2010-03-25

Formant Changes in Amateur Singers After Instruction in a Vowel Equalization Technique

Emily Mullins Heaton
Brigham Young University - Provo

Formant Changes in Amateur Singers after Instruction

in a Vowel Equalization Technique

by

Emily M. Heaton

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Christopher Dromey, Chair
J. Arden Hopkin
Ron Channell

Department of Communication Disorders

Brigham Young University

April 2010

ABSTRACT

Formant Changes in Amateur Singers after Instruction

in a Vowel Equalization Technique

Emily M. Heaton

Department of Communication Disorders

Master of Science

Vowel equalization is a technique that can be used by singers to achieve a more balanced vocal tone. The technique balances corresponding *front* and *back* vowels, which share approximate tongue heights, and also balances *high* and *low* vowels in a more neutral or centralized lingual posture. Formants are resonance peaks that define each specific vowel. This study measured shifts in the first and second formants (F_1 and F_2) of the vowels /e, i, ɑ, o, u/ following training in vowel equalization. Prior to the training, the vowel formants were measured in amateur 15 college-aged singers. They sang the first two stanzas of "Somewhere Over the Rainbow" and then sustained each vowel for approximately 2 seconds. Following a 15-minute instruction in the vowel equalization technique, the singers repeated the exercises and the formants were re-measured. Shifts in F_1 and F_2 represent changes in lingual placement within the oral cavity. Vowel equalization pulls the lingual posture of a particular vowel to a more neutral or central position. While singing, a neutral placement is perceived as a pleasing balance between bright and dark tones.

This study showed that following training the singers' formant values changed in a manner reflective of a more central tongue posture. These findings support the suggestion that the vowel equalization technique does indeed alter the articulation of sung vowels, shifting the formants to produce the desired *chiaroscuro* or balance between bright and dark sounds.

Keywords: formant, singing, vowel equalization, chiaroscuro

ACKNOWLEDGMENTS

I would like to thank my husband, Benjamin Heaton, for his support to continue to work on this thesis. Your example and persistent encouragement enabled me to accomplish this project. I love you with all my heart. I also thank our little Joshua for his patience as I played on the computer and would not allow him to do so.

I appreciate my father, Dr. Richard Mullins, for helping me to love music and instruments of all kinds, as well as my mother, Suzanne Mullins, for her lifelong value of receiving an education and for her constant support. I am grateful for the enthusiasm of Terry and Ginger Heaton, as they helped me rejoice in small accomplishments along the way.

I would not have been able to complete this thesis without the understanding and guidance of Dr. Christopher Dromey as my thesis advisor and teacher, and friend. Thank you so much for all your advice and direction. Also, the enthusiasm and interests of Dr. J. Arden Hopkin allowed my understanding of this topic to expand.

A special thanks to each of you.

Table of Contents

List of Tables

List of Figures

List of Appendixes

Introduction

The voice can be one of the most striking instruments on stage, capable of both carrying a message and expressing emotion though the skilled performer's mastery of a flexible and powerful instrument. This instrument has three main components: the respiratory apparatus, the oscillating vocal folds, and the vocal tract (Sundberg, 1977). With careful practice and preparation, a singer is able to use these three components in a complementary way, producing beautiful music.

Classically trained singers strive for projection and a balanced resonance, or *chiaroscuro*, when they sing. The challenge of achieving this balance increases when singers must negotiate difficult leaps and long phrases while trying to make the words intelligible. Classical singers seeking balance between bright and dark timbres soon learn that these features of resonance relate to the distribution of high and low frequency spectral energy that defines the vowels (Hopkin, 1997).

A method for equalizing vowels, developed by Hopkin (1997), provides a way for the singer to achieve a more balanced resonance between bright and dark timbres for each of the vowels, /ɑ, e, i, o, u/. Through this technique, a singer is able to modify the actions of the vocal tract in such a way as to enhance the production of both the bright and dark sounds by altering the tongue position for each vowel. The purpose of the present study was to evaluate the impact of vowel equalization on the acoustic properties of vowels produced by a group of undergraduate and graduate amateur singers as they sought to balance the opposite vowels (both front/back and open/closed opposites) in a more neutral articulatory position. In the literature review that follows, themes from several areas of scholarship will be explored in order to put in context the analysis of the singing voice from a vowel acoustic perspective.

The Production of Voice

In order to understand how singers can balance their opposite vowels, it is important to first consider some basic aspects of vocal physiology. The respiratory system, including the lungs, associated muscles, and trachea, provides the air pressure necessary to set the vocal folds into motion as air is exhaled. Early voice researchers in the 1950s and 1960s explained vocal fold oscillation with the myoelastic-aerodynamic theory. According to this theory, air pressure from the lungs builds up underneath the closed folds. Once this pressure becomes high enough, the folds are blown outward, thus opening the glottis and releasing a single pulse of air. The lateral movement of the vocal folds continues until the natural elasticity of the tissue takes over, and the vocal folds move back towards their original, closed position. This closing movement is enhanced by Bernoulli forces (negative intraglottal pressure) causing the vocal folds to be sucked together. Then, the cycle begins again. Each cycle produces a single pulse of air; the sound of the human voice originates with hundreds of these pulses of air being released every second and then resonated by the vocal tract (Titze, 2008).

The structures of the vocal tract provide the articulation that forms the words of a song and also shape the resonant quality of the voice. This resonance is a key factor in the strength and quality of a singer’s voice, and it relies on the configuration of the resonating cavities of the vocal tract. Thus, the resonances of the vocal tract are responsible not only for the identity of the individual vowels and consonants, but also the aesthetic quality of the singing.

Acoustic Analysis of Sounds

The following review of some elementary aspects of acoustic analysis sets the stage for a consideration of the specific approaches used in evaluating the data from the present study. The voice is a complex sound that is propagated through the air as a series of waves. Any sound wave, complex or simple, can be decomposed into a series of sinusoids or can be described as a combination of pure tones. Acoustic analysis focuses on physical aspects of the sounds emitted from the vocal tract. It is a noninvasive method from which inferences can be drawn from the sounds a person produces, while not interfering with those sounds or the movements that produce them. Acoustic analysis reveals features relating to the movement of the vocal folds as well as the activity of the vocal tract. According to the Fourier Theorem, all sounds are made up of a combination of sine waves. These sine waves will have varying amplitudes, frequencies, and phase angles. There are two main ways that these complex sounds can be visually represented and analyzed. In time domain analysis, a waveform represents sound pressure or amplitude changes over time, while in frequency domain analysis a spectrum shows the frequency components of a sound for a given point in time. The latter display is particularly helpful in the present study of singing vowel acoustics, because it reveals features of the voice that relate to resonance adjustments.

With the increasing availability of digital analysis tools, use of the fast Fourier transform (FFT) has grown (Ferrand, 2006). FFT converts mathematically from a time domain waveform to a frequency domain spectral display. A line spectrum, also referred to as a spectral slice, is essentially a snapshot in time. It provides information regarding both frequency and amplitude but does not show changes over time. In a spectrum produced by FFT, all harmonics of the voice are represented in detail. A different spectrum is produced by a linear predictive coding (LPC) process, which reveals the spectral envelope of a vowel. This display shows the resonant peaks associated with the shape of the oral cavity, but the individual harmonics are not visible. Using this method to display speech, the spectrum has a smoother appearance compared to the more jagged FFT display. Thus, the sound spectrum of a complex tone is shaped by resonant peak frequencies (Miller, 1986; Sundberg, 1977). The frequencies of these peaks would be predicted to change as singers adjust their vocal tracts during the process of vowel equalization explored in this study.

Multiple spectral slices can be lined up sequentially to produce a spectrogram, which shows how frequency and amplitude change over time as a series of sounds is produced in a sentence or phrase. A spectrogram's x-axis represents time, the y-axis represents frequency, and darkness reflects the amplitude. *Formants* appear as wide, dark horizontal stripes, reflecting the concentrations of more intense acoustic energy at those harmonic frequencies that have been amplified by the vocal tract resonances (Miller, 1986).

Basic Elements of the Source-Filter Theory

Although the singer's voice originates in the larynx, it is important to consider the contributions of the vocal tract to the sound that reaches the audience. In the 1960s, Swedish scientist Gunnar Fant proposed the source-filter theory of vowel production (1961). In this theory, the three elements involved in vowel production—the glottal sound source, the vocal tract resonator, and the resultant sound that ultimately emerges from the lips—are presented on three spectra (Ferrand, 2006).

The glottal sound source. The first of the three spectra, the glottal spectrum, shows the sound produced in the larynx prior to any modification by the vocal tract. This spectrum consists of the fundamental frequency (F_0) which has the greatest amplitude, as well as higher harmonics, which decrease in amplitude, at the rate of 12 dB/octave (Sundberg, 1977). Sound from the glottal spectrum would be perceived as a buzz rather than a recognizable vowel. The glottal spectrum is known as the source function because the glottis is the source of the as yet non-specific vowel sound.

The vocal tract transfer function. The second spectrum of Fant's theory does not represent a sound but is a resonance curve representing the frequency response of the adult male vocal tract (approximately 17.5 cm, or 7 inches in length) positioned for the schwa, the neutral /ə/ vowel. This model has formant resonances at 500, 1500, and 2500 Hz (Sundberg, 1977). Each articulatory configuration results in its own distinct resonance patterns. These resonance patterns, when combined, are commonly referred to as the *transfer function*. Because speech is so rapid, the articulatory configurations also change rapidly and are seldom in a static position for very long. Sung vowels, because they are typically much longer, have been used to study and measure the resonance patterns more easily than spoken vowels. The harmonics of the glottal sound wave that are at or near the formant spectral peaks of the transfer function of the vocal tract are resonated and amplified; those distant from the resonant frequencies of the vocal are attenuated (Sundberg, 1977). Minifie points out that the overall shape and size of the vocal tract determines the nature of the filtering properties (Miller, 1986; Minifie, 1973).

The output sound. The third spectrum of the source-filter theory reflects sound as it emerges from the lips and is referred to as the *output function*. This spectrum displays how the glottal sound of the source function has been filtered according to the frequency response of the vocal tract. The same fundamental frequency and harmonics are present in the output as in the glottal source, but the amplitudes of the harmonics have been modified by the formants, which results in a specific sound quality more easily recognized as a vowel.

The three spectra of Fant's source-filter theory illustrate how the sound generated by the vocal folds is modified by the resonances of the vocal tract. These spectra display graphically how the fundamental frequency and its selectively amplified partials contribute to a given vowel.

Application of the Source-Filter Theory to Speech and Song

With an understanding of the conceptual components of the source-filter theory, it is appropriate to now consider the application of this model to the specifics of voice production in speaking and singing.

Fundamental and partials. As is the case with many complex sounds, the voice consists of a fundamental frequency and also additional waves, called *partials*, of that fundamental frequency. Briefly defined, a partial may be a simple tone whose frequency is typically higher than the basic frequency, and it contributes to the timbre, or quality, of the complex sound (Appelman, 1967). A *harmonic* is defined as a partial whose frequency is an integer multiple of the fundamental frequency. The vocal tract allows the fundamental frequency and its partials to resonate in the open cavities of the mouth and pharynx. In these cavities, certain harmonics of the voice source are amplified. At the relatively low frequencies of speech, the close spacing of the harmonics, as multiples of the fundamental frequency, assures that the formants, or resonant peaks of a vowel, are never far from a harmonic (Miller, Sulter, Schutte, & Wolf, 1997; Wolf, Stanley, & Sette, 1935).

The number, intensity, and distribution of the partial waves that compose a sound determine its quality or timbre. In voiced sounds, the relationships between these factors depend on the nature of the laryngeal vibration as well as changes made in the sound as it passes through the resonating system (Appelman, 1967). These two factors cause some partials to be reinforced, while others are weakened. Singing, as well as speech, is produced by constantly adjusting the oral and pharyngeal cavities during articulation, thereby causing variations in the overtone structure.

Articulation and formants. As the oral and pharyngeal cavities are adjusted during articulation, the shaping of the resonator tube produces prominent distributions of acoustic energy, a phenomenon that has led to the identification of frequency maxima or formants for each vowel sound (Luchsinger & Arnold, 1965). Although there are multiple formants present for each vowel, the first two formants are sufficient to distinguish one vowel from another. As mentioned previously, a speaker can move the lips, tongue, and jaw, which will alter the size and shape of the resonating cavities. Any change of vocal tract configuration alters the frequencies at which the cavities resonate and therefore the frequencies of the formants (Baer, 1979; Borden, Harris, & Raphael, 2002; Sundberg, 1977).

When the vocal tract changes in shape, the relationships between the oral and pharyngeal spaces change (Baer, 1979). Formant frequencies are related to the volumes of the oral and pharyngeal spaces, due to the fact that all containers of air will resonate at particular frequencies directly related to their volume. Generally, containers with a larger volume will resonate at lower frequencies, while those with a smaller volume will resonate at higher frequencies.

Kantner and West describe how resonance patterns produce recognizable vowels:

> All vowels, per se, have resonance but each vowel has its own distinct pattern of resonance that is the result of the number, frequencies and energy distribution of the overtones that are present. It is by means of these differences in the overall patterns of resonance that we are able to hear and discriminate one vowel from another. These changing resonance patterns are produced by altering shape and size of the discharging orifice (Miller, 1986, p. 50).

The frequency of the first formant, F_1, is related to the volume of the pharyngeal cavity as well as to how tightly the vocal tract is constricted, often discussed in terms of tongue height.

The frequency of the second formant, F_2, is directly related to the length of the oral cavity, and its overall volume, which is discussed in terms of tongue advancement (Sundberg, 1977). The volumes of the pharyngeal and oral cavities are indirectly related and are altered as the tongue and the jaw move forward and back, up and down.

For example, if one raises the tongue toward the palate, as would be done to produce the vowel /i/, the individual's pharyngeal cavity behind the tongue would be enlarged, while the volume of the oral cavity in front of the tongue constriction would decrease. F_1 will be lower, because the greater volume of the pharyngeal cavity resonates more strongly to lower harmonics. F_2 will be higher due to the shorter length of the oral cavity, which naturally amplifies higher frequencies.

With each placement of the tongue and jaw to produce a specific vowel, the relationship between formants F_1 and F_2 changes, resulting in changes to the center frequencies of each. Thus, the formants F_1 and F_2 of different vowels vary systematically, depending on the tongue height and advancement (Hillenbrand, Getty, Clark, & Wheeler, 1995; Peterson & Barney, 1952). It is also important to remember that formant frequencies change depending on the length of the vocal tract. Fant's source-filter theory is based on an adult male's vocal tract, while adult females have a shorter vocal tract than adult males, and children's vocal tracts are shorter still. Because of these differences in size, a woman's voice will naturally resonate at higher frequencies than a man's, while a child's voice is even higher than a woman's (Maurer, Cook, Landis, & D'Heureuse, 1991).

Lindblom and Sundberg (1971) proposed an articulatory model for deriving formant frequencies from measures of lip and jaw position, tongue shape and height, and larynx height. They stated that simulated larynx lowering corresponded to effective lengthening of the pharynx tube. This resulted in the lowering of all formant frequencies, most especially those associated with back cavity resonance (Detweiler, 1994; Lindblom & Sundberg, 1971; Sundberg, 1977).

Quantification of Formant Relationships

Several approaches have been devised in order to describe the relationships between the vowel formants in individuals and groups, in order to better understand and interpret the acoustic data from studies of speech and singing. Research has provided a general idea of where the first and second formant frequencies lie for each vowel. Some variation in these values may be due to the different sample groups and sizes that were used by the different researchers. Consequently, the actual values of formant frequencies will vary with each individual, and will result in wide variability across subjects. For these reasons, one group of researchers studied preferred to list formant frequencies as ranges (see Table 1), rather than as specific values (Maurer et al., 1991).

F_1/F_2 plots. F_1/F_2 plots are often used to graph formant information, and thereby visually diagram the relationship between the first and second formants for each vowel. This plot, often referred to as a vowel space, vowel quadrangle, or vowel quadrilateral, is a graph with F_1 on the horizontal axis (corresponding to tongue height), while F_2 is plotted on the vertical axis (relating to tongue advancement). The vowel space can be read in the same manner as finding coordinates on any graph, by looking at a point's position in relation to each axis. Peterson and Barney (1952) produced a F_1/F_2 plot using the vowels /i/, /ɑ/, /u/, and /æ/ for averages of values obtained from men, women, and children. However, it should be noted that the relationship of the first two formants to tongue position is only an approximation, and reflects in general terms the vocal tract's response, rather than reflecting the individual resonating cavities in any level of anatomic detail.

Vowel space area. The vowel space area (VSA) is a single number calculation (in Hertz squared) that quantifies the area enclosed within the space of a vowel quadrilateral. The VSA is an acoustic index commonly used in clinical research to indirectly assess the normalcy of vowel articulation (Kent & Kim, 2003; Kuhl et al., 1997; Vorperian & Kent, 2007). Generally, as a more extreme or exaggerated articulatory placement is used, the VSA value increases. Therefore, the VSA allows researchers to compare articulatory activity across individuals or in the same speaker under different conditions.

Formant ratios. Ladefoged (1967) achieved a closer match between formant frequencies and actual tongue positions and shapes by plotting F_1 against the difference between F_1 and F_2 of a target vowel, rather than by simply plotting F_1 directly against F_2, as happens when VSA is used. What results from Ladefoged's method is a more physiologically accurate placement of the back vowels. Ladefoged (1967) pointed out that the impressions of vowel height are related more closely to the frequency of the first formant than to the actual height of the tongue (Borden et al., 2002; Ladefoged, 1967). Similarly, impressions of tongue advancement can be simply understood by referring to the difference between the first and second formant frequencies rather than to any measurement of the actual horizontal position of the tongue (Ladefoged, 1967). At the same time, Fant (1961) has observed that the highest point of the tongue is not as important as the maximum constriction and length of the tract from the glottis to this point.

Vowel articulation index. In more recent findings, evidence of vowel space expansion, where more pronounced articulation of the vowels occurs, has been documented following manual circumlaryngeal therapy in individuals with muscle tension dysphonia. Roy, Nissen, Dromey, and Sapir (2009) recently introduced a new measure for calculating an index of articulatory function. The Vowel Articulation Index-4 (VAI-4) is derived from a formula based

Table 1

Five Vowels and Their Corresponding Formant Ranges

Vowel	F_1 (min-max Hz)	F_2 (min-max Hz)
/i/	200-400	1900-3200
/e/	300-600	1800-2800
/ɑ/	550-850	900-1600
/o/	300-600	650-1200
/u/	200-400	600-1000

on the values of the formants of the 4 point vowels. The VAI-4 is expressed as (F2i+F2ae+F1ae+F1a) / (F1i+F1u+F2u+F2a). The vowel-formant elements are arranged such that during more distinct articulation, the elements in the numerator will increase and elements in the denominator will decrease with larger articulatory movements (Roy et al., 2009). Therefore, for exaggerated articulation, which is often characteristic of singing, the VAI-4 would likely display a higher value than in normal speaking for a given individual.

Evolution of the Acoustics of Singing

While acoustic analysis can be applied to any type of sound, the focus of the present study on the impact of training on vowel acoustics in singers necessitates a consideration of approaches that have been employed to understand the performing voice.

Impressionistic approaches. The first attempts to systematically describe the singing voice relied on purely perceptual approaches. The articulation of vowels in singing differs from the way vowels are produced during speech. Generally the vowels are prolonged and some consonant clusters are reduced while singing. However, to understand the study of sung vowels, spoken vowels must first be considered. The earliest traditional descriptions of basic vowels, /i/, /ɑ/, and /u/, are now well over 100 years old and were based almost entirely on the impressionistic, introspective evidence of phoneticians who viewed vowel articulation primarily as functions of tongue shape and tongue position (Borden et al., 2002). Lip posture, since it was so easily observable, also formed one of the bases for description. With the use of more objective techniques for describing articulation, other aspects became more important, such as the tongue root and jaw position.

Early acoustic studies. The development of acoustic analysis led investigators to explore the relations among articulation, the dimensions of the resonating cavities, and the acoustic features of speech, especially with regard to formant frequencies (Borden et al., 2002; Sundberg, 1977). Acoustic analysis of the voice often includes making measurements of an individual's fundamental frequency, formant frequencies, amplitude, and perturbation. The most widely cited experiment on the acoustics and perception of vowels was a surprisingly simple study conducted at Bell Telephone Laboratories by Peterson and Barney (1952) shortly after the introduction of the sound spectrograph. The researchers recorded two repetitions of ten vowels in a /hVd/ context (where V is the one of the ten vowels) spoken by 33 men, 28 women, and 15 children. Acoustic measurements from narrow-band spectra consisted of formant frequencies (F_l-F_3), formant amplitudes, and fundamental frequency (F_0). The measurements were taken at a single time slice that was judged to be the steady state. The /hVd/ signals were also presented to listeners for identification. The results of the measurement study showed a strong relationship between the intended vowel and the formant frequency pattern. However, there was considerable formant frequency variability from one speaker to the next, and there was a substantial degree of overlap in the formant frequency patterns among adjacent vowels. The listening study showed that the vowels were highly identifiable. The overall error rate was 5.6%, and nearly all of the errors involved confusions between adjacent vowels (Hillenbrand et al., 1995; Peterson & Barney, 1952).

Wolf, Stanley, and Sette (1935) were among the first to apply the newly developed technology of acoustic analysis to the singing voice. Some of the factors these researchers focused on included intensity as a function of time and pitch, vibrato, vibrato-tremolo, and tremolo-quality. The researchers plotted the power of a singer's voice as a function of the pitch range because they considered the ability to sustain a high level of intensity over a wide range of pitches more valuable than measuring a loud sound on a few vocal tones (Wolf et al., 1935).

Bartholomew (1934) undertook another of the early attempts at acoustic analysis of the trained singing voice. He recorded sample phonations representing various voice qualities from over forty participants. He then analyzed these samples with the intent of isolating those characteristics consistently present in the samples judged by his experts to be considered good quality. He concluded that good vocal quality in the male voice included a regular and even *vibrato*, adequate intensity, the presence of a strong low formant around 500 Hz, and a strong high formant between 2400 and 3200 Hz. (Bartholomew, 1934; Detweiler, 1994).

Sundberg's contributions. Much of the pioneering work in the field of musical acoustics was conducted in Stockholm at the Royal Swedish Academy of Music by Johan Sundberg. In a 1973 study he compared singers' sung vowels with those produced during normal and loud speech. He concluded that singing and speech produced by trained singers do not differ markedly in so far as the source spectrum envelope is concerned. He found that an acoustic model of the vocal tract including volumes simulating the expanded ventricular spaces and pyriform sinuses was adequate to account for all the resonance attributes observed in the trained singing voice (Detweiler, 1994; Sundberg, 1973).

In later work by Sundberg (1977), the researcher discussed how during singing, the voice is often perceived as *darker*, similar to how one sounds when yawning and speaking at the same time. Many voice teachers use the imagery of a yawn sensation to describe the desired placement to their students, and this is referred to as *covered* singing with the larynx lowered. X-ray images reveal that when a larynx is in this covered position, it is accompanied by an expansion in the lower portion of the pharynx. In other words, the evidence suggests that placing the larynx in a lowered position and the pharynx in an open position is judged favorably in singing (Sundberg, 1977).

Methods in the analysis high pitch. At high pitches, the harmonics of the voice are spaced widely apart and they do not naturally align with the formants of the vocal tract. Through various types of vowel modification, a singer is able to project the fundamental frequency and its harmonics more effectively regardless of the choice of vowel. This modification is achieved through tongue, lip, and velar adjustments, so that the formants are moved to better align with the fundamental or its harmonics and thus transfer energy more efficiently (Miller, 1995; Titze, 1995).

At the higher frequencies of singing, especially that of higher pitched voices, the proximity of the formants (widely spaced) to harmonics is a prominent and even a critical consideration in the singer's resonance. The same wide spacing of harmonics which makes the locations of the formants along the frequency spectrum critical makes it difficult for researchers to precisely determine formant frequencies of sung tones at higher pitches (Erickson, 2004; Miller et al., 1997). The complexity of determining formant frequencies of sung tones grows as F_0 increases.

Miller et al. (1997) tested whether an informal method could be useful in measuring formants in singers and found that substituting a nonperiodic sound source (such as vocal fry, ingressive phonation, or a vibrator applied to the exterior of the neck) while a singer maintained the specific posture of the vocal tract was successful in revealing where the formant frequencies *might be*. The researchers reported that nonperiodic imitations can never give a definitive determination of formant frequencies during singing (Miller et al., 1997).

Vocal Tract Adjustments Traditionally used by Singers

With experience, singers learn to make subtle adjustments to the larynx and vocal tract in order to achieve the most efficient radiation of sound without risking damage to the delicate vocal folds, while also delivering the most pleasing aesthetic balance of light and dark tones.

Mastering vocal tract tuning. Vocal tract tuning is often used by singers in order to amplify a given note regardless of the identity of the vowel and its associated resonance patterns (Caldwell, 2002; Sundberg, 1977). In short, a singer will adjust the resonances (or cavity volumes) to match the harmonics, which are determined by the composer, who chose the pitch for the sung vowel. The result of this vocal tract tuning is a louder output of the voice, with a mild distortion of the vowel that the listener perceives. In preparing to sing a piece, a singer is not at liberty to alter the assigned pitch of a given note, but can adjust the vocal tract to resonate more fully the fundamental frequency and its harmonics.

Vocal tract tuning can be accomplished in a number of ways. All frontal vowels may be sung with more extreme lip-spreading and with an extreme frontal tongue position at *pianissimo* and *piano* levels (very soft and soft volumes, respectively). As intensity (or loudness) is increased above the *mezzo piano* level (medium soft), the size of the oral cavity and its opening also increases, along with a passive increase in pitch (Erickson, 2004; Fletcher, 1993; Miller, 1995). This is a natural, although passive, adjustment in the acoustics to accommodate the rising fundamental and for an increase in loudness. The opening of the mouth permits vowel differentiation and vowel recognition to remain intact (Miller, 1995).

The fundamental frequency of a given note may lie at a higher frequency than the normal first formant of the vowel being sung. In this case, the amplitude of the fundamental is not enhanced by the first formant and the sound is weak. If the singer opens the jaw wider, the frequency of the formant is raised. Therefore, when the frequency of the first formant is raised to match or approach that of the fundamental frequency, the formant enhances the amplitude of the fundamental and the resultant sound is louder (Sundberg, 1977). According to Sundberg (1977), a singer can use this method of vocal tract tuning to boost the voice in order to be heard above the orchestra. In a study addressing this concept, analysis of formant frequencies confirmed that the articulation was being varied in such a way as to raise the first formant frequency close to the frequency of the fundamental frequency of the tone being sung (Sundberg, 1977).

Opening the jaw, however, is not the only way to raise the first formant frequency. Shortening the vocal tract by drawing back the corners of the mouth serves the same purpose, and that may be why some teachers tell their students to smile when they sing high notes (Miller, 1986; Sundberg, 1977).

According to Titze (1995), increasing vocal intensity often results in vowel modification to some degree. This vowel modification entails altering true or pure vowels slightly towards a more neutral vowel. Louder vocalization, along with its consequential vowel modification, simulates the *megaphone effect* where the jaw is lowered to open the mouth as widely as possible in an attempt to match the acoustic impedance of the oral cavity to the free space surrounding the mouth. Additionally, the lips are moved forward, producing a fish-mouth, which ultimately widens and lengthens the vocal tract. The megaphone effect results in the shifting of the first formant higher and the second formant lower.

According to Appelman (1967), increasing the jaw opening results in a change in the cavity coupling. This change is largely due to moving the point of constriction of the tongue and palate slightly downward and backward, thereby creating a larger frontal cavity and less constriction at the inner orifice. This action causes the first formant to rise and the second formant to lower and the ear hears this cavity alteration as a migration of each frontal vowel to or toward a phoneme directly below it, depending upon the increase of intensity and the lowering of the mandible. For example, an /i/ may be perceived to approach the slightly lower front vowel /ɛ/. This same action occurs with the singing of back vowels, where the perception of /u/ may approach /o/. An increase in the volume of the cavities and the separation of the point of constriction at the inner orifice causes these vowels to migrate toward the neutral vowels [Λ] and [ə] (Appelman, 1967).

Applying the vowel quadrangle in pedagogy. Training singers to make adjustments to the vocal tract to enhance the aesthetic quality of their singing often involves instruction in some basic principles of vowel production. One way to visually represent the relative position of each vowel within the mouth is by using a vowel quadrangle. The vowel quadrangle is widely used in such areas as phonetics and language diction and is based on the association of major vowels with each other. The vowel quadrangle consists of five vowels, /i/, /e/, /ɑ/, /o/, and /u/. It displays the vowels in a physiologic relationship in terms of tongue height (also jaw opening) and tongue advancement. As shown in Figure 1, jaw position is on the vertical axis, while tongue advancement, described as *front* or *back,* is on the horizontal axis. The left edge of the chart is considered the opening of the mouth, while the right edge represents the glottal end of the oral cavity. The top of the chart represents the roof of the mouth, while the bottom of the chart reflects the base of the jaw.

According to general pedagogical practices, the vowels at the top of the triangle, /i/ in the left (forward) corner and /u/ in the right (back) corner, are called *closed* because, in order to say them correctly, the jaw must be more closed than open (Hopkin, 1997; Miller, 1995). In contrast, the vowel /ɑ/ is considered an *open* vowel because the jaw is lowered in its position and the

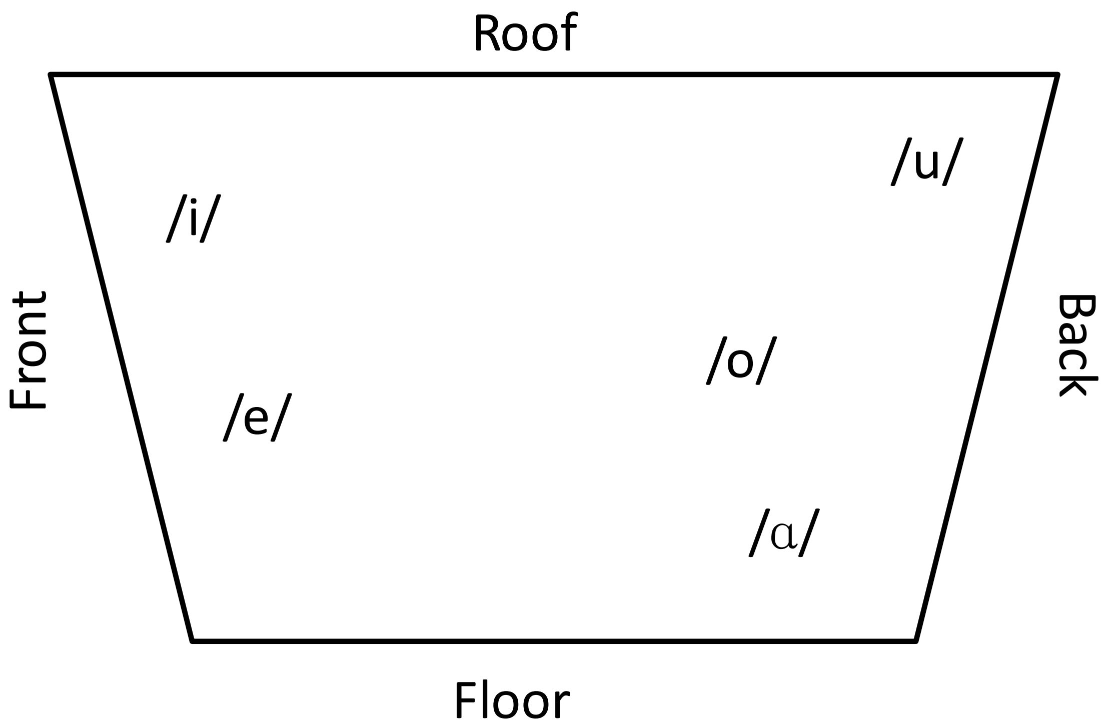

Figure 1.

Vowel Quadrangle.

mouth is more open than closed. Vowels on the left of the chart, /i/ and /e/, are said to be in a *forward* position because the tongue arches forward in the mouth, while the vowels on the right of the chart, /u/ and /o/, are *back* vowels because the tongue arches toward the back edge of hard palate (Detweiler, 1994; Ferrand, 2006; Kantner & West, 1960; McKinney, 1982; Miller, 1986; Sundberg, 1977; Titze, 1994).

Four Pedagogical Approaches in Vowel Singing

According to general pedagogical theory, there are four main approaches for teaching and obtaining the most desirable vowel sound while singing. These four methods comprise the basic tools taught to voice students in one combination or another. Therefore, most professional singers will use elements of these approaches as they perform. Each of these approaches will be discussed in order to differentiate the underlying theories, as well as to compare them.

Adjustamento. This approach suggests that as the sung scale increases in pitch, the same basic postures of lips, tongue, and the musculature of the zygomatic, or cheekbone, region remain consistent. However, the mandible may lower slightly as different vowels are shaped. Because of the change in the oral cavity as the jaw opens (as the mandible is lowered), the vowel is modified from its typical form.

According to proponents of this approach, there is no single ideal position of the vocal tract for singing. Rather it is the vowel, the *tessitura* (or the range in which a given type of voice presents its best-sounding texture or timbre), and the intensity-level (or loudness) that dictate the degree of oral opening and the extent of mandible lowering. The balance of these elements permits a constant subtle adjustment of the vocal tract necessary to produce the desired vowel.

Adjustamento is often described as *covering* the vowel. The idea of cover comes from the Italian concept of *chiaroscuro*. *Chiaroscuro* literally means light and shadow (Nix, 2004).

Classical singers must find a careful balance between these two elements to sing freely. The *scuro*, or shadow, is achieved by a*djustamento*, or cover. Cover is an integral part of healthy and beautiful singing (Nix, 2004).

In the *adjustamento* approach, the modification process is a gradual one of adjusting the acoustic energy in that part of the spectrum which defines the vowel, as the vowel is largely determined by the shifting energies located between the first, second, and third formants of the singing voice. Again, as the pitch increases, the articulatory position is maintained for each vowel, as the jaw opens to accommodate the higher notes.

Vowel modification. A second pedagogical approach is termed *vowel modification* and is similar to the a*djustamento* method. Vowel modification is used in Northern European schools of singing and, like a*djustamento,* adopts alterations at a specific pitch ranges. However, vowel modification emphasizes specific migrations towards the neutral vowel, schwa /ə/, at those specific pitch ranges as the scale ascends (Miller, 1977).

Pure vowels. A key to this approach lies in the preference of *vowel purity* rather than allowing the vowel modification of the first and second approaches. However, as higher pitches are sung, the oral position is maintained; the resultant sound quality will diminish. Although the vowels will be *pure*, they will be perceived as shrill in timbre because of the joining of high pitch and concentrated acoustic energy centered in the upper regions of the spectrum. The first formant will not be permitted to offer a sufficient counterbalance to the high concentration of upper harmonics and there will be a marked change in the *chiaroscuro* aspects of the balanced sound (Miller, 1995). The vowels may be perceived to consist of more *chiaro* than *oscuro* (more light than dark). This approach seeks to retain the *pure* vowel without modification throughout all pitches sung.

As a singer's pitch is raised the intensity or loudness of her singing also increases. Various hypotheses have been proposed to explain this increase of intensity as pitch is raised. One such theory addresses how the fundamental frequency passively increases (the pitch increases) as the voice becomes louder. In order for a singer to increase vocal intensity, tension in the laryngeal muscles also increases. This tension requires greater subglottal pressure from the muscles in the thorax to push air through the vocal folds. Therefore, indirectly the mechanism employed to obtain a greater intensity also results in a passive rise in the fundamental frequency. Furthermore, the converse is also true: as the fundamental frequency (pitch) rises, the tension on the vocal folds will increase. More pressure will then be needed to drive the stiffer vocal folds. The increase in breath pressure needed under these circumstances can partly account for the rise in vocal intensity at higher pitches. It is well known that a great deal more exertion is involved in singing a high tone than in producing a lower one (Wolf et al., 1935).

Idiot jaw. The final theoretical approach greatly differs from the three approaches previously discussed. It recommends cancellation of most vowel recognition in upper register singing as all vowels modify to a universal neutral vowel. The jaw is perpetually open and often fixed, while the tongue and lips are primarily responsible for forming the vowels. This is often referred to as the *dumb jaw, slack jaw,* or the *idiot jaw* (Miller, 1995). Although this method will not result in the perceived brassiness of the pure vowel approach, the vowel recognition is sacrificed. Additionally, conscious lowering of the larynx is also associated with this method. The vowels sung using this method seem to have *chiaro* outweighed by the *oscuro* (more dark than light). In other words, this approach emphasizes the depth and warmth of *oscuro* in favor of the bright *chiaro* timbres. This method of singing is currently one of the most commonly used by amateur singers in the United States. However, because passages sung with this method frequently display a radical change in the vowel at specific locations in the ascending scale, this approach is often avoided by professional singers.

The task of the singer is to determine the appropriate degree of vowel modification for the voice within the rising scale while still retaining language recognition. *Pure* and *modified* vowels are essential to the accomplishment of an evenly balanced scale. Beauty in singing depends on both the assigned melody and use of text, as well as the chiaroscuro (balancing between the light and dark aspects of the sound) which the trained voice develops and portrays. Additionally, singers strive to harness the ability to adequately project their voices as they balance the resonance. As singers attempt to present difficult passages they often rely on subjective sensations, which are difficult to discuss in precise terms (Hopkin, 1997).

Implementation of Hopkin's Vowel Equalization Approach

To help singers achieve the optimal balance of light and dark tones, Hopkin has developed a pedagogical technique that replaces the abstract terms of placement, resonance, and projection with more concrete terms and defined articulatory placements. This method, termed *vowel equalization*, offers an effective way to establish chiaroscuro balance in a sung tone (Hopkin, 1997). Vowel equalization is heavily associated with the use of the vowel triangle, focusing on the five vowels /i/, /u/, /e/, /o/, and /ɑ/. Hopkin indicates that the closed vowels /i/ and /u/ are produced with contrasting tongue positions (front vs. back), while the jaw position associated with producing each of these vowels is comparable with the opening of approximately 2 cm during singing. With this similarity, Hopkin indicates that these *opposite* (front-back) vowels are closely related. Similarly, the same relationship exists between /e/ and /o/. The tongue arches in opposite directions for each of these vowels, but the jaw position is shared at approximately 2.5 cm between the upper and lower front teeth while singing (Hopkin, 1997).

In order to equalize these opposite vowels using Hopkin's technique the singer will quickly glide from one vowel, such as /i/, to the matching vowel, in this case /u/. Only the singer's lips and tongue will move, while the jaw position will remain constant. The vowel equalization technique will result in a placement of the tongue in relation to the jaw in a manner that is not typically used while producing either vowel in isolation. According to Hopkin, rapid alternation between /i/ and /u/ helps eliminate any unnecessary movement between the two vowels (Hopkin, 1997). Through this technique vibrations will stabilize in the middle of the mouth as a result of the singer juggling the vibrations between the front and back of the mouth. This stabilization will indicate that the vowels are equalized. Vowel equalization allows for a singer to balance the *bright* of front open vowels and the *warm* of back closed vowels. Acoustically speaking, this method (as well as others mentioned previously) adjusts the filter of the singer's instrument. Therefore, Hopkin's vowel equalization technique alters the formants of the singer's vowels.

The present study was designed to measure the changes that occur as individuals of varying amateur vocal ability were trained to equalize their vowels. It was anticipated that the research would reveal a difference in the formant frequencies of the vowels following the vowel equalization training when compared with the initial productions.

Method

Participants

Participants were recruited through classroom announcements on the Brigham Young University campus. Non-audition university choral class students were invited to sign up to participate in the study. These individuals were then contacted by the researcher to be informed more specifically about what the study entailed. An interview was conducted with each participant to discover whether he or she met the requirements of the study, as well as to determine their age, year in college, and the length and type of vocal training they had received. Participants included 16 individuals, 7 males and 9 females, who engaged in recreational singing at least one hour per week (such as in a volunteer choir). Participants ranged in age from 18 years to 25 years and had developed varying levels of amateur vocal skill.

All participants were required to pass a hearing screening, to have no history of voice disorders or persistent articulation disorders beyond primary school, and were in good general health. Each participant signed an IRB-approved consent form (see Appendix), which indicated his or her willingness to volunteer as a participant in this research.

Materials

Recordings of the singers' voices were made in a sound booth using a head-mounted microphone (AKG C-420). A lip-to-microphone distance of 4 cm was maintained throughout the collection of data. To measure vocal intensity, a sound level meter (Larson-Davis 712) was positioned 100 cm from the singers' lips. An electroglottograph (EGG; Glottal Enterprises EG2) was used to measure changes in vocal fold contact area during phonation. The signals from the microphone, EGG, and sound level meter were routed into a multi-channel analog-to-digital conversion system (Windaq 720) on a laboratory computer. The microphone signal was low-pass filtered at 12 kHz and digitized at 25 kHz.

Procedures

Participants were scheduled in groups of 3-4 on five different occasions. Following several minutes to allow warm-up exercises (specific activities chosen by the singer), each singer was recorded separately as they sang the initial two stanzas of "Somewhere Over the Rainbow" (referred to as *the passage*) three times, on a pitch chosen by the participant at a comfortable loudness level. This passage was selected because of the occurrence of each of the five vowels, /ɑ/, /e/, /i/, /o/, and /u/, within the first two stanzas. After singing the passage, each participant also sustained the *isolated* vowels for five seconds, three times each. The groups of participants received a brief training in the vowel equalization technique. First, the singers were coached as they balanced the opposite front and back vowels (/i/ to /u/; /e/ to /o/) by holding a steady jaw position and reducing the extent of front/back tongue movement as they sustained phonation while moving between the two sounds in each set. Next, the participants balanced the opposite open and closed vowels (/i/ to /e/; /u/ to /o/). Following these equalizing exercises, the singers were guided as they attempted to balance /ɑ/ with /e/ and /o/, paying attention to vibratory sensations of the vowels in relation to tongue placement. A more complete description of the vowel equalization training is found in the original article by Hopkin (1997). After the training each singer was asked to sing the passage three times and sustain the isolated vowels, again three times.

Data Analysis

Vowels from the singing passage were segmented by visual inspection of the microphone waveform and spectrographic display. A one-second window at the midpoint of each sustained vowel token was used in the data analysis. Praat acoustic analysis software (version 5.0.47) was used to create formant tracks from these vowel segments. These tracks were then saved as text files and imported into Matlab, where custom analysis routines were used to perform additional numerical analyses of the acoustic signals. This included computing the average first and second formants for the middle 50% of each vowel segment. These formant values were averaged across the three repetitions and imported into SPSS 16.0 for statistical analysis by way of repeated measures ANOVA. In these analyses the independent variable was the pre-instruction or post-instruction condition. The dependent variables were the formant values of the vowels from the passage and the sustained vowels.

Results

Due to technical difficulties during one of the recordings, data for only 15 of the 16 singers are reported here. The sustained vowel and passage vowel analysis involved a comparison between the pre-instruction and post-instruction measures. The descriptive statistics and repeated measures ANOVA results are summarized in Table 2 for the passage vowels and in Table 3 for the sustained vowel tokens. These results reveal a greater number of changes in the formants of the vowels coupled front and back (/i/ to /u/; /e/ to /o/) than for those coupled vertically. The data do reveal changes in the vowels that were vertically coupled (/i/ to /ɑ/; /o/ to /ɑ/), but fewer of the results were statistically significant. Only those results that were found to reach statistical significance ($p < .05$) are reported here in detail. General trends and statistical differences between pre-instruction and post-instruction measures are reported in Table 4 for both the passage vowel and the sustained vowel tokens.

Vowels from the Singing Passage

Means for both F_1 and F_2 for the vowel /ɑ/ decreased significantly following the training in vowel equalization. This suggests that compared to the position prior to the training, the singers' tongues approached a position higher and further back in the mouth. F_2 for the vowels /e/, /i/, and /u/ also decreased significantly following training, indicating a retracted tongue

Table 2

Descriptive Statistics and ANOVA Results for First and Second Formants for the Vowels in the Singing Passage

	Mean Pre	*SD* Pre	Mean Post	*SD* Post	*F*-ratio	*p*-value
/ɑ/ F_1	701.7	133.9	632.7	113.6	13.40	0.003**
/ɑ/ F_2	1228.6	159.0	1122.6	107.9	19.32	0.001**
/e/ F_1	482.5	70.4	485.5	74.9	0.058	0.813
/e/ F_2	1954.0	233.3	1816.2	231.1	15.583	0.001**
/i/ F_1	328.5	56.7	331.6	52.8	0.041	0.843
/i/ F_2	2256.5	255.2	2126.0	207.9	10.464	0.006**
/o/ F_1	610.2	154.2	586.3	142.4	3.832	0.071
/o/ F_2	1102.6	182.5	1080.4	182.5	0.555	0.469
/u/ F_1	307.3	44.6	309.0	37.5	0.039	0.847
/u/ F_2	1231.1	195.2	1163.3	151.7	7.247	0.018*

Note. Degrees of freedom are 1, 14 for ANOVA.

*p < .05. **p < .01.

Table 3

Descriptive Statistics and ANOVA Results for First and Second Formants for the Sustained Vowels

	Mean Pre	*SD* Pre	Mean Post	*SD* Post	*F*-ratio	*p*-value
/ɑ/ F_1	683.3	135.2	622.0	122.0	4.161	0.061
/ɑ/ F_2	1169.8	140.6	1131.4	146.9	1.121	0.308
/e/ F_1	513.2	94.6	503.1	97.6	0.159	0.696
/e/ F_2	1942.9	266.8	1832.4	240.1	19.989	0.001**
/i/ F_1	305.9	48.1	321.5	44.0	4.180	0.060
/i/ F_2	2245.6	246.2	2121.6	222.2	9.168	0.009**
/o/ F_1	502.4	68.4	501.8	82.0	0.001	0.976
/o/ F_2	965.5	104.8	999.9	94.3	1.942	0.185
/u/ F_1	321.6	56.8	347.7	50.7	5.797	0.030*
/u/ F_2	1048.7	171.4	1166.8	227.2	5.785	0.031*

Note. Degrees of freedom are 1, 14 for ANOVA.

*p < .05. **p < .01.

Table 4

Trends in First and Second Formants from Pre-Instruction to Post-Instruction Measures for Vowels in the Singing Passage and the Sustained Vowels

	Passage vowels	*p-value*	Sustained vowels	*p-value*
/ɑ/ F_1	↓	.003**	↓	.061
/ɑ/ F_2	↓	.001**	↓	.308
/e/ F_1	↑	.813	↓	.696
/e/ F_2	↓	.001**	↓	.001**
/i/ F_1	↑	.843	↑	.060
/i/ F_2	↓	.006**	↓	.009**
/o/ F_1	↓	.071	↓	.976
/o/ F_2	↓	.469	↑	.185
/u/ F_1	↓	.847	↑	.030*
/u/ F_2	↓	.018*	↑	.031*

$^*p < .05$. $^{**}p < .01$.

position toward a more neutral placement for vowels /e/ and /i/. F_1 for the vowel /o/ decreased, but this change did not reach significance at the $p < .05$ level. See Figures 2-6.

Isolated Vowels

F_2 for the sustained vowels /e/ and /i/ decreased significantly following the vowel equalization training, bringing the tongue to a more neutral position following training. F_2 for the vowel /u/ increased significantly following the training, also bringing the tongue to a more neutral position. F_1 for the vowels /u/ and /i/ increased following training, while F_1 for the vowel /ɑ/ decreased, but this latter change did not reach significance at the $p < .05$ level. The changes in F_1 in vowels /u/ and /i/ represent bringing the height of the tongue into a more neutral vertical placement. See Figures 7-10.

Reliability Measures

The original and rechecked values for the two singers whose data were checked by the researcher were highly correlated, in that Pearson r was computed to be above .99 at $p < .001$. The mean absolute difference between the original and re-measured values was 0.96%.

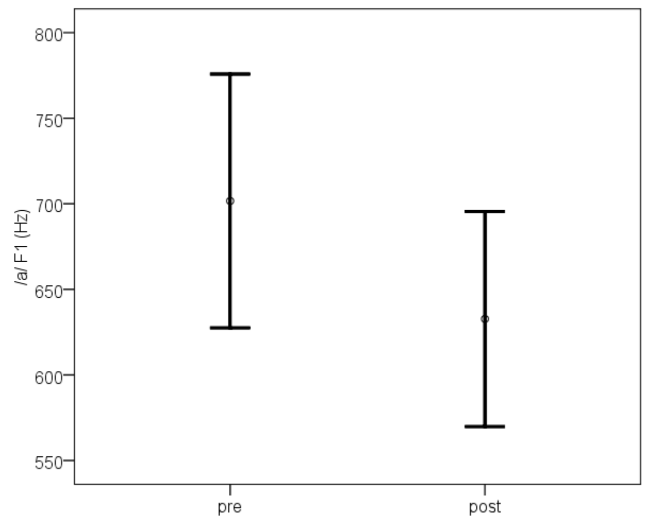

Figure 2.

The changes in the mean of F_1 of /ɑ/ during the passage following the instruction of the Vowel Equalization technique.

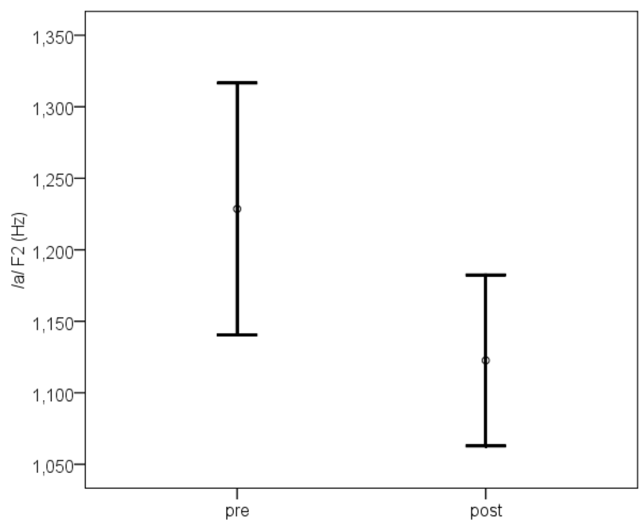

Figure 3.

The changes in the mean of F_2 of /ɑ/ during the passage following the instruction of the Vowel Equalization technique.

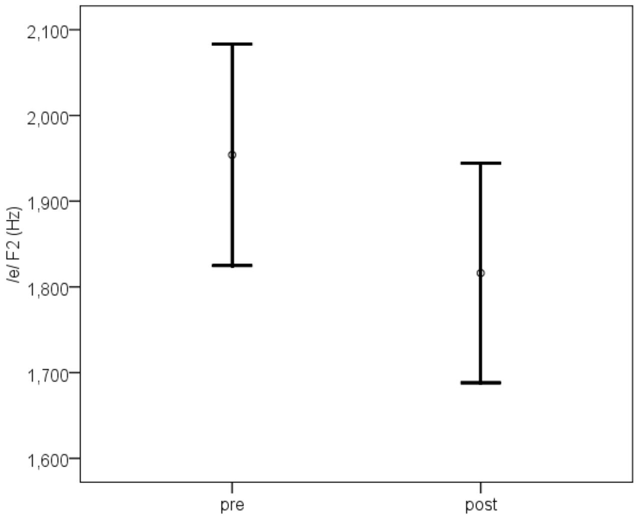

Figure 4.

The changes in the mean of F_2 of /e/ during the passage following the instruction of the Vowel Equalization technique.

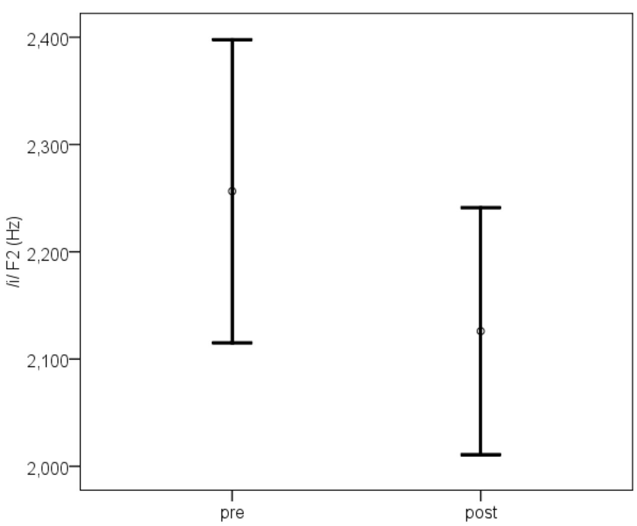

Figure 5.

The changes in the mean of F_2 of /i/ during the passage following the instruction of the Vowel Equalization technique.

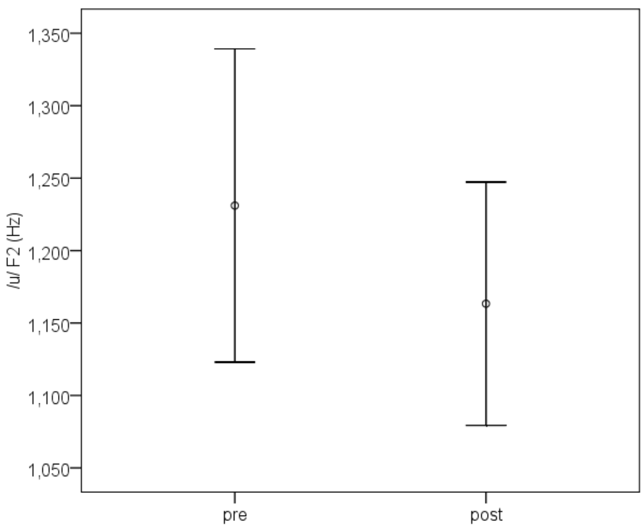

Figure 6.

The changes in the mean of F_2 of /u/ during the passage following the instruction of the Vowel Equalization technique.

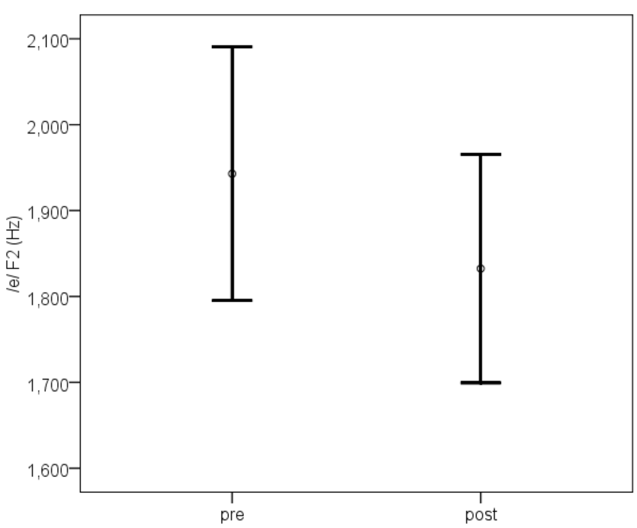

Figure 7.

The changes in the mean of F_2 of /e/ during sustained vowels following the instruction of the Vowel Equalization technique.

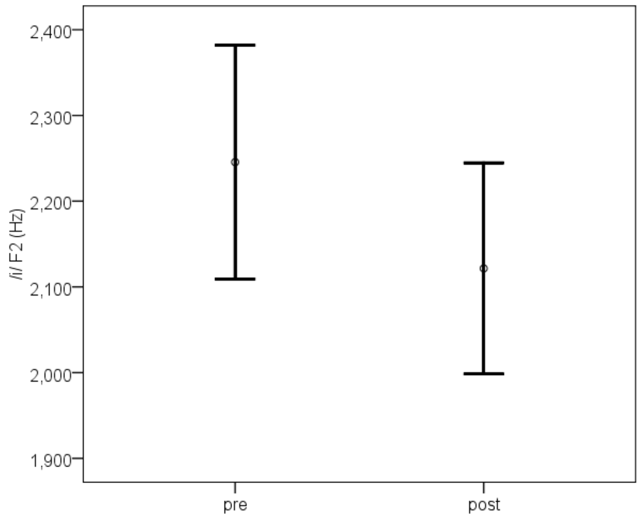

Figure 8.

The changes in the mean of F_2 of /i/ during sustained vowels following the instruction of the Vowel Equalization technique.

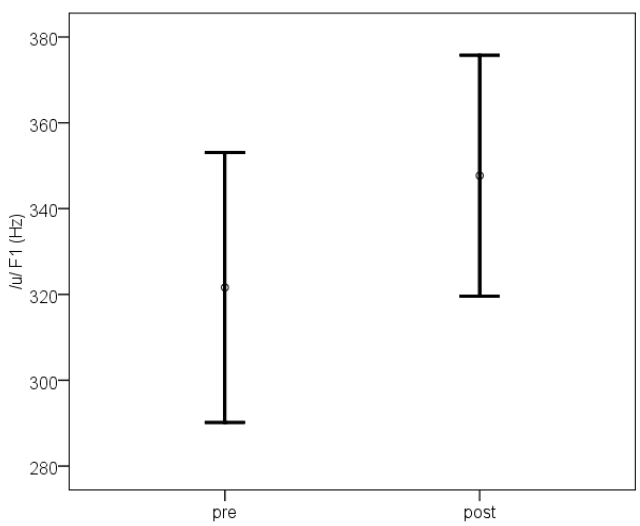

Figure 9.

The changes in the mean of F_1 of /u/ during sustained vowels following the instruction of the Vowel Equalization technique.

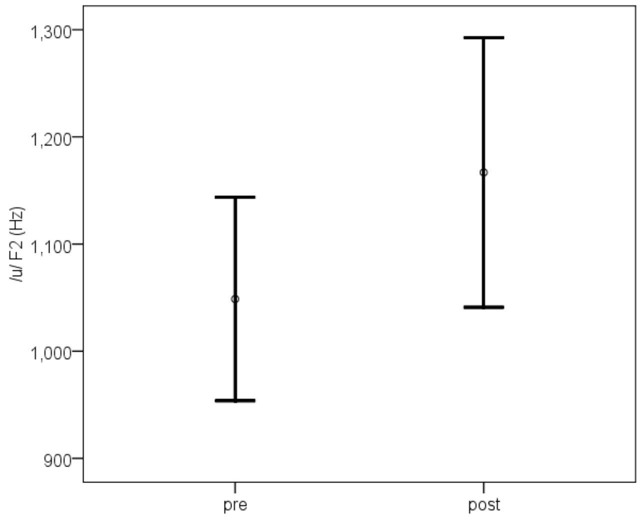

Figure 10.

The changes in the mean of F_2 of /u/ during sustained vowels following the instruction of the Vowel Equalization technique.

Discussion

Since singing, like speaking, requires a constant shifting in the configuration of the articulators, the formant frequencies also change continually (Appelman, 1967; Luchsinger & Arnold, 1965). The vowel equalization technique modifies the behavior of the vocal tract and thus its resonating cavities, as the tongue is placed in a more neutral position to balance two opposite vowels. At the same time the jaw remains stable in its position so that movements of the tongue are responsible for changing the vowels that are sung. The results of the present study are consistent with previous accounts that have reported the association between tongue movements and changes in the first and second formants (Peterson & Barney, 1952). The first formant (F_1) of a vowel is linked to tongue and jaw height.

F_1 increases as the jaw and tongue are lowered, since this movement constricts the pharyngeal cavity where F_1 primarily resonates; this smaller space then resonates a higher frequency. Tongue advancement is linked to the second formant (F_2). As the body of the tongue is pushed toward the front of the mouth F_2 will be higher as the length of the oral cavity is shortened. Likewise, as the body of the tongue is pulled towards the back of the mouth the length of the oral cavity is increased and this results in a lower F_2. As discussed in greater detail below, as the singers equalized their vowels, the formants for the most part changed in the anticipated direction.

Vowel equalization, as discussed by Hopkin (1997), relies on adjustments to the *opposite* vowels. One set of opposite vowels has a similar tongue height, but the vowels differ substantially in their tongue advancement. As /i/ was balanced with /u/ the tongue advancement for /i/ decreased; for the same pair the tongue for /u/ was retracted less. As this more neutral vowel placement was approached, the F_2 for /i/ decreased, while F_2 for /u/ increased. The equalization of the other front to back coupled vowels would be expected to result in similar trends. When equalizing /e/ and /o/, F_2 of /e/ should decrease, while F_2 of /o/ would be predicted to increase. The results of the present study showed that in sustained vowels, the F_2 of /e/ did indeed decrease significantly. The F_2 of /o/ increased, but not to an equivalent extent (see Table 3).

The opposing vowels in the second set generally have a similar tongue advancement, while their differences are manifest in tongue height. As /i/ and /e/ are both front vowels, they are balanced by decreasing the tongue height for /i/ and increasing the tongue height for /e/. Therefore, F_1 for /i/ was increased, while F_1 for /e/ was expected to decrease. The data reveal that in the passage F_1 for /e/ increased slightly, but it decreased in the sustained vowel measures. Coarticulatory effects from the sounds preceding and following the production of /e/ may have interfered with the equalization of this vowel. It was expected that these trends would also be found for the posterior opposite vowels /o/ and /u/. Indeed, the F_1 of /u/ did increase significantly, while the F_1 of /o/ also decreased (to a degree that nearly reached statistical significance), as seen in Tables 2 and 4.

Even though a number of the first and second formant values changed significantly following equalization training, the vowels were still perceived correctly as the intended sounds. This suggests that the singers were able to alter their tongue placement to sing the target vowels using a more neutral tongue position with less extreme articulatory postures. While singing, less extreme articulatory excursions can allow quicker articulation shifts as well as a more consistent presentation of vowels. This technique may, therefore, be used while singing passages at a quick tempo.

The present results indicate that F_2 of the sustained vowel /u/ increased significantly following the equalization instruction. This would suggest that the body of a singer's tongue was positioned more forward within the mouth following vowel equalization coaching. In contrast, F_2 for the vowel /u/ from the singing passage decreased significantly from pre- to post-instruction, suggesting that the body of the tongue was more retracted within the mouth after the training. Although this finding represents the opposite trend when compared to sustained vowel production, it is possible that this finding could be attributed to coarticulation. While singing the passage, production of the vowel /u/ was influenced by other sounds neighboring the target sound. As coarticulation occurs, positioning of the tongue to produce vowels is altered by other phonemes either preceding or following the particular vowel. Therefore, the changes in F_2 of /u/ in the passage might have been caused by lingual adjustments for neighboring sounds.

This study only partially addressed the full scope of Hopkin's vowel equalization. While this paper looks closely at *front* and *back* opposite vowels (where tongue advancement is the primary difference), *open* and *closed* opposite vowels (which differ in terms of tongue height) may also change significantly. One limitation of this paper is that both sets of opposite vowels were not explored equally.

As suggested by Hopkin (1997), vowel equalization offers a singer a means of developing *chiaroscuro* of her tones. The brightness of the forward vowels blends with the darkness of the back vowels, producing an artistic mixture of the two. Hopkin also noted the value of incorporating vowel equalization into the training of choral groups, allowing the director to obtain more balanced vowels from the individual singers in a choir when they perform together.

Vowel equalization may grant a singer a degree of flexibility in maneuvering through passages. A singer will typically produce vowels with a fairly open jaw. Singing vowels with more extreme lingual excursions may put more tension on the laryngeal musculature because of the tongue's attachment to the hyoid bone, from which the larynx is suspended. The equalization technique promotes relaxation of the hyolaryngeal region as front and back vowels are balanced more centrally within the singer's oral cavity. The body of the tongue may thus become looser as it moves less vigorously forward and backward, higher and lower, as the vowels are equalized. A potential benefit, therefore, is that singing may be perceived as less strained and more natural.

This technique may also prove to be useful to singers as they attempt to sing at a higher intensity. The loosening of the hyolaryngeal muscles may allow the singer to increase the airflow through the vocal folds, which results in a less pressed quality when singing more loudly. Because increasing loudness traditionally involves a passive widening of the jaw, as a singer maintains relaxed musculature, the singing may be noticeably different and will sound less pressed or forced. In future research, measuring the tension of the hyolaryngeal musculature before and after the vowel equalization training may provide more information regarding if this technique results in less strain on those muscles while singing. An electroglottograph (EGG) was used with each participant to measure changes in the vocal fold contact area during phonation. Although these data were not incorporated into the present study, they could be analyzed in future research to address quality changes relating to vocal fold adduction.

Additionally, future research could outline the parallels between sung and spoken vowels. Vowel equalization in speech articulation could be valuable in accent reduction. One of the most distinguishing differences between many languages is the variability in lingual postures while producing vowels. Therefore, this may be a helpful technique to individuals learning another language as they attempt to adjust their own articulation to match the patterns of a new language. However, some aspects of Hopkin's method may not be pertinent to speech, as vowels are typically not sustained or emphasized as they will be during singing.

In Hopkin's original article vowel equalization was proposed to equip singers to more readily access a ringing quality in their voices (Hopkin, 1997). This singers' ring, sometimes referred to as the singers' formant, is a quality that enhances a singers' instrument while performing. It has also been observed that this quality can be attained in a speaking voice in such contexts as public speaking and motivational speaking. This technique may allow public speakers a *commanding* tone in much the same way that performing singers may achieve a *ringing* tone. It may be productive to conduct further research with trained listeners to evaluate whether sung and spoken vowels attain the desired *rings*.

In the present study post-training performance was recorded a few minutes after the singers learned to equalize their vowels. In future research it may be valuable to learn whether the changes in the first two formants are maintained over time. Follow-up measurements could be taken weeks or even months later to determine whether there are lasting effects from this technique. Additional research could also investigate the impact of a weekly instruction session to reinforce use of the technique and allow the singers multiple exposures to coaching.

Given the relatively short training period for this technique and the degree of change in both F_1 and F_2, vowel equalization may prove to be a valuable tool for singers of all skill levels. The present study only included 15 participants. Future research could involve a larger sample of singers as well as people with more varied singing experience. The present study only included amateur or recreational singers. It would be valuable to learn whether similar results might be obtained with professional singers or individuals with no prior singing experience.

In order to more accurately identify the nature of the lingual movement changes, articulatory kinematic tools, such as the articulograph, could be used to directly measure tongue movements before and after training. Alternatively, a palatometer could reveal lingual contact patterns within the oral cavity. Future work could also evaluate the impact of biofeedback on the process of vowel equalization. For example, a real-time formant tracking system that plots F_1 against F_2 could provide individual singers the visual feedback necessary to monitor the progress of their equalization while practicing.

Hopkin's vowel equalization technique incorporates the idea of volleying the vibration energy associated with the front of the mouth to the back of the mouth, and from the upper to the lower regions of the oral cavity, depending on where the tongue is positioned, to a more neutral and centralized sensation. Future research could allow for participants of this study to self-report the changes in sensation that they notice, as they describe their experiences in equalizing their vowels.

It would also be informative to involve expert listeners to qualitatively evaluate singers' vowels before and after the vowel equalization technique is taught and adopted. Any qualitative changes in singing performance would be important to understand as singers and their vocal teachers strive to develop a more natural and pleasing sound.

References

Appelman, D. R. (1967). *The science of vocal pedagogy: Theory and application*. Bloomington: Indiana University Press.

Assmann, P. F. (1996). Modeling the perception of concurrent vowels: Role of formant transitions. *Journal of the Acoustical Society of America, 100*(2 Pt 1), 1141-1152.

Assmann, P. F., & Summerfield, A. Q. (1990). Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *Journal of the Acoustical Society of America, 88*(2), 680-697.

Baer, T., Bell-Berti, F., & Rubin, P. (1979). Articulation and voice quality. *Transcripts of the Seventh Symposium: Care of the Professional Voice, Part 1*, 48-53.

Bartholomew, W. (1934). A physical definition of "good voice quality" in the male voice. *Journal of the Acoustical Society of America, 6*(1), 25-33.

Beddor, P. S., & Hawkins, S. (1990). The influence of spectral prominence on perceived vowel quality. *Journal of the Acoustical Society of America, 87*(6), 2684-2704.

Borden, G. J., Harris, K. S., & Raphael, L. J. (2002). *Speech science primer: Physiology, acoustics, and perception of speech*. Baltimore: Lippincott Williams & Wilkins.

Caldwell, R. (Writer) (2002). Human speech source filter theory [Motion Picture]. United States: Caldwell Publishing Company.

Carlsson, G. (1992). Formant frequency tuning in singing. *Journal of Voice, 6*(3), 256.

Detweiler, R. F. (1994). An investigation of the laryngeal system as the resonance source of the singer's formant. *Journal of Voice, 8*(4), 303-313.

Erickson, M. L. (2004). The interaction of formant frequency and pitch in the perception of voice category and jaw opening in female singers. *Journal of Voice, 18*(1), 24-37.

Fant, G. (1961). *Acoustic theory of speech production*. The Hague, Netherlands: Mouton & Co.

Ferrand, C. T. (2006). *Speech science: An integrated approach to theory and clinical practice*. Boston: Allyn and Bacon.

Fletcher, S. G. (1993). Jaw position in vowel production and perception. *Journal of Medical Speech-Language Pathology, 1*(3), 171-189.

Heuer, R. J., Dean, C. M., Sataloff, R. T., & Barrichelo, V. M. O. (2001). Comparison of singer's formant, speaker's ring, and LTA spectrum among classical singers and untrained normal speakers. *Journal of Voice, 15*(3), 344-350.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*(5), 3099-3111.

Hopkin, J. A. (1997). Vowel equalization. *Journal of Singing, 53*(3), 9-14.

Kantner, C. E., & West, R. (1960). *Phonetics*. New York: Harper & Brothers.

Kent, R., & Kim, Y. (2003). Toward an acoustic typology of motor speech disorders. *Clinical Linguistics and Phonetics, 17*, 427-445.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L. et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684-686.

Ladefoged, P. (1967). *Three areas of experimental phonetics: Stress and respiratory activity, the nature of vowel quality, units in the perception and production of speech*. London: Oxford University Press.

Lindblom, B., & Sundberg, J. (1971). Acoustical consequences of lip, tongue, haw and larynx movement. *Journal of the Acoustical Society of America, 50*(4), 1166-1179.

Luchsinger, R., & Arnold, G. E. (1965). *Voice-Speech-Language* (G. Arnold & E. R. Finkbeiner, Trans.). Belmont, CA: Wadsworth Publishing.

Maurer, D., Cook, N., Landis, T., & D'Heureuse, C. (1991). Are measured differences between the formants of men, women and children due to F0 differences? *Journal of the International Phonetic Association, 21*(2), 66-79.

McKinney, J. C. (1982). *The diagnosis & correction of vocal faults*. Nashville, TN.: Broadman Press.

Miller, D. G., Sulter, A. M., Schutte, H. K., & Wolf, R. F. (1997). Comparison of vocal tract formants in singing and nonperiodic phonation. *Journal of Voice, 11*(1), 1-11.

Miller, R. (1977). *National schools of singing: English, French, German, and Italian techniques of singing revisited.* Oxford, England: Scarecrow Press.

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

Miller, R. (1995). What about the "pure vowel" and "vowel modification?" *Journal of Singing, 52*(1), 37-40.

Minifie, F. D. (Ed.). (1973). *Speech acoustics*. Englewood Cliffs, NJ: Prentice Hall.

Nix, J. (2004). Vowel modification revisited. *Journal of Singing*, *61*(2), 173-176.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*(2), 175-184.

Roy, N., Nissen, S., Dromey, C., & Sapir, S. (2009). Articulatory changes in muscle tension dysphonia: Evidence of formant expansion following manual circumlaryngeal therapy. *Journal of Communication Disorders*, 42, 124-135.

Sundberg, J. (1973). The source spectrum in professional singing. *Folia Phoniatrica, 25*, 71-90.

Sundberg, J. (1977). The acoustics of the singing voice. *Scientific American, 236*(3), 82-91.

Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs N.J.: Prentice Hall.

Titze, I. R. (1995). Speaking vowels versus singing vowels. *Journal of Singing, 52*(1), 41-42.

Titze, I. R. (Producer). (2008) Models of vocal fold oscillation. *Tutorials*. Retrieved from http://www.ncvs.org/ncvs/tutorials/voiceprod/tutorial/model.html

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research, 50*, 1510-1545.

Wolf, S. K., Stanley, D., & Sette, W. J. (1935). Quantative studies on the singing voice. *Journal of the Acoustical Society of America, 6*(4), 225-266.

Appendix A

Informed Consent – The measurement of formant alterations in amateur singers after instruction of a vowel equalization technique is taught.

Introduction

You have been invited to participate in a research study about the effect that a vowel equalization technique has on the formant frequencies of the five cardinal vowels (/a/, /ae/, /ei/, /o/, /u/). This study is being conducted by Emily Heaton, a graduate student at Brigham Young University, under the direction of Christopher Dromey, PhD and J. Arden Hopkin, PhD, who are members of the faculty in the Communication Disorders Department and BYU School of Music, respectively. You have been invited to participate because you are enrolled in a non-audition BYU choral group, and have no history of a previous speech or language disorder.

Procedures

You will be asked to attend one recording session lasting approximately one to two hours. Before the recording you will receive a complimentary hearing evaluation, and be asked to fill out a short questionnaire that will be used to develop a demographic profile of the participants in this study. You will then be asked to describe where you feel the sensations of each of the cardinal vowels as you produce them during singing.

After warming up in a manner that is most comfortable to you, and while sitting in a sound booth in the speech research lab, you be asked to sing the first two phrases of “Somewhere Over the Rainbow” at a comfortable loudness level, and at any pitch you prefer. You will be asked to sing each of the five cardinal vowels at any pitch, and sustain the vowel for at least two seconds. You will then be asked to repeat each of these measures twice, so that you have sung the short passage and each vowel a total of three times. These recordings will be analyzed with a computer program.

Following these recordings, you will join a small group of students, and will participate in a group instruction of vowel equalization. This training will be conducted by Dr. Hopkin, a music professor at BYU. After the training, you will be personally coached in producing equalized vowels by Dr. Hopkin. Following this coaching, you will be asked to repeat the recordings that occurred prior to the instruction: three recordings of “Somewhere Over the Rainbow” and three recordings of each of the cardinal vowels sustained for at least two seconds each.

Risks/Discomforts

There are minimal risks associated with participation in this study. The equipment used in this study has been used previously here and elsewhere with no adverse effects. If you have been very vocally active prior to your recording, it is possible that you may feel fatigue. If this occurs, you may rest and if necessary, postpone your recording.

Benefits

Aside from a complimentary hearing evaluation, you will receive no direct benefits from participating in this study. However, the technique taught in connection with this study may prove to be beneficial to you in your singing. The overall results of this study may provide valuable information to music teachers and choral conductors.

Confidentiality
An identification number instead of a name will be used in storing and analyzing the recordings of each singer. Your name and other identifying information will not be used in print or electronic records of this study. Only summary data without reference to names will be reported when the study is complete.

Participation
Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate entirely without any impact on your grades or your status at Brigham Young University, the BYU Music Department, or in any of its courses.

Questions about the Research
If you have any questions about this study, you may contact Dr. Christopher Dromey at (801) 422-6461.

Questions about Your Rights as a Research Participant
If you have questions you do not feel comfortable asking the researcher, you may contact Sandee Muñoz, IRB Administrator, at (801) 422-1461.

Signatures
I understand what is involved in participating in this research study. My questions have been answered and I have been offered a copy of this form for my records. I understand that I may withdraw from participating at any time. I agree to participate in this study.

____________________________________ ______________________
Signature Date

____
Printed Name