2006-08-12

Laryngeal-Level Amplitude Modulation in Vibrato

Lorie C. Reese
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd
Part of the Communication Sciences and Disorders Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/767

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
LARYNGEAL-LEVEL AMPLITUDE MODULATION

IN VIBRATO

by

Lorie Reese

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Communication Disorders

Brigham Young University

December 2006
GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Lorie Reese

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

<table>
<thead>
<tr>
<th>Date</th>
<th>Christopher Dromey, Chair</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>J. Arden Hopkin</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Ron W. Channell</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
BRIGHAM YOUNG UNIVERSITY

As chair of the candidate’s graduate committee, I have read the thesis of Lorie Reese in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Christopher Dromey
Chair, Graduate Committee

Accepted for the Department

Ron W. Channell
Graduate Coordinator

Accepted for the College

K. Richard Young
Dean, David O. McKay School of Education
ABSTRACT

LARYNGEAL-LEVEL AMPLITUDE MODULATION

IN VIBRATO

Lorie Reese

Department of Communication Disorders

Master of Science

Research in vocal vibrato has established that vocal tract filtering is primarily responsible for the amplitude modulation (AM) present in Western classical vibrato. Using electroglottography (EGG) and the EGG speed quotient, which is sensitive to fluctuations in the amplitude of vocal fold vibration, AM was detected at the laryngeal (source) level, in addition to the subsequent AM which results from vocal tract filtering. Seventeen classically-trained opera singers sang vowels in three pitch and loudness conditions. EGG and microphone measurements of FM and AM and their rates, extents, and periodicity were made. Airflow was also measured, and the samples were rated by voice professors for vibrato consistency, speed, and width. Physiologic and acoustic data revealed that AM from vocal tract filtering, or the resonance-harmonics interaction (RHI) described by Horii and associates, was present throughout the vibrato samples. Laryngeal-level AM was also present throughout, with soft conditions having the highest
mean extents. Singers with lower degrees of laryngeal-level AM were also those rated highest for vibrato consistency. Vibrato rate increased as pitch increased, and, to a lesser extent, as intensity increased. These findings document, in addition to the AM resulting from the RHI, the concurrent presence of laryngeal-level AM in a group of singers representing a range of training and experience.
ACKNOWLEDGEMENTS

My deepest gratitude goes to my children, Annie, Rhetta, John, and Janie, for their resilience and unconditional love during the past six years. I am indebted to my mother, father, sisters, and brother for their great love. The genuine friendship offered in particular by Ann Dorais and Kathy Pierce as well as old friends and fellow students has cheered and encouraged me.

Thanks go to the opera singers in our study, for their beautiful voices and willingness to participate, and also to Drs. Hopkin and Channell, who served faithfully and helpfully as thesis committee members.

And to Dr. Dromey, lasting thanks for quietly and consistently holding himself to the highest standards, which inspires absolute trust in both his scholarly judgment and personal character.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>List of Appendices</td>
<td>x</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>History of Western Classical Vibrato</td>
<td>2</td>
</tr>
<tr>
<td>Seashore’s Contributions</td>
<td>3</td>
</tr>
<tr>
<td>Acoustic Measurements of Vibrato</td>
<td>4</td>
</tr>
<tr>
<td>Effects of Classical Training on Vibrato</td>
<td>10</td>
</tr>
<tr>
<td>Mechanisms of Frequency Modulation in Vibrato</td>
<td>14</td>
</tr>
<tr>
<td>Mechanisms of Amplitude Modulation</td>
<td>19</td>
</tr>
<tr>
<td>Measuring Laryngeal-level Amplitude Modulation during Vibrato</td>
<td>35</td>
</tr>
<tr>
<td>Method</td>
<td>37</td>
</tr>
<tr>
<td>Singers</td>
<td>37</td>
</tr>
<tr>
<td>Equipment</td>
<td>39</td>
</tr>
<tr>
<td>Procedure</td>
<td>39</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>40</td>
</tr>
<tr>
<td>Results</td>
<td>49</td>
</tr>
<tr>
<td>Physiologic and Acoustic Data</td>
<td>49</td>
</tr>
<tr>
<td>Perceptual Ratings</td>
<td>60</td>
</tr>
<tr>
<td>Discussion</td>
<td>67</td>
</tr>
<tr>
<td>Physiologic and Acoustic Data</td>
<td>67</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>38</td>
</tr>
<tr>
<td>Table 2</td>
<td>57</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.5 second sample of modulation induced by abdominal pressure pulses.</td>
<td>42</td>
</tr>
<tr>
<td>Two-second sample of the EGG first derivative, speed quotient, and FM (first, second, and fifth panes, respectively), and mic AM and FM measures (third and fourth panes, respectively) during abdominally-induced laryngeal amplitude modulation. EGG in arbitrary units; see text for definition of EGG speed quotient.</td>
<td>43</td>
</tr>
<tr>
<td>Sensitivity of the EGGSQ, taken from the EGG first derivative, to laryngeal-level amplitude modulation resulting from abdominally-driven pressure pulses. EGG in arbitrary units.</td>
<td>44</td>
</tr>
<tr>
<td>FM peaks (upper pane) and Fast Fourier Transform (lower pane) during laryngeal-level amplitude modulation produced by subglottal pressure pulses.</td>
<td>45</td>
</tr>
<tr>
<td>Singers’ mean EGG AM during soft, comfortable, and loud singing conditions.</td>
<td>50</td>
</tr>
<tr>
<td>Singers’ mean EGG AM for each pitch condition.</td>
<td>51</td>
</tr>
<tr>
<td>Individual singers’ mean EGG AM across pitch and loudness conditions.</td>
<td>52</td>
</tr>
<tr>
<td>One of Singer 10’s comfortable G4 tokens, during which maximum EGG AM was present (0.8812). EGG in arbitrary units; see text for definition of EGG speed quotient.</td>
<td>53</td>
</tr>
<tr>
<td>One of Singer 4’s comfortable A5 tokens, during which minimum EGG AM (second pane) was present (0.0273). EGG in arbitrary units; see text for definition of EGG speed quotient.</td>
<td>54</td>
</tr>
<tr>
<td>The effect of loudness and pitch conditions on vibrato rate.</td>
<td>56</td>
</tr>
</tbody>
</table>
One of Singer 12’s comfortable D5 tokens, during which double peaks in amplitude modulation (second pane) occurred within single frequency cycles, visible in the third (mic FM) and fourth (EGG FM) panes. .................................................................58

FM peaks (upper pane) and FM FFT (lower pane) from one of Singer 17’s loud G4 tokens during which mic FM COV was notably high (16.8). ..................................................59

FM peaks (upper pane) and FM Fast Fourier Transform (lower pane) from one of Singer 9’s comfortable A5 tokens, during which mic FM COV was notably low (1.4). ..........61

One of Singer 16’s comfortable A5 tokens, during which mic AM (upper pane) and mic FM (lower pane) had a low correlation (-0.009). ....................................................62

One of Singer 4’s comfortable D5 tokens, during which mic AM (upper pane) and mic FM (lower pane) have a high negative correlation (-0.953). .........................................63

One of Singer 5’s loud G4 tokens, during which mic AM (upper pane) and mic FM (lower pane) have a high positive correlation (0.975). ..................................................64

Participants’ mean airflow in each pitch and loudness condition....................................65
List of Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professors’ Mean Ratings and Singers’ Mean EGG AM Extents</td>
<td>85</td>
</tr>
<tr>
<td>Singers Rated Highest and Lowest Overall and for Consistency, Speed, and Width</td>
<td>86</td>
</tr>
<tr>
<td>Consent to Be a Research Participant</td>
<td>87</td>
</tr>
</tbody>
</table>
Introduction

The study of vocal vibrato among Western classically-trained singers has received attention from researchers in several fields, including physiology, acoustics, voice pedagogy, and speech-language pathology. Vibrato, which is the modulation of the frequency, amplitude, and timbre of a singer’s voice, can develop naturally, yet can also be further shaped volitionally through training. Much acoustic and physiological research has focused on the mechanisms of frequency modulation (FM). Likewise, research has established that another component of vibrato, amplitude modulation (AM), occurs passively when the singer’s fundamental frequency and harmonics coincide with the upward and downward slopes of the vocal tract transfer function for any given articulatory configuration. This process takes place above the larynx in the vocal tract, and is known as the resonance-harmonics interaction or RHI (Horii & Hata, 1988).

However, the possibility remains that amplitude modulation (AM) also arises at the level of the vocal folds, prior to this filtering process. The purpose of this study is to determine whether amplitude modulation does indeed occur at the laryngeal level in addition to the AM known to subsequently arise as a result of the RHI. Another aim of this study is to determine the presence and degree of laryngeal-level AM among beginning to advanced singers, as a function of different pitch and loudness conditions. These findings will then be compared to expert perceptual ratings of the singers’ vibrato to detect any associations between laryngeal-level AM and quality ratings of vibrato. The effect of changing pitch and loudness conditions on singers’ vibrato rates will also be examined.
History of Western Classical Vibrato

The aesthetic standards regarding beauty in vibrato have been developed and refined historically among different cultural traditions worldwide. One such tradition is that of Western classical music, which originated in the sixteenth-century Italian bel canto (“beautiful singing”) style and subsequently spread throughout Europe (Sundberg, 2000; Troup, 1982). It was characterized by the near-constant presence of a high-quality vibrato, and was cultivated through extensive training in respiratory and vocalizing exercises designed to improve the singer’s agility, power, and control of vibrato. Another highly-prized feature was the ability to crescere e scemare la voce or to “swell and diminish” the voice (Troup, 1982). This emphasis on intensity control during pitch changes was the precursor of messa di voce (“placing the voice”), considered the soul of music, and the fundamental quality of classical singing (Mason, 2000; Titze et al., 1999). Wrote Nathan, “This swelling and dying of the voice is the most important to practice....on it depends the principal art of Singing, for it sweetens, enriches, and gives the delicious roundness and fullness to the tone” (Nathan, 1836, p.145). Singers trained to make increasingly subtle and gentle alterations to both their intensity and pitch could then convey layers of meaning, express nuanced feelings, and tell complicated stories, integral to the opera. Vocalists were encouraged to rely on the ear, imagery, and their own muscle movements to “discover,” rather than be told, how to place their voices correctly and make improvements in agility, power, and range (Mason, 2000). In the late nineteenth and early twentieth centuries, concert halls and orchestras became larger, requiring singers to project their voices with more power using la voix sombre, or “the covered voice,” as opposed to the earlier, lighter, castrato voice (Reinders, 1995). In this
technique, used today, the larynx is lowered, with more fluctuations in air pressure under the vocal folds, creating larger laryngeal oscillations and more changes in both pitch and loudness. These changes permit “greater acoustical potential in voice production” (Reinders, 1995, p. 142).

**Seashore’s Contributions**

Scientific advances in the nineteenth and twentieth centuries then allowed for more specific measurements of the acoustic parameters and physiological mechanisms of vibrato. For example, in the 1930s, Seashore attempted to answer, artistically and scientifically, what vibrato is, what makes it good or bad, and how it can be improved (Seashore, 1947). He began by offering what is still considered the classic definition of optimal vibrato: “A good vibrato is a pulsation of pitch, usually accompanied by synchronous pulsations of loudness and timbre, of such extent as to give a pleasing flexibility, tenderness, and richness to the tone” (Seashore, 1947, p. 55-56). Good vibrato was considered so essential in Western classical music, wrote Seashore, that “to the singer it is not a question of whether or not he will use the vibrato; the question is: What kind is to be cultivated and tolerated?” (Seashore, 1947, p. 61).

To achieve the desired vibrato, he recommended that the singer obtain correct information about the mechanisms and properties of good vibrato, learn to recognize those elements and how to produce them, and then practice producing them proficiently and effectively. He asserted that pitch, loudness, and timbre could be “isolated and demonstrated, and that they are the only kinds of modulation possible” (Seashore, 1947, p. 56). Further, he suggested that based on recordings of accomplished classically-trained singers, the optimal pitch vibrato was approximately 0.5 tone in extent, equal to one
semitone, and 6.5 Hz in rate; that optimal intensity vibrato was about 2-3 dB in extent, and that it should occur at the same rate as the pitch vibrato (Seashore, 1932). These standards are essentially unchanged today. A good vibrato, he wrote, should have the traits of regularity, moderation, and smoothness. He also noted that the pulsations found in vibrato are similar to those found in the discharge and refractory phrases of both laughing and crying; they take the form of a sine wave, and are evidence of the “organic basis for a trembling in tone production” (Seashore, 1947, p. 61).

At the crossroads between a rich history of beautiful vocal performances resulting from aesthetic, intuitive musical training, and the emerging scientific frontiers of recordable, visible physiological processes and acoustic spectra, Seashore recognized the valuable contributions of each approach. Accordingly, he felt that “the modulation of the vibrato will gradually take care of itself when the ear and esthetic judgment have been trained,” and that in turn, this training would occur through “exact knowledge of its nature and variables, and by the acquisition of skills for the hearing, evaluation, and feeling of each variable in turn” (Seashore, 1947, p. 70).

*Acoustic Measurements of Vibrato*

To better evaluate the nature of vibrato and its variables, researchers have studied the acoustic parameters which characterize vibrato. In addition, they have studied the interactions between these parameters and their possible physiological mechanisms.

*Pitch, Intensity, and Timbre*

As Seashore wrote, vibrato consists of pitch, intensity, and timbre modulations. In other words, it is the modulation of frequency and amplitude (FM and AM), measured in terms of presence, onset, rate, extent, and periodicity. Seashore alluded to the interactions
of amplitude and frequency modulation with his inclusion of a third identifiable component of vocal vibrato, that of timbre modulation (Seashore, 1947). Timbre, the “periodic modification of partials,” reflects the manner in which the vocal folds open and close, which alters the shape, speed, and power of the air pulses flowing through the glottis (Horii & Hata, 1988, p.304). The spectrum of timbre ranges from darker, mellower tones, produced when vocal fold closure is even, to brighter, brassier tones, which occur when the vocal folds close faster than they open. Schutte and Miller explained that “timbre and amplitude vibrato can result from the periodic movement of the harmonics of the voice source among the more stationary formants of the vocal tract” (Schutte & Miller, 1991, p. 221). Lebon (1999) described timbre vibrato as periodic fluctuations in the amplitude or strength of the individual harmonics, and the resulting combination of proportional tonal qualities.

**Onset, Presence, and Periodicity**

Onset refers to the amount of time after a tone has begun until vibrato modulations occur. In one study, expert judges rated singers’ vibrato as significantly lower if the onset occurred more than 0.5 seconds after vowel onset, and late onset also made it more likely that judges lowered their ratings of other perceptual characteristics as well (Elkholm, Papagiannis, & Chagnon, 1998). The presence of vibrato is considered to be the percentage of singing time that frequency modulation is present. FM among highly-trained classical singers was found to occur in 95% of singing time on average, while AM was present approximately 50-70% of the time (Mason, 1965). In a study with ten opera singers participating, no presence of amplitude vibrato was found at all (Shipp, Leanderson, & Sundberg, 1980). This could have been due to measurement techniques,
however (Horii, 1989a). The periodicity of vibrato describes its rhythmicity or evenness, and Lebon found that highly-trained vocalists’ vibrato contained an average of 6-7 modulations per second (Lebon, 1999).

**Frequency Modulation**

Frequency modulation is typically measured in terms of its presence, rate, extent, and periodicity. Frequency modulation was found to occur among classically-trained singers during 95% of singing time (Mason, 1965). The rate of frequency modulation is measured in Hz, or cycles per second, and 6.5 Hz (Seashore, 1932; 1947) and 5.5-7.0 Hz (Sundberg, 2000) are considered optimal FM rates. Two tenors, Caruso and Pavarotti, had mean vibrato rates of 6.5 Hz and 5.5 Hz, respectively (DeJonckere, 1995; Titze, Story, Smith, & Long, 2002). In their analysis of vibrato rate among ten opera singers, Shipp et al. (1980) found female rates ranged from 4.9-6.6 cycles per second (cps), while males’ range was 4.7-6.3 cps, with the averages of 5.9 cps (female) and 5.4 cps (male) being significantly different. They further noted that “waveshape within and between subjects had considerable variability” (Shipp et al., 1980, p. 21), while Horii (1989b) determined that an FM cycle was always initiated with an increase, rather than a decrease, in pitch. The historical consensus is that rate change is typically not dependent on intensity or where in a singer’s range the task note is, that there does not appear to be a relationship between rate and the singer’s voice type, and that rate is influenced by the setting: operatic performances tend to have the highest rates, followed by classical recordings, and then laboratory recordings.

The extent of frequency modulation is measured in two ways: the semitone (ST), equal to a musical half-note, and the cent, with 100 cents being equal to one semitone
(100 cents = 1 ST). Seashore considered optimal FM extent to fall within the range of 30-70 cents, with 50 cents (= 0.25 tone or 0.50 ST) being ideal (Seashore, 1932, 1947); it was reported recently that an extent of 50 cents (= 0.50 ST, or ~3% variation of the pitch) is still considered aesthetically acceptable (Arroabarren, Zivanovic, Bretos, Ezcurra, & Carlosena, 2001). Extents among classically-trained singers in another study averaged 1.20 ST (Shipp et al., 1980). Current optimal ranges have been reported as 0.50-2.00 ST (Horii, 1989a), 0.80-2.00 ST (Shipp et al., 1980), or 50-150 cent (= 0.50-1.50 ST), depending on pitch and loudness (Prame, 1997; Sundberg, 2000). Michel and Myers (1991) studied the effects of increase (crescendo) and decrease (decrescendo) of amplitude on the extent of vocal vibrato. Going from pianissimo (very soft) to fortissimo (very loud), vibrato extent (or width) was found to linearly increase; however, going back from fortissimo to pianissimo, vibrato extent tended to remain constant. Vibrato extent was the greatest at the middle frequencies. FM periodicity tends to be more regular than AM periodicity, with classically-trained singers averaging 5-7 oscillations per second (Mason, 1965; Seashore, 1947).

Amplitude Modulation

Amplitude modulation, or fluctuation in the intensity of the vocal signal, is typically measured in terms of its presence, extent, and periodicity. It is known specifically as the periodic increases and decreases of intensity relative to the mean (Horii, 1989a). Amplitude modulation has been reported as being present approximately 50-70% of the singing time among classically-trained singers (Mason, 1965). The extent of amplitude modulation is usually measured in decibels (dB), but has also been reported as a percentage of fluctuation relative to the mean; Seashore reported that an AM extent
of 2-3 dB was considered most pleasing (Seashore, 1932, 1947), while Mason found that among accomplished singers, the average AM extent was 2.4 dB (Mason, 1965). Others found that AM extents of 4-7% were associated with the highest-ranked voices (Rothman, Rullman, & Arroyo, 1990).

AM periodicity is a measure of the regularity and consistency of its waveform. Seashore (1932, 1947) urged singers to aim for 6-7 modulations of amplitude per second so that FM and AM could be synchronous; this was the same conclusion of Mason (1965) and Lebon (1999), while in another study classical singers had rates of 5-6 pulses per second (Rothman et al., 1990). The periodicity of amplitude modulation has been found to be less regular than that of frequency modulation (Horii, 1989a; Mason, 1965). Vennard found some singers’ AM rates to be either equal to or double the FM rate (Vennard, 1967).

Non-vibrato

There are some degrees of modulation, or combinations of them, that fall outside the established standards outlined above. On a spectrum of modulation, vibrato can be placed alongside wobble, bleat, tremolo, and various forms of vocal tremor. Wobble is characterized by a wider extent and slower rate (~5 Hz), while bleat is recognized by its narrower extent and faster rate (>7 Hz) than vibrato (Horii, 1989a; Sundberg, 2000; Troup, 1982). Wobble is typically associated more with senescence, while bleat is more likely to occur in the untrained singer, or the younger singer in whom the thyroarytenoid muscle is not fully formed (Titze et al., 2002). A frequency extent of greater than 6% of the mean fundamental frequency sounds too expansive to the classically-trained ear, while an extent smaller than +/-50-150 cents is associated more with choral or pop music
styles (Sundberg, 2000). Tremolo is considered the modulation of amplitude only, and is relatively rare.

Vibrato’s related counterpart, tremor, also involves modulation of the voice, and occurs in several forms as a result of different pathological conditions. Winholtz and Ramig (1992) found significantly higher levels of frequency and amplitude modulation in both vocal tremor patients and vibrato singers than in their control participants. In their study of opera singers and patients with vocal tremor, Ramig and Shipp (1987) analyzed sustained phonations of /a/ for type, rate, and extent of oscillation, as well as jitter and shimmer. They found that six of the nine singers, and all six patients, had both amplitude and frequency oscillations, while three of the singers had frequency oscillation only, and none had amplitude oscillation only. The tremor samples had slightly higher oscillation rates, extent ranges, and frequency ranges, but were not considered statistically significant. However, the rates of both jitter and shimmer were significantly higher in tremor than vibrato. They suggested that because both vibrato and tremor occurred at approximately 5-6 Hz, this could be evidence of a common central modulation generator, “which in patients is released from suppression by disease and in singers is selectively recruited” (Ramig and Shipp, 1987, p. 166).

These views are consonant with the EMG evidence reported by Koda and Ludlow (1992) of activation patterns in extrinsic and intrinsic laryngeal musculature during vibrato and tremor which displayed striking similarities. Others have speculated that “perhaps reflex gains can become too large and no longer able to be inhibited voluntarily,” leading to nearby anatomical structures experiencing vibration too (Titze et al., 2002, p. 2282). In one acoustic analysis of vocal tremor, the partials had stronger
vertical movement than the fundamental did, which suggested to the authors that
fluctuations in amplitude or in the voice source were more important factors in tremor
than fundamental frequency fluctuations (Imaizumi, Saida, Shimura, & Hirose, 1993).
Work done by Horii suggested that amplitude modulations in tremor may come from
sources such as respiratory alterations (Horii, 1989a).

Effects of Classical Training on Vibrato

Vocal pedagogues have experimented with various training techniques to better
understand which aspects of vocal oscillations during vibrato can be changed, how they
can be changed, and to what degree. Many substantial changes to vibrato quality due to
classical training have been documented. Several researchers have noted that among their
participants, for example, the number of years of training and the experience of the
singers affected the parameters of vibrato in the results of their studies (King & Horii,
1993; Shipp et al., 1980). Benefits of classical vocal training include progression from
non-existent or minimal vibrato in the beginning student to present but aperiodic
modulation of frequency and pitch extent, to eventual development of pitch excursions
decreasing from a whole tone to less than a semitone (Troup, 1982). Other changes
accrued from vocal training include improved control, efficiency, and strength of the
respiratory and laryngeal muscles and increased vital capacity; longer, more gentle, and
more efficient vocal fold closure phase; increased resonance in the vocal tract, richer
upper harmonics, and more regularity in vibrato, resulting in more pleasing timbre (Cook,
2001; Sundberg, 2000; Troup, 1982). McCann (1995) compared extracted vowels from
the spoken utterances and sung passages of trained and untrained singers. The trained
singers’ fundamental and second formants had significantly higher intensity levels during
speech and singing than the untrained singers’ did. Further exploration is needed regarding whether these and other trained singers have learned to specifically manage their amplitude at the laryngeal level, as well as in the vocal tract, for greater power, and possibly for more controlled amplitude modulations during vibrato. Also of interest is how amplitude modulation varies among singers, especially as it corresponds to length and type of training.

**Implications for Vocal Pedagogy**

Clinically-demonstrated change in vibrato due to training has implications for voice teachers and singers who wish to make systematic improvements to vocal vibrato that are well-informed by research evidence. Most Western classical vocalists are instructed in a variety of drill, exercise, imagery, and physical sensation techniques designed to cultivate awareness of and precise control over vocal structures for optimal resonance, power, and vibrato. Vocalists work to alter five harmonics of their fundamental frequency in particular: the first two harmonics, which determine the vowel shape and are affected by jaw (first harmonic) and tongue (second harmonic) movements, and the third, fourth, and fifth harmonics, which collectively form the range of the singer’s harmonic, or the ‘ring’ in a singer’s voice (produced by widening the pharynx and lowering the larynx) (Sundberg, 2000). The ‘ring,’ a frequency range in which the harmonics of the singer’s fundamental frequency align best with the vocal tract’s corresponding resonant peaks, creates tones that have been measured to radiate with maximum richness and power at the lips.
Volitional Control of Aspects of Vibrato

An example of singers’ learned control of selected vibrato parameters was Folger’s (2002) study, in which classically-trained singers each sang a passage three times alone, and then six times in different standing positions within a chorus to measure any changes in vibrato rate and extent due to choral singing effects. Folger found general chorus effects for both parameters when the singers sang as part of a choral group, suggesting that vibrato modulation is susceptible to volitional or even sub-conscious attempts to match other singers. Classically-trained singers have been found to be able to make changes in frequency or amplitude vibrato ‘on the vibrato,’ or during the peak or trough of the frequency cycle (Stanley, 1931; Troup, 1982). Duncan, Williams, and Troup (2000) compared three recordings of Dame Joan Sutherland singing Delibes’ ‘Flower Duet,’ each done with a different partner, to determine if partners could sing in ‘phase lock’ with each other, or in other words, to volitionally match periodic vibrato characteristics. One partner sang in phase lock, another sang in ‘antiphase,’ and the other singer’s vibrato periodicity was in ‘phase wander’ compared to Ms. Sutherland’s. These phases resulted in decreasing levels of pleasantness to the ear. The authors suggested that singing in phase lock is not only volitionally possible, but occurs regularly in choral singing with training. Similarly, trained singers demonstrated their ability to maintain a targeted pitch, but still change their vibrato FM rate to match a slower or faster target rate, in chest and mixed registers (Dromey, Carter, & Hopkin, 2003). These findings suggested that “…vibrato rate can be modified, and does not seem to be ‘hard-wired’ on the basis of an individual’s phonatory anatomy” (Dromey et al., 2003, p. 176).
Singers judged to have the most beautiful voices do appear to have several acoustic characteristics and volitional techniques in common in their production of vocal vibrato. In a study conducted by Robison, Bounous, and Bailey (1994), male classical and female belt singers’ voices were ranked by voice experts, and the highest-ranking voices were then analyzed acoustically to determine which properties were most essential to vocal beauty. The highest-ranked voices had many similar features, primarily a smooth, even vibrato, with an even rate (averaging 5.4 Hz), a moderate extent (averaging 22% of the median overtone frequency), and less than 4% variation in the rate and extent of vibrato. These singers had the smallest noise-to-signal ratios through the duration of the vibrato wave, and their vibrato was present at least 75% of singing time. In addition, the most beautiful voices had a characteristic timbre, which the authors defined as “the comparative strengths of the singer’s formant, vowel formant, and fundamental pitch,” or in other words, the proportional amplitudes of these formants at the lips (Robison et al., 1994, p. 22). Respiratory factors contributing to vocal beauty included the use of ventricular breathing (using primarily the diaphragm and abdominal muscles) over costal breathing (using primarily the rib cage muscles), and more significant abdominal closing at onset and throughout a tone.

While overall breathing patterns were observed, it is unknown whether those breathing patterns were characterized by pulsations that might correspond to or contribute to amplitude modulations. Also, the singers’ proportional formant amplitudes were measured at the lips, following alteration due to the vocal tract transfer function. Gaining a better understanding of how highly-ranked singers are altering supralaryngeal structures to create greater resonance and power during vibrato is useful to other singers who wish
to adopt similar techniques. It could also be helpful, however, to investigate fundamental and harmonic amplitudes at the laryngeal source before vocal tract shaping, to understand their contribution to overall amplitude as well. In addition, there may be modulations in the source-level amplitude occurring prior to the vocal tract transfer function. If source-level amplitude modulations are present in the vibrato of some singers, it could be beneficial to know whether and how these are contributing to overall vocal beauty, and what is causing these modulations. These factors could provide more detailed information about what well-trained singers are doing at the laryngeal level to enhance and regulate their vibrato amplitude and its modulation, to complement research evidence regarding vocal tract transfer function contributions to vocal beauty.

Mechanisms of Frequency Modulation in Vibrato

Acoustic measurements can provide detailed descriptions of the characteristics of vibrato, including those characteristics singers seek to develop through training. Likewise, physiological research offers explanations about the underlying mechanisms of vibrato generation and control that produce those perceptual and acoustic characteristics. Both vibrato and tremor consist of fluctuations of frequency and amplitude overlaid on the voice signal, with corresponding, observable oscillations in respiratory, laryngeal, and articulatory musculature. However, “the source of these oscillations is not well understood” (Leydon, Bauer, & Larson, 2003, p.1575). Laukkanen, Vilkman, and Unto (1992) suggested that singers produce vibrato using laryngeal musculature, or subglottal pressure (Psub) pulses, or a combination of the two. In some popular and non-Western styles, vibrato is indeed achieved primarily through modulation of Psub using rhythmic abdominal muscle contractions, resulting in air flow pulses that interact with the vocal
folds and create momentary F0 fluctuations (Sundberg, 1995). In Western classical music, however, vibrato has traditionally been thought of as a series of neurologically-driven laryngeal and respiratory adjustments, being mainly characterized by rapid increases and decreases in force of the cricothyroid (CT) muscle; this is the muscle primarily responsible, when contracted, for lengthening the vocal folds and thereby increasing pitch (Horii, 1989a; Titze et al., 2002). These contractions cause corresponding frequency fluctuations of the singer’s fundamental contour.

Shipp et al. (1980) explained a possible physiological model for achieving vocal vibrato. They postulated that the strength and rate of CT contractions determine frequency modulation; each CT contraction lengthens and stretches the vocal folds, which increases F0. This explains the increase in CT activity over a straight-tone condition. Neural impulses, they suggested, sent down the superior laryngeal nerve (SLN) to the cricothyroid (CT) muscle must be of sufficient strength to adjust the length of the vocal folds to match the desired pitch. This pitch is then overlaid with a second set of neural impulses traveling along the recurrent laryngeal nerve (RLN) which serve as an alternating current (AC) to modulate this frequency, “that is, to increase and decrease CT muscle contraction in a rhythmic pattern that results in systematic changes above and below the target pitch” (Shipp et al., 1980, p.23). This AC, they suggested, could be turned on or off, and could be immediate or delayed, “either for artistic effect or to establish the target sound initially” (Shipp, Sundberg, & Haglund, 1984, p. 23). It could be that the central and autonomic nervous systems working in concert are able to establish both target pitch and modulations relative to it, because the autonomic nervous system has been shown to modify impulses generated by the central nervous system.
Well-trained singers, they suggested, are probably better able to monitor and control the AC impulses, which may travel in synchrony with the vibratory cycle to the CT, but also to other structures as well, including the tongue, jaw, palate, pharyngeal walls, and larynx; this extra modulation can result in greater amplitude vibrato. The researchers concluded that “vibrato is produced by selective activation and inhibition of neural impulses in the critical muscle groups throughout the vocal tract” (Shipp et al., 1980, p.24).

Another subsequent hypothesis was that vibrato is a “stabilized physiologic tremor” (Titze, Solomon, Luschei, & Hirano, 1994, p. 215) of laryngeal muscles, driven by central and peripheral oscillators in a “reflex resonance model” (Titze et al., 2002, p. 2272). In this model, the singer begins by recalling the stored auditory and kinesthetic sensations associated with a desired pitch. This stimulates a central oscillator to create low-amplitude, wide-band oscillations at 4-6 Hz. These oscillations are sent via the midbrain as muscle commands to the CT and thyroarytenoid (TA) muscles, which act antagonistically in maintaining vocal fold length and tension. The midbrain acts as an intermediary, also receiving and incorporating sensory feedback from laryngeal muscles into the cortical commands. Stretch receptors in the TA monitor its length, and interact with the muscles that are changing and regulating its length, such as the CT. The activity of these stretch receptors stimulates a peripheral oscillator, which then stabilizes the oscillations into a narrower-band, periodic vibrato. This interaction may provide the nearly-immediate feedback necessary for a closed-loop reflex that creates its own oscillations. An increase in vocal fold length would then lead to a periodic, delayed decrease in vocal fold length, creating a phase delay and resulting in a regular, fluctuating
cycle of vocal fold length adjustments. In this way, the laryngeal system can continue the oscillation on its own without constant, direct neural input. The gains and delays that occur as a result of the reflex loop resonate, so that the “broad spectrum of central tremor frequencies” is refined into “a narrower band, which is known as vibrato” (Titze et al., 2002, p. 2274).

Within the laryngeal system, one possibility for the location of the peripheral oscillator was suggested to be in the thyroid and cricoid cartilage rotations involved in length adjustments of the membranous vocal folds. Other possibilities for the peripheral oscillator, wrote the authors, could be an oscillation of the arytenoid cartilages, vertical travel of the larynx, or its muscular attachments to the sternum or hyoid.

A potential expansion to their model, Titze et al. noted, would be to include the role of auditory feedback. Some researchers consider the ear so essential to vibrato control as to be its primary adjustor, or at least an integral part of an automatically-adjusted control loop (Deutsch & Clarkson, 1959). Notably, the frequencies in the core of a singer’s formant happen to correlate with a range in the human ear which is best-suited for auditory perception, not only giving the listener a finer perception of this aspect of vocal beauty, but also providing critical feedback to the singer as well (Cook, 2001). Deutsch and Clarkson (1959) demonstrated that introducing a delay in auditory feedback as well as lowering its perception threshold could affect both the frequency and amplitude of participants’ vocal oscillations. Leydon and colleagues (2003) explored the possibility that a pitch-shift reflex (PSR) within the auditory system might be sustaining the oscillations found in vocal vibrato. PSRs were described as bi-directional, sinusoidal modulations or “closed loop negative feedback reflexes” similar to a stretch reflex.
It was hypothesized that PSRs could be elicited while having untrained singers simultaneously produce and listen to their own 5-second straight-tone phonations of /u/. On these straight tones, the researchers overlaid periodic 5 Hz frequency modulations and then sent this new oscillating signal back, with 8-20 msec delays, to the singer as auditory feedback via headphones. Approximately 100-150 msec after hearing a 5 Hz increase or decrease of 0.25 ST, the singers compensated by adjusting their pitch downward or upward by 5 Hz, respectively. This pitch-shift reflex repeated itself throughout the phonation, creating a sustained vibrato of 5 Hz and 0.5 ST, which replaced the straight tone the singers were intending to produce. These findings suggest that a control loop which incorporates auditory feedback may be implicated in the modulation of frequency found in both vibrato and tremor. The authors reported that the latency of a PSR most likely included the nerve conduction and muscle contraction times involved in decoding the pitch, recognizing the error, and creating a motor command to compensate; this resulted in oscillations “at frequencies close to twice the predicted latency of the PSR loop” (Leydon et al., 2003, p. 1578). Although the authors believed that auditory system reflexes contributed to both frequency and amplitude modulations, they did not report whether the PSRs were found to be associated with any modulations in amplitude. However, they noted that adding frequency modulations to the signal introduced possible modulations in amplitude as well, because fluctuating the fundamental frequency then allows harmonics that fluctuate near vocal tract formants to be boosted further. If so, they suggested, participants may have been reacting to intensity changes as well as frequency changes in order to sustain vibrato. It is known that singers typically rely on auditory feedback to adjust their overall amplitude during vibrato, but
what is less clear is what role it appears to play in modulating amplitude. Obtaining measures of amplitude modulations at both the output and laryngeal levels helps pinpoint when those modulations are first occurring, a step toward understanding how they are centrally and/or peripherally generated and controlled.

*Mechanisms of Amplitude Modulation*

While the above models offer explanations for the mechanisms of frequency modulation, still to be explored are the mechanisms involved in amplitude modulation, and how amplitude and frequency modulation influence each other. Vennard (1967) observed that “much of what we seem to hear as variation in intensity is really our ears’ interpretation of the pitch variation. However, there is some true intensity vibrato, at least part of the time,” which “results in fluctuation in timbre, which is not quite so noticeable” (Vennard, 1967, p. 193). Horii noted that while frequency modulation is considered the main component of vibrato, with amplitude modulation being mostly derived from it, the exact mechanisms of frequency and amplitude modulations are less clear (Horii, 1989a). Accordingly, some physicists recommended that researchers not overlook the importance of amplitude variation, which they asserted has “an enormous relevance in the perception of the vibrato” (Arroabarren et al., 2001, p. 1533). They cited experiments in the literature using synthesized voice, in which resynthesized vibrato does not sound complete until amplitude variations resembling the human voice are properly included, because they are “crucial to produce warmth and natural quality in” vibrato (Arroabarren et al., 2001, p. 1533).
Horii (1989a) suggested that scientists and vocal pedagogues must come to a clear understanding of the origins of amplitude modulation during vibrato; he considered that two options could be possible. First, amplitude modulation could be resulting from a ‘resonance-harmonics interaction’ (RHI) due to modulation of the fundamental frequency and therefore its proportional harmonics. Second, amplitude modulation could be occurring independently of frequency modulation as a result of a separate, active oscillation generator (Horii, 1989a; Horii & Hata, 1988, p. 303). Horii and Hata showed that most amplitude modulations could in fact be explained in terms of the resonance-harmonics interaction: if the resonant properties of the vocal tract are held constant, they wrote, overall amplitude is, “in essence, the sum of the height (strictly speaking, the sum of the square of the height) of the harmonics” which they termed the “intensity function” (Horii & Hata, 1988, p. 304). For a given vocal tract configuration, for example, during sustained /a/ at a specific frequency, there is a corresponding resonance contour of intensity peaks and valleys which occur along a spectrum of frequencies. This is known as the intensity function of that vocal tract configuration. Some frequencies will be boosted if they occur at an upward slope or peak of the contour, while other frequencies will be attenuated if they encounter a downward slope or valley of the resonance contour in the vocal tract. As the fundamental and each harmonic encounter a resonant peak or valley in the singer’s current vocal tract configuration, they are each either enhanced or dampened in amplitude. This resonance-harmonics interaction alters not only the amplitude of the F0 and each harmonic, but also the combined amplitude of the entire sound as it radiates from the lips. In addition, during vibrato, F0 is oscillating regularly
rather than remaining relatively fixed; as it fluctuates, its harmonics also systematically and proportionally fluctuate in frequency. As F0 and its harmonics oscillate, they travel back and forth along the contour of resonant peaks and valleys occurring at particular frequencies. The amplitude of a given oscillating F0 or harmonic is then continually moved either upward or downward, depending on the slope of the contour at which it is oscillating.

Applying the intensity function to six fundamental frequencies from 100-450 Hz, Horii and Hata (1988) found that it could accurately predict the overall amplitude modulation of nearly all frequencies. While amplitude modulation may also be affected by Psub, there is general consensus that this FM-AM interaction is the major contributor to amplitude modulation. However, it is still unknown whether the amplitude modulation at the level of the lips is solely generated through the RHI, or if it could have been occurring independently at the laryngeal level prior to vocal tract influences.

Phase relationships. In optimal vibrato, the phases of frequency and amplitude modulation should occur in synchrony with each other, Seashore observed (Seashore, 1932, 1947). Mason (1965) found that in some instances, called parallel vibrato, frequency and amplitude vibrato occurred synchronously, meaning the peaks and valleys of both occur simultaneously. In other cases, known as opposite vibrato, the two occurred in antiphase, meaning that as frequency is rising, amplitude is falling (Mason, 1965). Still other singers experience AM that is out of phase with FM, because AM is typically less regular than FM (Prame, 1997).

Horii and Hata (1988) postulated that during vibrato, the intensity function of the resonance-harmonics interaction could predict not only overall amplitudes of a variety of
fundamental frequencies, but also the phase relationships between FM and AM, or between fundamental frequency and fundamental intensity (I0). During vibrato, F0 is periodically increasing and decreasing, so its harmonics are as well. For example, if a singer is singing at 200 Hz, F0 might be periodically oscillating from 195 to 205 Hz, with its harmonics oscillating proportionally. As F0 rises and falls, it may either be traveling upward toward a resonant peak or downward toward a resonant valley along the intensity function contour. Horii and Hata found that if a frequency oscillation occurred on a positive slope of the intensity function, the frequency modulation and resulting amplitude modulation would be synchronous, or in phase; in this case, overall intensity was increased. Conversely, if a frequency oscillation occurred on a downward slope of the intensity function, the frequency modulation and resulting amplitude modulation would be asynchronous, or out of phase, leading to a decrease in overall intensity. If F0 happened to oscillate at a frequency where the variation in the intensity function slope was relatively high, boosting the signal, the magnitude of amplitude modulation was 2-3 dB. If F0 was oscillating in a relatively low range of variation in the intensity function, where the signal was dampened, the AM magnitude was minimal. Horii and Hata concluded that the RHI and the intensity function it generated did accurately predict the phase relationships between FM and AM, and explained why FM and AM can occur in phase or out of phase.

However, there were still cases reported in the literature in which singers experienced one FM cycle for every two AM cycles (Vennard, 1967). This could also be explained by the RHI; if F0 increased or decreased exactly across a peak of the resonant contour, it would have to travel first up to the peak and then down the other side, creating
a rise and then fall in amplitude, or in other words, two changes or modulations in amplitude during a single frequency modulation. Likewise, if F0 increased or decreased across a valley of the resonant contour, it would travel first down to the valley and then up the other side, resulting in a fall and then rise in amplitude, and again, two amplitude modulations during the single FM cycle. This helped clarify why singers sometimes experience 2:1 AM-FM phase relationships (Horii, 1989a; Horii & Hata, 1988). Vennard (1967) suggested that perceptually, this 2:1 phase relationship may highlight the peak and valley of the vibrato cycle, rather than the mean pitch.

*Limits of the resonance-harmonics interaction.* Horii (1989a) concluded that among his samples, the RHI could account for nearly all of the amplitude modulations, and could even accurately predict the magnitude of those amplitude modulations and their phase relationships with frequency modulations, even when the vocal tract was lengthened or shortened. This led him to maintain that “the vocal tract resonance” is a “major cause of the presence of amplitude modulations in vocal vibrato” (Horii, 1989a, p. 39). However, there were occasions when the RHI hypothesis was not in perfect agreement with the data. Horii acknowledged that the RHI does not account for how frequency modulations are generated in the first place, which other models have addressed, and that Psub oscillation is a possible contributor as well.

The few phonations the RHI could not account for were very high or low in either frequency or intensity. Horii (1989a) explained this may be due to the differences in speech and vibrato, perturbation effects from movement of the larynx, soft palate, and/or jaw, and the pitch and intensity levels of the tasks used. Horii wrote that the spectral envelope “used in generating the present intensity function” for sustained /a/ may not
match that of a singer using the singer’s formant. “Sung vowel phonations, in addition, may have greater amplitudes of higher harmonics,” he wrote, “than normal phonations at the level of laryngeal sound source” (Horii, 1989a, p. 41). Since Horii’s samples were cut off at 2000 Hz, this may explain part of the discrepancy; however, as Horii noted, including frequencies up to 5000 Hz only increased overall amplitude by 0.5 dB. Thus, measuring amplitude modulation of the voice source spectrum during sung vibrato at the laryngeal level may help clarify part of the discrepancy. In addition, Horii’s model assumed that at the laryngeal level, the only direct change taking place is in the modulation of frequency, not of amplitude. However, it could be that laryngeal mechanisms may also be directly involved in the modulation of amplitude, the modulation of timbre through adjustments in the quality of vocal fold closure, and/or the control of subglottal pressure and airflow contributions to vibrato.

**Other Possible Contributors to Amplitude Modulation**

Arroabarren et al. (2001) contended that while vibrato is essentially considered a modulation of the fundamental frequency, this alone does not completely account for changes in amplitude of F0 or the partials during vibrato. Further, “decomposing a signal into its AM and FM parts is ill-posed, in the sense that there exist an unlimited number of possible combinations in such components;” because of this, “the appropriate decomposition, if it exists, will depend on the context,” and information regarding other vibrato parameters and mechanisms should be included in that context (Arroabarren et al., 2001, p. 1529). Other researchers likewise questioned the “previously-accepted dichotomy” of FM and AM, suggesting that these categories need further clarification or adjustment (King & Horii, 1993, p. 158; Rothenberg, Miller, & Molitor 1988).
Rothenberg and colleagues suggested that “there may be a number of physiological factors present that might cause a vibratosynchronous variation in acoustic amplitude” (Rothenberg et al., 1988, p. 158-159). Horii (1989a) acknowledged the necessity of delineating amplitude modulations passively produced through the RHI, and those “actively produced by singers,” suggesting that the RHI “does not preclude” Psub as an additional mechanism (Horii, 1989a, p. 42; p. 40). He raised the possibility that still other mechanisms could also be at work; for example, he reiterated Ladefoged’s suggestion (as cited in Horii, 1989a) that pitch modulation is a function of the positioning of the arytenoid cartilages, and intensity modulation is a function of the spacing between them, which contributes to the amount of force used as they come together.

Shipp, Doherty and Haglund (1990) suggested two mechanisms of vibrato, each producing different degrees of amplitude modulation: in the first, termed laryngeally-mediated vibrato, F0 oscillations are produced by periodic contractions of the CT muscle; this can only occur when the vocal folds are properly placed at midline for optimum vibration. Vibrato produced by signals restricted to the CT muscles, as opposed to extralaryngeal structures in general, is considered more efficient and desirable in the Western classical tradition. In this view, while “more effective singers are somehow able to inhibit the AC neural activation transmitted along the RLN,” less-skilled singers exhibit undulations of pharyngeal and oral cavity structures as well, contributing to increased amplitude vibrato (Shipp et al., 1990, p. 24).

In the second mechanism, known as abdominally-mediated vibrato, F0 oscillations are produced when the intrinsic laryngeal musculature provides continuous contraction “against which abdominally induced Psub pulses act” (Shipp et al., 1990, p.
This creates pulses of upward airflow, resulting in corresponding fluctuations in both amplitude and fundamental frequency. This type of vibrato is found primarily in non-Western styles, but may also be found in Western classical vibrato to a lesser extent (Rothenberg et al., 1988). It is considered more voluntary, and would logically result in corresponding amplitude changes. On the other hand, laryngeally-mediated vibrato is considered more dependent on specific minimal neuromuscular conditions, and may involve only small amplitude changes. Whether these two mechanisms are perceptually distinguishable, and whether they may both occur within the same singer at times has not been thoroughly examined.

More recently, Sundberg asserted that there are three possible sources of amplitude modulation in vibrato: first, the “amplitude of the strongest spectrum partials, i.e. the partial closest to the first formant” is typically equal to overall amplitude, so that frequency oscillations of particular harmonics toward and away from the first and other formants create proportional amplitude variations (Sundberg, 1995, p. 46). Second, the voice source is not necessarily held constant, as previously considered, and “in certain types of vibrato, the voice source varies in amplitude,” due to, for example, fluctuations in Psub and glottal configuration (Sundberg, 1995, p. 47). Finally, the shape of the vocal tract, and therefore its formants, may oscillate, creating another source of varying amplitude. Rothenberg et al. (1988) noted that there could be different neural sources for the various types of activation and oscillation which the laryngeal and respiratory muscle groups experience during vibrato. These sources could then combine to create synchronous oscillations in amplitude as well as frequency, which contribute to the perception of amplitude modulation.
In general, research involving amplitude modulation during vibrato has examined three factors in addition to the resonance-harmonics interaction: the respiratory, voice source, and vocal tract contributions to amplitude modulation. In some cases, the effects of pitch and register on amplitude and its modulation are also discussed.

Respiratory contributions to amplitude modulation. Davis and colleagues discussed the “biological time-sharing” arrangement that the respiratory and phonatory systems use for breathing and voicing, pointing to the larynx as a “variable resistor,” with its rich array of sensory endings monitoring minute changes in both respiration and phonation parameters (Davis, Bartlett, & Luschei, 1992, p. 59). While King and Horii (1993) noted that many acoustic studies of vibrato are limited by not making simultaneous physiological and/or aerodynamic measurements, which can be invasive and costly, some researchers have studied the roles of subglottal air pressure and air flow in vibrato production as well. In a rare, invasive study, Rubin, LeCover, and Vennard (1967) used tracheal punctures and naso-esophageal catheters for intra-esophageal pressure recordings, among other measures, to determine $P_{sub}$ and airflow at different pitches and amplitudes during vocal vibrato. They found that among trained singers, when pitch was held constant, and amplitude increased, $P_{sub}$ increased, in order to sustain the phonation. Airflow stayed the same, rose slightly, or fell, depending on the singer. On the other hand, when amplitude was held constant but pitch was increased, $P_{sub}$ again rose, but airflow fell, due to increased vocal tension to increase pitch. When the singers then used breathy voice, first amplitude was held constant, which resulted in increased $P_{sub}$; then $P_{sub}$ was held constant, and amplitude decreased. During both breathy tasks, airflow predictably increased significantly. The researchers concluded that
glottal resistance is more crucial than airflow in sustaining a tone of increasing intensity, but also that poor breath support introduces “secondary interfering glottal tensions” (Rubin et al., 1967, p. 411). Measurements of Psub did not necessarily correspond with overall sound pressure levels, leading Rubin and colleagues to conclude that other factors besides Psub and airflow were also influencing output. The researchers did not measure the modulation of either amplitude or frequency and any corresponding fluctuation in Psub, glottal resistance, or airflow.

Later, Large and Iwata (1971) reported that airflow was indeed synchronous with amplitude modulations in vibrato. Their findings with Western classical singers indicated that, compared to glottal airflow during straight-tone singing, airflow during vibrato is approximately 10% higher, regardless of register. In addition, glottal resistance is lower during vibrato, because there are rhythmic periods of contraction and relaxation, rather than constant contractions, of both respiratory and intrinsic laryngeal muscles. Large and Iwata concluded that these differences could not be ignored; even though laryngeal control plays the primary role in the Western classical style, vibrato should still be considered as a function of the balanced interaction between laryngeal musculature control and respiratory control. During vibrato, they explained, air flow “consists of alternating fluctuations of air flow (AC factor) and constant air flow (DC factor),” and in their study, “the magnitude of fluctuations of air flow was very large and clearly correlated with the amplitude vibrato” (Large & Iwata, 1971, p. 55).

Bretos and Sundberg (2002) examined ten sopranos’ recordings of two long notes sung with crescendo only. The authors found that the intensity levels of the two notes increased linearly at the average rate of 4.1 dB/s. At the same time, most of the singers’
long notes were characterized by a slight increase in pitch as well, which they attributed
to “the increase of Psub producing the increase of sound level” (Bretos & Sundberg,
2002, p. 44). They noted previous findings that every 10 dB increase in sound level was
accompanied by a corresponding doubling of Psub and proportional increase in mean F0
(Bretos & Sundberg, 2002). Titze suggested this ratio was approximately 2-6 Hz/cm H$_2$O
(Titze, 1989). In addition, he demonstrated that regularly increasing and decreasing Psub,
or, “more specifically, translaryngeal pressure” by ~10% generated frequency
modulations of 1 ST, considered an optimal extent in Western classical vibrato (Titze,
1989, p. 901). However, he noted, these rapid fluctuations would have to be controlled by
muscles with fast response times, such as the internal intercostals. This suggests the
presence of a separate modulation generator that initiates Psub modulations, in addition to
the CT contractions which are known to passively modulate glottal resistance (King &
Horii, 1989) and alter fundamental frequency. Because Psub affects F0 as well as
amplitude modulation, and F0 also alters amplitude modulation through the RHI, Psub is
an important contributor to amplitude modulation.

*Voice source contributions to amplitude modulation.* Psub and glottal
configuration, resulting from specific muscle positioning, are the two main contributors
to the characteristics of voice at the source level, and play a significant role in the
modulation of amplitude. To increase vocal intensity, singers use increased Psub; as a
result, each of the harmonics of the source spectrum also increases in intensity, “more so
for the higher partials than for the lower partials” (Sundberg, 1995, p. 38). Also, some
amplitude oscillations in the voice source can be attributed to the “adductive force in the
glottis” (Sundberg, 1995, p. 47). For example, using inverse filtering to measure source-
level amplitude, Sundberg found amplitude modulations that corresponded to pulses of
glottal leakage. If glottal leakage is extensive, some of the sound is absorbed down into
the trachea, reducing intensity available to resonances in the vocal tract. Glottal leakage is
greater when glottal adduction is decreased, and this “attenuation can be expected to vary
with the vibrato,” also reducing the intensity levels of the formant peaks (Sundberg, 1995,
p. 49). If, on the other hand, this “adduction-based vibrato” occurs without glottal
leakage, the intensity levels of the lowest harmonics will vary proportionally, depending
on their phase relationship with frequency modulation (Sundberg, 1995, p. 50). In
addition, vertical movements of the larynx affect amplitude; Western operatic singers
today are trained to use the covered voice, in which the larynx is lowered for increased
resonance and power. If the larynx is raised, Sundberg and Askenfelt (1981) noted, the
amplitude of both F0 and its harmonics decreases due to higher glottal resistance. Each of
these factors contributes to the intensity levels of F0 and its harmonics, and their
amplitude fluctuations.

Schutte and Miller (1991) used simultaneous naso-esophageal catheter,
electroglottography, and microphone recordings to measure supraglottal pressure, vocal
fold contact area, and voice source characteristics simultaneously. They analyzed voice
samples twice: once, with a SPL integration time of 250 msec (in “real time,”) and again
with an integration time of 8 msec (covering ~3 glottal cycles). The 250 msec sample
revealed amplitude modulations of very few decibels, but the 8 msec sample revealed
amplitude modulations of up to 10 dB within one cycle, “a far more dramatic effect
which is fully perceptible only on a slow playback” (Schutte & Miller, 1991, p. 218).
They defined “realized resonance” as the level of harmonic sound actually briefly aligned with a formant, as opposed to resonances “which continue to be present regardless of the degree to which they are ‘realized’” (Schutte & Miller, 1991, p. 218). Explaining that while either the first or second harmonic typically makes up the bulk of the realized resonance, singers experience a more complex realized resonance because F1, F2, and the singer’s formant were all shown to contribute periodically to the overall amplitude. During vibrato, these harmonics “sweep in the vicinity” of the formants, (Schutte & Miller, 1991, p. 222), creating a rich array of enhanced intensity levels among the harmonics. Both amplitude and timbre vibrato were demonstrated to be the product of this interaction, the researchers wrote, rather than resulting from other “independent mechanisms,” with F1, F2, and the singer’s formant all making “substantial separate contributions to the total SPL” (Schutte & Miller, 1991, p. 222).

Vocal tract contributions to amplitude modulations. In addition to the amplitude modulations resulting from the resonance-harmonics interaction, researchers have demonstrated that amplitude modulations arising in the vocal tract can also be a result of periodic oscillations of its structures. For example, amplitude modulations corresponding to tongue movements have been documented; also known to affect amplitude modulations are rhythmic movements of the jaw and pharyngeal walls (Shipp et al., 1984).

Proportional roles of Psub, F0, source spectrum, and formants. Imaizumi et al. examined the interactions between four of the previously-discussed contributors to vibrato and its overall amplitude: Psub, F0, the voice source spectrum or levels of the harmonics, and the formant frequencies. The researchers suggested three hypotheses:
first, F0 modulation might be the primary factor in vibrato, with its fluctuations causing corresponding movement of the harmonics nearer to or further away from the resonances, following along the vocal tract transfer function contours. This is consonant with the resonance-harmonics interaction proposed by Horii and Hata (1988). In this case, phase relationships would interact with each other and be either parallel or opposite. Second, it could be that not only F0 but also Psub are both primary components of vibrato and its overall amplitude, so SPL at the lips would be a direct result of both of these. In this case, as Psub rises, F0 would generally increase, as would the level of the harmonics. Because of this, the harmonics would vary “across the vocal tract transfer function, or traverse it,” boosting the level of the harmonics beyond the resonant contours predicted by the RHI alone, rather than aligning with them. In this case, they postulated, AM and FM would be in phase (Imaizumi et al., 1993, p. 197). Third, it is possible that F0 modulation and the voice source spectrum variation are both strong components, in which case the authors speculated that lower harmonics would follow along the transfer function, but higher harmonics would traverse it, because the voice source spectrum has been observed to have a shallower slope (stronger harmonics) at higher frequencies than at lower frequencies during singing. In addition, these equations became more complex when vocal tract vibrations, which in these samples occurred at the same rate as the vibrato, were factored in; these vibrations would also conceivably cause fluctuation in formant frequencies, creating movements in some of their levels.

The authors found that soprano vibrato samples most closely matched the third scenario, in which variations in the fundamental and also in the voice source spectrum contributed to the vibrato; this was made clear because the lower formants varied along
the transfer function, but the higher formants traversed it. This voice source fluctuation, the authors noted, was due to the periodic increases in F0 because of vocal fold lengthening. Also documented were several instances of mezzo-soprano vibrato that could be accounted for by the second scenario, in which F0 and Psub fluctuations were the primary contributors. In these instances, individual harmonics had larger vertical movements which the authors suggested may have been caused by vibrations in the pharyngeal walls which corresponded to the FM of the vibrato. The authors concluded that in addition to fundamental frequency and Psub variations, “voice source spectrum variations, vocal tract wall vibrations and formant frequency variations may also contribute to the detailed acoustic structure of vibrato tones” (Imaizumi et al., 1993, p. 197). Their study of the “level-frequency variability of harmonics” allowed them to consider oscillations along the spectrum of harmonic amplitude levels at the sound source as contributors to vibrato and its overall amplitude, in addition to F0 oscillations and the resulting RHI (Imaizumi et al., 1993, p. 197). Whether these various factors might have been a function of register or individual vibrato technique is unclear.

*The effects of pitch and register on amplitude.* The sound pressure level (SPL) of vibrato amplitude is affected by the singer’s register, which typically corresponds with a certain range in pitch. Dromey et al. (2003) found that as trained singers moved from chest to mixed to head registers, SPL increased. They reasoned that this might have been due to increased respiratory effort required when the vocal folds stiffen, as occurs at higher pitches. This is consonant with Titze’s observation that as fundamental frequency increases, so does phonation threshold pressure (Titze et al., 1994). Another possibility suggested was that in the mixed register, the frequency of the first harmonic would be
relatively close to the first formant, while in the head register, the F0 would be relatively close to the first formant. In either case, the proximity would increase SPL at the output due to “the alignment of the resonant frequency with strong components in the source spectrum” (Dromey et al., 2003, p. 177). Singers in Horii’s 1989 study also reported increased exertion at higher frequencies, which Horii anticipated would require increased Psub.

Similarly, Isshiki (1965) found that raising intensity levels involved different mechanisms depending on the register. For example, in the lower pitches, he found that to increase intensity, singers used more glottal resistance, rather than more air flow. However, in mixed and head registers, singers’ increased intensities were accompanied by proportional increases in flow rate. Isshiki concluded that to effectively increase intensity, well-trained singers relied on glottal resistance at lower pitches, and on increased flow rate at higher pitches. Likewise, Large and Iwata (1971) found that “the magnitude of air flow fluctuations in vibrato increases from chest to head,” and that this increase “is directly related to the amplitude fluctuations of the vibrato and probably to the Psub fluctuations as a function of intensity” (Large & Iwata, 1971, p. 62). In their study, in which amplitude vibrato and pulses of air flow were shown to be correlated, they noted that the correlation was especially strong at higher frequencies, while in the middle register (440 Hz), oscillations of air flow, frequency, and amplitude were in phase.

As Titze explained, Psub is the main factor in amplitude control and has also been shown to affect F0 (Titze, 1989). In his study, its effect on F0 varied depending on register. In chest register, the effect of Psub was moderate (2-6 Hz/cm H₂O), perhaps due to glottal resistance. In head register, the effect was small (1-3 Hz/cm H₂O), possibly
because of elongated, stiffened vocal folds. In falsetto register, the effect was largest (5-10 Hz/cm H₂O), due perhaps to the diminished control experienced in this register because only the anterior third of the vocal folds are vibrating along their edges (Lebon, 1999). Titze concluded that the Psub contribution to F0 was more essential at lower pitches, with its largest effects occurring “when the vocal folds are very short and lax” (Titze, 1989, p. 906-907). He commented, “it is evident that intensity and frequency are not controlled independently in the human vocal system,” which he referred to as “intensity-F0 dependence” (Titze, 1989, p. 901).

*Measuring Laryngeal-level Amplitude Modulation during Vibrato*

While the modulation of F0 is still viewed as the primary contributor to vibrato through the RHI, as Seashore observed, there is more than one way to achieve vocal vibrato, and variations exist among singers, even within the Western classical style, in terms of their natural mechanisms and learned skills. There may be, for example, several physical sources and/or learned techniques contributing to modulations in amplitude during vibrato. Identifying these, and measuring their relative contributions, may increase singers’ awareness of AM present in their vibrato, so they can cultivate it for maximum effect. Horii (1989a) recommended that any amplitude modulations not explained by the RHI be examined in more detail, along with their perceptual characteristics. Exploring the possible respiratory and laryngeal contributions to AM, in addition to the supralaryngeal contributions made by F0 through the RHI, allows a more complete model of its entire process, similar to the more comprehensive models of FM.

The purpose of this study is to explore in more detail whether amplitude modulations occur independently at the laryngeal level, or are solely derived from
frequency modulations through the RHI as demonstrated by Horii and Hata (1989). If laryngeal-level AM is detected, another aim is to measure its presence and degree under differing task conditions and among a variety of singers. Obtaining reliable measures of AM and other components of vibrato and their interrelationships ultimately helps singers know which characteristics of their vocal performance can be adjusted, how and to what degree they can be modified, and how they affect each other. If, within the Western classical tradition, AM is identified as a contributor to vibrato beauty, identifying its mechanism(s) and interactions leads to better-informed vocal pedagogy.
Method

Singers

Seventeen female college-aged singers trained in Western classical opera participated in this study. Fourteen of the singers were recruited from the School of Music in the College of Fine Arts and Communications at Brigham Young University, and three from the Emerson-Smith College Theater Company who were also enrolled in private classical voice lessons. Five singers were advanced opera students (BYU graduate students); five were intermediate-level students (BYU juniors and seniors), and six were beginning-level students (BYU freshmen and sophomores, and theater company students). All participants reported an absence of voice or hearing disorders, and were in good general health. The participants were first given a brief overview of the study, and each then signed an IRB-approved consent form indicating her willingness to volunteer as a participant in this research.

The singers were interviewed to determine their age, year in college, length and type of vocal training, and age of first vibrato presence. They were also questioned regarding any changes in their vibrato that they attributed to vocal training, and their perceived change in vocal effort required to sing without vibrato. The singers’ characteristics are listed in Table 1.

The singers reported that vibrato improvements resulting from vocal training include more consistent, relaxed vibrato; increased ability to manipulate vibrato and to shape it “on the air”; and better abdominal breath support with decreased thoracic tension. Some singers reported their vibrato becoming narrower, and others looser; some noticed a
<table>
<thead>
<tr>
<th>Singer</th>
<th>Age</th>
<th>University level</th>
<th>Years of Training</th>
<th>Type of vocal training/lessons</th>
<th>Age of first vibrato</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27</td>
<td>graduate</td>
<td>9</td>
<td>private classical and group choral</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>junior</td>
<td>9</td>
<td>choral group training</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>graduate</td>
<td>7</td>
<td>weekly private classical</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>23</td>
<td>senior</td>
<td>8</td>
<td>weekly private classical</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>freshman</td>
<td>9</td>
<td>private classical and group choral</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>sophomore</td>
<td>6</td>
<td>private classical and group choral</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>graduate</td>
<td>9</td>
<td>weekly private classical</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>senior</td>
<td>9</td>
<td>weekly private classical</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
<td>junior</td>
<td>8</td>
<td>weekly private classical</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>23</td>
<td>graduate</td>
<td>5</td>
<td>weekly private classical</td>
<td>14</td>
</tr>
<tr>
<td>11</td>
<td>27</td>
<td>graduate</td>
<td>9</td>
<td>private classical and group choral</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>21</td>
<td>senior</td>
<td>8</td>
<td>weekly private classical</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>junior</td>
<td>9</td>
<td>weekly private classical</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>22</td>
<td>theater company</td>
<td>8</td>
<td>private classical and group choral</td>
<td>8</td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>sophomore</td>
<td>13</td>
<td>weekly private classical</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>theater company</td>
<td>0.5</td>
<td>weekly private classical</td>
<td>14</td>
</tr>
<tr>
<td>17</td>
<td>19</td>
<td>theater company</td>
<td>0.6</td>
<td>private classical and jazz</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>22.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>8</td>
<td>7.5</td>
<td></td>
<td></td>
<td>12.5</td>
</tr>
<tr>
<td>SD</td>
<td>2.35</td>
<td>3.09</td>
<td></td>
<td></td>
<td>3.18</td>
</tr>
</tbody>
</table>
faster, and others a slower rate. To describe attempting to sing without vibrato, singers used terms such as *unnatural, restrictive, conscious, breathy, tense, pressed, tighter, and difficult.*

**Equipment**

Recordings of the singers’ vibrato samples 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. An airflow mask (Glottal Enterprises MA-2) with a wide band differential pressure transducer (Glottal Enterprises PTW) measured oral airflow. A digital audiotape (DAT) recorder (Panasonic SV-3800) recorded acoustic signals from the head-mounted microphone and the sound level meter. These signals, along with EGG recordings and flow mask data, were also routed into a multi-channel analog-to-digital conversion system (Windaq 720) on a laboratory computer.

**Procedure**

Following several minutes of warm-up exercises (specific activities were chosen by the singer), each singer was asked to produce five-second tokens of sustained /a/ with vibrato in each of nine conditions: three pitches, each at three levels of loudness. Three tokens in each condition were obtained, for a total of 27 samples per singer. This process was then repeated while the participant wore a flow mask, for a total of 27 more samples. The three pitches, which were requested in randomized order from each singer, were pre-selected by BYU voice professor Dr. J. Arden Hopkin to represent low, medium, and
high pitches in the singer’s upper middle register. These pitches were: G4 (392 Hz), D5 (587 Hz), and A5 (880 Hz). Each participant sang each pitch first at a comfortable level of loudness, and then either at soft then loud levels, or loud then soft levels. The requested sequence of loudness levels was randomized across participants. These loudness levels corresponded to the musical dynamic ranges of piano, mezzoforte, and fortissimo.

Data Analysis

The middle two seconds of each token were used in the data analysis. Praat acoustic analysis software (version 4.4.07) was used to create amplitude and frequency modulation contours from the head-mounted microphone signal. These data were then saved as text files and imported via Microsoft Excel into Matlab, where custom analysis routines were used to perform additional numerical analyses of the EGG and acoustic signals. Upward and downward peaks in the first derivative of the EGG waveform were automatically identified, and a ratio of the closing to the opening peak was calculated. This measure has previously been defined as the EGG speed quotient (Dromey, Stathopoulos, & Sapienza, 1992), and varies in proportion to vocal intensity. It was reasoned that any rhythmic fluctuations in the EGG speed quotient would be evidence of amplitude modulation at the level of the vocal folds. This fluctuation was operationally defined as the EGG amplitude modulation (AM). To test the premise that laryngeal AM could be measured by this means, a non-singing individual induced rhythmic subglottal pressure pulses by manually applying pressure on the abdominal muscles. These pulses drove the resulting modulation in amplitude at the laryngeal level, which then led to corresponding passively-induced modulations in the F0 as well, as seen in Figures 1 and
2. These figures depict the same vocal sample; Figure 1 shows a smaller portion (0.5 seconds) in greater detail, while Figure 2 shows the entire 2-second sample.

**Dependent Measures**

Several physiologic and acoustic dependent measures were used to obtain additional information about both laryngeal- and output-level vibrato characteristics.

*Sound pressure level (SPL) and fundamental frequency (F0).* SPL is the mean sound pressure level during the two-second sample. F0 is the mean fundamental frequency during the two-second sample.

*EGG AM extent.* This measure is the average variation in the amplitude of the EGG speed quotient cycles, as represented in Figure 3. This figure illustrates the clear modulation of the EGG speed quotient when laryngeal amplitude modulation is present. A larger EGG AM extent indicates, on average, more variation in the vocal folds’ closed-open ratios.

*EGG AM frequency.* This refers to the rate of amplitude modulation derived from a Fast Fourier Transform (FFT) spectral analysis of the smoothed EGG speed quotient.

*EGG FM frequency.* This measure is the rate of frequency modulation derived from an FFT spectral analysis of the F0 of the EGG signal. This F0 was calculated from the period of the EGG signal. Figure 4 illustrates the FM peaks produced during a sample of abdominally-pulsed modulations, along with the FFT spectral analysis of the F0 from the EGG signal.

*Mic AM frequency.* This refers to the rate of amplitude modulation derived from an FFT spectral analysis of the amplitude trace from the Praat analysis.
Figure 1. A 0.5 second sample of modulation induced by abdominal pressure pulses. The EGG first derivative (upper pane) is shown with upward (vocal fold closing) spikes and downward (opening) spikes identified by circles. The EGG speed quotient (second pane), EGG FM (lower pane), mic AM (third pane) and mic FM (fourth pane) were all derived from an abdominally-induced laryngeal amplitude modulation recording. Laryngeal-level rhythmic changes to the vocal folds’ closed-open ratios are visible in panes 1 and 2, with corresponding passive F0 adjustments visible in panes 4 and 5. EGG in arbitrary units; see text for EGG speed quotient definition.
Figure 2. Two-second sample of the EGG first derivative, speed quotient, and FM (first, second, and fifth panes, respectively), and mic AM and FM measures (third and fourth panes, respectively) during abdominally-induced laryngeal amplitude modulation. EGG in arbitrary units; see text for definition of EGG speed quotient.
Figure 3. Sensitivity of the EGGsQ, taken from the EGG first derivative, to laryngeal-level amplitude modulation resulting from abdominally-driven pressure pulses. EGG in arbitrary units.
Figure 4. FM peaks (upper pane) and Fast Fourier Transform (lower pane) during laryngeal-level amplitude modulation produced by subglottal pressure pulses.
**Mic FM frequency.** This refers to the rate of frequency modulation derived from an FFT spectral analysis of the F0 trace from the Praat analysis. This measure was compared with both mic AM frequency and the EGG spectral analyses of FM and AM frequency to more specifically determine changes in the singers’ vibrato rate in each condition. As previously defined in the literature (Horii, 1989a; Horii, 1989b), however, mic FM frequency will for the present study be the primary measure used when discussing vibrato rate.

**EGG AM COV.** This is a measure of the variability in the duration of each modulation cycle in the EGG speed quotient. First, the modulating peaks in the signal were identified, then the duration between peaks was measured; from this list of durations, the standard deviation (SD) was calculated, and a coefficient of variation (COV) was computed by dividing the standard deviation by the mean modulation period. The COV has been applied here to both EGG and mic signals of amplitude and frequency modulation to compare their degrees of periodicity across singers as well as pitch and loudness conditions.

**EGG FM COV.** This is a measure of variability in the duration of each modulation cycle in the F0 of the EGG signal.

**Mic AM COV.** This is a measure of the variability in the duration of each modulation cycle in the amplitude trace from the Praat analysis.

**Mic FM COV.** This is a measure of the variability in the duration of each modulation cycle in the F0 trace from the Praat analysis.

**AM mic-EGG correlation.** This measure is the Pearson correlation between the EGG speed quotient and the amplitude trace from the Praat analysis.
FM mic-EGG correlation. This measure is the Pearson correlation between the F0 traces from the EGG signal and the Praat analysis.

AM mic-FM mic correlation. This measure is the Pearson correlation between the amplitude trace and F0 trace, both from the Praat analysis of the microphone signal.

Mean vowel airflow. Mean vowel airflow was measured in liters per second for the same two seconds that were analyzed in the unmasked conditions.

Perceptual Ratings

Three Brigham Young University voice professors from the School of Music in the College of Fine Arts and Communications rated vibrato samples from each of the 17 participating singers. J. Arden Hopkin, DMA, Professor of Voice and Division Coordinator, Clayne W. Robison, DMA, Professor of Voice, and Ruth M. Christensen, DMA, Assistant Professor of Voice, listened to de-identified samples from each singer presented in random order. In addition, repeats of eight of the singers’ vibrato samples were randomly inserted between the 17 samples to assess intra-rater reliability. This resulted in a total of 25 vibrato samples, each of which contained three tokens: the singer’s second soft D5, second comfortable D5, and second loud D5 tokens. These samples included the entire token, including the onset and offset of voicing, rather than the middle two seconds only. Using a scale of 1-5, the raters were asked to judge three aspects of vocal vibrato: consistency, speed, and width. For consistency, 1 = inconsistent vibrato, and 5 = most consistent vibrato. For speed, 1 = slow vibrato, 3 = optimal speed, and 5 = fast vibrato. For width, 1 = narrow vibrato, 3 = optimal width, and 5 = wide vibrato.
**Intra-rater reliability.** To determine intra-rater reliability, the raters’ first and second ratings of the repeat samples were compared to each other. This was done to determine whether the raters gave a sample the same rating both times it was presented, or, if not, how much the two ratings differed. These repeat ratings were categorized as being the same, or as being within one, two, or three points of each other. More than 82.5% of the ratings were identical or within one point.

**Inter-rater agreement.** A mean across all three professors’ ratings was calculated; the standard deviation around this mean, 0.79, was a reflection of rater agreement.
Results

Physiologic and Acoustic Data

Dependent Measures

SPL and F0. The mean SPL (at 100 cm) was 76.8 dB for soft, 81.6 dB for comfortable, and 84.4 dB for loud vowels. The mean F0 levels for the three pitch conditions were 388.7 Hz (G4), 586.9 Hz (D5), and 883.4 Hz (A5).

EGG AM extent. The degree of EGG AM present during each loudness condition is illustrated in Figure 5. As the singers’ loudness increased, EGG AM generally decreased. The degree of EGG AM present during each pitch condition is shown in Figure 6; as pitch increased, EGG AM decreased slightly. The soft condition for D5 had the highest mean magnitude of EGG AM (0.517), which was more than double the amount for the lowest rate of EGG AM, occurring in the loud condition for D5 (0.242). The three soft conditions had the highest degrees of EGG AM, while comfortable and loud conditions had the lowest.

Among individual singers, as shown in Figure 7, the degree of EGG AM varied. The range of individual singers’ mean EGG AM across conditions was 0.187-0.493. However, there was considerable variation in the degree of EGG AM within each singer’s 54 tokens. For example, Figure 8 illustrates the token during which the highest degree of EGG AM was present (0.881), while Figure 9 illustrates the token during which the lowest degree of EGG AM was present (0.027). Those singers who had the highest mean EGG AM tended to also have the widest ranges of EGG AM, such as Singer 12, whose mean EGG AM across conditions was 0.493, and whose range of EGG AM was 0.148-
Figure 5. Singers’ mean EGG AM during soft, comfortable, and loud singing conditions.
Figure 6. Singers’ mean EGG AM for each pitch condition.
Figure 7. Individual singers’ mean EGG AM across pitch and loudness conditions.
Figure 8. One of Singer 10’s comfortable G4 tokens, during which maximum EGG AM was present (0.8812). EGG in arbitrary units; see text for definition of EGG speed quotient.
Figure 9. One of Singer 4’s comfortable A5 tokens, during which minimum EGG AM (second pane) was present (0.0273). EGG in arbitrary units; see text for definition of EGG speed quotient.
0.818. Those with the lowest mean EGG AM tended to also have the narrowest ranges of EGG AM, such as Singer 13, whose mean EGG AM across conditions was 0.187, and whose range of EGG AM was 0.053-0.373. Of 153 unmasked sets of 3 tokens (9 each from 17 singers), 12 sets had an average EGG AM of <0.1, nine of which occurred during the A5 pitch condition, and three in the D5 pitch condition. Six sets of tokens had an average EGG AM of >0.8; five occurred in the soft D5 condition, and one occurred in the comfortable G4 condition.

*Mic FM frequency.* To better understand the effects of pitch and loudness on vibrato rate, mic FM frequency data were examined in each pitch and loudness condition. As shown in Figure 10, as loudness level increased, mean rate typically increased. In the unmasked G4 condition, however, mean rate remained approximately the same for both the comfortable and loud conditions; likewise, in the masked G4 and A5 conditions, mean rate stayed approximately the same for both the soft and comfortable conditions. In one condition (D5), mean vibrato rate fell as singers increased from the comfortable to the loud level.

As shown in Figure 10, as pitch increased by fifth intervals from G4 to D5 to A5, vibrato rate typically increased as well. The range, however, of vibrato rate across singers also tended to be the widest at the highest pitch. One exception was that in the soft condition, mean vibrato rate decreased from G4 to D5 before increasing again during A5. Out of 153 samples, 29 were outliers (19%); those with higher rates (10/29) occurred at G4 and A5. Those with lower rates (19/29) occurred nearly equally during all three pitches. Outliers were considered to be those which, upon examination of the scatter plots for each condition, were clearly well-removed from the others.
Figure 10. The effect of loudness and pitch conditions on vibrato rate.
Among individual singers, three participants’ vibrato rates increased along with increased pitch and loudness, while ten singers’ rates stayed generally consistent, and four singers’ rates varied in each condition. Singer 2 had the widest range of vibrato rate change across conditions (3.0 Hz), while Singer 7 had the smallest range across conditions (0.2 Hz). The singers’ average change in vibrato rate was 0.96 Hz.

*Mic AM frequency.* Mean mic AM rate was high for some singers due to instances of double-peaked AM. Figure 11 illustrates one of these samples. Half of the double-peaked AM samples were nearly equally divided between Singers 12, 14, and 16, while Singers 3 and 9 had none. Each of the nine conditions contained at least some double-peaked AM samples, from 5 (soft D5) to 18 samples (loud D5).

*COV.* The mean, SD, and range for each COV dependent measure are presented in Table 2. Both AM measures have approximately double the COV value of the FM measures, indicating lower periodicity.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGG AM COV</td>
<td>26.1</td>
<td>9.5</td>
<td>8.1</td>
<td>43.8</td>
</tr>
<tr>
<td>Mic AM COV</td>
<td>20.1</td>
<td>11.4</td>
<td>5.2</td>
<td>43.9</td>
</tr>
<tr>
<td>EGG FM COV</td>
<td>8.7</td>
<td>6.2</td>
<td>3.5</td>
<td>26.9</td>
</tr>
<tr>
<td>Mic FM COV</td>
<td>6.5</td>
<td>4.5</td>
<td>2.9</td>
<td>19.7</td>
</tr>
</tbody>
</table>

It was predicted that sinusoidal modulation would most likely occur in FM measures, which it did. Figure 12 illustrates a sample in which the singer’s mic FM COV
Figure 11. One of Singer 12’s comfortable D5 tokens, during which double peaks in amplitude modulation (second pane) occurred within single frequency cycles, visible in the third (mic FM) and fourth (EGG FM) panes.
Figure 12. FM peaks (upper pane) and FM FFT (lower pane) from one of Singer 17’s loud G4 tokens during which mic FM COV was notably high (16.8).
was notably high (16.8), while Figure 13 shows a sample in which it was notably low (1.4).

AM mic-EGG correlation. AM mic-EGG correlations of >+0.6 occurred in three singers’ samples in both the soft G4 and D5 conditions. Correlations of >-0.6 occurred in three singers’ samples in the loud D5 condition.

AM mic-FM mic correlation. Figure 14 shows a sample which had a low AM mic-FM mic correlation (-0.01). Figure 15 shows a sample which had a high negative correlation (-0.95), and Figure 16 shows a sample which had a high positive correlation (0.98). AM mic-FM mic correlations of >±0.7 occurred in 31/153 samples (20.3%). Those correlations were most likely to occur in the comfortable D5 and A5 conditions, each with six such correlations. The conditions least likely to have strong correlations occurred at G4 pitch: soft (0), comfortable (2), and loud (3).

Mean vowel airflow. As shown in Figure 17, participants’ mean airflow increased as loudness condition increased. As pitch increased, however, participants’ mean airflow stayed relatively consistent from G4 to D5, then increased for A5, except in the loud condition, when mean airflow decreased slightly from G4 to D5.

Perceptual Ratings

Participants’ mean EGG AM and consistency ratings are presented in Figure 18. Higher consistency ratings were associated with lower EGG AM; a Pearson correlation was found to be significant ($r = -0.538$, $p = .03$). No other correlations involving perceptual ratings were significant. For participants’ consistency, speed, width, and overall ratings, see Appendix A.
Figure 13. FM peaks (upper pane) and FM Fast Fourier Transform (lower pane) from one of Singer 9’s comfortable A5 tokens, during which mic FM COV was notably low (1.4).
Figure 14. One of Singer 16’s comfortable A5 tokens, during which mic AM (upper pane) and mic FM (lower pane) had a low correlation (-0.009).
Figure 15. One of Singer 4’s comfortable D5 tokens, during which mic AM (upper pane) and mic FM (lower pane) have a high negative correlation (-0.953).
Figure 16. One of Singer 5’s loud G4 tokens, during which mic AM (upper pane) and
mic FM (lower pane) have a high positive correlation (0.975).
Figure 17. Participants’ mean airflow in each pitch and loudness condition.
Figure 18. Participants’ mean EGG AM and consistency ratings. Numbered squares refer to singer number.
Discussion

The purpose of this study was to measure amplitude modulation during Western classical vibrato, and determine whether any AM might be occurring at the source (laryngeal) level in addition to the AM known to arise at the subsequent filter (vocal tract) level through the resonance-harmonics interaction (RHI). Several dependent measures provided data regarding the presence, extent, and periodicity of AM during vibrato. These measures included the use of the EGG speed quotient, which made it possible to detect source-level vocal amplitude contours and fluctuations. Modulations in amplitude detected by the EGG speed quotient were considered to be evidence of laryngeal-level AM. It was concluded that these source-level amplitude modulations were generated independently from the subsequent amplitude modulations occurring in the vocal tract as a result of the RHI. Laryngeal AM, or EGG AM, was further investigated to determine during which pitch and loudness conditions it was most likely to occur, and to what degree. This EGG AM was evaluated in light of any associations with singers’ age, length and type of training, class, and ratings for consistency, width, and speed. Also investigated were vibrato rate changes as a function of pitch and loudness conditions.

Physiologic and Acoustic Data

Resonance-Harmonics Interaction

As a singer’s F0 and its harmonics coincide with the formants of a particular vocal tract configuration, their individual amplitudes are either attenuated or enhanced. During vibrato the F0 and harmonics rhythmically and proportionally fluctuate at a rate of approximately 4-7 Hz, sweeping back and forth across the formants, which causes their individual amplitude levels to fluctuate accordingly. This rhythmic attenuating and
enhancing of harmonic amplitude levels as they coincide with formants constitutes the AM that results from vocal tract filtering. It was theoretically modeled and also identified in singers’ voices by Horii and Hata (1988) and is known as the resonance-harmonics interaction (RHI). This process is considered the main source of amplitude modulation in vibrato. AM resulting from the RHI was present extensively in the vowels recorded in the present study. In instances in which very little laryngeal-level AM was detected in the EGG speed quotient, output-level AM was present to a greater degree in the microphone trace, revealing that additional AM was arising as a result of the RHI.

As predicted by the RHI model, as F0 traverses an upward contour of the vocal tract intensity contour, FM and AM will be in phase, and amplitude will increase. As F0 travels along a downward slope of the contour, FM and AM will be in anti-phase, and amplitude will decrease. The FM-AM phase relationships predicted to occur as a result of the RHI were indeed present among these singers. Instances of AM and FM both in phase and out of phase with each other were present in all conditions.

In cases where the F0 happens to sweep across a peak, or in other words travels up to the peak and then down the other side, amplitude will first increase and then decrease. Conversely, as F0 sweeps across a valley, amplitude will first decrease and then increase. These occurrences during the RHI result in two amplitude modulations for each frequency modulation (2:1 ratio), or double-peaked AM, an example of which is seen in Figure 11. Approximately 10% of the vibrato tokens in this study contained instances of double-peaked AM, occurring across singers, pitches, and loudness levels.
**Dependent Measures**

*EGG AM extent.* In addition to the AM which occurs as a result of the RHI, the presence of EGG AM, or AM which has been detected by the EGG speed quotient, can be considered direct evidence of laryngeal-level AM. In Figure 3, which depicts abdominally-induced Psub fluctuations, corresponding laryngeal-level rhythmic changes to the vocal folds’ closed-open ratios are also visible. These changes are evidence of periodic fluctuation in the intensity contour at the laryngeal level, which the EGG speed quotient can be seen to detect. EGG AM extent, or the extent of the variation in the EGG speed quotient cycles, is higher when there is more amplitude fluctuation present at the laryngeal level.

As measured by the EGG speed quotient, laryngeal-level (or EGG) AM was present in each of the 918 vibrato tokens, across singers and across pitch and loudness conditions. EGG AM extent was greatest in all three soft conditions. The D5 comfortable and loud conditions had, on average, the smallest mean EGG AM extents, approximately half the size of the greatest mean EGG AM extent, found in the soft D5 condition. Three-fourths of the tokens with the lowest EGG AM extents occurred at pitch A5, while almost all of the tokens with the highest EGG AM extents occurred at the soft D5 pitch. These findings suggest that modulation of amplitude as well as frequency is occurring at the laryngeal level, and to a greater extent during soft conditions. Contrary to the assumption that the only laryngeal-level fluctuation is in the singer’s fundamental frequency, the EGG speed quotient did detect fluctuations in amplitude occurring in addition to those arising from the RHI alone, and these instances of laryngeal-level AM occurred regularly.
It is possible that laryngeal-level AM may be driven by respiratory factors such as Psub; however, the present study lacked the necessary invasive procedures to determine with certainty if respiratory forces are directly generating laryngeal-level AM. Given the higher EGG AM extents associated with softer singing conditions, another possibility is that decreased vocal fold tension, associated with lower muscular resistance, may introduce more susceptibility to fluctuations in amplitude at the laryngeal level. Further, it is known that as a consequence of CT activity during pitch fluctuations, tension in the thyroarytenoid muscles of the vocal folds increases and decreases. It may follow that as tension in the vocal folds changes, the amplitude of vocal fold excursion may passively change accordingly. If vocal folds are stretched tightly, vibratory amplitude will be smaller for the same level of respiratory drive, but when vocal folds are looser, the amplitude of vibration may be greater.

The singers in this study were relatively close in age, and shared similarities in the types and lengths of voice training received, especially once admitted to the university voice program, and yet their levels of EGG AM were fairly unpredictable. There was no apparent association between a singer’s class, age, or years or type of training, and the presence or extent of EGG AM in her vibrato samples. Singers with differing levels of EGG AM were found across all levels of training and experience, and each singer had a range of EGG AM extents which varied under different conditions, rather than remaining at a relatively stable level across conditions. These findings suggest that EGG AM may be an unconsciously or passively-occurring laryngeal phenomenon, subject to pitch and loudness demands and dependent more on individual physiological makeup than on deliberate training or technique.
Within the classical style, there appear to be several ways to achieve pleasing vibrato. The presence, to a greater or lesser degree, of EGG AM across pitch and loudness conditions and among singers in fluctuating degrees supports the theory of motor equivalence. According to this theory, the different singers are most likely achieving vibrato through the execution of slightly different motoric patterns which nonetheless result in the desired perceptual standards for traits such as extent, rate, and periodicity.

**Mic FM frequency.** As previously documented in the literature, (Dromey et al., 2003; Duncan et al., 2000; Folger, 2002; Troup, 1982), vibrato rate can be volitionally modified when the singer attempts to blend with other choral singers, match duet partners, or meet task demands. In the present study, in which singers were simply asked to sing tokens by themselves at certain pitches and loudness levels without reference to vibrato rate, individual vibrato rates nonetheless passively fluctuated up to 3.0 Hz as a function of changes in pitch or loudness. As pitch increased, mean vibrato rate nearly universally increased. As loudness level increased at the two lower pitches, G4 and D5, mean vibrato rate stayed relatively consistent, while at the highest pitch, A5, it increased. The increases in vibrato rate as pitch increased may have been due to the increasing elongation and tension of the vocal folds associated with higher frequencies, creating faster vibrato cycles. At the lower two pitches, increasing amplitude alone did not stiffen the vocal folds enough to significantly increase rate, but at the highest pitch, when vocal folds were already stiffened, increasing amplitude as well may have added additional tension to increase rate. In addition, the fact that the jump in frequency between the higher pitches (approximately 300 Hz) was greater than that between the two lower pitches (approximately 200 Hz) may have been a factor.


**Mic AM frequency.** Because AM is considered to be mostly a derivative of FM through the RHI, it is not as consistently periodic in nature as FM. The amplitude of F0 and each harmonic are continually changed as they encounter peaks and valleys along the resonance contour in the vocal tract, leading to more flux and less sinusoidal modulations than those found in FM. The instances of double-peaked AM predicted by the RHI, in which amplitude both rises and falls within one frequency cycle as it travels over both sides of a peak or valley in the resonance contour, also contribute to reduced periodicity in AM. Vennard (1967) suggested that double-peaked amplitude modulations may highlight the peak or valley, rather than the mean pitch; in this study, half of the instances of double-peaked AM were attributed nearly equally to three of the four singers who were ranked highest for vibrato width. The predominance of double-peaked AM in their samples may have highlighted the parameters of their vibrato widths so that they became perceptible in the ears of experienced judges.

**COV.** The EGG and microphone measures of AM had COVs approximately twice as large as those of FM, indicating lower periodicity; this is consistent with previous research findings of AM irregularity (Mason, 1965; Seashore, 1947). The two FM measures tended to be fairly periodic across conditions. Samples with higher AM COVs also had greater mean EGG AM extents, such as the soft A5 condition, while conditions with low FM COVs, as found in the comfortable D5 condition, also had lower mean EGG AM extents than most conditions. Singers with the highest COVs tended to be rated lower for consistency, and singers with the lowest COVs were also rated among those with the most consistent vibrato.
**AM mic-EGG correlation.** There were relatively few strong correlations between mic and EGG AM measures. Those few strong correlations occurred nearly equally across conditions and were equally positive and negative. These findings suggest that source-level AM goes through pronounced changes as a result of the FM-AM phase relationships and amplitude filtering in the RHI before being measured again at the output level. This finding may also be consistent with the work of Schutte and Miller (1991), in which they discovered, on slow playback of vibrato samples, that amplitude modulations of up to 10 dB per cycle were occurring as the F0 and harmonics swept across formants; this led them to conclude that during vocal tract filtering, interactions of a more complex nature and greater magnitude were occurring than the smoothed output-level amplitude, or final product, revealed. This complex, dynamic activity taking place in the vocal tract filtering process led them to distinguish the marked differences between the original resonances of the F0 and its harmonics, and the greatly altered “realized,” or ultimate, composite resonance at the lips (Schutte & Miller, 1991, p. 218).

**AM mic-FM mic correlation.** Large and Iwata (1971) found that frequency and amplitude were more likely to be in phase during higher pitch conditions, consistent with results of the present study, in which the highest correlations between mic AM and FM were found in the higher ranges (D5 and A5), with the fewest number of strong correlations occurring at G4 in the lower middle range. It could be that AM and FM are more likely to be in phase, or in other words, F0 and its harmonics are more likely to be traveling along an upward slope of the vocal tract’s intensity contour, during higher frequencies when the vocal folds are more elongated and tense and intensity is greater.
During these conditions, EGG AM extents were also the lowest, suggesting less variation in amplitude when the vocal folds are more tightly stretched.

*Mean vowel airflow.* There did not appear to be any association between increase in airflow and increase in EGG AM; instead, airflow increased linearly as loudness increased, and in most cases, loud conditions had among the lowest rates of EGG AM.

While it was not measured in the present study, Psub could potentially be a main contributor to amplitude modulation, because it influences glottal configuration, resistance, and closure type and speed, which may also affect EGG speed quotient values. Rothenberg et al. (1988) suggested that instances of abdominally-induced vibrato, which are driven by volitional Psub fluctuations, can be found even among Western classical singers who have been trained to use laryngeally-mediated vibrato, in which the CT is primarily responsible for FM, with AM derived from it. The use of an alternate means of generating AM, such as abdominally-induced Psub pressure fluctuations, should not be excluded as a possibility among classically-trained singers, especially those with less experience. Increased EGG AM extents may not always be a function of pitch or loudness condition; they may at times be influenced by Psub fluctuations as well, which may be contributing to the fluctuations in laryngeal-level amplitude found in the EGG speed quotient.

*Perceptual Ratings - Consistency*

Singers’ speed, width, and consistency were judged by voice professors, and then compared to their mean EGG AM extents. Lower EGG AM at the laryngeal level was associated with higher consistency ratings at the perceptual level. There were no significant correlations between singers’ ratings and their age, class, or length or type of
training; both high- and low-ranking singers were found in all classes. Singers with the least EGG AM were also the highest-ranked for consistency, suggesting that the degree of laryngeal AM may be a factor in the perception of vibrato consistency and its overall beauty.

Directions for Future Research

Evidence of the phase relationships and AM predicted by the RHI was found among the participants’ vibrato samples in this study. In addition, the EGG speed quotient was considered to be sensitive to vocal fold closure and opening, allowing measurement of laryngeal-level amplitude, and providing a more complete model of the entire process of modulations in amplitude which includes laryngeal contributions, as well as the supralaryngeal contributions of the RHI. While Psub is generally regarded as the main component of amplitude control, it was not possible in the present study to use invasive procedures such as tracheal puncture to better understand the role of Psub in driving amplitude modulation. More in-depth study is needed regarding the roles of not only Psub, but of glottal configuration and resistance in AM during vibrato as well. Similarly, electromyography, a direct but invasive measure of muscular activity, was not possible in the present study, but might have revealed important details of laryngeal muscle activation.

Additionally, vibrato rate was found to passively increase as pitch increased, and at the highest pitch, as amplitude increased as well. It may be useful to assess the effects of pitch and loudness in other registers, among male singers, or during singing styles other than Western classical opera to better understand underlying physiological mechanisms involved in passive vibrato rate change.
Although they were not given a time limit during which to judge the samples, and they were allowed to listen to samples more than once, the judges in this study were asked to rate three aspects of vibrato quality (consistency, speed, and width) simultaneously, which may have influenced their ratings. Rather than following students longitudinally, the authors relied on interviews to ascertain singers’ changes to vibrato as a result of training and maturation. Perhaps if the singers were to be followed longitudinally, pattern changes in their EGG AM may become apparent over time as a function of age or training, leading to a better understanding of whether laryngeal-level AM is ‘hard-wired,’ or may be shaped over time through training or maturation. Whether changes to laryngeal AM were occurring passively, subconsciously, or volitionally might be considered. If a relatively homogenous group of singers such as this one were to be compared to other singers who had been trained with distinctly different techniques or emphasis, more distinct patterns of EGG AM extent might emerge.

While it is uncertain that an optimal level of AM, either at the source or filter level, would in itself be consistently perceptible to audiences, researchers have previously emphasized the crucial role that AM plays in the perception of vibrato beauty and richness, encouraging further direct study of how it is produced. It is unclear whether AM, because it appears to be mostly a passive by-product in the Western classical style, is a vocal attribute that could be directly trained or manipulated to the same degree as other more volitional aspects of vibrato quality. It may be more likely that a desirable level of AM may be achieved indirectly through training in breath support, locating and singing within the singer’s ‘ring,’ and other techniques.
In his vibrato samples, Horii (1989a) found instances of AM that were not explained by the RHI; these instances were either very low or high in pitch or loudness. While it is not clear which of the instances of laryngeal-level AM in the present study would be considered as unexplained by the RHI alone, it is clear that both laryngeal-level and RHI-generated AM occurred regularly. The next step would be to determine, in the output-level AM, its laryngeal (source) and vocal tract (filter) contributions.

Similarly, Schutte and Miller (1991) found amplitude fluctuations of up to 10 dB as a result of the RHI; another step would be to translate EGG AM extents derived from the EGG speed quotient into decibels of AM change. For a given mean EGG AM extent, what range of fluctuation in dB does it represent? Whether AM always begins with an increase, rather than a decrease, in amplitude, as Horii (1989b) discovered with initiation of FM, is another aspect of the mechanism of AM that has yet to be fully understood.

It is generally acknowledged that amplitude is not controlled independently from frequency; understanding more specifically the mechanisms of laryngeal-level AM would provide information for the entire laryngeal context in which it occurs. This context includes neural and muscular activity, rapid adjustments to glottal configuration, and fluctuating Psub, F0, harmonics, and timbre. This laryngeal context is also where vocal tremor occurs. Since tremor is thought to be characterized by a greater degree of AM than vibrato (Imaizumi et al., 1993), further study of laryngeal-level AM and its interactions with other contributing factors may have implications for the study and treatment of tremor as well.

The role of auditory feedback in inducing laryngeal-level AM could be studied by introducing amplitude stimulation similar to the frequency stimulation used to elicit the
pitch-shift reflex (Leydon et al., 2003). If a slight modulation in amplitude is introduced in the singer’s ear, would it induce the singer to adjust his or her amplitude correspondingly, and if so, would this change take place laryngeally or supralaryngeally, and be mediated centrally or peripherally?

The pursuit of questions such as these and others would offer glimpses into the intricate activity at the vocal folds, diminutive tissues from which a vast array of spoken exchanges, beautiful songs, and compelling stories flow.
References


Appendix A

Professors’ mean ratings of the singers along with participants’ mean EGG AM are shown in Table A1.

Table A1

<table>
<thead>
<tr>
<th>Singer</th>
<th>Consistency</th>
<th>Speed</th>
<th>Width</th>
<th>Mean EGG AM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.00</td>
<td>3.67</td>
<td>3.17</td>
<td>0.395</td>
</tr>
<tr>
<td>2</td>
<td>2.67</td>
<td>2.33</td>
<td>2.83</td>
<td>0.274</td>
</tr>
<tr>
<td>3</td>
<td>3.67</td>
<td>3.67</td>
<td>3.00</td>
<td>0.263</td>
</tr>
<tr>
<td>4</td>
<td>3.83</td>
<td>3.50</td>
<td>3.50</td>
<td>0.274</td>
</tr>
<tr>
<td>5</td>
<td>1.83</td>
<td>3.00</td>
<td>3.33</td>
<td>0.311</td>
</tr>
<tr>
<td>6</td>
<td>2.67</td>
<td>3.33</td>
<td>3.17</td>
<td>0.259</td>
</tr>
<tr>
<td>7</td>
<td>2.83</td>
<td>2.67</td>
<td>3.17</td>
<td>0.290</td>
</tr>
<tr>
<td>8</td>
<td>2.00</td>
<td>2.83</td>
<td>3.50</td>
<td>0.456</td>
</tr>
<tr>
<td>9</td>
<td>3.50</td>
<td>3.00</td>
<td>3.33</td>
<td>0.383</td>
</tr>
<tr>
<td>10</td>
<td>3.33</td>
<td>3.00</td>
<td>3.33</td>
<td>0.329</td>
</tr>
<tr>
<td>11</td>
<td>2.50</td>
<td>3.00</td>
<td>3.67</td>
<td>0.450</td>
</tr>
<tr>
<td>12</td>
<td>2.33</td>
<td>3.00</td>
<td>3.00</td>
<td>0.493</td>
</tr>
<tr>
<td>13</td>
<td>3.67</td>
<td>3.33</td>
<td>2.83</td>
<td>0.187</td>
</tr>
<tr>
<td>14</td>
<td>2.00</td>
<td>4.00</td>
<td>3.00</td>
<td>0.439</td>
</tr>
<tr>
<td>15</td>
<td>2.83</td>
<td>2.33</td>
<td>3.17</td>
<td>0.333</td>
</tr>
<tr>
<td>16</td>
<td>2.50</td>
<td>3.33</td>
<td>3.00</td>
<td>0.284</td>
</tr>
<tr>
<td>17</td>
<td>2.00</td>
<td>3.50</td>
<td>2.00</td>
<td>0.374</td>
</tr>
</tbody>
</table>
Appendix B

Those singers who were rated as the top or bottom five singers in at least two categories were considered the highest- and lowest-rated singers, respectively, as shown in Table B1.

Table B1

_Singers Rated Highest and Lowest Overall and for Consistency, Speed, and Width_

<table>
<thead>
<tr>
<th>Singers</th>
<th>Consistency</th>
<th>Speed</th>
<th>Width</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest-rated</td>
<td>4</td>
<td>5</td>
<td>12</td>
<td>3, 9, 10, 12, 13</td>
</tr>
<tr>
<td>Lowest-Rated</td>
<td>5</td>
<td>15</td>
<td>17</td>
<td>8, 14, 17</td>
</tr>
</tbody>
</table>
Appendix C

Consent to Be a Research Participant

Introduction: You have been invited to take part in a research study which will be conducted to learn more about amplitude modulations during vocal vibrato. Your participation will help us understand how singers physically increase and decrease the loudness of their voices several times each second during vibrato. This study is being conducted by Lorie Reese, a graduate student at Brigham Young University, under the direction of Dr. Christopher Dromey, an associate professor in the Audiology and Speech-Language Pathology Department. Because you are a healthy, Western-classically trained singer admitted to the Vocal Studies Division of the BYU Music Department, you were selected for participation.

Procedures: You will be asked to attend a recording session lasting approximately one hour. While standing in a sound booth in 106 TLRB, you will sing 5-second samples of the vowel /a/ in combinations of three different pitches and loudness levels, for a total of 27 samples. You will then be asked to repeat these samples while wearing a flow mask, which measures the volume of air you are exhaling while singing. While singing, you will wear a comfortable head-mounted microphone and a small, adjustable, velcro electroglottography (EGG) band on your neck, which holds two electrodes about the size of a quarter. These detect vocal fold vibration. Audio recordings will be made, which will then be analyzed with a computer program.

Risks/Discomforts: There are no known 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: You will receive no direct benefits from participating in this study. However, the results of this study are expected to provide valuable information regarding how amplitude modulations are generated during vibrato. This information may eventually help us understand how fluctuations in amplitude contribute to vocal beauty.

Confidentiality: An anonymous identification number will be used to identify and organize the recordings of each singer. Your name and other identifying information will not be used in print or electronic records of this study, other than on this consent form.

Participation: Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate entirely without jeopardy to your status at
the university, in the Vocal Studies Division of 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 Dr. Renea Beckstrand, IRB Chair, at (801) 422-3873, or 422 SWKT, or renea_beckstrand@byu.edu.

**Signatures:** I have read the above and 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                                            Age