ARTICULATING A FEMALE VOWEL: MALE TO FEMALE TRANSGENDER VOICE THERAPY

Caroline Menezes, Ph. D., CCC-SLP, Marisa Lucarelli, M.A., Taylor Koesters, B.S., Kristen Ruta, Alexis Rymers, Kassidy Turshon

University of Toledo, Toledo, Ohio, USA
caroline.menezes@utoledo.edu, marisa.lucarelli@rockets.utoledo.edu, taylor.koesters@rockets.utoledo.edu, kristen.ruta@rockets.utoledo.edu, alexis.rymers@rockets.utoledo.edu, kassidy.turshon@rockets.utoledo.edu

ABSTRACT

Achieving true feminine-voice quality requires manipulation of phonetic features like fundamental frequency and resonance. This study focuses on resonatory changes and plots the changing vowel space for two male-to-female transitioning patients who received therapy at the university clinic. Acoustic and articulatory output of 4 English vowels were recorded using the 3-D electromagnetic articulograph (EMA). Transgender subjects were recorded at the beginning, middle, and end of their treatment period. Acoustic data was analyzed using PRAAT and articulatory data was analyzed using the VisArtico software. The acoustic and articulatory vowel spaces of transgender subjects were further compared on a continuum with biological female subjects. Both subjects showed an increase in F2 values when sounding more feminine. Furthermore, F1 was also increased but mostly for the lower vowels resulting in larger vowel spaces. Both subjects had different articulatory strategies which is explained in the paper.

Keywords: transgender, vowel-space, articulatory vowel-space, resonance, speech-therapy

1. INTRODUCTION

Transgender (TG) individuals believe that their biological sex does not match their psychological gender assigned at birth [5, 9]. Of those who identify as TG, a percentage will choose to "transition" to the gender with which they identify. This may include all or some of the following: dressing as the desired gender, changing their name, hormone therapy, voice therapy, or sexual re-assignment surgeries [9]. For a male-to-female (MTF) TG person, the ability to successfully adopt female-like communication behaviors is often desirable [7]. This study follows two TG persons at three stages within their transition from male to female.

Research findings focus on the importance of pitch in gender attribution to a speaker [8]. This is hardly surprising given that pitch appears to be critical in gender differentiation in cisgender speakers (speakers whose gender identity is defined by their biological sex) [10]. Therefore, the most common type of therapeutic intervention with MTF clients involves changing fundamental frequency [8]. For biological males to be successfully perceived as female, studies show fundamental frequency must be increased to at least 155-165Hz and sometimes must increase to as much as 180Hz, which falls in the gender ambiguous fundamental frequency range [11]. The higher the pitch is increased, the greater is the chance of being perceived as female [7].

Altering fundamental frequency alone is not sufficient in achieving the true perception of a female voice [3, 4, 5, 7, 8]. In a study analysing synthetic speech, after appropriately adjusting fundamental frequency and formants, Klatt and Klatt reported preliminary findings that women's voices tended to be slightly breathier than those of men indicating changes in acoustic measures and perception [10]. Mount and Salmon in their case study reported that their client was never perceived as a female while speaking on the phone with a habitual pitch of 210 Hz. In this case, female perception occurred only when the subject spoke using F2 values closer to the female range [12]. This study highlights the effect of vocal tract size and resonance on the perception of gender. Resonance is described as the modification of the glottal sound source in the vocal tract [7]. Most women have shorter vocal tracts than men, resulting in differences in formant frequencies in which women's voices are approximately 20% higher than those of males [4, 10]. The shape and length of the vocal tract may be altered by rounding or retracting the lips and/or changing the place where the vocal tract is constricted, which affects the vowel formant frequency values [2]. Women also tend to speak towards the front of their vocal cavity, and therefore, some speech-language pathologists teach the MTF client to speak with the tongue positioned more forward in the oral cavity [6].

In this study, we seek to examine the resonatory changes of the two TG speakers as they progress through therapy and compare them with age-matched female controls.

2. METHODS

2.1. Subjects

Two MTF TG subjects received voice therapy at the University of Toledo speech-language clinic. The clients were JJ, age 22 at time of recording, and EE, age 24 at time of recording. Six female speakers of American English and of normal hearing from the Northwest Ohio area aged 18-20 were recorded as control female speakers. Subjects and control speak the same dialect of American English (Midwest Standard). Patients were treated at the clinic at different times. Both patients were receiving hormone therapy simultaneously.

2.2. Treatment Paradigm

The two TG speakers had slightly different approaches to therapy. For JJ, treatment first focused on increasing fundamental frequency, and upon achieving female f0 range, it changed to using a forward articulation to decrease the front resonance chamber. EE, however, worked simultaneously on increasing fundamental frequency and forward resonance. Therapy for both increasing pitch and changing resonance were first initiated on vowels in isolation moving through the four extreme vowels in the American English vowel quadrilateral. On achieving 90% accuracy at the vowel level therapy targeted accuracy at the basic syllable level (CV) working on a small set of voice and voiceless onsets. Finally, words and phrases containing a predominance of the same vowels were also introduced in the practice routine. The clients were asked to maintain on average the target pitch or F2 value through the entire practice unit while maintaining natural speech prosody. Normal intonation declinations were expected while the habitual target was practiced. Simple instructions like “lets raise the pitch”, “speak forward”, “articulate in the front” were used. Auditory and visual feedback using pitch analysers and PRAAT spectrograms were used.

2.3. Experimental Paradigm

The Electromagnetic Articulograph (EMA) AG 500 and AG 501 were used to record acoustic and articulatory data. Reference coils were placed on the left and right mastoid processes and the upper jaw - between the two maxillary central incisors. Additional coils were placed on the cupids bow of the upper lip, the center of the vermillion border of the lower lip, lower jaw - between the two mandibular central incisors and three points on the tongue corresponding to tongue tip, tongue dorsum and tongue back. Subjects read the stimuli from a computer screen that was placed approximately five feet in front of the subject. The stimuli comprised of natural English utterances with the target word occurring once in either the initial, medial or final position of the utterance. For this study, we focused on target words that contained four English cardinal vowels /i ɑ u æ/. Each vowel was produced 15 times per recording – five times in each of the three sentence positions. Here is an example of stimuli for the vowel /i/:

1. *They **flee** past the mob.*
2. *They caught a big **flea**.*
3. ***Flee** from the mad riots.*

Each subject was recorded three times over the duration of treatment: once at the start of treatment, once about halfway through therapy, and once prior to discharge from therapy. Acoustic variations for the transitioning subjects were compared to the average data recorded from six female controls recorded a single time.

2.4. Analysis

PRAAT software was used for labeling the acoustic stimuli and for the subsequent formant analysis [1]. First formant (F1) and second formant (F2) values were measured at the most stable point in the vowel. Articulatory and acoustic signals were imported into Visartico, a visualization software [13]. Articulatory measurements were made at the point of maximum deviation for Jaw Z, Tongue Blade, Tongue Tip, Tongue Dorsum Z (vertical displacement) and Tongue Blade, Tongue Tip, Tongue Dorsum X (horizontal displacement) within the PRAAT labelled vowel unit.

3. RESULTS & DISCUSSION

3.1. Acoustic Analysis

Both subjects were successful in transitioning from male voice to female voice and were discharged after a year of therapy. Therefore, it was crucial to understand how they executed forward resonance in the acoustic and articulatory domains. Vowel spaces were plotted for each subject as a scatterplot of mean F1 and F2 values for the four cardinal vowels. TG subjects were separately compared at the three recording stages and then subsequently compared to the mean F1 and F2 values for all control females. See Tables 1-3 for a descriptive of formant values for a numeric comparison. See Figures 1 and 2 for Subject JJ and EE, respectively.

Figure 1: Female Control vs. JJ Acoustic Vowel Space at Three Recording Intervals

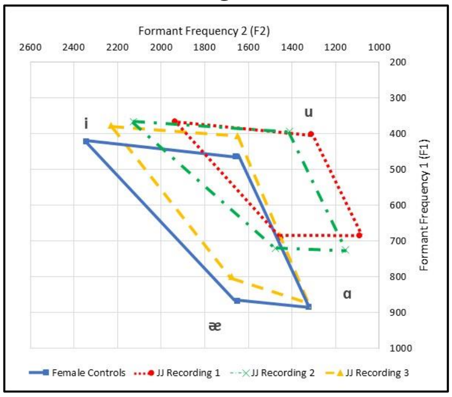
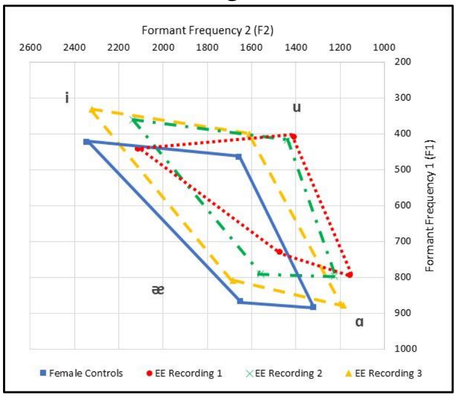


Figure 2: Female Control vs. EE Acoustic Vowel Space at Three Recording Intervals



Both TG speakers show an increased conformity to the female control vowel space as the recordings progressed, with the third recording most resembling the controls. Figures 1 and 2, show that both subjects increased F2 values for all vowels as would be expected if the tongue was positioned forward. Interestingly, there was also a concomitant increase in F1 indicating a downward positioning of the jaw/tongue to match the female control. It is clearly evident in the acoustic analyses that subjects had a gradual but accurate transition from male to female vowels. So, the next question to ask is how did they do it?

Table 1: Healthy Female Acoustic Descriptive Means

	Healthy Females	
	F1 (Hz)	F2 (Hz)
i	423	2343
ɑ	884	1321
u	464	1656
æ	868	1651

Table 2: JJ Acoustic Descriptive Means

	JJ					
	Recording 1		Recording 2		Recording 3	
	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
i	365	1935	366	2126	376	2228
ɑ	684	1090	723	1154	877	1329
u	402	1313	394	1412	405	1648
æ	690	1451	719	1474	800	1678

Table 3: EE Acoustic Descriptive Means

	EE					
	Recording 1		Recording 2		Recording 3	
	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
i	442	2111	360	2134	331	2321
ɑ	797	1153	796	1225	877	1186
u	408	1408	418	1437	400	1610
æ	729	1472	790	1560	808	1685

3.2. Articulatory Analysis

To understand if both subjects used the same articulatory strategy to reach the same acoustic targets, a similar scatter plot was created by plotting the tongue dorsum vertical position against its horizontal position. The scope of this paper prevents us from discussing other crucial articles such as the jaw and other points on the tongue. Mean Tongue Dorsum Z and X values were averaged for each TG subject and plotted on an x-y axis. See Figures 3 and 4 for Subject JJ and EE, respectively. Negative numbers on the x-axis correspond to tongue placement posterior into the vocal tract from resting position. Similarly, negative numbers on the y-axis correspond to tongue positions inferior in the vocal tract from the resting position.

Articulatory vowel spaces reveal a completely different picture from the acoustic vowel spaces. First, at every recording the subject articulated the vowels in a different quadrant indicating that there are different strategies that a person can use to achieve the same acoustic goal. For this reason, we do not compare articulatory changes with control females.

Figure 3: JJ Articulatory Vowel Space at Three Recording Intervals

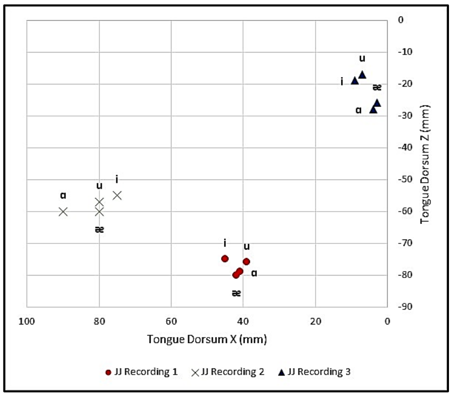


Figure 4: EE Articulatory Vowel Space at Three Recording Intervals

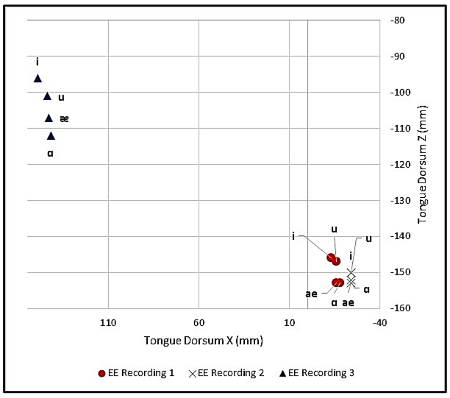


Figure 3 reveals that there is a forward articulation of the tongue from Recording 1 to Recording 2, but in Recording 3, the subject has deviated further back into the vocal tract. There is also a drastic decrease in tongue height which allows for this posterior articulation.

JJ was exposed to a treatment paradigm where fundamental frequency was targeted first, and then resonance therapy was integrated towards the last six months of therapy. Therefore, it is surprising to see that therapy focusing on fundamental frequency showed more changes on the x-axis. In the third recording, her tongue positions were consistently further back in the vocal tract than her previous two recordings. This was successful in increasing the acoustic values to conform to the female controls.

Subject EE (see Figure 4) used exactly the opposite strategy from JJ in that her first and second recordings continued to employ posterior articulation. Towards the end of therapy in Recording 3, EE has dramatically moved forward on the horizontal plane.

In contrast to JJ, EE (Figure 4) was exposed to a treatment paradigm where fundamental frequency and resonance were simultaneously targeted. In her articulatory vowel space progression, we see a clear forward movement of the articulators in Recording 3 well beyond that of Recording 1 & 2.

The similarity between both subjects was the decrease in jaw or tongue height as treatment progressed. A higher jaw or tongue position would result in lower F1 values, but this is not congruent with the higher F1 values in the acoustic domain. In addition, both subjects showed minor variations in vowel articulation from recording to recording, in such a way that front vowels were sometimes articulated more posteriorly than back vowels as in the case of the second and Recording 3 of JJ and the high vowel being articulated lower than /ɑ/ in Recording 2 of EE. [Note: Variation in articulatory position could be influenced by variation in coil placement from Recording to Recording. Since maximum articulatory values were extracted from the acoustic labelling this could also have some effect on the articulatory data.]

4. CONCLUSION

This study, comparing the acoustic and articulatory vowel spaces, reveals several important features of phonetics. First, there is less variance in the acoustic vowel space when compared to the articulatory vowel space. Here, subjects were trained to change their speech articulation to impersonate a vowel system of a relatively shorter vocal tract. Both speakers used different articulatory strategies to achieve near similar acoustic targets.

Second, higher tongue placements can produce high first formant values as was evidenced by both subjects just before they were discharged from therapy. This point needs to be further investigated especially with normal gender and healthy controls.

The treatment strategy of changing resonance along with fundamental frequency resulted in a perceivably more feminine voice for both these clients. However, what we psychologically perceive as more feminine might be a resonance system that habitually matches a given fundamental frequency range and that might be the only purpose for resonance therapy. Regardless, our next step is to conduct a perception study to compare the voice at different stages of TG treatment to understand the categorical boundaries in TG male to female voice change.

Finally, more experiments need to be conducted but we are constrained by the number of persons seeking such treatment at the speech clinic. Similarly, treatment and analysis needs to be done on female-to-male transitions to completely understand how articulating at one end of the vocal tract can change resonance and change the perceived voice of the speaker.

5. REFERENCES

[1] Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345.

[2] 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. doi:10.1016/j.jvoice.2006.05.005

[3] Dacakis, G., Oates, J. & Douglas, J. 2012. Beyond voice: Perceptions of gender in male-to-female transsexuals. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 20(3), 165-170. Doi: 10.1097/MOO.0b013e328530f85.

[4] Freidenberg, C. B. 2002. Working with Male-to-Female Transgendered Clients: Clinical Considerations. *Contemporary Issues in Communication Science & Disorders*, 29, 43-58.

[5] Gelfer, M. P., & Van Dong, B. R. 2013. 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. doi:10.1016/j.jvoice.2012.07.008

[6] Hancock, A. B., & Garabedian, L. M. 2012. Transgender voice and communication treatment: a retrospective chart review of 25 cases. International *Journal of Language & Communication Disorders*, 48(1), 54-65. doi:10.1111/j.1460-6984.2012.00185.

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

[8] Hardy, T., Boliek, C., Wells, K., Dearden, C., Zalmanowitz, C., & Rieger, J. 2016. Pretreatment acoustic predictors of gender, femininity, and naturalness ratings in individuals with male-to-female gender identity. *American Journal of Speech-Language Pathology*, 25, 125-137.

[9] Harris, B. C. 2015. Likely Transgender Individuals in U.S. Federal Administrative Records and the 2010 Census. *CARRA Working Paper Series*, 1-36.

[10] Hillenbrand, J. M., and Clark, M. J. 2009. The roled of f0 and formant frequencies in distinguishing the voices of men and women. Attention, Perception & Psychophysics, 71(5), 1150-1166.

[11] Mordaunt, M. 2006. Pitch and intonation. In R.K. Adler, S. Hirsch, & M. Mordaunt (Eds), Voice and clinical guide (pp. 168-208). San Diego, CA: Plural.

[12] Mount, K. H., and Salmon, S. J. 1988. Changing the vocal characteristics of a postoperative transsexual patient: A longitudinal study. *Journal of Communication Disorders*, 21(3), 229-238.

[13] Ouni, S., Mangeonjean, L., & Steiner, I. 2012. VisArtico: A visualization tool for articulatory data. [13th Annual Conference of the International Speech Communication Association - *InterSpeech* 2012.