THE SOUND OF THE SNOW QUEEN:
AN ACOUSTIC ANALYSIS OF VOWEL CLARITY IN “LET IT GO”

A Thesis Presented to the Honors Tutorial College of Ohio University

In Partial Fulfillment of the Requirements for Graduation from the
Honors Tutorial College with the Degree of
Bachelor of Science in Communication Sciences and Disorders

Megan Smith
This thesis has been approved by
the Honors Tutorial College
and the Department of Communication Sciences and Disorders

Dr. Chao-Yang Lee
Associate Professor, Communication Sciences and Disorders
Thesis Advisor

Dr. Chao-Yang Lee
Associate Professor, Communication Sciences and Disorders
Director of Studies

Jeremy Webster
Dean, Honors Tutorial College
Abstract

Vowels can become difficult to understand when sung at exceptionally high frequencies. This is because the resonances that make up each vowel’s unique identity, called formants, are sometimes adjusted in order to maintain good voice quality and volume in such high fundamental frequencies. According to the current literature, professional singers are often trained to intentionally adjust their formants. Additionally, current research shows extensive differences between the male and female singing voice, which make formant adjustment likely in female singers. Using Disney’s Frozen, this thesis examines the relationships between formant adjustment and singer experience and singer sex. The hypotheses of this thesis include: 1) professional singers engage in formant tuning significantly more than amateur singers, and 2) female singers engage in formant adjustment significantly more than male singers. The procedure designed to test these hypotheses includes five tasks: 1) pronouncing eight vowels in the “hVd” context, 2) reading the lyrics to “Let it Go” without music, 3) reading the lyrics to “Let it Go” while listening to the instrumental music, 4) singing scales in two differing keys with two different vowels, for a total of four scales, and 5) singing “Let it Go” in two differing keys. Once each participant was recorded, the data was analyzed acoustically and statistically. The results indicate significantly more formant adjustment in professional and female singers across various tasks. These results confirm the hypotheses of this thesis. Future research should implement larger sample sizes and a more stringent definition of ‘professional’, and future areas of interest include the “singer’s formant”, the impact of different singing styles, and the intentional lack of formant adjustment as a professional stylistic choice.
Acknowledgements

I would like to thank the Honors Tutorial College (HTC) and the College of Health Sciences and Professions (CHSP) for providing me with the following grants which have made this research possible: the Provost’s Undergraduate Research Fund and the Dean’s Research and Travel Grant from HTC, and the Student Research Grant and the Student Research and Scholarly Activity Presentation Award from CHSP.

I would also like to thank my thesis advisor, Dr. Lee, for his motivation, encouragement, and calming reassurance in times of panic. This thesis is significantly improved through his involvement, and my sanity is significantly more secure through his expert guidance.

Finally, I would like to thank my fiancé, family, and friends for encouraging my Frozen obsession, and for patiently listening to more information about formants than they ever cared to know.
# Table of Contents

Abstract...........................................................................................................................................3

Acknowledgements.........................................................................................................................4

Introduction.......................................................................................................................................7

Method...............................................................................................................................................18

Participants......................................................................................................................................18

Materials and Procedure..................................................................................................................18

Data Analysis.....................................................................................................................................21

Results...............................................................................................................................................24

hVd Task..........................................................................................................................................24

Scales Task........................................................................................................................................27

Reading and Singing Tasks................................................................................................................35

Discussion.........................................................................................................................................43

Interpretations....................................................................................................................................43

Limitations.........................................................................................................................................46

Future Directions...............................................................................................................................47

Conclusion.........................................................................................................................................48
References...........................................................................................................................................50

Appendix A: Consent Form..................................................................................................................53

Appendix B: “Let it Go” Lyrics..........................................................................................................56
Introduction

In 2013, Disney released its top grossing princess movie, *Frozen*, in which main character Elsa sings “Let it Go” as an expression of freedom from the fear she’s experienced for most of her life. The movie’s popularity has been astronomical, and viewers’ love of Elsa has made “Let it Go” one of the most successful Disney songs in history. While the theatrical version of the song by Idina Menzel has been the most popular, a second version was also recorded and released by Demi Lovato. It is from this version that this thesis takes its inspiration. Throughout the song, Lovato’s voice changes drastically. At the beginning she sings softly, and by the end she is belting every note. What caught my attention was the difference this change made to the vowels in the song. The sounds “ee” (phonetically, /i/) and “aw” (phonetically /a/) are common vowels in the lyrics, and their sound differs greatly throughout the song. When Lovato tries to sing a very clear /i/ at a particularly high frequency, her voice quality suffers. When she modifies the vowel slightly, she is able to maintain strong voice quality. These differences demonstrate the basic principles of the source-filter theory of speech production.

The source-filter theory (Fant, 1960) describes speech as a combination of two main parts: the source and the filter. The source when discussing vowels is the vibration of the vocal folds, which determines the pitch that our ears hear, otherwise known as the fundamental frequency (F0; see Figure 1). All speech sounds are complex sounds, meaning they are made up of more than one frequency, but the pitch we perceive is primarily dependent on F0. The other frequencies that create the sound, harmonics (see Figure 2), are dependent on the F0, which is also sometimes referred to as the first harmonic. According to Fourier theorem (Johnson, 2012), the second harmonic is twice the frequency of the first, the third is three times the frequency of
the first, and this pattern continues. These harmonics contribute to the timbre, or quality, of the sound, but do not affect the pitch our ears perceive.

The filter, however, is determined by the shape of the vocal tract, including tongue and lip configuration and vocal tract length. The filter creates unique spaces in the oral cavity for the sound to resonate and amplifies specific ranges of harmonics. These amplified ranges are called formants (see Figure 3), and they distinguish which vowel our ears hear. The first two formants are those most highly correlated with vowel identity (Hollien, 2000). The first formant (F1) is inversely related to tongue height (Small, 2012). As the tongue height increases, or moves up toward the roof of the mouth, the frequency of F1 decreases. The second formant (F2) is directly related to how far forward the tongue is within the oral cavity, or tongue advancement (Small, 2012). As the tongue moves toward the front of the mouth, F2 also increases. Each vowel requires a unique vocal tract configuration, and thus consists of a specific set of formants. If a singer changes their articulatory posture for any given vowel, the formants will change and modify the vowel. This is what happens in some of the later vowels in Demi’s recording of “Let it Go”.

Figure 1. Vocal fold vibrations, or the “source”, create complex periodic waves that characterize sound. Reprinted from Sound Waves by the University of Manitoba, 2015, retrieved from http://home.cc.umanitoba.ca/~krussll/138/sec4/acoust1.htm
Figure 2. Discrete Fourier analysis (DFT) of /i/, with harmonics shown as tall, specific spikes, and the overarching formants as broad peaks.

Figure 3. Vocal tract, or “filter”, forms specific configurations which the vocal fold vibrations resonate from in different ways, resulting in discrete vowels. Reprinted from *Speech Waveforms*, by SW Phonetics, 2015, retrieved from [http://swphonetics.com/praat/tutorials/understanding-waveforms/speech-waveforms/#ex3](http://swphonetics.com/praat/tutorials/understanding-waveforms/speech-waveforms/#ex3)
A study conducted by Hollien, Mendes-Schwartz, and Nielson (2000) investigates how vowels are perceived at high frequencies, and concludes that accuracy is significantly correlated with fundamental frequency. In this study, professional singers perform various vowels at high and low frequencies in order to determine if vowels become more difficult to accurately perceive at high pitches. Four groups of listeners including voice teachers, phoneticians, speech-language pathology students, and a control group of undergraduates observe recordings of these vowels and try to determine which vowel they are hearing. When isolated, the vowels sung at a fundamental frequency higher than the first formant frequency were more often incorrectly identified than those sung at a lower pitch. If vowels sung at a frequency approaching or surpassing F1 are more difficult to understand, and formants are what make vowels distinguishable, it would make sense to suggest that the formants are changing at such high frequencies.

Since the source and the filter are relatively independent, the F0 and the formants can be changed independently. Because formants are significantly higher in frequency than a typical F0, keeping formants constant while changing F0 is easily accomplished. However, when the frequency of F0 exceeds that of F1 (as in the Hollien et al., 2000 study) neither F0 nor its accompanying harmonics can align with the formant to be amplified, resulting in reduced volume or vocal quality. People vary in their techniques concerning the production of these problematic high-pitched vowels, but many professional singers are encouraged to modify their vowels through formant adjustment in order to improve their resonance, as this results in a better overall sound (Callaghan, 2000). In the old bel canto tradition, professional singing instructions included “opening the throat”, “loosening the neck”, and “singing the tone forward, at the lips” (Callaghan, 2000). All of these techniques aim to reduce restrictions, which are a necessary
component of vowels. Vowels are unique from each other due to the location and degree of restrictions in the vocal tract, and these adjustments recommended by the masters of old illustrate a history of formant tuning among professionals. These teachings are still emphasized today. John Nix, a specialist in vocology with the National Center for Voice and Speech in Denver, goes so far as to say “Elite singers…become experts at making subtle changes in vowels as they sing, or they do not have consistent careers” (Nix, 2004). In his 2004 paper, Nix outlines several vowel modification concepts already widely accepted by the vocal pedagogy community. Vowel modification has been encouraged in the professional singing realm for centuries, and the tradition is still evident today.

Lovato, though a professional, does not adhere to these rules consistently. At the 1:15 point in her recording, she embraces the reduction in volume in order to maintain a clear /i/. Since /i/ has a high tongue body position and inversely low F1, F0 could easily exceed F1, creating the aforementioned problem of reduced volume. Later, at 3:02, Lovato changes tactics, and chooses to maintain her voice quality and volume, but sacrifice a certain level of vowel clarity. In both of these approaches, Lovato sacrifices something when singing at such a high pitch: her voice quality and/or volume or her vowels’ integrity. As an avid Frozen enthusiast, I listened to this song for countless hours, and after a while these small but significant differences piqued my curiosity.

I began to wonder about the reasons for the changes throughout “Let it Go”. If one professionally trained singer applies such drastically different approaches concerning these difficult high notes within the same song, is there any pattern amongst professionals, and if there is, how might it differ from that of amateur singers? A study by Joliveau, Smith, and Wolfe (2004) uses broad-band acoustic excitation to measure the formants of a variety of vowels sung
softly by trained soprano singers. This method allows for the formants to be measured in conjunction with and independent of the note’s harmonics. This can prevent the results from being influenced by potentially misleading factors, but requires that the vowels be sung softly. However, in a true singing scenario, professional singers are encouraged to increase their volume, so the requirement of singing softly could have a significant impact on the study’s capacity for generalization. Regardless, this research finds that as F0 surpasses F1, these singers adjust F1 to match F0. F2 also adjusts, but Joliveau et al. (2004) predict that this is as an involuntary response to F1 adjustment, not a separate intentional alteration. The authors conclude that it is this adjustment that results in a lack of vowel intelligibility. This confirms that professional soprano singers actively participate in formant tuning, and suggests that a significant pattern of behavior exists for this specific group of singers. However, it does not address any potential disparities between professionals and amateurs.

The comparison of professional and amateur singers is addressed in Murray’s 1990 study comparing pitch-matching accuracy between singers and non-singers. Murray concludes that professionals show greater control over fundamental frequency than non-singers. When given a variety of notes within their range, participants were asked to match the pitch as quickly as possible and sustain the note for two seconds. Professionals demonstrated significantly more accurate performance both in initial pitch and sustained pitch. These results depict professionals as having increased control over F0 in comparison to amateurs, and while it is not directly addressed, could imply a greater control of formant frequencies among the professionals as well. In another study which investigates a related area of research, Saitou and Goto (2009) attempt to control for within-group variance by studying the same three amateur male participants both before and after receiving nine hours of vocal training over the course of three days. The results
indicate that vocal training improves the presence of the “singer’s formant”, a resonance that is consistently present at 3kHz regardless of the vowel being sung (Krieman & Sidtis, 2013). While this thesis is not concerned with the singer’s formant specifically, any change in formants due to vocal training indicates that experience can lead to a more significant level of formant adjustment among professional singers. While previous studies suggest that professionals have greater general vocal control than amateurs, and indicate that they do adjust their formants to accommodate high notes, none specifically address whether they participate in formant tuning more than amateurs. Based on reliable research from a diverse group of professionals, I hypothesize that professionals are significantly more likely to adjust their formants than amateurs.

Another interesting area in the field of singing is the effect of singer sex. Many studies regarding vowel intelligibility include both males and females, but solely as a control, not to statistically compare the two groups, despite the widely acknowledged fact that men and women differ in many areas of voice. Source-filter theory provides explanations for differences in both the source and the filter. In general, men and women differ most prominently in their average F0. Female F0 is approximately double that of male, with men averaging 100-120 Hz and women averaging 200-220 Hz (Simpson, 2009). This can be attributed mainly to the source, specifically the lengthening and thickening of the vocal folds during puberty which affects mostly males (Kahane, 1978), and causes the vocal folds to vibrate at a significantly lower rate and produce a lower frequency. The filter also has a significant impact, specifically through vocal tract length. On average, men’s vocal tracts are 17.5 cm long, while the average female vocal tract is 14.5 cm (Simpson, 2009), and this difference in length produces lower formant frequencies in males (Fant, 1970).
These differences, among others, are also seen in a study by Hillenbrand, Getty, Clark, and Wheeler (1995). The authors replicated and extended the classic 1952 study by Peterson and Barney on vowel acoustics by studying multiple aspects of vowel acoustics in men, women, and children and conducting listening tests to determine the intelligibility of each group. In the perceptual component of this study, the authors found that vowel identification was not impaired by high F0. This is likely because the study utilized speech instead of singing, and even a high F0 does not approach or surpass the F1 in typical speech. Additionally, through LPC spectrum analysis, this study maps the contours of the first four formants of 12 vowels in the English language (/i,ɪ,æ,a,e,ɔ,o,ʊ,u,ʌ,ɝ/) when spoken in the hVd context. This study does not include singing, but does imply marked differences between male and female vocalization. Males consistently show lower frequencies, in fundamental and all measured formants.

Lower formants would seem to increase the likelihood of males engaging in formant adjustment when singing high notes. However, males also have a lower fundamental frequency than females, and this relationship between fundamental and formant frequencies is a fundamental distinction between the sexes. As previously discussed, each fundamental frequency has accompanying harmonics, which are multiples of the F0. Below is an example of a male pronouncing the vowel /i/ in typical speech (see Figure 4). Because F0 is low, there are more harmonics in a given space than if F0 were higher. An example of /i/ pronounced at a higher frequency by a female speaker is also shown below (see Figure 5). Because F0 is high, there are fewer harmonics compared to the male speaker. This illustrates why voice quality can suffer at high fundamental frequencies. Formants amplify distinct ranges of frequencies, and when F0 is particularly high, there are fewer harmonics to be amplified and voice quality suffers as a result. This is especially problematic when F0 becomes higher than F1 because then there are no
frequencies in the formant’s amplification range at all. This causes substantial detriment to voice quality and volume, and because females have a naturally higher F0, they are more likely to face this problem.

Figure 4. DFT analysis depicting a male participant producing the vowel /i/ in a normal speaking voice. F0 is low, around 150 Hz, which creates numerous harmonics in the areas of F1 and F2.

Figure 5. DFT analysis depicting a female participant producing the vowel /i/ in a normal speaking voice. F0 is high, around 250 Hz, which results in fewer harmonics in the areas of F1 and F2.
Put another way, the difference between fundamental frequency and the first formant in male speakers is significantly greater than that of females. Males have a generally lower F0 than females, but they also have lower formant frequencies. Because of this, it could be assumed that males and females would have the same problem with F0 approaching or surpassing F1. However, this is not the case because the decrease in frequency for male formants is not equal to the decrease in F0. Continuing with the vowel /i/, F1 in males is typically around 342 Hz, and in females it is typically around 437 Hz (Hillenbrand et al., 1995). Mathematically, this means the typical male F1 is only 22% lower than the typical female F1. In F0, however, the difference between males and females is much more prominent. Based on the average F0s previously provided by Simpson (2009), the average male F0 is 50% lower than that of the average female F0. This places the male F0 far lower than the male F1, and creates a larger gap between the two than is observed in females. This large gap makes it incredibly unlikely that the male F0 would ever come close enough to F1 to create any problems with voice quality, even while singing at exceptionally high frequencies. In females, however, their F0 is naturally higher, and thus closer to F1. When asked to sing at exceptionally high frequencies, F0 can surpass F1 and significantly diminish voice quality.

While it is widely accepted that the frequencies of males and females differ, the research on how these differences effect formants and vowel clarity in singing is less definitive. The available research in this area is provided primarily through voice teachers, vocologists, and singers themselves. In his previously mentioned paper, Nix (2004) describes various accepted truths about vowel modification in the singing world, one of which is that males and females inherently differ in their formant tuning. Males often tune F1 to one of the harmonics of F0, while females typically tune F1 to F0 itself in order to give the fundamental frequency more
strength, especially in particularly high notes (Nix, 2004). This difference is likely because when
the target is a particularly high frequency, females do not have a harmonic between F0 and F1.
Because males have a considerably lower F0, they can boost volume and vocal quality by
making a smaller adjustment and raising F1 to a harmonic of F0 instead of lowering it to the
fundamental frequency itself. Additionally, males sometimes focus less on F1 and more on F2,
opting to tune the second formant frequency to a higher harmonic by moving the tongue toward
the teeth, thus lowering the second formant.

A gap in knowledge is apparent when studying the current literature. There are studies
comparing male and female formant adjustment in addition to studies addressing innate
differences in fundamental and formant frequencies in males and females. However, there is
little research available that directly addresses the impact of singer experience and singer sex on
formant adjustment at high frequencies between amateurs and professionals, or between males
and females. Based on past and current literature, I hypothesize that female singers engage in
formant tuning significantly more than males.

Vowel clarity and voice quality are both top priorities in singing. Barrows and Pierce
(1932) describe the vowel as “the primary vocal element in speech and song” and assert that “the
voice depends largely upon it for beauty … power and expressiveness”. When vowels lose their
clarity, the voice also loses much of its power. However, the long history of the bel canto
tradition and contemporary voice pedagogy demonstrate that voice quality is often improved
through modification of the articulators, which results in vowel distortion. Due to this history of
articulation adjustment, professional singers are more likely to engage in formant tuning.
Furthermore, female professionals are the group most likely to modify vowels sung at high
frequencies due to the combination of formant tuning training and the physical limitations
imposed by shorter vocal tracts and thinner vocal folds in comparison to males. Because the vowel is an integral aspect of singing, it is important to explore formant tuning, a technique which purposefully modifies vowels in order to improve other aspects of the voice.

**Method**

*Participants*

20 adult singers from Athens, Ohio participated in this study, ten males and ten females. Of each of the ten, five were professional and five were amateur. Professional as defined in this study requires the singer to have a minimum of two years formal individual singing training. Amateur is defined as anything that does not meet the criteria for professional. Participants were recruited through flyers and music faculty contacts. All but one participant attend Ohio University; the exception attends Hocking College. In order to be included in the study, all participants were required to be healthy and be comfortable singing “Let it Go” given only the lyrics and instrumental music. After signing a consent form (see Appendix A) each participant was compensated $40 for taking part in the study. Prior to any participant involvement, this study was approved by the Ohio University Internal Review Board.

*Materials and Procedure*

Every participant performed the tasks in a sound treated room containing a television displaying the necessary visual aids, a two-way radio through which the researcher communicated with the participant, a Shure SM81-LC microphone and microphone stand with
pop filter, and Audio-Technica ATH M50 headphones through which the music and pitch notes were played for the participant. A Shure microphone and USBPre II external sound card were used to record each session, and the Brown Lab Interactive Speech System (BLISS) software utilized throughout the study recorded each session at a sampling rate of 44,100 Hz and 16-bit quantization, and later down-sampled to 11,200 Hz for acoustic analysis. The StatView statistical analysis program was used to analyze the acoustic data. An iPhone was used to play the pitch notes and instrumental music used throughout the experiment.

This experiment consists of five tasks: (1) reading a list of baseline words in the “hVd” style (“heed”, “hid”, “head”, “had”, “hod”, “hawed”, “hood”, and “who’d”) twice, (2) reading the lyrics to “Let it Go” (see Appendix B) without pitch or rhythm, (3) reading the lyrics to “Let it Go” while listening to the instrumental music but without changes in pitch, (4) singing major one-octave scales one semitone apart, given two different vowels for a total of four scales, and (5) singing “Let it Go” in two different keys.

Each task provides the researcher with information necessary to properly analyze each singer’s formants and any adjustments they might make. In order to determine if the formants change when sung at high frequencies, the formants’ natural position while speaking and while singing low notes must be determined. While each formant has a generally consistent placement, each individual is different, and the only way to accurately measure adjustment is through degree of change based on the individual’s formants when not effected by high frequencies. The first task does this through a set of baseline words in the hVd style, meaning the word begins with an ‘h’ sound, then a vowel, then ends with a ‘d’ sound. In the procedure, a list of these words (“heed”, “hid”, “head”, “had”, “hod”, “hawed”, “hood”, and “who’d”) appeared on the TV screen and the participant received instructions from the researcher through the two-way radio,
instructing them to read the list twice. Analyzing the vowels in each of these words provides a standard for each individual, against which the other tasks can be compared.

The lyric reading tasks provide a more precise comparison to the singing tasks by placing the vowels in the same context. The lyrics to “Let it Go” appeared on the TV screen for the participant and they were instructed by the researcher to read the lyrics. Once that was completed, they were instructed to put on the headphones in the booth and to read the lyrics again, in rhythm with the music, but without changing their pitch. In analysis, it is more accurate to compare the sung vowels with the vowels in the reading task instead of the hVd task because the consonants surrounding the vowel are the same.

The hVd task provides an accurate frame of reference for the vowels sung in the scale tasks. The scales demonstrate how formants change as the fundamental frequency increases. The vowels used are /i/ (“ee”) and /a/ (“aw”), which correlate to the vowels sung as high notes in “Let it Go”. The two scales begin on E-flat (622.25 Hz) and E (659.25 Hz), which was intended to correlate to the high notes the two original recordings by Demi Lovato and Idina Menzel respectively, since Lovato’s recording is one semitone lower than Menzel’s. However, due to misinterpretation, these notes are both one semitone lower than the frequencies used in the original recordings. Regardless, the scales are still useful in providing information about how singers would adjust vowel quality in two contrasting keys. The researcher instructed the participant to put on the headphones, then played a pitch note representing the first and last note in the scale to be sung. The researcher then proceeded to sing a scale to provide an example, and played the pitch note once more before asking the participant to sing the scale. This procedure was repeated for each of the four scales, /a/ and /i/ at E-flat, then /a/ and /i/ at E.
Finally, singing “Let it Go” provides a true singing scenario. The variety of settings in which each voice was recorded provides a robust and detailed understanding of how vowel identity shifts in relation to changes in fundamental frequency. Singing the song in two different keys was intended partially to relate back to the inspiration for this work, the two different recordings of “Let it Go”, but also to allow for contrast between the two high notes. If singers generally adjust their formants for one key more than the other, that would indicate a relationship between frequency and formant adjustment, and furthermore, could suggest a relationship between singer experience or singer sex and formant adjustment.

Data Analysis

Using BLISS, each instance of /i/ and /a/ was isolated from surrounding consonants and analyzed for F0, F1, and F2. F0 was measured by examining a visibly stable portion of the vowel on the waveform and applying a hamming window of 25.58 ms. Discrete Fourier Analysis (DFT) analysis was then applied to the window to generate a power spectrum for inspection of harmonics and formants. Linear predictive coding (LPC) was then used to estimate the formant frequencies. With a sampling rate of 11,200 Hz, one can expect to find four or five formants, depending on the sex of the singer. Males tend to have lower fundamental frequencies, thus having more harmonics and more formants in a given range while females tend to have fewer. In order to find five formants, the program analyzes the sample looking for 10 “poles”, or points of interest (Ladefoged, 2003). This allows for one buffer formant that may not appear in the given frequency range, but does not force the program to look for so many points that it assumes each harmonic is a formant. An example of this is shown below, using /i/ as in “heed”. In Figure 6 the program is set to 10 poles. It finds four formants with one presumably appearing outside of the displayed frequencies. In Figure 7 the program is set to 16 poles and it finds seven formants.
Figure 6. Linear predictive coding (LPC) analyzes DFT analysis of harmonics to determine formants; specifically, this image shows LPC analysis of /i/ set at 10 poles.

Figure 7. LPC analysis of /i/ set at 16 poles

Figure 7 shows an overabundance of “formants” in places where they logically cannot exist in order for a vowel to sound like /i/. This is because the program believes it needs to find approximately eight formants, so it gives credence to sections of harmonics that may form a slight climax, but not one truly indicative of a formant. Even when the program is set to an appropriate number of poles, computer programs do not have the same intelligence as humans.
As such, the formant estimates were verified individually by the experimenter based on expected values from the literature (Hillenbrand et al., 1995). This method was chosen over others because it has proven to have similar levels of accuracy and higher levels of consistency (Ladefoged, 2003).

Once each recording was acoustically analyzed and reviewed, the measured F0, F1, and F2 values were organized by participant. Descriptive and statistical analyses were then conducted to test the main hypotheses regarding the effects of singing experience and singer sex on vowel production. The five tasks were studied first descriptively, then statistically as needed. In the hVd task, it is expected that professionals and amateurs will present similar vowel spaces within their sex, and that males and females will differ in overall frequencies of both formants, the males being generally lower. In comparing the two keys used in the scales task, expected results would indicate formant tuning at an earlier onset for the higher key than for the lower one, which would confirm that vowel distortion is correlated to fundamental frequency. The reading tasks were fundamentally designed to be a contextually consistent comparison point for the singing tasks, and this study anticipates a significant difference between reading and singing for the female and professional groups, as they are expected to demonstrate formant tuning. Between the two keys used within the singing task, a similar result to the scales task is expected. The key that is one semitone higher is anticipated to show greater formant adjustment, which would then confirm a positive correlation between changes in F0 and changes in formants.
Results

*hVd Task*

Data from the hVd reading task provides information on the vowel space for each group represented in this study. Each vowel can be plotted on a graph with tongue height, represented by the frequency of F1, on the abscissa (horizontal axis) and tongue advancement, represented by the frequency of F2, on the ordinate (vertical axis; Small, 2012). When each vowel from the task is plotted, it creates a defined acoustic space which reflects the articulatory vowel space; for example, when a vowel is plotted at a low F1 and high F2, this demonstrates that the vowel is produced when the tongue is close to the roof of the mouth and advanced forward, close to the teeth. This vowel could easily be /i/. The acoustic vowel space is a direct product of the shape of the articulators, which reiterates the concept of source-filter theory, on which this thesis is built. Acoustics is not solely the product of vocal fold vibrations, but is also influenced by articulatory factors which have a powerful impact, especially in determining the vowels our ears hear.

By comparing the vowel spaces of each group, it is possible to understand each group’s articulatory maneuvers and evaluate whether there are any significant differences between groups. For example, as Figures 8-11 demonstrate, the male vowel space is characterized by a smaller range in F1 frequencies than the female vowel space, which supports current literature indicating that females are inclined to a greater difference in formant frequencies between vowels than are males (Diehl, Lindblom, Hoemeke, & Fahey, 1996; Simpson, 2002). This is a fascinating area of research which suggests that the large difference between vowels in female speakers is due to an innate unintelligibility caused by a high fundamental frequency. The graphs
shown below coincide with this hypothesis, as the male vowel space is more central while the female vowel space is more dispersed.

Figure 8. Values of formant frequencies spoken in the hVd task and averaged among the amateur male group plotted with F1 on the x-axis and F2 on the y-axis, to create a visual vowel space.

Figure 9. Values of formant frequencies spoken in the hVd task and averaged among the professional male group plotted with F1 on the x-axis and F2 on the y-axis, to create a cleaner, more precise vowel space than in the amateur singers.
Figure 10. Values of formant frequencies spoken in the hVd task and averaged among the amateur female group plotted with F1 on the x-axis and F2 on the y-axis, to create a visual vowel space much more diffuse than in the male singers.

Figure 11. Values of formant frequencies spoken in the hVd task and averaged among the professional female group plotted with F1 on the x-axis and F2 on the y-axis, to create a visual vowel space.
**Scales Task**

Descriptive data also provides interesting insights to the scales task. A total of 16 graphs (Figures 12-15) were created based on the four scales in the task and the four groups of participants. The F0, F1, and F2 values obtained from singing these scales are displayed and discussed below.

![Amateur Female Graphs](image)

Figure 12. Averaged values among the amateur female group show a pattern of adjustment in both formants and both scales, though to varying degrees.
Figure 13. Averaged values among the professional female group in both scales show consistent and predictable adjustment in F1, and visible but less predictable adjustment in F2.
Figure 14. Averaged values among the amateur male group show minor fluctuations in both formants in the E-flat Major scale, but show more pointed adjustment of F2 in the E scale.
It is clear that formant tuning in /i/ scales and /a/ scales differs due to the nature of their different formants. In particular, /i/ is characterized by a low F1 frequency (indicating relatively high tongue body position) and high F2 frequency (indicating relatively front tongue body position). The vowel /a/, on the other hand, is characterized by a high F1 frequency (indicating relatively low tongue body position) and low F2 frequency (indicating relatively back tongue body position). In addition to these intrinsic properties, there are additional differences based on experience, sex, and key. The strongest apparent differences are between males and females.
According to Simpson (2002), this is expected due to inherent differences in vocal tract length, oral-to-pharyngeal cavity length ratio, and even behavioral factors. Despite the assumptions and hypothesis of this study, the changes based on experience and on key are more subtle and less likely to be significant.

In each of these graphs, it is expected that F0 will steadily increase because participants were asked to sing an ascending scale, in which F0 would increase by a set amount for each note. F1 and F2, however, are not related to the notes in the scale, but to the vowel being sung, and therefore should not change as frequency increases. The hypothesis of this study posits that females and professionals will not fit this mold, and the descriptive graphs lend mixed support to this idea.

For the /a/ scales, the male graphs show a relatively steady F1 and F2 with minimal changes as F0 increases with each note, meaning the vowel is sustained throughout each scale with little to no adjustment. However, the female graphs, both amateur and professional, show moderate variability in F1 and F2. Across experience levels in females, F1 consistently decreases in the last two notes of each /a/ scale. In doing this, the singers amplify F0 by taking advantage of the natural resonances of the vocal tract. The precise frequency that corresponds with /a/’s first formant is so close to the fundamental frequencies at the end of each scale that the female singers in this study raised their tongue slightly in order for F1 to match the fundamental frequency. A similar pattern is observed for F2, specifically in amateur females. In an attempt to increase resonance in the fundamental frequency, the second formant is lowered by pulling the tongue toward the back of the mouth. Professional females make very slight adjustments to F2, if they make any at all. Finally, the two keys, one semitone apart, demonstrated little variance between them for all participants.
Adjustments in the /i/ scales are more universal; males, females, amateurs, and professionals all show some level of adjustment. Professional males show the least, with F1 increasing only very slightly, and F2 decreasing infinitesimally. Amateur males show a more prominent decrease in F2 and increased overlap between F0 and F1. Except for more pronounced overlap, amateur females produce similar results. F2 shows the greatest decrease in professional females, who also demonstrate almost total overlap of F1 and F0. Each of these groups drops their jaws to various degrees as F0 increases, which effectively lowers tongue height (increases F1), decreases tongue advancement (decreases F2), and potentially distorts the intended vowel, /i/.

To statistically verify these observations, analyses of variance (ANOVA) were conducted separately on F1 and F2, with note (1 through 8) and vowel-scale (/i/-E♭, /i/-E, /a/-E, /a/- E♭) as within-subject variables, singing experience (amateur/professional) and singer sex (female/male) as between-subject variables, and participants as a random factor. Table 1 shows the results of the ANOVA for F1, and of these results, the significant main effects and interactions are highlighted in blue.
### ANOVA Table for F1

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F-Value</th>
<th>P-Value</th>
<th>Lambda</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>1</td>
<td>8094.026</td>
<td>8094.026</td>
<td>1.17</td>
<td>.248</td>
<td>0.13</td>
<td>0.095</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>986664.026</td>
<td>986664.026</td>
<td>22.775</td>
<td>.002</td>
<td>0.997</td>
<td>0.095</td>
</tr>
<tr>
<td>Experience * Sex</td>
<td>1</td>
<td>51015.306</td>
<td>51015.306</td>
<td>1.179</td>
<td>.2937</td>
<td>1.179</td>
<td>0.167</td>
</tr>
<tr>
<td>Subject (Group)</td>
<td>16</td>
<td>692454.275</td>
<td>43278.392</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vowel</td>
<td>3</td>
<td>10657915.019</td>
<td>3519305.006</td>
<td>244.433</td>
<td>&lt;.0001</td>
<td>733.300</td>
<td>1.000</td>
</tr>
<tr>
<td>Vowel * Experience</td>
<td>3</td>
<td>1445.625</td>
<td>481.875</td>
<td>0.333</td>
<td>.9917</td>
<td>0.100</td>
<td>0.055</td>
</tr>
<tr>
<td>Vowel * Sex</td>
<td>3</td>
<td>37156.625</td>
<td>123858.756</td>
<td>8.60</td>
<td>0.4682</td>
<td>2.581</td>
<td>0.218</td>
</tr>
<tr>
<td>Vowel * Subject (Group)</td>
<td>48</td>
<td>691094.475</td>
<td>14397.802</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Note</td>
<td>7</td>
<td>1054149.969</td>
<td>150202.138</td>
<td>23.358</td>
<td>&lt;.0001</td>
<td>163.505</td>
<td>1.000</td>
</tr>
<tr>
<td>Note * Experience</td>
<td>7</td>
<td>4574.350</td>
<td>65.3479</td>
<td>0.102</td>
<td>.9981</td>
<td>0.711</td>
<td>0.076</td>
</tr>
<tr>
<td>Note * Sex</td>
<td>7</td>
<td>36933.150</td>
<td>5133.307</td>
<td>0.798</td>
<td>0.5903</td>
<td>5.588</td>
<td>0.326</td>
</tr>
<tr>
<td>Note * Subject (Group)</td>
<td>112</td>
<td>720212.525</td>
<td>6430.469</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vowel * Note</td>
<td>21</td>
<td>1087309.906</td>
<td>51760.472</td>
<td>18.045</td>
<td>&lt;.0001</td>
<td>378.937</td>
<td>1.000</td>
</tr>
<tr>
<td>Vowel * Note * Experience</td>
<td>21</td>
<td>41000.7200</td>
<td>1952.724</td>
<td>0.660</td>
<td>.8524</td>
<td>14.290</td>
<td>0.543</td>
</tr>
<tr>
<td>Vowel * Note * Sex</td>
<td>21</td>
<td>758965.800</td>
<td>37569.684</td>
<td>13.092</td>
<td>&lt;.0001</td>
<td>274.942</td>
<td>1.000</td>
</tr>
<tr>
<td>Vowel * Note * Subject (Group)</td>
<td>336</td>
<td>964178.325</td>
<td>2989.578</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Results of ANOVA for F1 in the scales task, with significant results highlighted in blue.

A significant difference between the sexes is expected \([F(1, 16) = 22.775, p = .002]\), due to the inherent difference between male and female vocal tracts as described in the introduction of this paper. Significance depending on the vowel being sung is also expected \([F(3, 48) = 244.433, p < .0001]\); vowels are defined by formants, so their formants must differ significantly in order to be distinct from one another. A significant difference between notes \([F(7, 112) = 23.358, p < .0001]\) is interesting because it confirms the existence of some level of formant adjustment. Its interactions with singer sex and singer experience, however, are not significant, which suggests that while F1 adjustment does appear to occur in high-pitched scales, it does not depend on the factors implicated in this study’s hypotheses. The significance of the interaction between vowel and note \([F(21, 336) = 18.045, p < .0001]\) serves to better describe the relationship between F1
adjustment and associated factors, specifically demonstrating that F1 adjusts for different vowels in different ways. Finally, the significant relationship between vowel, note, and singer sex \[ F(21, 336) = 13.092, p < .0001 \] indicates that the nuanced differences in F1 based on the relationship between vowel and note are further complicated by a significant difference in those patterns of adjustment based on the sex of the singer.

Table 2, shown below, displays the significant test results for F2 adjustment.

<table>
<thead>
<tr>
<th>ANOVA Table for F2</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F-Value</th>
<th>P-Value</th>
<th>Lambda</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>1</td>
<td>570971.025</td>
<td>570971.025</td>
<td>1.688</td>
<td>.2123</td>
<td>1.688</td>
<td>.220</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>3122295.006</td>
<td>3122295.006</td>
<td>9.230</td>
<td>.0001</td>
<td>9.230</td>
<td>.826</td>
</tr>
<tr>
<td>Experience * Sex</td>
<td>1</td>
<td>511890.625</td>
<td>511890.625</td>
<td>1.513</td>
<td>.2644</td>
<td>1.513</td>
<td>.202</td>
</tr>
<tr>
<td>Subject(Group)</td>
<td>16</td>
<td>5412702.275</td>
<td>338253.952</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vowel</td>
<td>3</td>
<td>151561478.531</td>
<td>50520492.844</td>
<td>381.519</td>
<td>&lt;.0001</td>
<td>1144.657</td>
<td>1.000</td>
</tr>
<tr>
<td>Vowel * Experience</td>
<td>3</td>
<td>210662.587</td>
<td>70220.862</td>
<td>.530</td>
<td>.6637</td>
<td>1.591</td>
<td>.148</td>
</tr>
<tr>
<td>Vowel * Sex</td>
<td>3</td>
<td>331309.631</td>
<td>110436.510</td>
<td>.834</td>
<td>.4818</td>
<td>2.502</td>
<td>.212</td>
</tr>
<tr>
<td>Vowel * Experience * Sex</td>
<td>3</td>
<td>48604.237</td>
<td>15601.412</td>
<td>.118</td>
<td>.9402</td>
<td>.353</td>
<td>.070</td>
</tr>
<tr>
<td>Vowel * Subject(Group)</td>
<td>48</td>
<td>9356128.925</td>
<td>132419.353</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Note</td>
<td>7</td>
<td>2941869.119</td>
<td>420240.731</td>
<td>10.819</td>
<td>&lt;.0001</td>
<td>75.731</td>
<td>1.000</td>
</tr>
<tr>
<td>Note * Experience</td>
<td>7</td>
<td>86430.000</td>
<td>12347.143</td>
<td>.316</td>
<td>.9446</td>
<td>2.225</td>
<td>.142</td>
</tr>
<tr>
<td>Note * Sex</td>
<td>7</td>
<td>1918169.459</td>
<td>274024.210</td>
<td>7.055</td>
<td>&lt;.0001</td>
<td>49.382</td>
<td>1.000</td>
</tr>
<tr>
<td>Note * Experience * Sex</td>
<td>7</td>
<td>155918.650</td>
<td>22274.093</td>
<td>.573</td>
<td>.7762</td>
<td>4.014</td>
<td>.236</td>
</tr>
<tr>
<td>Note * Subject(Group)</td>
<td>112</td>
<td>4350613.325</td>
<td>38843.869</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vowel * Note</td>
<td>21</td>
<td>3996885.544</td>
<td>190327.883</td>
<td>5.459</td>
<td>&lt;.0001</td>
<td>114.645</td>
<td>1.000</td>
</tr>
<tr>
<td>Vowel * Note * Experience</td>
<td>21</td>
<td>637288.687</td>
<td>30347.080</td>
<td>.870</td>
<td>.6301</td>
<td>10.280</td>
<td>.615</td>
</tr>
<tr>
<td>Vowel * Note * Sex</td>
<td>21</td>
<td>1193044.094</td>
<td>56811.624</td>
<td>1.630</td>
<td>.0410</td>
<td>34.221</td>
<td>.963</td>
</tr>
<tr>
<td>Vowel * Note * Experience * Sex</td>
<td>21</td>
<td>947246.987</td>
<td>45106.999</td>
<td>1.294</td>
<td>.1756</td>
<td>27.170</td>
<td>.092</td>
</tr>
</tbody>
</table>

Table 2. Results of ANOVA for F2 in the scales task, with significant values highlighted in blue.

Many of the significant effects are similar to the results from the F1 analysis, such as sex \[ F(1, 16) = 9.230, p = .0078 \], vowel \[ F(3, 48) = 381.519, p < .0001 \], note \[ F(7, 112) = 10.819, p < .0001 \], the interaction between vowel and note \[ F(21, 336) = 5.459, p < .0001 \], and the interaction between vowel, note, and sex \[ F(21, 336) = 1.630, p = .0410 \]. However, there is one
additional interaction in the case of F2, which is the interaction between note and sex \[F(7, 112) = 7.055, p < .0001\]. This relationship relates directly to the hypothesis that claims male singers demonstrate less formant adjustment than females. F2 changes as the notes increase, and the degree of this change is dependent upon the sex of the singer. Both descriptive and statistical data shows that females adjust F2 more than male singers do in the scales task. However, this study measured formant adjustment in the setting of scales and “Let it Go”. In order to accurately reflect on the hypotheses of this study, descriptive and statistical data from the reading and singing tasks must be evaluated.

Reading and Singing Tasks

Descriptive data provides insight to formant tuning in each of the four reading and singing tasks. Participants were asked to complete four tasks: read the lyrics to “Let it Go” (task 1), read the lyrics in rhythm with the instrumental music but without changing pitch (task 2), sing the lyrics in the F major key (task 3), and sing the lyrics in the A-flat major key (task 4). F0, F1, and F2 values were obtained from these tasks for the three high-frequency instances of the vowel /i/ and the one high-frequency instance of the vowel /a/ throughout “Let it Go”. The three instances of /i/ were averaged for each participant, then for each category of singer represented by this study. These values and those of the vowel /a/, also averaged for each group of singers, are shown in the following eight graphs (Figures 16-23).
Figure 16. Average male amateur values for F0, F1, and F2 for the vowel /i/ in the reading and singing tasks. F0 is expected to be higher for tasks 3 and 4 (the singing tasks), but the slight changes in F2 are an indication of formant tuning.

Figure 17. Average amateur female values for F0, F1, and F2 for the vowel /i/ in the reading and singing tasks. F0 is expected to be higher for tasks 3 and 4 (the singing tasks) but the slight increase in F1 and significant decrease in F2 are indications of formant tuning.
Figure 18. Average professional male values for F0, F1, and F2 for the vowel /i/ in the reading and singing tasks. F0 is expected to be higher for tasks 3 and 4 (the singing tasks) but the slight increase and decrease in F1 and F2, respectively, indicate formant tuning.

Figure 19. Average professional female values for F0, F1, and F2 for the vowel /i/ in the reading and singing tasks. F0 is expected to be higher for tasks 3 and 4 (the singing tasks) but the slight increase in F1 and the significant decrease in F2 indicate formant tuning.
Figure 20. Average amateur male values for F0, F1, and F2 for the vowel /a/ in the reading and singing tasks. F0 is expected to be higher for tasks 3 and 4 (the singing tasks). F1 and F2 remain stable across tasks, which suggests a lack of formant tuning.

Figure 21. Average amateur female values for F0, F1, and F2 for the vowel /a/ in the reading and singing tasks. F0 is expected to be higher for tasks 3 and 4 (the singing tasks). F1 and F2 both decrease in the singing task with the higher key, indicating formant adjustment dependent upon increased F0.
Figure 22. Average professional male values for F0, F1, and F2 for the vowel /a/ in the reading and singing tasks. F0 is expected to be higher for tasks 3 and 4 (the singing tasks). F1 and F2 increase slightly in the singing tasks, but minimal formant tuning is apparent.

Figure 23. Average professional female values for F0, F1, and F2 for the vowel /a/ in the reading and singing tasks. F0 is expected to increase for tasks 3 and 4 (the singing tasks). F1 and F2 both decrease in the singing task with the higher key, indicating formant adjustment dependent upon increased F0.
These graphs illustrate the average F0, F1, and F2 values of the vowels /i/ and /a/ in “Let it Go” across the four reading and singing tasks. Because the vowel is consistent across the tasks, the formants should not vary depending on the task being performed. Changes in formant values indicate formant tuning. By examining the descriptive data, the reading and singing tasks appear to indicate greater formant adjustment in females and professionals.

For /i/, both amateur and professional males demonstrate a slight increase in F1 and a slight decrease in F2 in the singing tasks (tasks 3 and 4) when compared with the reading tasks (tasks 1 and 2). However, these differences are more pronounced in female singers. Females significantly increase F1 and significantly decrease F2 for the singing tasks. In general, professionals show more adjustment to F1 than their amateur counterparts, and females show more adjustment to F2 than their male counterparts.

For /a/, male singers, both amateur and professional, demonstrate very little difference in formant frequencies from one task to the next. Despite the distinct jump in F0 from speaking to singing, and the small jump in F0 from the first singing task to the second, male formants remain relatively constant. Contrastingly, female formants vary greatly depending upon the task being performed. Amateur and professional females both increase F1 and F2 in the first singing task, though professionals demonstrate a more prominent increase. Then all female singers decrease F1 and F2 slightly from the first singing task to the second. While the formant frequencies decrease, they remain higher than their typical values.
As with the scales task, significant differences based on sex \(F(1, 16) = 27.388, p < .0001\) and vowel \(F(3, 48) = 145.247, p < .0001\) are expected. A significant difference in the reading/singing factor \(F(3, 48) = 81.623, p < .0001\) suggests an inherent difference in F1 based on the task being performed (reading without music, reading with music, singing in the first key, singing in the second key). Because formants are the defining characteristic of vowels, they should remain stable across all contexts; thus, the significant differences imply that some level of tuning is occurring for F1. This significance does not implicate any one group of singers as engaging in formant tuning, but the significant interaction between reading/singing and experience does \(F(3, 48) = 4.438, p = .0078\). The significance of this interaction means that F1 differs between the reading/singing tasks based on experience, and upon deeper examination of the statistics, professional singers show more F1 tuning than amateurs, which supports one of the main hypotheses of this paper.
In the interaction between vowel and the reading or singing task being performed \([F(9, 144) = 10.720, p < .0001]\), F1 changes depending on which vowel is being pronounced, and that difference between vowels is dependent upon which task is being performed. There is a greater difference between /i/ and /a/ in the frequency of F1 for the reading tasks than there is in the singing tasks, likely because there is little need to adjust the first formant of /a/. F1 of /i/ is naturally lower than that of /a/, which is apparent in the reading tasks when there is presumably no formant adjustment occurring. However, when singing /i/ at such high frequencies, F1 is raised to match the pitch, which places the first formants of /a/ and /i/ closer together. The significance of the relationship between vowel, task, and sex of the singer \([F(9, 144) = 3.302, p = .0011]\) is similar, but shows that females show an even more significant change, which lends more support toward the hypothesis that female singers adjust formant frequencies more than male singers.

The significant factors involved in the adjustment of F2 are exhibited in Table 4, shown below.

### ANOVA Table for F2

<table>
<thead>
<tr>
<th>Experience</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F-Value</th>
<th>P-Value</th>
<th>Lambda</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>1</td>
<td>14960.450</td>
<td>14960.450</td>
<td>.050</td>
<td>&lt;.0005</td>
<td>.005</td>
<td>.055</td>
</tr>
<tr>
<td>Experience * Sex</td>
<td>9</td>
<td>4485939.200</td>
<td>4485939.200</td>
<td>14.873</td>
<td>.0014</td>
<td>14.873</td>
<td>.964</td>
</tr>
<tr>
<td>Subject (Group)</td>
<td>16</td>
<td>248645.000</td>
<td>248645.000</td>
<td>824</td>
<td>.3774</td>
<td>.824</td>
<td>.132</td>
</tr>
<tr>
<td>Vowel</td>
<td>3</td>
<td>43936096.675</td>
<td>14645365.558</td>
<td>232.905</td>
<td>&lt;.0001</td>
<td>697.486</td>
<td>1.000</td>
</tr>
<tr>
<td>Vowel * Experience</td>
<td>3</td>
<td>183376.425</td>
<td>61125.475</td>
<td>97.0</td>
<td>.4145</td>
<td>2.911</td>
<td>.242</td>
</tr>
<tr>
<td>Vowel * Sex</td>
<td>3</td>
<td>6849.025</td>
<td>2283.008</td>
<td>0.36</td>
<td>.9996</td>
<td>.109</td>
<td>.056</td>
</tr>
<tr>
<td>Vowel * Experience * Sex</td>
<td>3</td>
<td>236929.175</td>
<td>78976.392</td>
<td>1.254</td>
<td>.3038</td>
<td>3.761</td>
<td>.306</td>
</tr>
<tr>
<td>Vowel * Subject (Group)</td>
<td>48</td>
<td>3023621.075</td>
<td>62992.106</td>
<td>&lt;.0001</td>
<td>39.299</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Reading/Singing</td>
<td>3</td>
<td>275796.825</td>
<td>919328.276</td>
<td>13.100</td>
<td>&lt;.0001</td>
<td>39.299</td>
<td>1.000</td>
</tr>
<tr>
<td>Reading/Singing * Experience</td>
<td>3</td>
<td>48115.325</td>
<td>16038.442</td>
<td>2.29</td>
<td>.0861</td>
<td>.686</td>
<td>.089</td>
</tr>
<tr>
<td>Reading/Singing * Sex</td>
<td>3</td>
<td>1678519.425</td>
<td>629066.475</td>
<td>8.920</td>
<td>&lt;.0001</td>
<td>26.750</td>
<td>.996</td>
</tr>
<tr>
<td>Reading/Singing * Experience</td>
<td>3</td>
<td>133229.925</td>
<td>44405.975</td>
<td>.633</td>
<td>.5975</td>
<td>1.898</td>
<td>.169</td>
</tr>
<tr>
<td>Reading/Singing * Subject (Group)</td>
<td>48</td>
<td>3366168.375</td>
<td>70176.549</td>
<td>&lt;.0001</td>
<td>77.519</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Vowel * Reading/Singing</td>
<td>9</td>
<td>4250889.650</td>
<td>473221.072</td>
<td>8.613</td>
<td>&lt;.0001</td>
<td>77.519</td>
<td>1.000</td>
</tr>
<tr>
<td>Vowel * Reading/Singing * Experience</td>
<td>9</td>
<td>393682.800</td>
<td>43740.311</td>
<td>7.98</td>
<td>.190</td>
<td>7.179</td>
<td>.380</td>
</tr>
<tr>
<td>Vowel * Reading/Singing * Sex</td>
<td>9</td>
<td>1340574.150</td>
<td>148952.683</td>
<td>2.716</td>
<td>.0059</td>
<td>24.447</td>
<td>.950</td>
</tr>
<tr>
<td>Vowel * Reading/Singing * Experience</td>
<td>9</td>
<td>378969.900</td>
<td>42107.767</td>
<td>.768</td>
<td>.6462</td>
<td>6.911</td>
<td>.365</td>
</tr>
</tbody>
</table>

Table 4. Results of ANOVA for F2 in reading and singing, with significant values highlighted in blue.
The significance of sex \([F(1, 16) = 14.873, p = .0014]\), vowel \([F(3, 48) = 232.495, p < .0001]\), the reading or singing task being performed \([F(3, 48) = 13.100, p < .0001]\), the interaction between sex the task being performed \([F(3, 48) = 8.920, p < .0001]\), vowel interaction with the reading or singing task being performed \([F(9, 144) = 8.613, p < .0001]\), and the interaction between vowel, task, and sex \([F(9, 144) = 2.716, p = .0059]\) on F2 adjustment is similar to the results describing the tuning of F1 established above. One major difference is that F2 is not affected by singer experience level in relation to which task is being performed, but is instead affected by the sex of the singer. While the statistical results of analyzing F1 support this study’s hypothesis that professional singers make more formant adjustments than amateurs, the significance of F2 supports this study’s other hypothesis that females tune their formants more than males.

**Discussion**

**Interpretations**

The descriptive and statistical data provided above indicate strong associations between formant adjustment and singer sex. In both the reading and singing tasks and the scales task, singer sex has a statistically significant impact on the value of F2. The consistency of increased F2 adjustment in female singers despite a less consistent increase in F1 adjustment is an unusual finding. Joliveau et al. (2004), mentioned in the introduction of this thesis, suggests that adjustments made to F2 are a byproduct of F1 adjustment. However, in this study, female singers adjusted F2 at a significant level more than they adjusted F1, which suggests that F2 adjustment
is independent of F1 adjustment. Female singers demonstrated decreased F2 as F0 approached F1, which would likely tune F2 to a harmonic of F0 in order to increase volume and voice quality. This is consistent with one hypothesis of this paper – female singers engage in formant adjustment significantly more than male singers.

The effect of singer sex on the adjustment of F1 does not reach statistical significance (p < .05) for either set of tasks, but in the reading and singing tasks, it approaches significance (p = .0649). Future studies could use a larger sample size to test whether this trend persists or eventually reach statistical significance. Singer experience, on the other hand, was found to be a significant factor in F1 adjustment for the reading and singing tasks. Upon examining the data more closely, it is evident that professional singers engage in formant adjustment significantly more than amateurs. Specifically, professionals increase F1 as F0 approaches or surpasses the typical frequency of F1. This supports the hypothesis that professional singers engage in formant adjustment significantly more than amateur singers.

Both F1 and F2 are integral factors in determining vowel identity, but it appears that the experience level of the singer has a more prominent effect on F1. This is a reasonable result when one considers current and longstanding teachings in voice. Singing training throughout the ages has promoted a more open vocal tract in order to improve resonance. This often includes lowering the jaw and other techniques designed to reduce restrictions in the vocal tract. Because this is a major aspect of many professional singing traditions, this study’s result indicating increased F1 in high notes is expected.

Although typical vocal pedagogy includes lowering the jaw, which raises F1 to create better voice quality, professionals are not the only singers applying these techniques. While experience level is a significant factor in F1 adjustment for the reading and singing tasks, it is not
significant for the scales task. When examining the descriptive data in the scales task, it is clear that singers tune F1 as F0 increases. However, there does not seem to be a significant difference between professionals and amateurs in the extent of the tuning. Both amateurs and professionals, especially females, demonstrate formant adjustment in the scales task. This may be due in part to the nature of singing a scale. Specifically, in scales, F0 increases incrementally, and it is natural for all singers to try to maintain or even increase volume throughout the scale. In order to accomplish this, small changes must be made to the formants as F0 increases. As F0 surpasses F1, F1 must be increased in order to amplify F0 or its harmonics. To do this, singers lower the jaw and tongue, which raises F1. This is a simple maneuver, which amateurs likely perform subconsciously, to maintain voice quality and volume. While this change may be more intentional in professionals, the resulting formant adjustment is the same. However, when F0 increases more suddenly, such as in the singing tasks, this intentionality may be required in order to successfully maintain voice quality.

Singer experience level is a significant factor in the singing tasks – professionals engage in formant tuning to a greater extent than amateurs. In the scales task it is natural to maintain volume as F0 increases, even if vowel clarity is sacrificed in the process. In the singing tasks, however, the high notes appear suddenly, which does not allow for slow and steady adjustment. Sudden formant adjustment is more likely to occur in a singer who has been trained in formant tuning techniques. This is reflected in the results of this study, as professionals show significant F1 tuning in the singing tasks specifically.

The results of this thesis indicate that female and professional singers prefer to enhance voice quality rather than vowel clarity when the two conflict. While formant tuning distorts the vowel and renders it incomprehensible in isolation, when sung in context the human brain is
capable of accurately identifying most vowels (Hollien, 2000). Thus, singers can reasonably adjust their formants and maintain a strong voice, knowing that their intended meaning will still be conveyed.

Limitations

All research has limitations, and it is important to acknowledge the factors that could have an impact on a study’s results in order to accurately interpret potential correlations. The biggest limitation of this thesis is the sample size. With 20 participants, only five could be included in each category of amateur male, amateur female, professional male, and professional female. It is likely that one participant could skew the data dramatically due to the limited sample size. An additional factor which could potentially limit the study is the definition of ‘professional’ used in this thesis. Previous to any recordings, the definition was set as ‘a singer who has had a minimum of two years of formal individual singing training’. ‘Amateur’ was defined as any singer who did not meet the definition of ‘professional’. However, once recruitment and recording began, that definition was revealed to be insufficient. Some participants had been involved in professional choirs for ten years, but were categorized as amateurs because they’d never had individual training. On the other hand, some participants had received two years of formal individual training five years prior and had done little performing since. It is possible that with a more stringent and specific definition of ‘professional’, the interactions between singer experience and formant tuning may be more statistically significant.
Future Directions

Upon conducting this research, several areas for future research have come to my attention, including: the singer’s formant, differences among singing styles, and the possibility of professionals making a conscious stylistic choice to diminish voice quality.

Through the process of reviewing the current literature on formant adjustment in singers, the “singer’s formant” was mentioned in many books and research articles. As previously stated in the introduction, the singer’s formant is a resonance consistently present around 3 kHz, regardless of the vowel being sung (Kreiman & Sidtis, 2013). The purpose of this additional formant, which is actually a merging of F3, F4, and F5, is to increase the singer’s volume without damaging the vocal folds. 3 kHz is the weakest part of the orchestral spectrum, thus, if a singer amplifies the harmonics in that range, they can boost their volume without straining the vocal folds (Sundberg, 1974). The singer’s formant is a widely acknowledged and studied aspect of the male voice; however, future research on its presence in the female voice could reveal further differences based on sex. Additionally, it would be interesting to determine whether the singer’s formant appears in singers without formal training. Since its existence is meant to help the singer be heard over an orchestra, amateur singers who do not typically compete with a full orchestra may not develop such a formant.

Another area for future exploration is the effect of singing style on formant adjustment. It could be enlightening to investigate whether different genres are associated with different levels of formant adjustment. Singers of the operatic style could be more likely to engage in formant adjustment, as their voice regularly competes with the volume of an orchestra. In contrast, pop singers may be less likely to tune their formants as they typically perform using a microphone, or
perform mostly in quiet recording booths. Additionally, the typical R&B voice is characterized by complicated riffs, and these rapid, often improvised changes to the melody could make formant tuning exceedingly difficult. Because of this increased difficulty, it is possible that the disparity between professional and amateur singers is more pronounced in this genre in particular.

Finally, further research should explore instances in which singers intentionally choose not to engage in formant tuning. As previously stated, the human brain is capable of correctly identifying vowels despite significant distortion (Hollien, 2000). Thus, when singers choose to diminish their volume or voice quality, the purpose is not likely preservation of the vowel. Alternative explanations could be explored in future research. One potential factor could be the role of emotion in singing, and its varied forms of expression. Softness is often associated with vulnerability, and it is possible that some singers purposefully reduce their volume in order to convey vulnerable emotions, such as sadness. Future research could involve interviews with singers of varying genres to determine why singers allow for a decrease in volume or voice quality in some instances. This could also outline further genre differences.

**Conclusion**

The data from the scales, reading, and singing tasks provide useful insight into the formant adjustment of a variety of singers. This research demonstrates a significant relationship between singer experience and formant tuning, but more research could provide further clarification. The relationship between singer sex and formant tuning is evident throughout this
study and is supported by previous and current literature on the topic. There is strong support for this study’s hypotheses that female singers and professional singers show greater levels of formant adjustment in high frequency vowels than male singers and amateur singers. This study ascertains that sex is a defining factor in the formant adjustment of high frequency vowels, and calls for more research to determine the precise role of experience level in formant tuning.
References


http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.158.1227&rep=rep1&type=pdf


Appendix A: Consent Form

Ohio University Consent Form

Title of Research: The Sound of the Snow Queen: An Acoustic Analysis of Vowel Quality

Researchers: Megan Smith and Chao-Yang Lee (faculty advisor)

You are being asked to participate in research. For you to be able to decide whether you want to participate in this project, you should understand what the project is about, as well as the possible risks and benefits in order to make an informed decision. This process is known as informed consent. This form describes the purpose, procedures, possible benefits, and risks. It also explains how your personal information will be used and protected. Once you have read this form and your questions about the study are answered, you will be asked to sign it. This will allow your participation in this study. You should receive a copy of this document to take with you.

Explanation of Study

The purpose of this study is to explore the link between speech and music. To that end, we will make a recording of your singing and speech and conduct acoustic analysis on the recording. If you agree to participate, you will be asked to read a list of words, read the lyrics of Frozen’s “Let it Go” with and without music, sing a basic scale twice with two different vowels, and finally, sing “Let it Go” in two different keys. Your participation in the study is expected to take approximately one hour in a sound booth in Grover Center.

Risks and Discomforts

There are no known risks associated with this study. No discomforts are expected.

Benefits

This study is important to science because music and speech/language are both distinctively human activities, and their relationship has intrigued scholars in a wide range of disciplines. The proposed project will use well-established scientific methods to explore the interaction between singing and speech. The findings will contribute to our understanding of the music-speech relationship, which could have a series of fascinating implications across a myriad of disciplines.
You may not benefit, personally by participating in this study other than to learn how an acoustic phonetics project is conducted.

**Confidentiality and Records**

Information about you will be kept confidential by assigning you and your recordings a participant number instead of using your name. A master code list connecting the participant number and your recordings will be securely stored. The master code list will be destroyed once the study is completed. Your recordings will not be destroyed, but will be kept securely for their potential use in further analysis. The scientific literature indicates that without personal acquaintance, the possibility of identifying a person from such recordings is quite low.

Additionally, while every effort will be made to keep your study-related information confidential, there may be circumstances where this information must be shared with:

- Federal agencies, for example the Office of Human Research Protections, whose responsibility is to protect human subjects in research
- Representatives of Ohio University (OU), including the Institutional Review Board, a committee that oversees the research at OU
- The OU Finance Office, whose funds supply compensation for this study

**Compensation**

You will receive $40 to compensate for your time/effort spent in the study, which is expected to take approximately one hour. If after being briefed on the study, you decide not to participate, your compensation will be prorated based on the time you have spent in the study. Per Ohio University regulations, information about your name, address, and amount of payment will be reported to Ohio University Finance.

**Contact Information**

If you have any questions regarding this study, please contact Megan Smith, (937) 212-8356 or ms432111@ohio.edu, or Dr. Chao-Yang Lee, faculty advisor, (740) 593-0232 or leec1@ohio.edu.
If you have any questions regarding your rights as a research participant, please contact Chris Hayhow, Director of Research Compliance, Ohio University, (740)593-0664 or hayhow@ohio.edu.

By signing below, you are agreeing that:

- you have read this consent form (or it has been read to you) and have been given the opportunity to ask questions and have them answered;
- you have been informed of potential risks and they have been explained to your satisfaction;
- you understand Ohio University has no funds set aside for any injuries you might receive as a result of participating in this study;
- you are 18 years of age or older;
- your participation in this research is completely voluntary;
- you may leave the study at any time; if you decide to stop participating in the study, there will be no penalty to you and you will not lose any benefits to which you are otherwise entitled.

Signature __________________________ Date __________________________

Printed Name __________________________

Version Date: 12/22/14
Appendix B: “Let it Go” Lyrics

The snow glows white on the mountain tonight
Not a footprint to be seen
A kingdom of isolation,
And it looks like I'm the queen.

The wind is howling like this swirling storm inside
Couldn't keep it in, heaven knows I tried!

Don't let them in, don't let them see
Be the good girl you always have to be
Conceal, don't feel, don't let them know
Well, now they know!

Let it go, let it go
Can't hold it back anymore
Let it go, let it go
Turn away and slam the door!

I don't care
What they're going to say
Let the storm rage on,
The cold never bothered me anyway!

It's funny how some distance
Makes everything seem small
And the fears that once controlled me
Can't get to me at all!

It's time to see what I can do
To test the limits and break through
No right, no wrong, no rules for me I'm free!

Let it go, let it go
I am one with the wind and sky
Let it go, let it go
You'll never see me cry!

Here I stand
And here I'll stay
Let the storm rage on!

My power flurries through the air into the ground
My soul is spiraling in frozen fractals all around
And one thought crystallizes like an icy blast
I'm never going back,  
The past is in the past!

Let it go, let it go  
And I'll rise like the break of dawn  
Let it go, let it go  
That perfect girl is gone!

Here I stand  
In the light of day  
Let the storm rage on,  
The cold never bothered me anyway!