IJRCS International Journal of Research in Choral Singing
choralresearch.org

Preferences for Strong or Weak Singer's Formant Resonance in Choral Tone Quality

J. Kevin Ford
The University of Tennessee-Chattanooga

Abstract

The purpose of this study was to assess the preferences of undergraduate college student auditors (N = 139) with respect to a choral sound produced in an anechoic chamber that included a fully resonate singer's formant and choral sound produced in the same chamber using a weaker singer's formant resonance. An ensemble comprised of graduate voice students (N = 8) sang four choral music excerpts. Each excerpt was recorded twice. Singers first employed a full soloistic placement that resulted in a tone with strong upper resonance in the singers' formant range (2 kHz – 4 kHz). The same singers then employed a greatly reduced singer's formant resonance. Recordings were analyzed, then compiled into a stimulus recording for this study.

Auditor participants were randomly assigned to one of two listening groups. One group listened to six pairings each of two of the choral music excerpts for a total of twelve trials. For each excerpt, the order of presentation of the non-resonant version versus the resonant version was presented three times in one order and three times in the reverse order. Each of the six trials was randomized with the six trials from the second choral music excerpt. The same procedure was followed for the second two choral music excerpts. In each trial, subjects were asked to indicate that excerpt performed with the tone quality that they liked the best.

Responses were cross-tabulated according to participants' musical training: (a) college undergraduate choral or vocal music majors with choral training (N = 49), (b) college undergraduate music majors with instrumental training but no choral training (N = 47), and (c) college undergraduates with no music training (N = 43).

Results indicated a significant difference (p < .000) between preferences for non-resonant and resonant tone quality. Examination of the mean scores revealed that auditors overall preferred a non-resonant tone quality (M = 7.95) over a resonant tone quality (M = 4.05). When the three training groups were used as factors, a significant difference of (p < .000) was observed between the choral training group and the instrumental training group. There was also a significant difference (p = .001) between the choral training group and the no music training group. The difference between the instrumental training group and the no music training group was not significant (p = .998).

Such results appeared to suggest that not only did most auditors prefer a non-resonant choral tone to a resonant one, but also that choral training seemed to increase that preference. Although results were limited to the particular conditions, participants, and choral music excerpts of this study, these findings were discussed in terms of possible ramifications for choral pedagogy and directions for future research.

Choral tone quality is a topic of both interest and debate among choir conductors and voice teachers. Whether it is preferable for choir singers to sing with a fully resonant or less resonant tone quality frames a part of this ongoing discussion, particularly with respect to use of those notably strong resonance frequencies within the vocal tract known as the singer's formant. According to Sundberg (1988b), singer's formant resonance is "a prominent spectrum envelope peak appearing in the range of 2 to 4 kHz in all vowel spectra sung by male singers and also by altos. It belongs to the typical features of a sung vowel" (p 14).

In 1863, Hermann L.F. Helmholtz (1821-1894) employed reference to a choral singing context to discuss what would later become known as the singer's formant. Helmholtz (1954), building on a vowel theory based on the length of reed pipes as advanced by Willis in 1830 and the vowel formant theory offered by Wheatstone in 1837, identified the frequency range of what he then termed "tinkling" or "partial tones" as between 2640 to 3168 vibrations per second (Hz). According to Helmhotz:

> In powerful male voices singing forte, these partial tones sound like a clear tinkling of little bells, accompanying the voice, and are most audible in choruses, when the singers shout a little. Every individual male voice at such pitches produces dissonant upper partials…If many voices are sounding together, producing these upper partials with small differences of pitch, the result is a very peculiar kind of tinkling, which is readily recognized a second time when attention has been once drawn to it… I have never heard it from any other musical instrument so clearly as from human voices (p 116).

Monohan's (1978) survey of writings on the teaching of singing between 1777-1927 indicated comparatively few references to resonance prior to Helmholtz' work. Fillebrown (1911), while recognizing proliferating opinions on the matter, stated that systematic research was needed: "the basic importance of resonance in the use of the voice is still too little recognized, though obvious enough in the construction of musical instruments" (p 43). Witherspoon (1925) perhaps summed up the frustration still occasioned by this topic when he said: "probably not even the question of breath has caused more dire confusion and uncertainty, not to speak of faulty emission of voice, than this comparatively new bugaboo, RESONANCE!" (p 21).

With respect to individuals' use of the singer's formant resonance in a choral singing context, the literature of the past five decades is replete with a rich variety of perspectives from choral conductors, choral methods texts, and vocal pedagogy materials. Such writings, however, are largely anecdotal, and may lack the credibility afforded by more systematic, controlled research.

Scientific Studies of Singer's Formant Resonance in Individual Voices

Scientific studies of the solo voice have indicated that full upper resonance in the singer's formant range is one of the components that is preferred for a solo voice quality to be judged as good (e.g., Bartholomew, 1934; Bloothooft & Plomp, 1986; Bloothooft, 1987). Among the early scientific investigations of formants, Paget (1922) explored different resonating cavities in the vocal tract. In that report, he referred to vocal formants, or, more specifically, vowel formants, as "resonance notes' (Paget, 1922). Bartholomew (1934) observed that a large concentration of energy in the frequency region between 2400 and 3200 Hz was characteristic of good voice quality. Albert (1951) concluded that the relative, as opposed to absolute, dimensions of the vocal tract determined formant frequencies. Arment (1960) found that perceived brightness of a vowel depended upon the strength of high partials and narrow formant bands, while perception of vowel darkness depended upon broad formant bands and the lack of high partials. Gunn (1960) indicated that brightness and darkness were directly impacted by the frequency of the first two formants, the intensity of the second and third formats, and the intensity of the harmonics in the range of 2800 cps (Hz).

Rzhevkin (1956) reported detecting singer's formants in the regions from 500 to 2500 cps (Hz). He found upper formants were not as clearly defined in an untrained singer as in a trained singer, and that the highest formant was missing in the untrained voice.

Delattre, Liberman, Cooper and Gerstman (1952) published an article entitled "An experimental study of the acoustic determinants of vowel color; observations on one- and two-formant vowels synthesized from spectrographic patterns." These researchers used a pattern playback machine to play the results of hand-painted spectrographic patterns in an attempt to determine the effect of various acoustical features on the vowel sound produced. They then modified various formants and reported the resulting changes in subjects' perceptions of the vowel and vowel color. Vennard and Irwin (1966) contrasted the acoustical aspects of singing to speech by comparing the sonograms of spoken and sung versions of a secco recitative. Among other findings, a difference in the formant in the vicinity of 3000 Hz was noted. Vennard and Irwin referred to this formant as the *2800*, due to the fact that 2800 Hz seemed to be an average location of the formant. The authors stated that this formant produced the "ring" in the singing voice. Teie (1976) concluded that the development of the singer's formant is related to the amount of vocal training. He also noted that even the untrained voices in his study possessed resonance in the singer's formant range when an [i] vowel was employed.

Magill and Jacobson (1978) examined various vowels at different pitch levels as produced by 22 singers of various voice types, both college students and professional singers. They concluded that the singer's formant may be present in all voice types, but that the intensity of the singer's formant was subject to training.

Schultz-Coulon, Battmer and Riechers (1979a, 1979b) explored the dependency of the relative amplitude of the singer's formant on pitch and intensity. They found that with untrained singers the relative amplitude of the singer's formant appeared to grow as vocal intensity increased and diminish as the pitch rose. This factor was more pronounced with the male voices than the female voices. Trained voices appeared to have more energy in the singer's formant range, though not for all pitches. In both trained and untrained voices, the singer's formant intensity appeared stronger in male voices than in female voices.

Seidner, Schutte, Wendler and Rauhut (1985) found that female singers in their investigation generally had two peaks in the spectrum in the singer's formant range. One occurred between 2500 and 3000 Hz and the other occurred between 3000 and 4000 Hz. They also found that the male singers in that study had a higher relative intensity and narrower bandwidth of the singer's formant

Bloothooft and Plomp (1986) sought to measure the sound level of the singer's formant in professional singers. They found almost no difference in sound level between male singers (4 dB). However, they did find a greater sound difference in female singers (24 dB).

Wang (1985) compared singers with specialties in western operatic singing, Chinese singing, and western early music singing. Singers appeared to evidence the spectrum peak in the range attributed to that of the singer's formant, but did so with varying larynx heights. Wang concluded that a low larynx was not the only method of producing a bright timbre.

Sengupta (1990) examined singer's formant in North Indian classical singing and found that the center frequency of the singer's formant and the resonance balance rose with the frequency of the fundamental. Nawka, Anders, Cebulla and Zurakowski (1997) observed a male "speaker's formant," i.e., a spectral peak near the level associated with the singer's formant.

The Sundberg Studies

Johan Sundberg has researched the singer's formant for over thirty years. The following studies, in particular, have contributed to current understandings of phenomena associated with singer's formant.

Sundberg (1968) investigated the formant frequencies in a bass singer. He suspected that the larynx was the source of the singer's formant in this individual. Sundberg (1972a) explored an articulatory interpretation of the singer's formant. Findings suggested that a lowered larynx, as revealed by tomogram data, increased the laryngeal ventricle, the sinus piriformes, and the cross-sectional area, which produced a resonating chamber that would appear to be the source of the singer's formant. Sundberg also suggested that the singer's formant was produced acoustically by a clustering of the third, fourth, and fifth formant frequencies. Sundberg (1972b) took an average spectrum from an orchestra playing alone and from an orchestra accompanying a singer. There was a pronounced peak in the 2800 Hz range when the singer was singing. This singing formant peak seemed to be the spectral component that enabled the human voice to project through the orchestral sound.

In exploring perceptual aspects of the solo singing voice, Sundberg (1979) noted that the singer's formant and pitch dependent formant frequencies were both responsible for audibility of the voice over loud orchestral sounds. He suggested that vocal economy allowed some resonatory phenomena to occur independently of vocal effort.

Sundberg (1994) investigated two qualities important to solo singing. Vowel quality, suggested Sundberg, was determined by the lower two formants. Vocal quality was determined by the higher frequency components, specifically the center frequency of the singer's formant. Sundberg stated that while singer's format could be found in male voices and alto voices, it appeared not to be present in the soprano voice. He did caution that the studies had used measurement tools that had fixed and narrow band-pass filters and that if the soprano singer's formant was much higher or much broader in frequency it might have gone undetected.

Sundberg (1999) suggested that sopranos used a method of formant tracking to boost the projecting power of their voices in absence of a singer's formant. This tracking occurs when the fundamental frequency approaches and passes that of the lowest formant frequency. The singer then adjusts her mouth opening wider and raises the lowest formant, keeping it in close proximity to the fundamental tone.

Sundberg stated that the lower voice categories of alto, tenor, baritone, and bass did not need to use the formant tracking procedure because they were able to cluster the third, fourth, and fifth formants and produce the singer's formant. According to Sundberg, this procedure of clustering the formants does result in the modification of the vowels but not in the same manner as the formant tracking procedure.

Scientific Studies of Solo Voices in a Choral Situation

Other studies have focused upon individual voices in a choral situation. Harper (1967) recorded individual subjects while they were singing in a choir by using close microphone techniques. Results, from spectrograms and auditor responses, indicated no consistent differences between solo and choral singing for the first two formant frequencies. There were, however, differences in the amount of energy located between formants. Goodwin (1977) found that sopranos instructed to sing vowels in a "blended mode" tended to have lower intensity levels, stronger fundamental frequencies, fewer and weaker upper partials, and stronger first formants than when singing in a "solo mode."

Rossing, Sundberg, and Ternström conducted two studies in 1985 that investigated the solo versus the choral voice. The researchers noted that sopranos sang with more intensity in the singer's formant range and with a slightly greater vibrato extent in solo mode compared to choral mode (Rossing, et al 1985a). They also found that with bass and baritone singers the strength of the singer's formant was greater and the fundamental weaker in solo singing than it was in choral singing. The lower-voice men seem to cluster their third, fourth and fifth formants in order to strengthen the singer's formant (Rossing, et al, 1985b).

Scientific Studies of Choir Acoustics

In his 1989 dissertation, Sten Ternström investigated acoustical features of choir singing with a series of studies. Among Ternström's findings were several items relevant to consideration of singer's formant in choral contexts.

One study investigated the intonation precision and the vowel articulation of singers in a choir. Ternström found that intonation was affected by the aural feedback that singers received from the physical environment, which included the singer's own voice and the voices of the rest of the chorus. He also found that singers from one choir used different vowel articulations in choral singing than they used in speech and, to a lesser extent, in solo singing.

Another study concluded that choirs adapted sound level and vocal technique to the acoustics of the room, and that room acoustics affected the long-time (or long-term) average spectrum (LTAS) of choir sound.

In a third study, Ternström investigated flutter by means of synthesized vocal sound. Flutter is the natural frequency variation that is present in the human voice. He found that without flutter, synthesized voices sounded more like a "mediocre electric organ" than a chorus; however, with the presence of flutter, the synthesized chorus sounded realistic. In discussing vocal tone quality, Ternström made reference to the singer's formant. Ternström stated that this "singer's formant …would defeat choral blend, unless it is used by most or all of the choir members." He also stated that through his research he had only rarely recorded an experienced chorus singer singing with a pronounced singer's formant.

In a separate study, Ternström investigated the formant frequencies of eight bass choir singers in an effort to discover whether or not these singers altered their formant frequencies from those used in normal speech and those used in an attempt to blend with other voices. None of these bass choral singers sang with a pronounced singer's formant. Ternström, however, noted a difference between the vowel formants used for speech and those used for singing.

In brief summary, the literature suggests that the particular resonance phenomenon known as the singer's formant: (a) occurs in vowel spectra of the 2 –4 kHz range among male and alto singers (while sopranos may use to similar effect a method of formant tracking); (b) has its physical source in a resonating chamber produced within the human vocal tract that may or may not be aided by laryngeal height; (c) is developed through practice and training as an efficient and economic means of vocal projection; (d) contributes to the vocal quality of solo voices as having "ring" or "ping;" and (e) does not appear to be ordinarily employed when singing in a choir by either amateur choral singers or those trained singers previously experienced in choral singing.

Despite such indications, voice teachers and choir directors continue to debate the use of singer's formant in choral singing, frequently confusing it with other phenomena and sometimes applying it to larger debates about so-called "straight tone" or "blended" singing. Such controversy often appears to have a life of its own, quite apart from scientific findings. In this sense, the issue has become a matter of some practical import, particularly for choir directors who may be told that they are impeding the singing of voice students asked to sing chorally with less resonance and for voice teachers who fear that their students may be harmed by such practice. Smith (2002), for instance, in an article that frames the debate from the perspective of a voice teacher, contends that audiences prefer a robust choral sound where all singers employ the singer's formant. To date, however, no empirical studies have investigated auditor preference with regard to a conglomerate, choral sound where all singers simultaneously employ a full singer's formant resonance (or, in the case of sopranos, formant tracking). The present investigation is framed to examine such contentions.

Purpose

The purpose of this study is to assess the preferences of undergraduate college student auditors (N = 139) with respect to a choral sound produced in an anechoic chamber that includes a fully resonate singer's formant and choral sound produced in the same chamber using a weaker singer's formant resonance. The following research questions guide this investigation:

1. Given choral excerpts recorded in an anechoic chamber, do undergraduate college students prefer choral tone quality that possesses a fully resonate tone complete with a pronounced singer's formant or a choral tone quality that is produced with the same voices using a much weaker singer's formant resonance?

2. Is there a difference in overall preference between (a) college music majors with experience in choral music, (b) college music majors with an instrumental background and no choral experience, and (c) college students with no experience in instrumental or choral music for an anechoic choral tone quality that possesses a fully resonate tone complete with a pronounced singer's formant to an anechoic choral tone quality that is produced with the same voices using a much weaker singer's formant resonance?

3. What is the self-reported degree of intensity of those preferences indicated?

Limitations of the Study

Choral sound phenomena are by nature rather complex. They involve not only particular groups of singers and particular music compositions, but also the interaction of these variables with particular acoustic venues and the perceptual capacities of particular audiences. In order to focus on the primary interest of this investigation, i.e., the effect of a pronounced, conglomerate singer's formant resonance on the preferences of auditors, there was a need to isolate and control some of the complexity ordinarily associated with choral sound phenomena. This need was all the more apparent given (a) a relative lack of previous scientific research investigating conglomerate singer's formant resonance; and (b) the difficulty of securing singers who could sing choral literature with a demonstrated ability to switch between singing with a pronounced singer's formant and singing with a weaker singer's formant, without substantially altering the dynamic level or the vibrato speed or amplitude between conditions.

Therefore, this study should be approached as an initial foray into this particular line of investigation. The following research decisions and the limitations they impose should be noted:

1. This study focuses on choral sound produced by eight adult singers singing four-part choral music (two singers per part) in an anechoic chamber. The decision to use an anechoic venue was prompted by the desire to investigate, insofar as possible, the effect on listener preference of pronounced and conglomerate singer's formant resonance per se, i.e., as a resonance produced within and by the human vocal tract, without the confounding variables potentially introduced by the interactions of this resonance with more reverberant venues. Eight singers (two singers per voice part) is the least number of singers that could be used to have a choir, and this array approaches the maximum number of singers that could be accommodated in the particular anechoic chamber used for this study.

The difficulty in controlling for all possible variables in the act of singing has led some researchers to use synthesized voices in their experiments. However, the quality of the sound, especially when the vocal production involves text, becomes an issue with artificial voices. To answer the research questions posed by this study, it was deemed preferable to employ human voices evaluated primarily by human auditors.

2. This study used four particular choral music excerpts selected from the Renaissance and Romantic periods. No attempt is made to suggest that the results of this study would be the same were other compositions employed.

3. All auditor participants in this study were selected from the undergraduate study body at the same comprehensive research university, according to their prior choral music experience or lack thereof. There is no suggestion that the preferences of these

particular auditors would necessarily be the preferences of any other group of auditors.

Future studies in this area may well employ more reverberant venues, a greater variety of choral music compositions, more singers, and a greater range of auditor ages and experiences. This particular inquiry, however, is not concerned with such possible variables. It is, rather, one way of beginning to investigate preferences of undergraduate student auditors regarding the use of conglomerate singer's formant. This focus, to date, has not been systematically addressed by the research literature.

METHODS AND PROCEDURES

Stimulus Recording

Singers. Eight university graduate students (2 soprano, 2 alto, 2 tenor, 2 bass) comprised the choral ensemble for this study. Singers were selected based on their ability to sing both with full singer's formant resonance and with reduced singer's formant resonance without greatly altering the vibrato or the dynamic level of their voices. Following task instruction, prospective singers sang into a Radio Shack PRO-3010 dynamic microphone connected to a personal computer, which was running the software package *Spectra Plus*™ version 2.32.01 by Sound Technology Inc (1998). *Spectra Plus*™ is a Fast Fourier Transform (FFT) spectral analysis system.[1] The software was running in real time mode displaying a spectrographic representation of the singer's vocal tone on the monitor. Each singer was able to see the prominence of his or her singer's formant and to experiment with altering the energy level in the corresponding frequency range (2 kHz-4 kHz). After further instruction and experimentation, singers performed an exercise singing once with full singer's formant resonance and then again while attempting to greatly reduce the singer's formant resonance. Those singers who were most successful with resonance change without significantly altering dynamics or vibrato speed were selected for the recording.

Music excerpts. The ensemble recorded eight music excerpts for this study, four of which were ultimately used. The first four examples were excerpted from psalm settings found in *The Hymnal 1982* (Butcher, 1985). These settings were basic four voice chord progressions, seven measures long, sung without text on a single "ah" vowel. The remaining four examples were selected from "real music," sung with the original text. Compositions selected included: (a) *If Ye Love Me* by Thomas Tallis (1505-1585); (b) *Adoramus Te* by Giovanni Gastoldi (1550-1609), commonly attributed to Giovanni Palestrina; (c) *Waldesnacht* (Op. 62 No. 3) by Johannes Brahms (1833-1897); and (d) *Locus iste* by Anton Bruckner (1824-1896). These pieces were selected because each had an opening phrase that was almost completely homophonic. Each of the music examples used for the study consisted of roughly 8s to 14s of the opening phrase of each piece performed at the same tempo. Following the first pilot study, the first four pieces with no text were judged as redundant and were discarded. The four pieces of "real music" were used for the main study.

Recording process. Recordings were made in an anechoic chamber at a major research university. A Brüel and Kjaer type 2669 microphone was suspended from the ceiling in the center of the chamber and the eight singers stood in a circle an equal distance (0.75 m) from the microphone. However, during the recording process it became evident that the tenors were louder than the other sections of the chorus so they were moved to a distance of 0.92 m from the microphone. This process of balancing the voices was viewed as less vocally inhibiting than asking the tenors to alter their vocal production in order to balance with the other voices. The recording was made on a Sony digital audio tape recorder, model DTC-75ES.

Each musical excerpt was recorded twice, once with a fully resonant tone in the singer's formant range and once with a tone with much less resonance. In addition to their prior training experience with the *Spectra Plus*™ software, singers used a hand-held Singer's Formant Analyzer,[2] which was monitored during the recording in order to further assure that there was a difference in the level of the singer's formant. The hand-held unit used a relative scale to indicate the intensity of sounds in the frequency range of the singer's formant.

Following the recording process, all examples from the digital audio tape were converted to monaural wave files at a sampling rate of 44,100 Hz at 16 bits per sample. They were then analyzed using the *Spectra Plus*™ Fast Fouier Transform (FFT) software package, running on a personal computer. The FFT size was set to 8192 points and

[1] The FFT is a spectral analysis system that uses a mathematical algorithm to convert a sound signal from time domain (amplitude versus time) into the frequency domain (amplitude versus frequency).

[2] The singer's formant analyzer was designed and built by Dr. Allen Goodwin of Advantage Showare, Inc. in Ringgold Georgia.

the decimation ratio was set at one, which produced spectral lines of 5.38 Hz. For additional clarity, the Blackman smoothing window was employed. These wave files were then used to make the stimulus recording for the study.

The composite choral sound of the four music excerpts was analyzed using the *Spectra Plus*™ software to investigate any difference of energy in the singer's formant range. The total root mean square (RMS) power output for both versions of the four music excerpts used is contained in Table 1.

Table 1. *Total RMS Power for Each Music Example and Differences Between Versions*)

Example	*Total RMS Power*	*Difference Between Resonant and Non-Resonant Version*
Tallis Resonant	-18.69 dB	-4.08 dB
Tallis Non-Resonant	-22.77 dB	
Gastoldi Resonant	-17.77 dB	-4.31 dB
Gastoldi Non-Resonant	-22.08 dB	
Brahms Resonant	-20.80 dB	-4.83 dB
Brahms Non-Resonant	-25.63 dB	
Bruckner Resonant	-18.71 dB	-4.84 dB
Bruckner Non-Resonant	-23.55 dB	

Further, average spectrum data were computed for each version (resonant/nonresonant) of each excerpt in intervals of 5.30 Hz from 10.80 Hz through 5076.50 Hz. While the *Spectra Plus*™ software program does not separate voices for individual comparisons of vocal spectra, it may indicate if a strong sound component or energy level exits in the singer's formant range of the total sound.

Analyses of the wave files indicated a difference in the relative strength of the singer's formant in each of the recordings (See, for example, Figures 1 and 2). The *Spectra Plus* ™ software package was used to compute average spectra for each of each of the wave files and to view spectrographic data on the wave file. The average spectrum plotted the average strength in decibels for specific frequencies over time throughout, in this case, the entire musical example (See, for example, Figure 3).

The recorded music examples were arranged into matched pairs of the same example with a resonant tone and without a resonant tone using *Cool Edit 96* software by Syntrillium Software Corporation (Johnston, 1996). The matched pairs were measured for total power using the *Spectra Plus*™ software. Initially the overall power level of the two examples was matched using a gain control in the *Spectra Plus*™ software that adjusted all aspects of the sound uniformly. In this way, individual spectral components of the sound were not altered in relation to the fundamental frequency.

A panel of six judges, consisting of music faculty and doctoral choral music students, rated the performances as being resonant or not resonant. Panel results were examined for interjudge reliability. Judges were not in agreement on which versions contained the strong singer's formant resonance and which did not. Upon examining the results and the comments made by the judges, it appeared that by adjusting the gain on the examples that lacked resonance, those recordings were perceived as louder and more sonorous than the resonant examples.

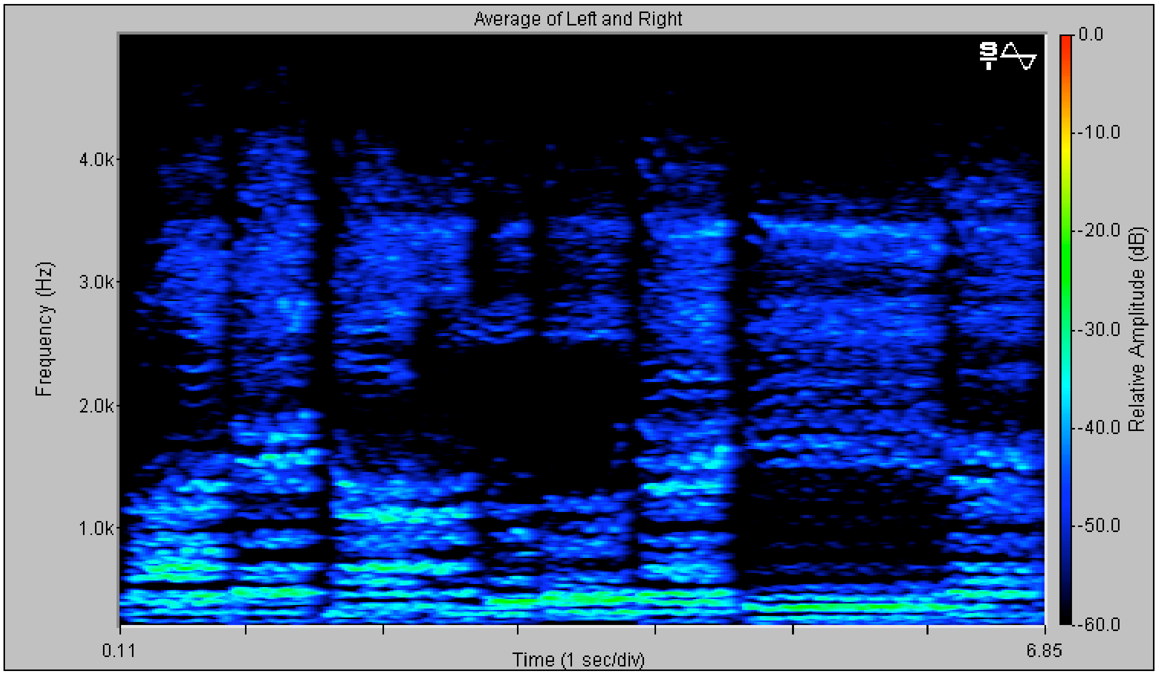

Figure 1. Brahms-resonant-average of left and right channel with spectrum added.

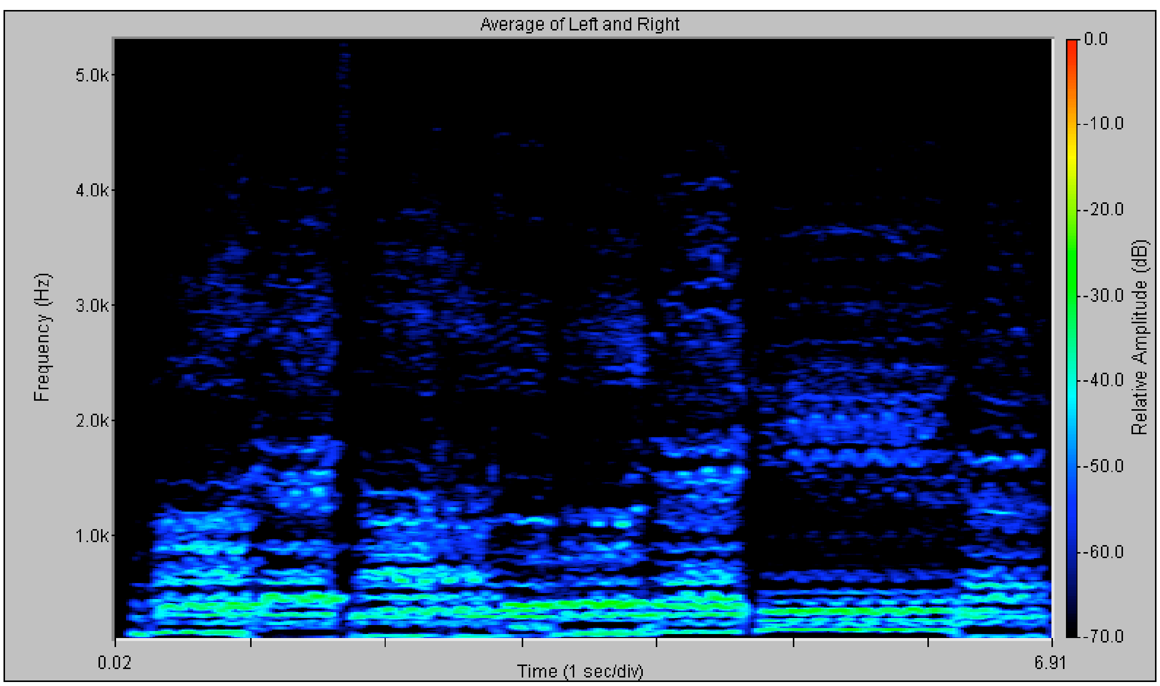

Figure 2. Brahms-non-resonant-average of left and right channel with spectrum added.

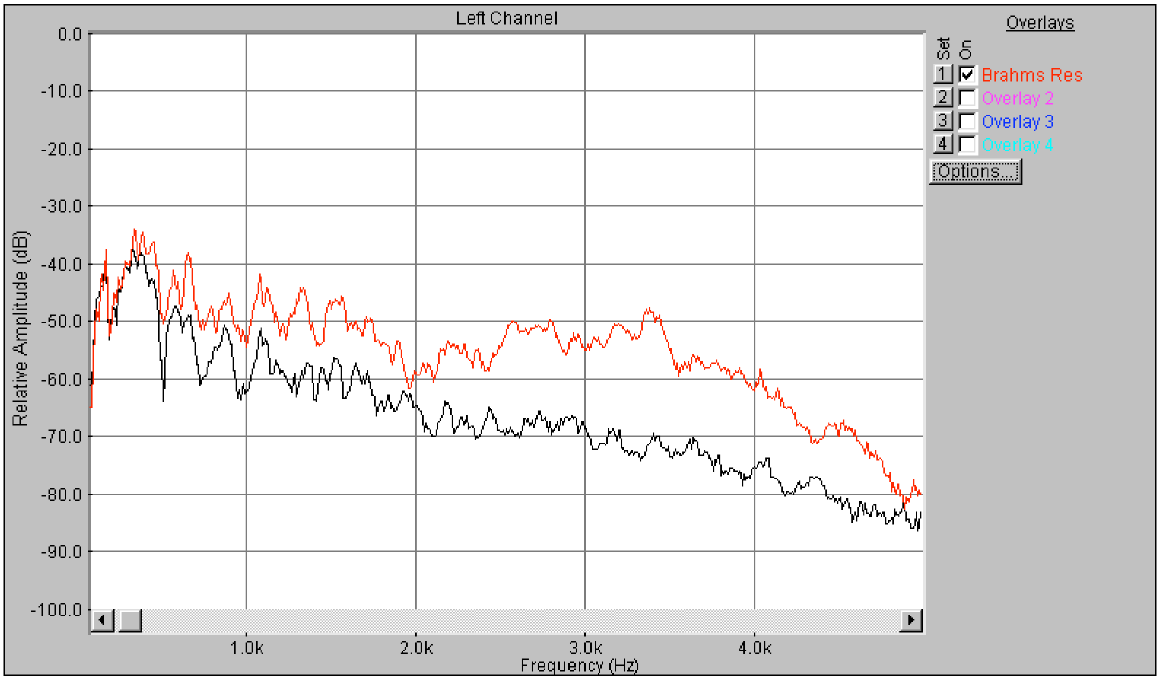

Figure 3. Brahms-average spectrum-black=non-resonant, gray=resonant.

Due to this effect, it seemed prudent to make another tape of unadjusted examples, then have six new judges rate those examples. This process produced a slight perceived volume difference between the two examples, but the judges were instructed to attend to the tone quality differences and to not consider volume in their decision. This time, judges were in 100% agreement on all eight of the examples. The adjusted recordings were discarded and the unadjusted recordings were used for the study. Furthermore, an instruction to disregard the volume difference was incorporated into the questionnaire used by auditor participants.

Pilot Studies

Pilot study one. Auditor participants (N = 38) in pilot study one were divided into two groups: (a) participants with little or no choral training (N = 11); and (b) participants with training in choral music (N = 27). The stimulus recording included eight choral music examples (four chord progressions and four "real music" examples), which were paired in resonant versus non-resonant, or non-resonant versus resonant, music examples of the same piece. Each pairing was used twice for a total of thirty-two possible combinations. Three different random orderings were employed, with participants randomly assigned to listen to one of three orderings.

An auditor questionnaire developed for these matched pair music excerpts included: (a) demographic questions; (b) a forced-choice response grid; and (c) a Likert type scale. Auditor participants (N = 38) in pilot study one were divided into two groups: (a) participants with little or no choral training (N = 11); and (b) participants with training in choral music (N = 27).

Participants were asked to indicate which of the two excerpts in each pair they preferred. They were also asked to indicate the intensity of their preferences by circling a number on a Likert type scale from one to five, with one being low intensity and five being high intensity preference.

Results of pilot study one indicated that the four chord progressions sung to a single vowel were redundant. They were therefore removed from future versions of the study. Responses indicated an overall difference between the choral training group and the non-choral training group.

However, participants appeared to change their minds considerably about preference for the resonant and non-resonant examples as the study progressed. Data were examined for potential biases due to fatigue and order effects. No apparent patterns emerged. Nevertheless, it was decided to limit the study so that each subject would be responding to either the two pieces from the Renaissance or the two pieces from the Romantic period. Repetitions were increased to three of each presentation order, which resulted in six paired examples of each piece. The questionnaire was rewritten and new stimulus presentations were made for a second pilot study.

Pilot study two. Auditor participants (*N* = 12) for a second pilot study were divided into two groups as in the first pilot study: participants with training in choral music (*n* = 10) and participants with no musical training (*n* = 2). The proportion of subjects who changed their minds from either resonant to non-resonant or non-resonant to resonant example was greatly reduced in the second pilot study.

After examining demographic data from both pilot studies, however, there appeared to be a distinct third group. That group consisted of instrumental musicians who had no training in choral music. Therefore, the final study was designed with three groups of auditor participants.

Main Study

Auditor participants. Auditor participants (*n* = 139) were undergraduate students from a comprehensive research university in the southeastern United States. Participants were divided into three groups as follows: (a) undergraduate music majors (*n* = 49) who had training in choral music (choral training), (b) undergraduate music majors (*n* = 47) who had training in instrumental music but very little or no training in choral music (instrumental training), and (c) undergraduate students (*n* = 43) who had no training in music (no music).

Participants in each category were randomly assigned to one of two listening groups. One of these groups listened to paired presentations of the two short choral music excerpts by Tallis and Gastoldi (choral training *n* = 28, instrumental training *n* = 23, no music *n* = 24) and the other group listened to an equivalent presentation of the two choral excerpts by Brahms and Bruckner (choral training *n* = 21, instrumental training *n* = 24, no music *n* = 19). Three different random orderings of each group of music were prepared. Auditor participants were randomly assigned to one of the resulting six random orderings.

Auditor questionnaire. The overall format of the auditor questionnaire for the main study was identical to that used in the two pilot studies. Participants were asked to listen to 12 randomly ordered pairs of choral music excerpts, each from 8 to 12 seconds in length. Following each pair, auditors were requested (a) to indicate the tone quality that they liked the best, and (b) to designate the intensity of that preference.

Auditor directions. Upon completion of the demographic information on the questionnaire, participants were instructed to put on their headphones and, during the recorded instructions, adjust the volume to a comfortable level. The written and recorded participant instructions were as follows:

> **As the following instructions are read to you, adjust the volume on the cord of your headphones to a comfortable level.**
>
> This study deals with a specific element in the tone quality or sound of a chorus. You are about to be asked to listen to 12 short pairs of recorded choral music. Each example is from 8 to 12 seconds long and both musical examples in each pair are identical. Every effort has been made to ensure that the only element that changes is the specific element in the tone quality or sound that we are interested in. However, a slight difference in volume did appear in the recording. Therefore we would like for you to **rule out volume** as an element in your preference decision. For instance, if one example is louder than the other, don't choose it because it is louder but listen for the tone quality or sound sung by the singers and make your decision based on that instead. The first example of each pair will be labeled "A" and the second example will be labeled "B." Your job is to <u>circle</u> the letter that corresponds with the pair member that you feel has the tone quality or the sound that **you like best.** Then, circle a number from 1 to 5 that represents the intensity of your preference. The scale is a five-point scale with 1 representing weak intensity and 5 representing strong intensity. For example, if in one of the pairs, you like example "A" just a little more than example "B" then you would circle "A" and then circle number 1, or 2 on the intensity scale. However, if you like example "A" a lot more than example "B" then you would circle "A" and then circle number 4 or 5. Once again, please do not let the volume or loudness of the example affect your judgment. Also, remember that there are no right or wrong answers, just mark the musical example that is performed with a sound that you like best.

Listening venue and equipment. The listening venue was sound dampened and had the approximate dimensions of 4 m by 5.5 m. Tables, desks, and chairs were arranged with blinders so that participants could not see each other's responses. A senior university acoustical engineer constructed cabling so that a total of 10 Koss gt/4 portable headphones could be connected to a single music source – a Pioneer SA1510 amplifier and a JVC XL-2441 compact disc player. Each headphone had an intensity control on the cord and had a frequency response of 20 to 20,000 Hz. To listen to a sample pair of excerpts (Gastoldi: Non-resonant vs Resonant), click here.

RESULTS

Results were organized according to the study's three research questions. All statistical tests were computed at a predetermined .05 alpha level.

Research Question One

The first research question sought to gauge participant preferences with respect to choral music excerpts sung with and without singer's formant resonance. Results of all four excerpts were combined and analyzed by means of a Paired Samples *t*-test, and according to frequencies, mean scores, and standard deviations. The mean for non-resonant preferences was 7.95, and the mean for resonant preferences was 4.05 (*SD* = 3.42). Standard deviations were identical due to the forced-choice design employed in the study.

Only 28 participants (20%) responded with a preference for non-resonant choral tone in every case, and only 3 participants (2%) responded with a preference for a resonant choral tone in every case. Remaining participants were less consistent in their judgments. As indicated in Figures 4 and 5, however, a strong trend toward a non-resonant preference was observable. Each figure reports frequencies on the six matched pairs with similar order of presentation of resonant versus non-resonant. For example, 4N - 2R would report the frequency of subjects who preferred non-resonant four out of the six possible times.

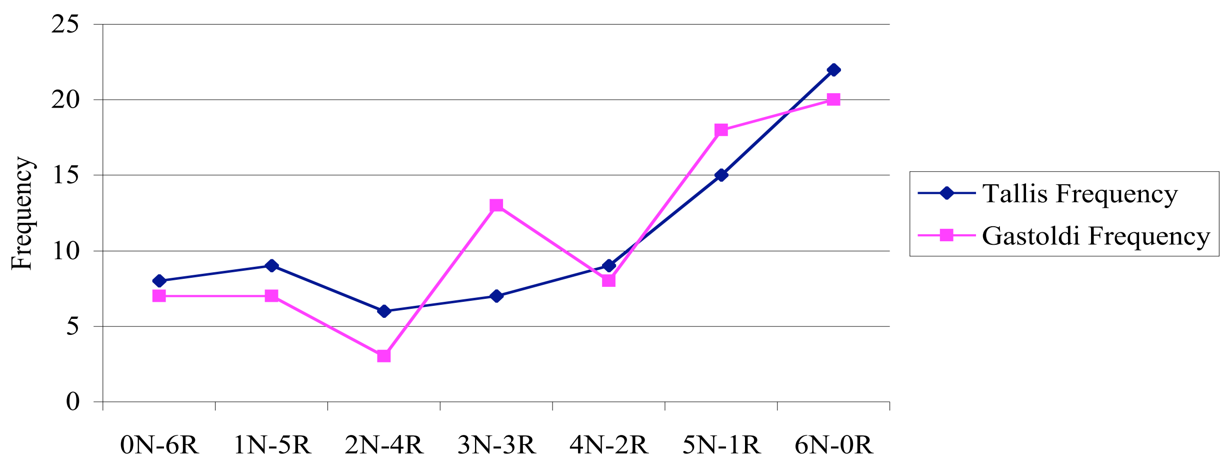

Figure 4. Frequency of resonant versus non-resonant responses to Renaissance excerpts.

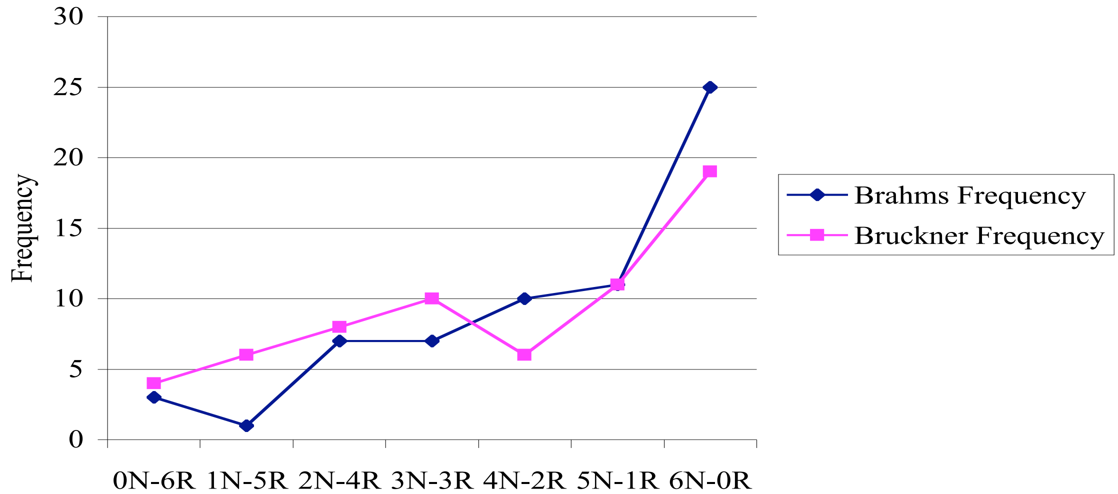

Figure 5. Frequency of resonant versus non-resonant responses to Romantic excerpts

Finally, a Paired Samples *t*-test indicated that the differences between the two preference choices were significant (two-tailed significance level of $p < .000$) (Table 2). All groups demonstrated a preference for non-resonant tone quality.

Research Question Two

The second research question asked if participants in the three training groups differed significantly in their preference for a resonant or a non-resonant choral tone quality across all four of the music examples. A One-Way Analysis of Variance test was used to answer this question. The test revealed that a significant difference (F = 11.101) existed between the three training groups. The F-ratios are identical in both the non-resonant and the resonant responses because of the forced-choice procedure (Table 3).

A Scheffé post hoc test revealed a significant difference ($p < .000$) between the group with choral training and the group with instrumental training but no choral training (Table 4). There was also a significant difference (p = .001) between the group with choral training and the group with no musical training. The difference between the group with instrumental training and the group with no musical training was not significant (p = .998).

Table 2

T-Test Results for All Subject's Preference for a Non-resonant or a Resonant Choral Tone Quality

	Paired Differences					
Music	*Mean*	*SD*	*SD error mean*	*t*	*df*	*Sig. 2-tailed*
Non-resonant – Resonant	3.90	6.84	.58	6.720	138	.000

Table 3

One-Way ANOVA for Training versus Resonance Preference

Factor	*Source of Variation*	*Sum of Squares*	*Mean Squares*	*df*	*F*	*Sig.*
Non-resonant	Between Groups	219.638	109.819	2	11.101	.000
	Within Groups	1395.009	10.257	136		
Resonant	Between Groups	219.638	109.819	2	11.101	.000
	Within Groups	1395.009	10.257	136		

Table 4

Multiple Comparisons for the Three Training Groups, Non-resonant Preferences (Scheffé Test)

Dependent Variable	*Training (I)*	*Training (J)*	*Mean Difference (I-J)*	*Standard Deviation Error*	*Sig.*
Non-Resonant	Choral Training	Instrumental Training / No Choral	2.653 [a]	.654	.000
		No Music Training	2.607 [a]	.669	.001
	Instrumental Training / No Choral	Choral Training	-2.653 [a]	.654	.000
		No Music Training	-.047	.676	.998
	No Music Training	Choral Training	-2.607 [a]	.669	.001
		Instrumental Training / No Choral	.047	.676	.998

Note. [a] The mean difference is significant at the .05 level.

Table 5

Multiple Comparisons for the Three Training Groups, Resonant Preferences (Scheffé Test)

Dependent Variable	*Training (I)*	*Training (J)*	*Mean Difference (I-J)*	*Standard Deviation Error*	*Sig.*
Resonant	Choral Training	Instrumental Training / No Choral	-2.653 [a]	.654	.000
		No Music Training	-2.607 [a]	.669	.001
	Instrumental Training / No Choral	Choral Training	2.653 [a]	.654	.000
		No Music Training	.047	.676	.998
	No Music Training	Choral Training	2.607 [a]	.669	.001
		Instrumental Training / No Choral	-.047	.676	.998

Note. [a] The mean difference is significant at the .05 level.

A Scheffé post hoc test for training versus a resonant preference yielded the same results due to the forced-choice design (Table 5). The mean scores for the non-resonant and resonant responses indicated that all three training groups preferred the non-resonant version to the resonant version. However, the group with choral training had a significantly higher mean score (9.653) for a non-resonant preference when compared to both the instrumental group (7.000) and the group with no musical training (7.047). The standard deviation also suggested a stronger consensus of opinion for the choral training group (2.521) than that of the instrumental training group (3.394) or the group with no musical training (3.651) (Table 6).

Research Question Three

The third research question asked about the intensity of preferences among auditor participants. The mean intensity rating for all participants was 3.059 (*SD* = 0.78) out of the possible five points.. Only a slight difference was noted in the mean scores of each of the three training groups. The choral training group had a mean score of 3.362 (*SD* = 0.755), the instrumental training group had a mean score of 2.88 (*SD* = 0.787), and the no music training group had a mean score of 2.842 (*SD* = 0.8).

Evaluation of Order Effect

A separate Paired Samples *t*-test was used to investigate order effect by comparing the three presentations of each ordering of each piece. For example, the three orderings of the Tallis piece that had the non-resonant version first were compared with the three orderings that had the resonant version first. No significant order effects were observed in the first three choral music excerpts. The Tallis piece produced a two-tailed significance of $p = .390$, the Gastoldi piece produced a two-tailed significance of $p = .260$, and the Brahms piece produced a two-tailed significance of $p = .717$. However, the Bruckner piece revealed a two-tailed significance of $p < .000$ (Table 7). This level of significance would normally suggest an order effect for this piece.

Table 6

Descriptive Statistics for Three Training Groups Across Resonant versus Non-Resonance Preferences

Resonance Factor	*Training*	*N*	*Mean*	*sd*	*sd error*	*Minimum*	*Maximum*
Non-resonant	Choral Training	49	9.653	2.521	.360	0	12
	Instrumental Training / No Choral	47	7.000	3.394	.495	0	12
	No Music Training	43	7.047	3.651	.557	1	12
	Total	139	7.950	3.421	.290	0	12
Resonant	Choral Training	49	2.347	2.521	.360	0	12
	Instrumental Training / No Choral	47	5.000	3.394	.495	0	12
	No Music Training	43	4.954	3.651	.557	0	11
	Total	139	4.050	3.421	.290	0	12

Table 7

T-Test Results for Order Effect

	Paired Differences					
Music	*Mean*	*sd*	*sd error mean*	*t*	*df*	*Sig. 2-tailed*
Tallis NR/R Renaissance Group	.08	.80	.09	.865	75	.390
Gastoldi NR/R Renaissance Group	-.13	1.01	.12	-1.134	75	.260
Brahms NR/R Romantic Group	-.05	1.02	.13	-.364	64	.717
Bruckner NR/R Romantic Group	-.58	.96	.12	-4.837	63	.000

DISCUSSION

The main finding of this study is that its particular auditor participants, including those with (a) prior choral music training, (b) instrumental music training but no choral training, and (c) no prior training in music, appear to prefer, among the choral sound stimuli created for this investigation, that sound produced with less resonant tone quality, i.e. group singing accomplished without a pronounced singer's formant resonance. The intensity of such preference overall appears to be modest to moderate (with means ranging from 2.84 to 3.36 on a 5 point Likert scale). Nonetheless, it is a clearly expressed, consistent, and statistically significant preference among these participants as a whole. A concomitant finding is that such preference appears most prominent, in terms of both consistency and degree, among those auditors with prior choral music training.

When the paired-samples *t*-test was used on the data, a difference was observed in the Bruckner music excerpt but not in the other three excerpts. The fact that a difference was not observed in three out of the four music excerpts may suggest that this phenomenon is not due to an order effect. The Bruckner excerpt was the only one of the four that had an instance where one vocal section (basses) was heard alone. The two exposed bass notes in the Bruckner excerpt could have contributed to the difference in the test for order effect.

When given an opportunity to comment on the study generally, three auditors commented on the bass section. One noted "Sometimes the lower voices were blaring which affected my decision," while another commented that "The basses seemed to stand out in a way that I disliked." A third participant apparently had a positive reaction to the bass section and stated that he or she "like[d the] powerful male voices." Because the difference in the singer's formant was pronounced in the bass voices, as measured in pre-recording sessions with the *Spectra Plus*™ software, this difference could have contributed to the discrepancy with the Bruckner excerpt.

Care should always be taken in interpreting preference when human voices are used. It is impossible to completely control for all aspects of vocal production when attempting to alter one specific aspect. Even with instructions to the highly skilled singers to alter only their resonance, other changes are inevitable. For instance, while vibrato was present in all of the recordings, it was impossible to keep the vibrato speed and amplitude exactly the same. Also, in some cases increasing the resonance altered the balance and blend in the voices. Some of the imbalance was counteracted by positioning the singers at different distances from the microphone for the resonant recordings. However, due to the limited size of the anechoic chamber, this was not always completely sufficient.

The act of increasing the resonance in the singer's formant range added <u>volume</u> to the sound. Although the sound level at the fundamental frequencies was relatively equivalent the increased energy in the upper frequencies, including the singer's formant, produced a louder perceived volume. The subjects were instructed to disregard the volume of sound as a factor in their decisions and based on the comments that the subjects made, most were able to do this. However, there were seven subjects that admitted this posed a difficulty for them.

The act of singing is an extremely complex process. When all of the aspects of vocal production are combined with the correct pronunciation of text and with the aspect of singing in an ensemble, the amount of variables is enormous. The difficulty in controlling all of the variables in singing has led some researchers to use synthesized voices in their experiments (Berndtsson, 1995; Karlsson, 1992; Sundberg, 1988; Ternström, 1989). However, the natural quality of the sound, especially when the vocal production involves text, becomes an issue with artificial voices. This researcher felt that to answer the research questions involved with this study, it was necessary to have human subjects evaluating human voices.

Researchers have found that a choral singer who is also an experienced solo singer will lower the energy level in the singer's formant range when switching from a solo singing mode to a choral singing mode (e.g., Goodwin, 1977; Rossing, 1985; Ternström, 1989). That is, human singers appear to have an ability to switch intuitively between methods of vocal production in various contexts. This study dealt with only one aspect of choral tone quality - the amount of resonance in the singer's formant range. Results seem to indicate that although training and experience in choral music appears to increase auditor preference for a choral tone with less resonance, training may not be the sole determining factor in this preference decision. A general preference for a choral tone quality with less resonance compared to one with strong resonance is present to varying degrees in most of the present auditor sample.

As spectrographic software becomes increasingly more sophisticated, it may have enormous potential to aid investigation of numerous phenomena associated with choral singing. A problem with using such analyses at present, of course, is that choral sound involves numerous, simultaneous sound sources. It is difficult to say with certainty exactly what spectrograms of conglomerate, choral sound are picking up, though it is possible to compare and contrast the data yielded by spectrographic analysis among controlled conditions.

This study is limited by its research conditions, its number of singers, its particular array of auditor participants, and its choice of choral music excerpts. Future research in this area may well consider use of more reverberant venues, a greater number of singers, different arrays of auditors, and a greater variety of choral literature.

The present investigation, however, may suggest at least two concepts worthy of consideration in ongoing discussions among choral directors and voice teachers. First, singing chorally with a less resonant tone quality than fully resonate solo singing may be preferred by auditors, more because of the context of choral singing itself than some universal preference for one correct or even one desirable mode of vocal production. Secondly, the ongoing debate about whether to allow singer's formant in choral singing may be, in certain respects, a misplaced debate. Human singers seem to be capable of employing some variety in their vocal production according to the context in which they sing. The matter of preference for singer's formant resonance, therefore, need not necessarily be an either/or question. In terms of pedagogy, both choir directors and voice teachers, as well as their students, may benefit from seeking greater understanding of how voices can work efficiently and in a healthy manner in a variety of singing contexts.

REFERENCES

(1998). Spectra Plus (Version 2.32.01) [software]. Campbell, CA: Sound Technology Inc.

Albert, V. F. (1951). Theoretical and experimental analysis of networks of resonators and their application to the production of sound by the vocal tract. *Dissertation Abstracts International, 22*, (01), 287. (University Microfilms No. AAG6004429)

Arment, H. E. (1960). A study by means of spectrographic analysis of the brightness and darkness qualities of vowel tones in women's voices. *Dissertation Abstracts International, 21,* (03), 638. (University Microfilms No. AAG6002989)

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

Berndtsson, G. (1995). Systems for synthesizing singing and for enhancing the acoustics of music rooms: two aspects of shaping musical sounds (reverberation). *Dissertation Abstracts International, 58*, (01C), 228. (University Microfilm No. AAGC522486)

Bloothooft, G., & Plomp, R. (1986). The sound level of the singer's formant in professional singing. *Journal of the Acoustical Society of America, 79*, 2028-2033.

Bloothooft, G. (1987). A note on voice quality in professional singing. *Journal of Research in Singing, 10*(2), 3-16.

Butcher, G., et. al. (Ed.). (1985). *The hymnal 1982*. New York: Publishing Incorporated.

Delattre, P. C., Liberman, A. M., Cooper, F. S., & Gerstman, L. J. (1952). An experimental study of the acoustic determinants of vowel color; observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word* 8, 195-210.

Fillebrown, T. (1911). *Resonance in singing and speaking*. Boston: Oliver Ditson Company.

Goodwin, A. W. (1977). An acoustical study of individual voices in choral blend. (Doctoral dissertation, North Texas State University). *Dissertation Abstracts International, 38*, (12A), 7201. (University Microfilms No. AAG7807833)

Gunn, J., George Henry. (1960). An acoustical analysis of Quality Variations in Sung Vowels. *Dissertation Abstracts International, 21,* (12), 3885. (University Microfilms No. AAG6006874)

Harper, A. H., Jr. (1967). Spectrographic comparison of certain vowels to ascertain differences between solo and choral singing, reinforced by aural comparison. (Doctoral dissertation, Indiana University, 1967). *Dissertation Abstracts International, 28*, (12A), 3067.

Helmholtz, H. (1954). *On the sensations of tone as a physiological basis for the theory of music.* (Alexander J. Ellis, Trans.). New York: Dover.

Johnson, A. R. (1973). The Christiansen choral tradition: F. Melius Christiansen, Olaf C. Christiansen, and Paul J. Christiansen. *Dissertation Abstracts International, 34*, (09A), 6023. (University Microfilms No. AAG7407389)

Johnston, D. (1996). Cool Edit 96 (Version 96). Phoenix: Syntrillium Software Corporation.

Karlsson, I. (1992). Evaluations of acoustic differences between male and females voices: a pilot study. *Speech Transmission Laboratory Quarterly Progress and Status Report, 1*, 19-31.

Magill, P. C., & Jacobson, Loren. (1978). A comparison of the Singing Formant in the Voices of Professional and Student Singers. *Journal of Research in Music Education, 26*(4), 456-469.

Monahan, B. J. (1978). *The art of singing: A compendium of thoughts on Singing published between 1777 and 1927*. Metuchen, N. J.: The Scarecrow Press, Inc.

Nawka, T., Anders, L. C., Cebulla, M., & Zurakowski, D. (1997). The speaker's formant in male voices. *Journal of Voice, 11*(4), 422-428.

Paget, R. S. (1922). Vowel resonances. In D. B. Fry (Ed.), *Acoustic phonetics: A course of basic readings* (pp. 95-103). Cambridge: Cambridge University Press.

Rossing, T., Sundberg, J., Ternström, S. (1985a). Acoustic comparison of soprano solo and choir singing. *Speech Transmission Laboratory Quarterly Progress and Status Report, 4*, 43-58.

Rossing, T., Sundberg, J., Ternström, S. (1985b). Voice timbre in solo and choir singing: is there a difference? *Journal of Research in Singing, 8* (2), 1-7.

Rzhevkin, S. N. (1956). Certain results of the analysis of a singer's voice. *Soviet Physics-Acoustics, 2*, 215-220.

Seidner, W., Schutte, H. K., Wendler, J., Rauhut, A. (1985). *Dependence of the high singing formant on pitch and vowel in different voice types.* Paper presented at the Stockholm Music Acoustics Conference, Stockholm: Royal Swedish Academy of Music.

Sengupta, R. (1990). Study on some aspects of the "Singer's Formant" in North Indian Classical Singing. *Journal of Voice, 4*(2), 129-134.

Smith, P (2002). Balance or blend? Two approaches to choral singing. *Choral Journal, 43* (5), 31-43.

Sundberg, J. (1970a). Formant structure and articulation of spoken and sung vowels. *Folia Phoniatrica, 22*, 28-48.

Sundberg, J. (1970b). The level of the 'singing formant' and the source spectra of professional bass singers. *Speech Transmission Laboratory Quarterly Progress and Status Report, 4*, 21-39.

Sundberg, J. (1972a). An articulatory interpretation of the formant'. *Speech Transmission Laboratory Quarterly Progress and Status Report, 1*, 45-53.

Sundberg, J. (1972b). A perceptual function of the 'singing formant'. *Speech Transmission Laboratory Quarterly Progress and Status Report, 2-3*, 61-63.

Sundberg, J. (1979). Perception of singing. *Speech Transmission Laboratory Quarterly Progress and Status Report, 1*, 1-48.

Sundberg, J. (1988b). Vocal tract resonance In singing. *National Association of Teachers of Singing Journal, 44*(4), 11-31.

Sundberg, J. (1994). Perceptual aspects of singing. *Journal of Voice, 8*(2), 106-122.

Sundberg, J. (1999). The perception of singing. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 793). San Diego: Academic Press.

Teie, E. W. (1976). A comparative study of the development of the third formant in trained and untrained voices. (Doctoral dissertation, University of Minnesota, 1976). *Dissertation Abstracts International, 37,* (10A), 6135.

Ternström, S. (1989a). Acoustical aspects of choir singing. *Dissertation Abstracts International, 51,* (04C), 609. (University Microfilm No. AAGC150108)

Ternström, S., Sundberg, J. (1989b). Formant frequencies of choir singers. *Journal of the Acoustical Society of America, 86*(2), 517-522.

Ternström, S. (1989c). Long-time average spectrum characteristics of different choirs in different rooms. *Speech Transmission Laboratory Quarterly Progress and Status Report, 3*, 15-31.

Vennard, W., & Irwin, J. W. (1966). Speech and song compared in sonograms. *The NATS Bulletin, 23*(2), 18-23.

Wang, S. (1985). *Singing voice: bright timbre, singer's formants and larynx positions.* Paper presented at the Stockholm Music Acoustics Conference, Stockholm.

Witherspoon, H. (1925). *Singing: A treatise for teachers and students*. New York: G. Schirmer, Inc.