Vowel modification in sopranos: COG as an acoustic measure for vowel changes at high pitches

ABSTRACT

One of the techniques that classical sopranos use to sing a wide range of frequencies beautifully is vowel modification. Due to the nature of harmonics spreading out at high fundamental frequencies, a traditional LPC analysis would not be able to accurately track formant changes in soprano voices at high pitches. In this paper, the possibility of using COG as an acoustic measure for sopranos' vowel modification is explored. Six sopranos were asked to sing a range of pitches of around two octaves using the vowels [a], [ɛ] and [ɔ], and the COG of the vowels at each f_o was recorded to interpret the vowels' change in frontness at high pitches. By comparing the COG measurements with the sopranos' mean F1 in speech, the vowels' change in height was also inferred and discussed. Results show that the vowel modification process differs by vowel quality. Furthermore, although sopranos must compromise vowel distinctiveness at high pitches, they still try to distinguish their vowels as much as possible in terms of frontness. Vowel modification appears to be conditioned by height before the second *passaggio* and is conditioned by frontness after the second *passaggio.* Results provide implications for its affect on vowel intelligibility at high pitches, and further research may explore the effects of vowel modification on vowel perception.

INTRODUCTION

Singers need to both vocalize specific pitches accurately and convey lyrics and text clearly. Western classical music covers a wide pitch range, which challenges singers to maintain clear diction (Callaghan, 2000). For sopranos specifically, some vowels are harder to sing at higher pitches than others because the f_o would be as high as vowel formant frequencies in speech. Although the non-linear source filter theory predicts that change in pitch (i.e., voice source) will not affect resonance space (filter) (Ladefoged & Johnson, 2015), empirical data revealed that vowel formants are adjusted in high pitched speech (Chládková, Boersma, & Podlipský, 2009) and vowels should be readjusted based on pitches in singing (Callaghan, 2000). In high pitched speech, the shift in vowel formants is to accommodate for pitch change and word intelligibility. Compared to speech, classical singers have two extra factors to consider, *timbre* and *voice projection*. First, *timbre* refers to the difference in tone quality that is perceptually noticeable when two singers sing the same frequency with the same amplitude (Campbell, 2001). Keeping a consistent timbre across pitch changes is not easy to achieve without vocal training. Even if a person has a refined speaking voice and sings well at low pitches, abrupt timbral changes often occur after reaching a certain pitch. This is because the human vocal range can be divided into different groups of adjacent pitches known as *registers* (Lewcock et al., 2001). When pitches are in the same register, they are sung in a similar way, when two pitches are in different registers, it is more difficult to sing them smoothly. When an untrained singer tries to sing through the registers, a sudden change in timbre may occur, which is known as register breaks (Lewcock et al., 2001). The transitions between registers are *passaggi*, and the *passaggi* ranges for individuals may differ (Callaghan, 2000). Second, classical singers should also consider *voice projection*, as they often sing without amplification technology while keeping themselves heard over the orchestra (Doscher, 1994). Such vocal strength is not controlled at the voice source in order to avoid straining the vocal folds. Instead, singers project their voices by controlling their resonance space. When voice projection is achieved via desirable resonance, singers are said to have a *vocal ring* (Doscher, 1994).

Literature on the acoustics of the singing voice have attributed the *vocal ring* to higher formant clusters at around 3000 Hz, known as the *singer's formant* (Doscher, 1994). The singer's formant has been found in altos, tenors, and basses, and it tends to remain constant even if the f_o or vowel changes. The act of manipulating formant clusters to achieve the vocal ring is called *formant tuning* (Doscher, 1994). Formant tuning is realized differently for sopranos. As they usually sing in higher pitches with widely spaced harmonics, there may not be harmonics around 3000 Hz to form formant clusters (Johan Sundberg, Lã, & Gill, 2013). Sopranos tend to keep their F_1 close to the f_o, reinforcing resonance often via jaw opening (Johan Sundberg et al., 2013). Nonetheless, sopranos must keep the F_1 above f_o, because formant formation is dependent on the component waves of f_o (Johan Sundberg, 1999).

Since sopranos often tune the first formant, which is a vowel determining formant, formant tuning often involves compromising a speech-like vowel space for better resonance (Callaghan, 2000). To sing through registers without abrupt timbral changes, and to maintain a "vocal ring", vowels qualities often need to be adjusted based on pitch ranges (Callaghan, 2000). In vocal pedagogy, the act of adjusting one's vowel space at different pitches is known as vowel modification (*aggiustamento)* (Callaghan, 2000), which has been characterized as singers' conscious vowel changes at different pitches. R. Miller (2000) describes vowel modification as a process where vowels subtly alter towards the schwa when the pitch rises, and proposes that vowels should be modified as follows when approaching the upper *passaggio*: [i] to [ɪ]; [ɪ] to [e]; [e] to [ɛ]; [ɛ] to [ɑ]; [ɑ] to [ɔ]; [ɔ] to [o]; [o] to [ʊ]; and [ʊ] to [u]. After the upper *passaggio*, the vowels are not adjusted towards the schwa; the front vowels are modified towards the neutral center, and the back vowels require more closure (Miller, 2000). This suggests that singers may approach different vowels with a different technique, and the pattern at which vowels change in higher pitches may differ. Bozeman (2013) differentiates two forms of vowel modification. First is R. Miller's method, called *Active Vowel Modification*, which is to deliberately change the shape of their vocal tract and altering the vowel (Bozeman, 2013). The second method is *Passive Vowel Modification*, where a singer maintains the shape of their vocal tract while letting the pitch and its harmonics move accordingly (Bozeman, 2013). If the singer continues to sing past the harmonic boundary without adjusting the vowel space, thus keeps the formant frequency stable, the vowel will sound like it is modified (Bozeman, 2013). Unfortunately, studies concerning vowel modification have often assumed a uniform technique across all voice types, overshadowing the vocal techniques that sopranos use in comparison to voice types with lower f_o ranges. Additionally, while research on resonance in soprano voices have been done (Echternach et al., 2015; Garnier, Henrich, Smith, & Wolfe, 2010), focus has not been on the acoustics of vowel modification patterns. One reason for this is that acoustic studies concerning formant tuning have only focused on voice projection, neglecting the compromises that vowel modifications make.

Another reason that vowel modification has not been well studied acoustically is because the study of vowels typically relies on formant measurements. Formant frequencies are resonant frequencies that amplify specific harmonics as a result of the shape of the vocal tract filter; F_1 is inversely proportional to vowel height, and the distance between F_1 and F_2 is inversely proportional to vowel backness (Zsiga, 2013). There are two ways in which formants could be understood. Traditionally, formants (e.g. F_1, F_2) involves both the source and the filter, and refers to the peaks in the output spectrum envelope (Titze et al., 2015). The most typical way to analyze this is via linear prediction analysis, which is an accurate prediction of formants in speech. However, since harmonics inevitably space out at higher fundamental frequencies, linear prediction analysis has been found to become inaccurate when the f_o exceeds 350 Hz (Monsen & Engebretson, 1983). This means that linear prediction analysis would be unsuitable for studying sopranos' sung formants. Alternatively, "formants" could also be understood as intrinsic resonance frequencies (e.g. R_1, R_2) of the vocal tract (Titze et al., 2015). Such a measurement does not include the voice source and considers the vocal tract itself by tracking resonance frequencies directly. Although studies of resonances in singing have been done and can provide an accurate presentation of singers' strategies (Henrich, Smith, & Wolfe, 2011), it is not representative of the entire acoustic output.

Since an LPC analysis is unsuitable for studying F_1 and F_2 in soprano voices, we propose using another acoustic measure, namely the center of gravity (henceforth COG) of the spectra to infer how vowels modify over an f_o continuum. Studies have used the spectral mean to compare fricatives like /s/ and /ʃ/ (Kris & S. Greg, 1997; Nittrouer, Studdert-Kennedy, &

McGowan, 1989). There difference in frontness is characterized by the COG where /s/, which is more front has a higher mean than /ʃ/. The spectral mean is suitable for the measurement of fricatives since there is a frontness distinction but no height distinction. Since vowels have both a height and frontness distinction, the COG value of vowel sounds should lie between F_1 and F_2, the two most prominent peaks in the spectrum and typically shows no uniform pattern. However, it has been found that sopranos avoid cases where the f_o becomes higher than the R_1, and tunes R_1 to f_o instead (Henrich et al., 2011; J. Sundberg, 1975). This suggests that at high pitch ranges, vowels tend not to differ much in height. If a height distinction between vowels is neutralized, we can infer that like fricatives, a COG measurement may be a predictor that reflects vowel frontness at such a pitch range and can thus be used to study vowel modification. COG measurements are more effective than linear prediction coefficients because although the spectral formant still require the existence of harmonics, the spectral mean considers the frequency's amplitude as well. For example, if F_2 is aligned with $2f_o$, the amplitude of $2f_o$ will be heightened; if F_2 is lower than $2f_o$, the amplitude of $2f_o$ will decrease, and the COG value will decrease too.

The use of COG to measure vowels is not unprecedented. The spectral integration hypothesis acknowledges that the F_1-F_2 model assumes F_1 and F_2 are always perceived as distinct peaks and that the F_1-F_2 model ignores the role of higher formants in speech perception (Hayward, 2016). The center of gravity effect thus argues that the perceived frequency peak is an integration of two formants, which is closer to the original peak that is stronger in amplitude (Hayward, 2016). As such, we propose using COG measurements to explore the acoustics of sopranos' vowel modification.

Against this background, the present study explores when and how sopranos change their vowel backness when singing through a wide pitch range. It is predicted that COG measurements will not be uniform in relatively low pitches. However, when R_1:f_o tuning occurs, the COG values will start to converge and reflect a change in vowel backness. The predictions will be tested through a voice production experiment, where sopranos sing 9 CV combinations across a pitch range of around two octaves in an ascending and descending manner, resembling vocal warm-up exercises. Acoustic data is gathered and analyzed to track the vowel modification process. Data formalization of the obtained acoustic data with specific reference to the interplay between f_o and the COG will follow.

METHODOLOGY

Participants

Six Cantonese native female speakers were recruited. All of them were sopranos who have been trained for classical voice ranging from four to ten years. They were in the age range between 20 and 22. Five of the participants were undergraduate students majoring in music with voice as their main instrument. One of the participants did not major in music but has obtained the Associate Performance Diploma from Trinity College London (ATCL) in voice, which is equivalent in standard to the first year of an undergraduate degree in musical performance. She has also been hired by private music centers as a voice teacher. Using Bunch and Chapman's (2000) singer classification, four of the singers were considered to be category 6.2 "singing teacher[s]", and two of the singers were considered to be category 7.2 "full-time voice student[s]". All participants warmed up their voices before they started the experiment.

Stimuli

The vowels chosen for the study include [a], [ɛ], and [ɔ]. These vowels are common across Italian, French, and German, which are commonly sung languages in both western art songs (e.g. mélodies and Lieder) and operatic repertoires. The three vowels also exist in Cantonese, so participants can read stimuli written in Chinese characters and pronounce the intended vowel naturally. [i] and [u] were not included in the study as they are in complementary distribution in Cantonese open syllables (Kwan, 2014), which does not allow the exhibition of same prevocalic conditions. Each vowel appeared in three CV contexts, with the consonants [p], [t] and [l], totaling up nine CV combinations (3Cs X 3Vs). The three consonants have the place of articulation closer to the front of the mouth and are not nasalized, which ensures that co-articulation effects are minimized. All the nine CV combinations were paired with tone one (high level tone). A level tone was chosen to avoid unnatural singing of flat pitches in a contour tone, as repertoires in complex tonal languages are rather rare. A high tone was chosen as opposed to mid-level (tone 3) or low-level (tone 6) tones because all nine CV combinations in high tone are real words in Cantonese and thus can be represented in Chinese characters. The CV combinations were 巴 [paː˥], 啤 [pɛː˥], 波 [pɔː˥], 吖 [taː˥], 爹 [tɛː˥], 多 [tɔː˥], 啦 [laː˥], 哩 [lɛː˥ and 囉 [lɔː˥]. One consonant was presented for each scalic passage, with the vowels being sung in the order of [a], [ɛ] and [ɔ]. The CV combinations were all sung across a pitch range of around two octaves, starting from B3 (247 Hz) to C6 (1047 Hz), rising by semitone. At every pitch, all three vowels were sung in the same breath group to encourage singers to keep the same level of jaw opening, minimizing F1 differences. The token at 1047 Hz marked the turning point of the ascending and descending scale. Since it was not repeated, this token was excluded from the analysis. Tokens at 247 Hz was also excluded as they were the first and last notes. This means that data from all 9 CV combinations were plotted at 24 different frequencies, starting at 262 Hz, which is within a normal pitch range for female adults' speaking voice.

To keep pitch accuracy and to control vowel length, an audio file was given to participants to sing along with. The audio file was generated using Version 2018.6 of Sibelius First (2006), where the notes were set to the sounds of a digital keyboard to maximize pitch accuracy, with three crotchets per bar, set at a tempo of 88 beats per minute, which corresponds roughly to 682 ms per syllable. Between each pitch change, there was a bar rest where the audio file played the note that was just sung, the next note to be sung, and a crotchet rest. This allowed the participants to expect the next pitch. The words that participants should sing were presented in the form of a color-coded musical manuscript to guide the singers along the audio file, it was also generated using Sibelius First. The Chinese characters were written below the musical stave as lyrics. Only the notes ranging up to 349 Hz were notated fully as guidelines, since the sequence of words repeated at every note, while guiding pitches were given via audio input.

For comparison, the participants' speaking voice for all the given CV combinations were also recorded. It was conducted after the singing task. The stimuli were presented on the screen with the carrier phrase "佢話______啩" *[kʰɵy˨˩ wa:˧ ____ kwa:˧]* (3sg say ___ SFP), meaning "I guess he/she said____". The words before and after the target word were both kept at a flat tone to minimize tonal influences.

Procedure

a. Singing Task All recordings took place at the sound booth of the authors' institute. For this task, participants stood next to a Sennheiser MD 46 microphone on a microphone stand. Another Sennheiser mini clip-on microphone was attached to the participant's collar. Two microphones were used as the volume of the participants' singing gets significantly louder at higher pitches. Two different microphones located at different areas ensured that the participant's voice was picked up through her entire pitch range. For each participant, the recording that has no clippings or better audio quality will be chosen for analysis. The microphones were set at 44.1 kHz sample rate and 16-bit sample depth. Audio was presented using Samsung EHS64 Wired Stereo Earbuds. Participants only used the earbuds on one ear of their own choice, so that they can hear their own voices during the experiment. Participants were first given the visual stimuli presented on paper and asked if they could read all the words out once; this was to ensure that the participants know the intended pronunciation of each character. A practice run using the stimulus set 巴 [pa:˥], 啤 [pɛ:˥], 波 [pɔ:˥] was sung once. This allowed the participant to get used to reading Chinese characters as stimuli, and to adjust the distance of the microphone if needed.

Six sets of stimuli with different consonant orders were used across the participants. For example, one participant sang in the order of [pa] [pɛ] and [pɔ], followed by [ta] [tɛ] and [tɔ] and at last [la] [lɛ] and [lɔ]. Whereas, another participant sang in the order of [t], [p], and then [l]. Each scale was sung in an ascending-descending order, such that there were two values for each CV combination at each pitch.

b. Speaking Task The recordings were taken at the same location as above immediately after the singing task. The same Sennheiser clip-on microphone was used, and the microphone was continued to be set at 44.1 kHz sample rate and 16-bit sample depth. Participants were seated in front of a computer, with a monitor placed at eye level at the participant's most comfortable distance. The stimuli in the carrier phrase were presented in a randomized order using the software Psychopy (Peirce & MacAskill, 2020), with two repetitions for each phrase. A 500 ms gap was left between each stimulus to indicate that the coming phrase was to be a new stimulus in case of consecutive stimuli repetition.

Data processing

In total, 2862 tokens were collected and analyzed with Praat Version 6.0.43 (Boersma & Weenink, 2018). After the exclusion of recordings at the lowest Hz (B3, 247 Hz) and the highest Hz (C6, 1047 Hz), 2700 sounds were left. For the sung stimuli, recordings from the Sennheiser mini clip-on microphone was chosen for participants 1, 2, 3, and 5, and recordings from the Sennheiser MD 46 microphone was chosen for participants 4 and 6. COG measurements for all tokens were taken at the mid-50 percent of the vowel's total duration, minimizing influences from consonantal formant transitions. The F_1 and F_2 values of the speech tokens were also measured at the vowel's mid-point using Praat's LPC analysis.

RESULTS

To see the pattern of COG change, the mean COG value by vowel and participant is plotted in Figure 1. The x-axis shows the target fundamental frequency, the y-axis shows the center of gravity. The diagonal grey lines show f_o to $4f_o$, with the rightmost grey line being the first harmonic (f_o). The participants' average F_1 in speech are included as vertical lines.

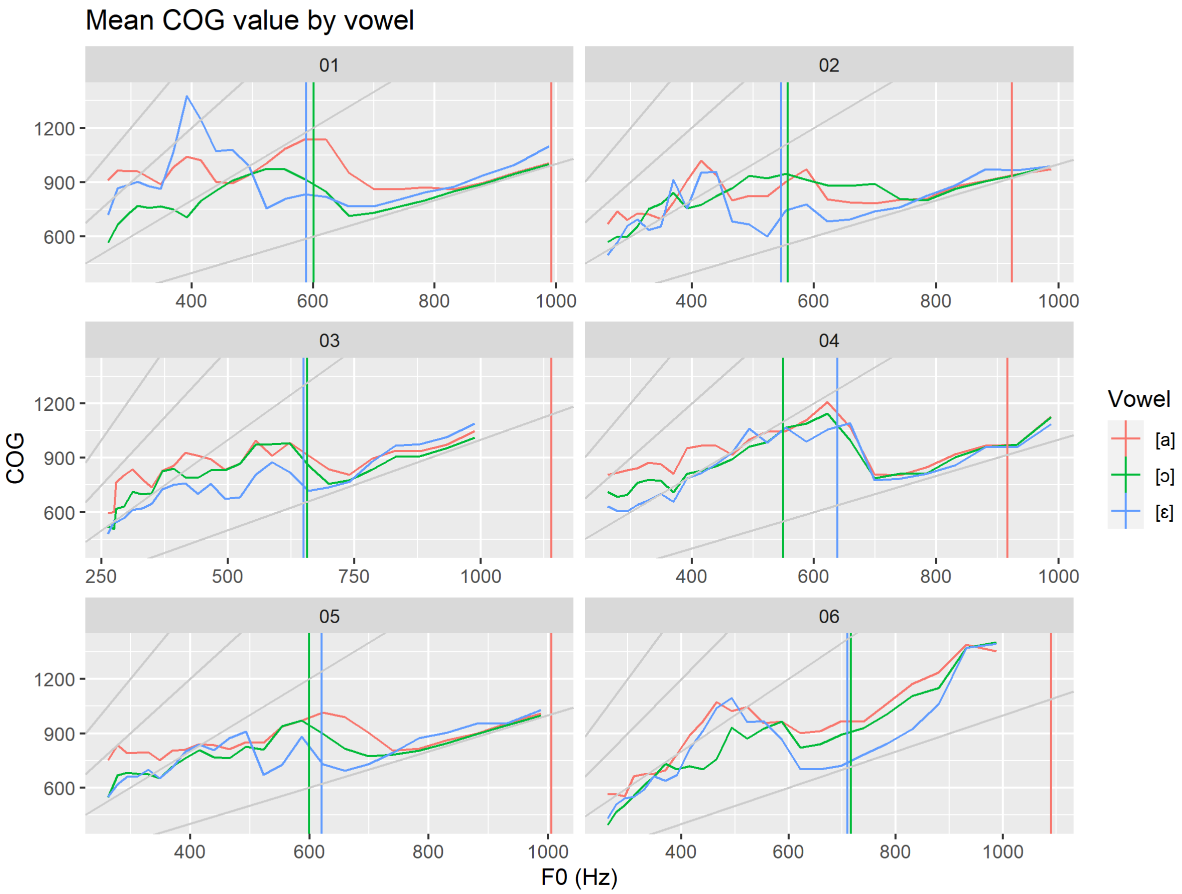

Figure 1. Mean COG (Hz) by participant and vowel. x-axis: target fundamental frequency (Hz), y-axis: COG (Hz) Vertical lines indicate the participant's mean F1 in speech.

Overall, all participants show the same general trend. In the lower part of the pitch range, the COG across vowels are not uniform and fluctuates. This is expected, since the COG should reflect both F_1 and F_2 values which would vary across vowels. To confirm the varying COG levels, the mean COG from spoken vowels are summarized in Figure 2.

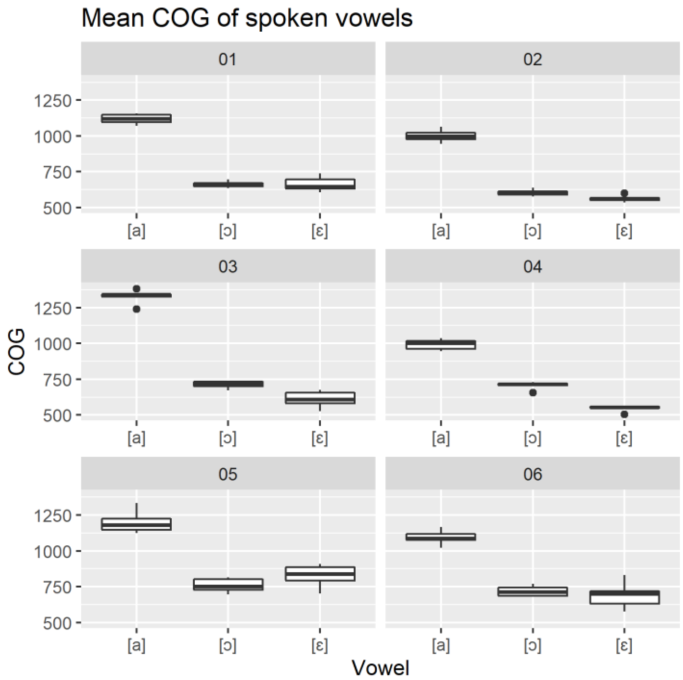

Figure 2. Mean COG (Hz) of spoken vowels by participant. x-axis: vowel, y-axis: COG (Hz)

From around 600 Hz to 700 Hz (around D#5 to F#5 in note value), the COG of the three vowels start to converge and rise in a steady slope. Apart from the measurements from Participant 6, the COG values stay mostly close to the first harmonic (f_o). This confirms that the singers' f_o is emphasized and reinforced by the R_1, such that the COG difference between vowels thereafter could be representative of vowel frontness. The notes at which the COG starts to raise steadily marks the note at which vowels no longer differ by height and only differ by frontness, and each vowel reaches this stage at a slightly different f_o. This is also when we can infer that R_1: f_o tuning might have begun. Usually, [ɛ] starts to modify first, followed by [ɔ] and [a]. Note also that there is a gap between the participant's mean spoken F_1 and the start of R_1: f_o tuning. Results show that the mean F_1 in speech for [ɛ] and [ɔ] are lower compared to the start of R_1: f_o tuning. On the other hand, the mean F_1 in speech for [a] is much higher, its mean F_1 is either at the end of the pitch range, or it surpasses the pitch range that the participants sung in the experiment.

Figure 3 compares the COG values of vowel pairs. In the lower half of the pitch range, the COG difference across the three vowels fluctuate as expected. After around 600 Hz to 700 Hz, the COG difference between [a] and [ɔ] become minimal. However, for pairwise comparisons that involve [ɛ], most participants still demonstrate a slight difference in COG, with [ɛ] having a slightly higher COG than [a] and [ɔ].

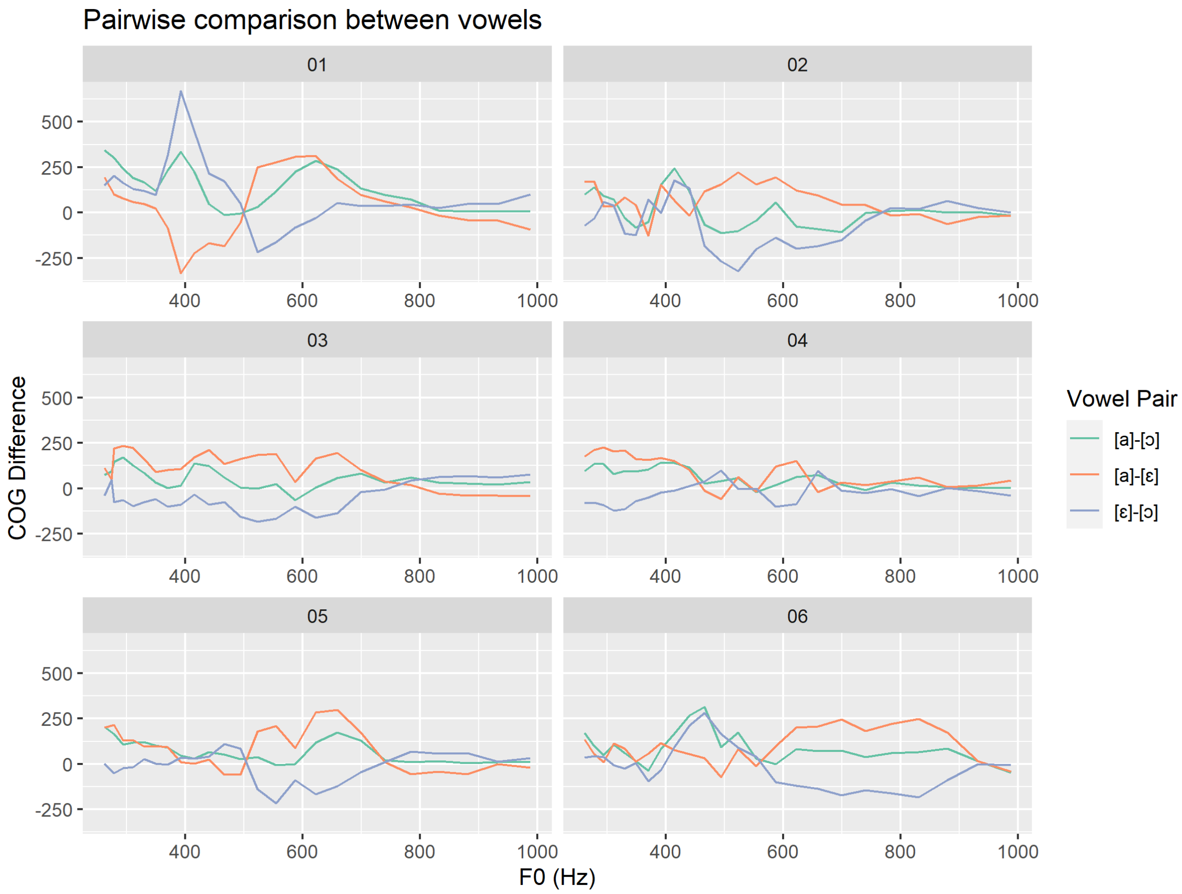

Figure 3. Pairwise comparison between vowels by participant. x-axis: target fundamental frequency (Hz); y-axis: COG difference (Hz)

DISCUSSION

One can infer that R_1 is tuned f_o when the COG increases smoothly and stays close to the first harmonic. In the current set of results, the COG starts increasing unidirectionally around 600 Hz to 700 Hz, which is around D#5 (622 Hz) to F#5 (740 Hz) in note value. This frequency range roughly corresponds to the second *passaggio* as defined by Miller (1986), which is around E5 (659 Hz) to F#5 (740 Hz) for sopranos. This suggests that at the second passaggio, vowel modification with a focus on vowel frontness starts occurring. Considering that [a] and [ɔ] are both back vowels and differ mostly in height during speech, it is unsurprising that their COG difference in the upper range is minimal, as their height difference is minimized by R_1: f_o tuning. Although the vowel height difference between the three vowels are minimized, [ɛ] still shows a COG difference, suggesting that the participants still attempt to differentiate the vowels as much as they could.

The pattern in which the vowels modify at and after the second passaggio may be explained by the dispersion theory of contrast (Lindblom, 1986). The main claim of the theory is that there are three goals of phonological contrast, including (1) maximize the number of contrasts, (2) maximize the distinctiveness of contrasts and (3) minimize articulatory effort. These three goals conflict with each other and phonological inventories seek to strike a balance between these goals. The need to balance between the three goals also appear to be relevant to the singers in question. First, the participants demonstrate that even after the second *passaggio*, a COG difference exist, suggesting that they try to keep a distinction between the front and back vowels. Second, although singers attempt to maximize the distinctiveness of vowels and the three vowels' COGs are not fully merged, vowel distinctions are still compromised. The distance between the vowels are reduced fully in vowel height and drastically in vowel frontness as a result of vowel modification. This process is argued to be systematic as outlined by Miller (2000) and Bozeman (2013), and is used as a strategy to ease singers' articulation when singing at a high register.

The vowel modification strategy is not unified across vowels and depends on vowel quality. The frontness of [ɛ] tends to change before [a] and [ɔ] across all participants. A possible reason for this is that the inherent frontness of the vowel may affect the magnitude of its modification. The trapezoidal shape of the vowel chart roughly represents the side view of our mouths, such that back vowels vary along a vertical line, and front vowels vary along a diagonal line. The fact that the anterior parts of our mouths have a greater range of space for tongue movement suggests that front vowels may have a greater potential for variation and modification compared to back vowels. As [ɔ] lies in the posterior part of our mouths, there is inherently less space for it to vary in vowel backness. While there is no clear consensus as to whether [a] in Cantonese is a front or back [a], since [a] is a low vowel, the small range of space implies that it cannot physically vary much in frontness regardless of its vowel quality. The present study shows that the vowel quality affects the vowel modification process, with the front vowel [ɛ] modifying before vowels [a] and [ɔ]. Further research may discern whether this pattern is widespread and why vowel certain vowels modify earlier than others.

It is also worth comparing the spoken F_1 with the onset of the COG slope change to infer how the vowel modification process differed before the second *passaggio*. As expected, the mid-vowels [ɛ] and [ɔ] have a relatively mid-ranged spoken F_1 value. It is not the case that R_1 tunes to f_o as soon as the f_o is high enough to meet the participants' spoken F_1. The F_1 value for speech tends to be lower than the start of the COG slope change, which suggests that the F_1 of the sung tokens might have risen before the second *passaggio*. The back vowel [a] has a spoken F_1 value that is near or beyond a 1000 Hz. Theoretically speaking, [a] would not have needed to change their F_1 value at all, since it would not coincide with the f_o. However, the vowel [a] still behaves like [ɛ] and [ɔ] and the COG stays close to the f_o around and beyond the second passaggio. This suggests that during the vowel modification process, the F_1 has been lowered. This finding partially supports Miller's (2000) description of the vowel modification process. He has noted that when approaching the upper passaggio [ɛ] modifies towards [ɑ]; [ɑ] towards [ɔ]; and [ɔ] towards [o]. Miller's proposal suggests that vowels modify systematically, in that front vowels like [ɛ] raises its F_1, and back vowels like [ɑ] and [ɔ] lowers its F_1. Our current set of results also suggest that vowels modify systematically, and behavior is determined by height. The mid-vowels [ɛ] and [ɔ] raised its F_1, and the back vowel [a] lowered its F_1 prior to or around the second *passaggio*. This suggests that the vowel modification behavior is systematic and vowel quality dependent, and future acoustic or articulatory research considering a wider range of vowels may outline the exact process of such modification.

CONCLUSION

This study explored the possibility of mapping sopranos' vowel modification process using the COG. Results show that the vowel modification process is systematic and vowel quality dependent. We can infer from the results that vowel modification occurs both before and after the second *passaggio*, and that although vowel distinctiveness is compromised at high pitches, sopranos still try to keep their vowels as distinct as possible. To summarize, the current results suggest that prior to the second *passaggio*, vowel modification is conditioned by height, with [ɛ] and [ɔ] behaving similarly; after the second *passaggio*, vowel modification is conditioned by frontness, with [ɔ] and [a] modifying similarly. This provides implications to studies on vowel intelligibility as a result of f_0 change, suggesting that vowel confusability at high pitches may be conditioned by the phonological properties of the vowels. Future research may pair the current set of data with a perception test to further explore the effects of vowel modification on vowel intelligibility.

Word Count: 4470

REFERENCES

Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer (Version 6.0.43). Retrieved from http://www.praat.org/

Bozeman, K. W. (2013). *Practical Vocal Acoustics: Pedagogic Applications for Teachers and Singers*. Hillsdale, NY: Pendragon Press.

Bunch, M., & Chapman, J. (2000). Taxonomy of singers used as subjects in scientific research. *Journal of Voice*, *14*(3), 363–369. https://doi.org/10.1016/S0892-1997(00)80081-8

Callaghan, J. (2000). *Singing and voice science*. San Diego, Calif: Singular Pub. Group.

Campbell, M. (2001). *Timbre (i)* (Vol. 1). https://doi.org/10.1093/gmo/9781561592630.article.27973

Chládková, K., Boersma, P., & Podlipskỳ, V. J. (2009). On-line formant shifting as a function of F0. *Tenth Annual Conference of the International Speech Communication Association*.

Doscher, B. M. (1994). *The functional unity of the singing voice* (2nd ed). Metuchen, N.J: Scarecrow Press.

Echternach, M., Birkholz, P., Traser, L., Flügge, T. V., Kamberger, R., Burk, F., … Richter, B. (2015). Articulation and vocal tract acoustics at soprano subject's high fundamental frequencies. *The Journal of the Acoustical Society of America*, *137*(5), 2586–2595. https://doi.org/10.1121/1.4919356

Garnier, M., Henrich, N., Smith, J., & Wolfe, J. (2010). Vocal tract adjustments in the high soprano range. *The Journal of the Acoustical Society of America*, *127*(6), 3771–3780. https://doi.org/10.1121/1.3419907

Hayward, K. (2016). *Experimental Phonetics: An Introduction*. Retrieved from http://www.vlebooks.com/vleweb/product/openreader?id=none&isbn=9781317887720

Henrich, N., Smith, J., & Wolfe, J. (2011). Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones. *The Journal of the Acoustical Society of America*, *129*(2), 1024–1035. https://doi.org/10.1121/1.3518766

Kris, T., & S. Greg, T. (1997). Spectral Properties of Fricatives in Amyotrophic Lateral Sclerosis. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1358–1372. https://doi.org/10.1044/jslhr.4006.1358

Ladefoged, P., & Johnson, K. (2015). *A Course in phonetics* (Seventh edition). Stamford, CT: Cengage Learning.

Lewcock, R., Pirn, R., Meyer, J., Hutchins, C. M., Woodhouse, J., Schelleng, J. C., … Sundberg, J. (2001). *Acoustics* (Vol. 1). https://doi.org/10.1093/gmo/9781561592630.article.00134

Lindblom, B. (1986). Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental phonology* (pp. 13–44). Orlando: Academic Press.

Miller, R. (1986). *The structure of singing: System and art in vocal technique*. Massachusetts: Schirmer.

Miller, R. (2000). *Training soprano voices*. New York: Oxford University Press.

Monsen, R. B., & Engebretson, A. M. (1983). The Accuracy of Formant Frequency Measurements: A Comparison of Spectrographic Analysis and Linear Prediction. *Journal of Speech, Language, and Hearing Research*, *26*(1), 89–97. https://doi.org/10.1044/jshr.2601.89

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R. S. (1989). The Emergence of Phonetic Segments. *Journal of Speech, Language, and Hearing Research*, *32*(1), 120–132. https://doi.org/10.1044/jshr.3201.120

Peirce, J., & MacAskill, M. (2020, May 23). Building Experiments in PsychoPy. Retrieved May 24, 2020, from SAGE Publications Ltd website: https://uk.sagepub.com/en-gb/eur/building-experiments-in-psychopy/book253480

Sibelius First (Version 2018.6). (2006). Retrieved from https://my.avid.com/get/sibelius-first

Sundberg, J. (1975). Formant Technique in a Professional Female Singer. *Acta Acustica United with Acustica*, *32*(2), 89–96.

Sundberg, Johan. (1999). 6—The Perception of Singing. In D. Deutsch (Ed.), *The Psychology of Music (Second Edition)* (pp. 171–214). https://doi.org/10.1016/B978-012213564-4/50007-X

Sundberg, Johan, Lã, F. M. B., & Gill, B. P. (2013). Formant Tuning Strategies in Professional Male Opera Singers. *Journal of Voice*, *27*(3), 278–288. https://doi.org/10.1016/j.jvoice.2012.12.002

Titze, I. R., Baken, R. J., Bozeman, K. W., Granqvist, S., Henrich, N., Herbst, C. T., … Wolfe, J. (2015). Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization. *The Journal of the Acoustical Society of America*, *137*(5), 3005–3007. https://doi.org/10.1121/1.4919349

Zsiga, E. C. (2013). *The sounds of language: An introduction to phonetics and phonology*. Hoboken: Wiley-Blackwell.