people may undergo facial feminisation surgery, a term covering a range of procedures that aim to alter the appearance of facial features, thereby potentially changing characteristics of the vocal tract. Effects of facial feminisation surgery on the voice are relatively understudied, however, so, little information on the vocal effects of these surgeries is available to people considering undergoing these procedures. In this single-case study, we present an acoustic analysis of speech and singing data collected from a transgender singer before and after facial feminisation surgery, alongside an examination of longitudinal interview data from the participant. Our quantitative results suggest facial feminisation surgery can have an impact on the voice, and our qualitative analysis suggests this may not only be as a result of the altered characteristics of the vocal tract, but also as a result of the altered social context. Several issues for future research are identified.

Index Terms: facial feminisation surgery, transgender speech, speech and gender, vocal dysphoria

1. Introduction

Transgender people are those whose gender identity differs from the gender they were assigned at birth [1]. They may experience gender dysphoria: a distress arising from the mismatch between gender identity and gender assigned at birth. For transfeminine people, the appearance of the face can cause dysphoria. Facial feminisation surgery (FFS) is a term covering a range of procedures which aim to make facial features appear more feminine [2]. While FFS does not aim to change the voice, this can happen as a side effect. One instance of FFS is chondrolaryngoplasty, which involves removing thyroid cartilage [2]. This procedure can lead to hoarseness, and lower mean and maximum F0 [3]. Another instance of FFS, rhinoplasty, which involves altering the shape of the nose, is known to impact nasal phones, causing an increase in frequency and decrease in amplitude of nasal murmurs (in presumably cisgender participants) [4, 5]. It is possible vocal changes can also happen as a result of other facial feminisation procedures, such as genioplasty and cranioplasty (which involve altering the shape of the chin and forehead respectively, removing cartilage and bone), as they change characteristics of the vocal tract, and it is known that, for example, the mechanical impedance of the vocal tract wall impacts formant frequencies and bandwidths [6].

The voice can also cause dysphoria for transgender people, as particular vocal traits are culturally associated with specific genders: lower mean F0, formants and centre of gravity of /s/ and less use of breathy voice are often associated with masculine voices in English [7, 8, 9, 10]. Speech therapy can help transgender people learn to alter these features [11]. Zimman highlighted that literature on gender differences in the voice often attributes acoustic differences to biological differences, even when no evidence for influence of biological differences exists [12]. In a study on transmasculine people beginning hormonal treatment, participants exhibited vocal changes influenced both by physical changes caused by the testosterone, and by conscious or subconscious efforts to change aspects of their voice, as more people began to perceive them differently [13].

To the best of our knowledge, no studies have investigated the effects of facial feminisation surgeries other than chondrolaryngoplasty and rhinoplasty on the voice. There is therefore a scarcity of information available to individuals considering FFS, regarding the potential vocal impact. We address this gap through an initial investigation of speech, singing and interview data of a singer, before and after undergoing FFS, using data originally collected for building a singing voice conversion system. As chondrolaryngoplasty can cause hoarseness, we hypothesise there may be differences in the participant's voice quality across recording sessions. We therefore analyse voice quality measurements H1-H2 and H1-A1, which correlate with hoarseness [14], and calculate long-term average spectra (LTAS). We hypothesise that cranioplasty, rhinoplasty and genioplasty may affect resonances, so we analyse F1, F2 and F3. To explore a potential (sub)conscious change the participant may make to her voice, we investigate centre of gravity (CoG) of /s/, as it correlates with gender in English (among other languages), and the observed patterns are believed to be socially rather than biologically driven [9]. Singing is analysed alongside speech because singers may be particularly concerned with the possible vocal impact of FFS. In line with the findings of [13] for transmasculine people, we hypothesise that any vocal changes observed may be a combination of direct and indirect results of the physical changes the surgery causes.

2. Methods

2.1. Participant

The participant, who is an adult transgender woman, is a professional singer. She underwent cranioplasty, rhinoplasty, genioplasty and chondrolaryngoplasty. She is a native speaker of Bulgarian and describes her English accent as 'international', having been influenced by American and British media. She was 25 at the time of surgery. Prior to the surgery and data collection, she had 4 gender-affirming speech therapy sessions.

2.2. Data Collection

Singing and speech data was collected at three time points: a month prior to surgery, 3 months post-surgery, and 7 months post-surgery. Recordings were made in the participant's record-

 10.21437/Interspeech.2024-1132

ing studio using a Sonotronics Aria microphone with a UA-610B (UAD plugin) pre-amp, with a 48kHz sampling rate and a 24-bit depth. Interview data was additionally collected shortly after each recording session.[1]

Speech data consisted of the participant reading the Rainbow Passage [15]. Singing data consisted of a subset of songs from the NUS Sung and Spoken Lyrics Corpus [16]. Only songs recorded in all three sessions were analysed, so phonemic material was consistent.[2] This resulted in 2 minutes of speech and 16 minutes of singing on average from each recording session.

Interview recordings were collected to track changes in the participant's feelings and perception regarding her voice. Employing a semi-structured interview approach, sessions explored the participant's rationale behind opting for FFS and engaging in this research, alongside her perception of her voice. She also described her experiences concerning the interplay between her voice, physical appearance, and sense of self-identity, as well as her anticipations and concerns regarding both her voice and identity. Each interview lasted roughly one hour.

2.3. Annotation

Speech recordings were automatically segmented using the Montreal Forced Aligner [17]. For the singing data, a singing-oriented forced aligner pretrained on English data was used.[3] For both, a dictionary for a US variety of English was used, as the researchers deemed this the closest available to the participant's variety. All forced aligner output boundaries were systematically checked, and errors were manually corrected.

2.4. Measurements

Formants and voice quality measurements were extracted from all vowel tokens (1305 tokens in speech and 3237 in singing), and CoG was extracted from all /s/ tokens (142 in speech and 438 in singing) using Praat scripts adapted from [18] and [19] (which uses the methods described in [20]) respectively. Measurements were taken at the midpoint of each token. For LTAS, recordings were first normalised and manually end-pointed to remove silence at the start/end. Mel-spectra were calculated over each sound file using Librosa's [21] melspectrogram function with an FFT window of 2048 samples, hop size of 512, and a Hann window. 80 Mel-frequency bands between 0 and 8 kHz were analysed. LTAS were calculated by averaging the power spectra over time before computing dB relative to peak power.

2.5. Analysis

To explore the effects of surgery on the formant and voice quality measurements, linear mixed effects regression (lmer) models were fit to the data using using lme4 [22] in R [23], for each dependent variable. Speech models had recording date as a predictor and vowel quality as a random intercept. For singing models, song was also given as a random intercept. To explore (indirect) effects of surgery on CoG of /s/ in singing, we fit a lmer model with CoG as the dependent variable, recording date as the predictor and song as a random intercept. For speech data, a linear regression model was fit, with CoG as the dependent variable, and recording date as the predictor. In all models, date was given as a categorical variable, to account for the possibility of measurements not following a single trajectory across the three sessions. For formants and CoG, frequencies were analysed in Hz. We did not control for the position of the token in the word/sentence, as phonemic material was matched across recording sessions. For each model, an ANOVA was used to compare it with a null model which did not specify date as a predictor, to quantify the effect of recording date on the dependent variables. Pairwise t-tests with Bonferroni correction were used to compare differences in measurement values between pairs of recording sessions.[4] For all statistical tests, the standard alpha level of 0.05 is adopted. We analysed LTAS plots descriptively, and conducted a thematic analysis on the interview data [24].

3. Results

3.1. Quantitative Results

3.1.1. Formants and Voice Quality Measurements

The left and middle panels of figure 1 show values for the formants and voice quality measurements across recording sessions.[5] For all regression models except F1 for speech and F2 for speech and singing, the effect of recording date was significant ($p < .001$ for each). For speech, H1-H2 and H1-A1 are significantly lower 3 months post-surgery compared with the other two recording sessions, and H1-H2 7 months post-surgery is also significantly lower than pre-surgery. For singing, both voice quality measures are significantly lower in the post-surgery recordings compared with the pre-surgery recording, and for H1-H2 the 7 months post-surgery recording is also significantly higher than the 3 months post-surgery recording.

In singing, mean F1 shows an increase between each consecutive recording session, and is significantly higher in the 7 months post-surgery recording than the pre-surgery recording. In speech, F3 is significantly lower in the 3 months post-surgery recording compared with the other two recordings. In singing, F3 is significantly lower in the post-surgery compared with the pre-surgery recordings, and is significantly higher in the 7 months post-surgery recording compared with the 3 months post-surgery recording.

3.1.2. Long-Term Average Spectra

Figure 2 shows the LTAS. There is little difference between the speech recordings, but a slightly shallower spectral slope in the 3 months post-surgery recording compared with the pre-surgery recording. The spectral slope of the 7 months post-surgery speech recording lies between those of the pre-surgery and 3 months post-surgery recordings. There is a larger difference in spectral slope in singing, with both post-surgery recordings having a shallower slope than the pre-surgery recording, and little difference between the two post-surgery recordings.

3.1.3. Centre of Gravity of /s/

The rightmost panel of figure 1 shows the CoG of /s/ across the three recording sessions. Recording session was a significant predictor of CoG in speech and singing ($p < .05$ for each). CoG values were overall high across all conditions. In speech, no pairwise comparisons were significant. In singing, mean values increased across consecutive recording sessions, and val-

[1] We do not make the data publicly available, since the participant did not consent to this.

[2] Namely: *Can You Feel the Love Tonight*, *I Have A Dream*, *Jingle Bells*, *Moon River*, *Silent Night*, and *You Are My Sunshine*.

[3] https://github.com/qiuqiao/SOFA/discussions/4

[4] We do not directly compare singing and speech, because the phonemic content is not matched across modalities, and there was insufficient data to control for this and still make a meaningful comparison.

[5] Since the data was phonemically matched across recording sessions, we do not analyse individual vowel phonemes separately here.

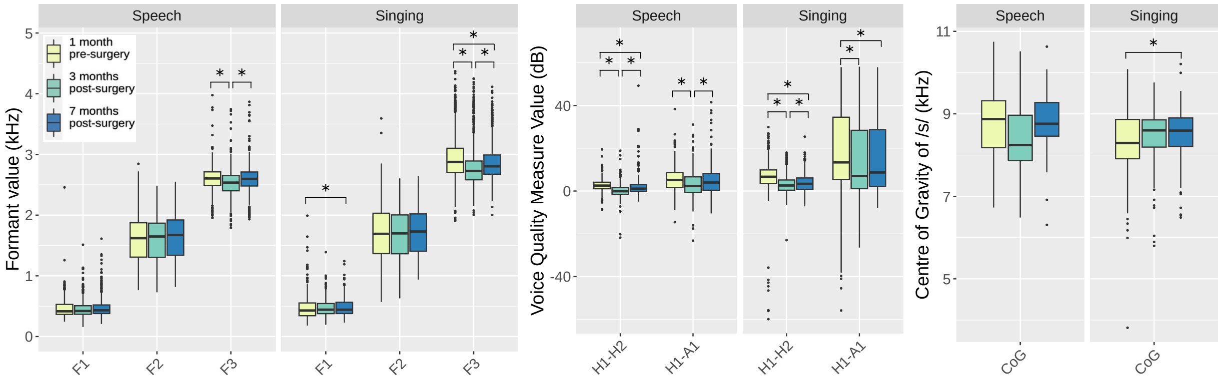

Figure 1: *Formants, voice quality and centre of gravity of /s/ measurements across recording sessions for speech and singing data. Statistically significant pairwise comparisons are indicated with *.*

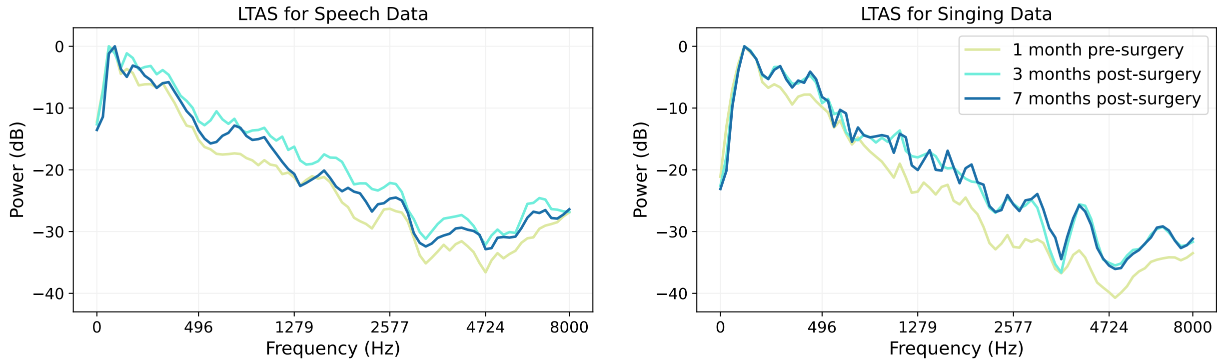

Figure 2: *Long-term average spectra (LTAS) for speech and singing data recorded in three different sessions.*

ues were significantly higher 7 months post-surgery than pre-surgery.

3.2. Qualitative Results

The following themes emerged from the interview data: i) changes to the voice and vocal practice; ii) emotional impact of surgery; iii) impact of surgery on work/life; iv) representation and community. Here, we focus only on the aspects of i and ii that relate most directly to the quantitative data. The participant noted few physical changes to her voice after surgery. She stated that she initially felt less in control when singing higher notes, but this feeling faded over time. She also spoke about ways the surgery affected her relationship with her voice. Before surgery, she stated she did not generally experience vocal dysphoria. She had speech therapy prior to this study, where she had learnt vocal feminisation techniques, but seldom felt the need to use them. However, following the surgery, she became conscious of an incongruence between her face and her voice, talking about how her face now appeared very feminine but her voice still sounded like “a voice touched by testosterone”. This led to an onset of vocal dysphoria, which in turn led to her putting the aforementioned vocal feminisation techniques into practice more often.

4. Discussion

4.1. Quantitative Results

4.1.1. Voice Quality

We hypothesised that the chondrolaryngoplasty may have an effect on the singer’s voice quality. H1-H2 and H1-A1 show differences in the pattern of results across singing and speech, and the differences observed are not in the expected direction: hoarseness correlates with higher measures, but the measures are lower 3 months post-surgery than they are pre-surgery. These differences across recording sessions are therefore unlikely to be due to physical changes caused by surgery. Instead, they may be due to artistic decisions made by the singer on how to perform the material on different dates. Specifically, she may have opted for a less breathy reading of the Rainbow Passage in the 3 months post-surgery recording session and a more breathy performance of the songs in the pre-surgery recording session compared with the other recording sessions. Alternatively, H1-H2 has also been found to correlate with vocal fatigue, with higher values corresponding with more fatigued voices [25]. The participant performed far more material before the analysed recordings in the pre-surgery recording session, so it is also possible the differing H1-H2 measures across sessions are due to differing levels of vocal fatigue.

The LTAS show shallower spectral slopes in post-surgery compared with pre-surgery recordings for singing and speech. Previous research finds women’s voices tend to have a shallower spectral slope than men’s [26]. This is attributed to greater aspiration noise in the higher frequencies, and thus the results may be due to a (sub)conscious effort from the participant to shift her voice quality, or it may also be due to artistic decisions or vocal fatigue, as discussed above.

4.1.2. Formants

We hypothesised that the cranioplasty, rhinoplasty and genioplasty would have an effect on the participant’s formants, through physically altering the vocal tract. However, in addition to being potentially influenced by an altered vocal tract shape, formants can be changed by the speaker through manipulation of the larynx position, and altering formant frequencies is often a focus of voice and communication therapy for transgender individuals, as higher formants are commonly associated with feminine voices [27]. Since the significant raising of F1 over time is gradual and found only in singing, it is likely a result of the participant making a (sub)conscious effort towards more typically feminine vocal features in her singing. For both singing and speech, F3 was significantly lower in the 3 months post-surgery compared with the other two sessions. By 7 months post-surgery, F3 had returned to pre-surgery levels in speech, while in singing it had risen to somewhere between the pre-surgery and 3 months post-surgery levels. These results are consistent with a hypothesis that the surgery directly affected F3 for this participant, through altering the vocal tract. The raising of F3 again by 7 months post-surgery may be caused by the participant learning, either consciously or subconsciously, to actively compensate for this effect (because it was not in the desired direction: lower formants are associated with masculine voices). This active compensation for surgical effects has been found for other facial surgeries (in presumably cisgender participants) [28, 29]. We note, however, that it is also possible that these results are due to the participant simply employing vocal feminisation practices relating to formant raising differently in the different recording sessions.

4.1.3. Centre of Gravity of /s/

CoG of /s/ is significantly higher in post-surgery recordings than pre-surgery recordings in singing. These results seem indicative of the participant making a (sub)conscious effort to employ more feminine features. The results for speech are less clear, with a significant effect of recording session but no significant differences between pairs of recordings. Further data would be needed to draw firm conclusions on the speech results.

4.1.4. General Discussion of Quantitative Results

The pattern of results for several measures in the speech data shows a difference in the 3 months post-surgery recording compared with pre-surgery, and either a return to pre-surgery levels or to an intermediate level by 7 months post-surgery. In singing, some measures follow the same pattern, but others indicate a move towards values typically associated with femininity, linearly over the three recording sessions. The speech results are consistent with prior research investigating impacts of both surgical procedures and gender-affirming speech therapy, whereby participants initially show a large change in acoustic measures but over time either return to pre-treatment levels or to an intermediate level [28, 29, 30]. In our results, the changes observed in the 3 months post-surgery recordings may be caused by the surgery directly, or indirectly as a result of the surgery causing the participant to employ vocal feminisation practices more regularly. The different pattern of results in the singing data may be due to the fact that singing is a method of performance: it is less casual than speech and used less frequently, and it involves a greater focus on producing sounds in a specific way.

4.2. Qualitative Results

In addition to informing our analysis of the quantitative results above, the participant's interview data provides valuable insights into how FFS can impact a person's relationship with their voice. This highlights important issues relating to evaluating the acoustic impact of the physical changes to a person's vocal tract caused by FFS. The fact that a person's relationship with their voice can change so much following FFS means that they may be making conscious changes to their voice as well at this time. Consequently, the acoustic impact of physical changes caused by FFS cannot be investigated in isolation. Furthermore, this highlights the inherent complexity in data collection for this type of research, as vocal dysphoria can make it less likely for a person to want to record their voice and consent to their recordings being used for research purposes.

4.3. Limitations of the Study

The singing and speech recordings were initially collected for the purpose of creating a singing voice conversion system, to convert the singer's post-surgery voice back to her pre-surgery voice, for her to use creatively in performance. Hence, our data is not ideally suited for fully evaluating all of the potential acoustic changes in her voice. For example, the data does not allow for analysis of changes to minimum and maximum F0, since we did not ask the singer to demonstrate the full extent of her range in the recordings. It is also important to note that there was far more singing data than speech data, again due to the recordings being collected for a singing voice conversion system, and this may have contributed to some of the differences in the patterns of results in the two modalities.

We highlight FFS as an area where disentangling social and physiological factors affecting the voice is extremely difficult, as the features expected to change with surgery (formants, voice quality) can also be changed consciously and are a target of gender-affirming speech therapy. As our interview data demonstrates, FFS can bring an unexpected change in a person's relationship with their voice, resulting in (sub)conscious vocal changes happening concurrently with physical changes. We have attempted to disentangle these explanatory factors in our discussion through comparing patterns of results across modalities, and stated what we believe to be the likely causes of significant results given the evidence, but we acknowledge there are several possible reasons for our results, and further research is needed to distinguish between these and alternative hypotheses.

Having only one participant of course limits the generalisability of the results to the wider population, and the fact that the types of facial feminisation surgeries are numerous and varied adds to this issue. Further research with more participants is needed in this area, to draw more generalisable conclusions.

5. Conclusion

This study has presented an acoustic analysis of the speech and singing data of a transgender woman before and after undergoing facial feminisation surgery, and given a brief description of relevant aspects of interview data with the participant around her changing perceptions and feelings towards her voice. We have highlighted that the participant's evolving relationship with her voice, brought on by the surgery, means that acoustic differences caused purely by changes to the vocal tract cannot be investigated in isolation. Since facial feminisation procedures are numerous and varied, there is much more work to be done in this area. We hope this initial study will motivate further research into the effects of facial feminisation surgeries on the voice, with more participants, so that in future people considering undergoing the procedures can have a fuller understanding of the likely vocal changes involved.

6. Acknowledgements

This work was supported by the Centre for Doctoral Training in Speech and Language Technologies (SLT) and their Applications funded by UK Research and Innovation [grant number EP/S023062/1]. Ethics approval for this study was granted by the Ethics Committee of the Music Department at the University of Sheffield.

7. References

[1] K. Becker, S. u. D. Khan, and L. Zimman, “Beyond binary gender: creaky voice, gender, and the variationist enterprise,” *Language Variation and Change*, vol. 34, no. 2, p. 215–238, 2022.

[2] K. Altman, “Facial feminization surgery: current state of the art,” *International Journal of Oral and Maxillofacial Surgery*, vol. 41, no. 8, pp. 885–894, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S090150271200197X

[3] B. Nuyen, Z. J. Qian, M. Rakkar, J. P. Thomas, E. Erickson-DiRenzo, and C. K. Sung, “Diagnosis and management of vocal complications after chondrolaryngoplasty,” *The Laryngoscope*, vol. 133, no. 9, pp. 2301–2307, 2023.

[4] M. Foroughian, K. Khazaeni, M. R. Haghi, N. Jahangiri, L. Mashhadi, and M. Bakhshaee, “The potential effects of rhinoplasty on voice,” *Plastic and Reconstructive Surgery*, vol. 133, no. 2, pp. 109e–113e, 2014.

[5] M. Nemati, A. Tahmasebi, H. Mohajerani, and R. Tabrizi, “Does open rhinoplasty alter voice quality?” *Journal of Oral and Maxillofacial Surgery*, vol. 77, no. 1, pp. 179.e1–179.e5, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278239118310838

[6] M. Fleischer, S. Pinkert, W. Mattheus, A. Mainka, and D. Mürbe, “Formant frequencies and bandwidths of the vocal tract transfer function are affected by the mechanical impedance of the vocal tract wall,” *Biomechanics and Modeling in Mechanobiology*, vol. 14, pp. 719–733, 2015.

[7] V. Cartei, R. Bond, and D. Reby, “What makes a voice masculine: Physiological and acoustical correlates of women’s ratings of men’s vocal masculinity,” *Hormones and Behavior*, vol. 66, no. 4, pp. 569–576, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0018506X14001639

[8] M. P. Gelfer and V. A. Mikos, “The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels,” *Journal of Voice*, vol. 19, no. 4, pp. 544–554, 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0892199704001729

[9] E. A. Strand, “Uncovering the role of gender stereotypes in speech perception,” *Journal of Language and Social Psychology*, vol. 18, no. 1, pp. 86–100, 1999.

[10] S. Whitling, H. M. Botzum, and M. R. van Mersbergen, “Degree of breathiness in a synthesized voice signal as it differentiates masculine versus feminine voices,” *Journal of Voice*, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0892199723001509

[11] E. Coleman, A. Radix, W. Bouman, G. Brown, A. de Vries, M. Deutsch, R. Ettner, L. Fraser *et al.*, “Standards of care for the health of transgender and gender diverse people, Version 8,” *International Journal of Transgender Health*, vol. 23, no. 1, pp. S1–S259, 2022.

[12] L. Zimman, “Transgender voices: Insights on identity, embodiment, and the gender of the voice,” *Language and Linguistics Compass*, vol. 12, no. 8, 2018, e12284 LNCO-0742. [Online]. Available: https://compass.onlinelibrary.wiley.com/doi/abs/10.1111/lnc3.12284

[13] ——, “Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/,” *Language in Society*, vol. 46, no. 3, p. 339–370, 2017.

[14] S. V. Narasimhan and K. Vishal, “Spectral measures of hoarseness in persons with hyperfunctional voice disorder,” *Journal of Voice*, vol. 31, no. 1, pp. 57–61, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0892199716000680

[15] G. Fairbanks, *Voice and Articulation Drillbook*. Harper & Row, 1960.

[16] Z. Duan, H. Fang, B. Li, K. C. Sim, and Y. Wang, “The NUS sung and spoken lyrics corpus: A quantitative comparison of singing and speech,” in *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 2013, pp. 1–9.

[17] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, “Montreal Forced Aligner: Trainable text-speech alignment using Kaldi,” in *Proc. Interspeech 2017*, 2017, pp. 498–502.

[18] C. Vicenik, “Phonation measurements script,” 2009. [Online]. Available: http://phonetics.linguistics.ucla.edu/facilities/acoustic/PraatVoiceSauceImitator.txt

[19] C. DiCanio, “Spectral moments of fricative spectra script in praat,” 2013. [Online]. Available: https://www.acsu.buffalo.edu/~cdicanio/scripts/Time_averaging_for_fricatives_4.0.praat

[20] K. Forrest, G. Weismer, P. Milenkovic, and R. N. Dougall, “Statistical analysis of word-initial voiceless obstruents: Preliminary data,” *The Journal of the Acoustical Society of America*, vol. 84, no. 1, pp. 115–123, 07 1988. [Online]. Available: https://doi.org/10.1121/1.396977

[21] B. McFee, M. McVicar, D. Faronbi, I. Roman, M. Gover, S. Balke, S. Seyfarth, A. Malek *et al.*, “librosa/librosa: 0.10.1,” Aug. 2023. [Online]. Available: https://doi.org/10.5281/zenodo.8252662

[22] D. Bates, M. Mächler, B. Bolker, and S. Walker, “Fitting linear mixed-effects models using lme4,” *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[23] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2019. [Online]. Available: https://www.R-project.org

[24] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa

[25] S. Narasimhan and M. Soumya, “Spectral and cepstral measures of vocal fatigue in Indian heavy metal vocalists,” *Journal of Indian Speech Language & Hearing Association*, vol. 34, no. 2, pp. 241–246, 2020.

[26] E. Mendoza, N. Valencia, J. Muñoz, and H. Trujillo, “Differences in voice quality between men and women: Use of the long-term average spectrum (LTAS),” *Journal of Voice*, vol. 10, no. 1, pp. 59–66, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0892199796800191

[27] L. Carew, G. Dacakis, and J. Oates, “The effectiveness of oral resonance therapy on the perception of femininity of voice in male-to-female transsexuals,” *Journal of Voice*, vol. 21, no. 5, pp. 591–603, 2007.

[28] J. Bowers, E. A. Tobey, and R. Shaye, “An acoustic-speech study of patients who received orthognathic surgery,” *American Journal of Orthodontics*, vol. 88, no. 5, pp. 373–379, 1985.

[29] A. S. Lee, T. L. Whitehill, V. Ciocca, and N. Samman, “Acoustic and perceptual analysis of the sibilant sound /s/ before and after orthognathic surgery,” *Journal of Oral and Maxillofacial Surgery*, vol. 60, no. 4, pp. 364–372, 2002.

[30] M. P. Gelfer and R. M. Tice, “Perceptual and acoustic outcomes of voice therapy for male-to-female transgender individuals immediately after therapy and 15 months later,” *Journal of Voice*, vol. 27, no. 3, pp. 335–347, 2013.