Registration Strategies of Professional Operatic Mezzo-Sopranos

D.M.A. Document

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Musical Arts in the Graduate School of The Ohio State University

By

Katherine Celine Osborne, B.M, M.V.P.

Graduate Program in Music

The Ohio State University

2015

D.M.A. Document Committee:

Scott J. McCoy, Advisor

Aharon E. Freud

David Huron

Robert J. Ward
Abstract

Operatic mezzo-sopranos are required to sing in their low to middle pitch ranges more frequently than other female voice types, with minimal perceptual evidence of tonal change or irregularities, at substantial intensity, and with musical finesse and flexibility. To observe register behaviors in professional operatic mezzo-sopranos, audio and EGG data were collected from two subjects who sang ascending and descending chromatic scales between C4 and G4 on vowels /a/, /i/, /o/, and /u/. General adductory and resonance patterns were observed through EGGW50 measurements and spectral analysis. Register transitions were located and defined through criteria based on perception, EGG signal intensity changes, and vibrato rate and extent pattern disturbances. Vibrato rate and extent pattern disturbances occurred in 79% of Subject 1’s samples with an identifiable transition (19 of 24 task tokens) and 80% of Subject 2’s samples with an identifiable transition (4 of 5 task tokens).

Subject 1 displayed audible, objectively identifiable evidence of register transitions in 24 of 24 samples. Subject 2 displayed audible, objectively identifiable evidence of register transitions in 5 of 24 samples. Transitions were located between C#4 and F4 and most frequently between D#4 and E4. Spectral analysis suggests that the subjects use different resonance strategies in this pitch range.
Contrasting EGGW50 value patterns suggest that the two subjects may have a different adductory strategy. The ANOVA for Subject 1 revealed significant interactions between phrase location, the position of the data point relative to transition, and EGGW50 values, indicating that Subject 1 may employ a different adductory strategy in ascending phrases as compared with descending phrases.

Keywords: vocal register, modal, falsetto, transition, glottal adduction, electroglottography
Dedication

This document is dedicated to my husband, James Jenkins, son, Ian Jenkins, and my parents, Carol and Richard Osborne.
Acknowledgments

My sincerest thanks to my graduate advisor at OSU and Westminster Choir College, Scott McCoy, who has steadfastly supported my academic and artistic goals with his guidance, expertise, patience, and humor. This document would not have been possible without the time he invested from the project planning and equipment setup stages through the editing and submission process.

My sincerest thanks to Dr. Ronald Scherer and Brittany Frazer at Bowling Green State University who spent hours with me over the past months refining my data analysis procedures, Voice Foundation presentation, and the preparation and editing of this document. Their commitment to scientific rigor and the use of precise language has been an education and inspiration to me.

My sincerest thanks to Dr. David Huron, Dr. Robin Rice, Dr. Michael Trudeau, and Dr. Robert Ward for serving on my graduate committee during the past three years. Thank you for supporting my academic, artistic, and research interests.

My sincerest thanks to Ziyue Chen, PhD Student in The Ohio State University Department of Statistics, who completed the statistical analysis for this project.

I am grateful for the project support provided through the 2014 Van Lawrence Fellowship, awarded by the Voice Foundation and National Association of Teachers of Singing (NATS).
And, finally, my immeasurable gratitude to my husband, James Jenkins, who gained a few wrinkles and gray hairs along with me during this process. His careful editing, personal support, and extra caregiving time given to our son, Ian, during the completion of this project and document was a true labor of love.
Vita

2013 – 2015.................................Graduate Teaching Associate, School of Music, The Ohio State University

2013-2014.....................................Adjunct Voice Instructor, Ohio Wesleyan University

2007-2012 ....................................Head of Voice Studies, Washington National Cathedral

2004-2012 .....................................Voice Instructor, Suitland High School Visual and Performing Arts Program

2010.............................................Master of Voice Pedagogy, Westminster Choir College

1997.............................................Bachelor of Music in Voice Performance, Stetson University

Fields of Study

Major Field: Music
Table of Contents

Abstract ........................................................................................................................................ ii
Dedication .................................................................................................................................... iv
Acknowledgments .......................................................................................................................... v
Vita ................................................................................................................................................ vii
List of Tables .................................................................................................................................. x
List of Figures ............................................................................................................................... xi

Part One: Introduction .................................................................................................................. 1
  Vocal Register Transitions: Voice Pedagogy ............................................................................. 3
  Vocal Register Definitions: Voice Science .................................................................................. 5
  Register Transitions: Singing Research ..................................................................................... 8
  Need for the Study ....................................................................................................................... 10
  Purpose of the Study .................................................................................................................. 11
  Research Questions .................................................................................................................... 12
  Limitations .................................................................................................................................. 12

Part Two: Review of Literature .................................................................................................. 13
  Voice Source ............................................................................................................................... 13
  Subglottal Resonance .................................................................................................................. 20
  Vocal Tract Resonance and Configuration .............................................................................. 21
  Nonlinear Source–Filter Interaction ........................................................................................... 23
Register Smoothing Considerations and Behaviors.......................................................... 24
Transition Pitch Locations................................................................................................. 27

Part Three: Method ............................................................................................................. 29
Subjects ............................................................................................................................. 29
Equipment ......................................................................................................................... 29
Protocol............................................................................................................................... 31
Observational Study Design............................................................................................... 32
Analysis............................................................................................................................... 33

Part Four: Results and Discussion ...................................................................................... 41
Transition Pitch Locations.................................................................................................. 41
Subharmonics ..................................................................................................................... 43
Descending Chromatic Scale Task: /a/ Vowel................................................................. 44
Ascending Chromatic Scale Task: /a/ Vowel.................................................................... 48
Descending Chromatic Scale Task: /i/ Vowel................................................................. 52
Ascending Chromatic Scale Task: /i/ Vowel.................................................................... 56
EGGW50 Analysis................................................................................................................ 60
Discussion........................................................................................................................... 65

Part Five: Conclusions ........................................................................................................ 72

Bibliography ...................................................................................................................... 76
List of Tables

Table 1. Registration Terms by Voice Range................................................................. 7
Table 2: Transition Pitch Locations by Subject and Task ............................................. 28
Table 3. Subject 1: ANOVA Effects with P–Values....................................................... 61
List of Figures

Figure 1. Recording System Schema ................................................................. 30
Figure 2. Ascending and Descending Chromatic Scale Tasks ....................... 31
Figure 3. Example of an EGGSI Trace with Transition in Brackets ............... 35
Figure 4. EGGW50 Definition ........................................................................... 35
Figure 5. Subject 1: EGGW50 Scatterplot, EGG Waveforms, and EGGSI Trace ... 37
Figure 6. DEGG Wavegram Construction Process ........................................ 39
Figure 7. Vibrato Rate and Extent .................................................................... 40
Figure 8. Subject 1: Transition Pitch Location Frequency Graph .................... 42
Figure 9. Subject 2: Transition Pitch Location Frequency Graph .................... 43
Figure 10. Subject 2: Alternating Adjacent Waveshapes (/a/ Descending Task) .... 44
Figure 11. Subject 1: Alternating Adjacent Waveshapes (/æ/ Descending Task) .... 44
Figure 12. Subject 1: /a/ Descending Chromatic Task (Token 2) .................... 46
Figure 13. Subject 2: /a/ Descending Chromatic Task (Token 1) .................... 47
Figure 14. FFT Spectral Slices (/a/ Descending Task) ..................................... 48
Figure 15. Subject 1 - /a/ Ascending Chromatic Task (Token 3) ................... 49
Figure 16. Subject 2 - /a/ Ascending Chromatic Task (Token 1) ................... 51
Figure 17. FFT Spectral Slices (/a/ Ascending Task) ....................................... 52
Figure 18. Subject 1 /i/ Descending Chromatic Task (Token 2) .................... 53
Figure 19. Subject 2 /i/ Descending Chromatic Task (Token 2) .................................. 54
Figure 20. FFT Spectral Slices (/i/ Descending Task).................................................. 55
Figure 21. Subject 1 /i/ Ascending Chromatic Task (Token 2) ................................... 57
Figure 22. Subject 2 /i/ Ascending Chromatic Task (Token 3) ................................... 58
Figure 23. FFT Spectral Slices (/i/ Ascending Task).................................................... 59
Figure 24. Subject 1: EGGW50 Values for Descending Tasks................................... 62
Figure 25. Subject 2: EGGW50 Values for Descending Tasks................................... 63
Figure 26. Subject 1: EGGW50 Values for Ascending Tasks..................................... 64
Figure 27. Subject 2: EGGW50 Values for Ascending Tasks..................................... 65
Part One: Introduction

The great art of the singer is to render imperceptible to the listener or watcher, the greater or lesser degree or difficulty with which he brings forth the two different registers of chest and head. This can be obtained only through endless refinement: but it is not easy to master this in a simple and natural method. One must use study, effort, and industry to correct the defects provided by the greater or lesser strength of the organs, and one reaps a management and economy which render the voice equally sonorous and pleasing, which few scholars reach, and of which few masters understand the practical rules, or how to execute them (Mancini, as cited in Stark, 1999, p. 101-2).

Giovanni Battista Mancini’s 1774 singing treatise, Pensieri e reflessioni pratiche sopra il canto figurato, was written primarily for the cultivation of the male castrato voice, but the above excerpt remains a concise description of arguably the most challenging and least understood issues of the female operatic singer throughout history: how to sing in the lower middle voice with sufficient loudness and resonance to be audible with simultaneous orchestral accompaniment and/or other operatic singers, while satisfying the public’s expectation of precise intonation, flexibility, tonal richness, expressive phrasing, intelligibility, and a smooth, inconspicuous register transition.
Considering the acoustic and physiological instabilities that occur in the 200-550 Hz range of the male and female voice, one might expect some allowances to be made for singers in difficult situations, but critics have not excused registration-related difficulties in singers. Leigh Hunt (as cited in Stark, 1999), who wrote performance reviews for *The Examiner* in the early 19th century, criticized tenor Thomas Phillips for having an “unpleasant gurgle to jump over” in “the passage between his natural voice and falsetto” which he has to “slip over…like a bump in the ice” (p. 63) Rossini’s biographer, Stendhal (as cited by Emerson, 2005), though greatly enamored with the artistry of legendary singer and Rossini specialist, Giuditta Pasta (1797 – 1865), acknowledged her critics’ views that her voice was “entirely uneven throughout its range” with “timbral variations” a “veiled, husky quality,” and “harshness,” but defended these very imperfections as a tool of her expressive and dramatic greatness (p. 113).

As the era of the castrato ended and opera composers conceived works with ever-larger orchestras in the 19th century, male and female operatic singers were obliged to develop new vocal strategies and pedagogies to compete effectively with more-substantial ensembles while meeting the audience’s demand for a more thrillingly visceral vocal quality. Concurrently, male singers increasingly were expected to extend the brilliant quality characteristic of the “chest” or “modal” voice upward, requiring them to relinquish the technique of transitioning to falsetto or voix mixte. The new, ringing high notes of the modern operatic male singer required crucial acoustic adjustments, but not a physiological adjustment at the voice source that would establish a new vocal fold vibratory pattern. This development left female singers and modern countertenors as the
major negotiators of both physiologic and acoustic register transitions in the 200 – 550 Hz range.

**Vocal Register Transitions: Voice Pedagogy**

At the crossroads of *bel canto* and a developing modern vocal technique for the operas of Giuseppe Verdi and Richard Wagner was famed pedagogue, Manuel Garcia II, who rejected the incompleteness of the two-register theory of the Italian era in favor of a three-register theory for female voice that included a separate “middle voice” designation. The key to maneuvering between registers, in Garcia’s opinion, was firm glottal closure. The female singer was admonished not to “give in to the tendency to aspirate the falsetto tones [middle voice] at the moment when she leaves the chest register” to avoid inefficiency, a dulled tone quality, and loss of vocal power (as cited by Stark, 1999, p. 72-73).

One of the first famous female pedagogues, Matilde Marchesi (1821-1913), who taught many of the greatest female singers of her generation, advised an acoustic strategy for blending the chest and middle registers in the female voice. Her method, to “close” or “darken” when singing descending passages and “open” or “make clearer” when singing ascending passages is in agreement with German pedagogue Julius Stockhausen (1826-1906), who also prescribed the *messa di voce* exercise on each note that employed both the “chest” and “head” registers to extend the intensity range (Coffin, 1989). Emma Seiler (1821-1877), who assisted German physicist Hermann von Helmholtz (1821-1894) in his vowel and registration studies, defined a series of 5 mini-registers based upon the physically perceived origin of tone, from the lower part of the lungs through the forehead.
Seiler believed that the modulated female speaking voice so often required smooth shifts between the “second chest” and “first falsetto” registers that the skill of register transition often did not require special training (Seiler, 1879).

In the 20th century, voice pedagogues advocated increasingly nuanced physiologic adaptations to address registration difficulties. William Vennard (1967), who taught famed mezzo-soprano Marilyn Horne, asserted that the most substantive skill to be acquired by a female singing student is the passage into and development of a “middle voice.” His recommendation for “the widest possible range without a break” was to make “the adjustment...heavy in the lower part of the voice...and shift [the balance] smoothly toward the lighter production as the scale is ascended” (p. 77). Richard Miller (1996) did not believe that there were two separate register mechanisms, but “changing, dynamic balances among the laryngeal muscles” that resulted in one combined register (p. 133). Miller (1997) also criticized female singers (particularly of the German School) who relied “almost entirely on head voice to negotiate the lowest extremities of the range” to avoid audible register transitions instead of working to achieve a “unified scale” (p. 149). Cornelius Reid (1990) also found register coordination to be a primary element of female voice development, advocating for each register to be identical in tone and intensity in order to “meet on terms of perfect equality” (p. 85).

The current generation of lower classified female singers performing actively on the operatic stage is expected to sing as uniformly through register transitions as singers of previous centuries. As with all opera singers, they must also sing at sufficient loudness and with sufficient resonance and spectral characteristics to compete with an operatic-
sized orchestra in a large performance space. Contralto Ewa Podleś (1952-) was praised in the *Wall Street Journal* in 1997 for her perfectly smooth register transitions without “a discernible break from the booming, tenor-like bottom to the burnished top” (Scherer, 1997). Mezzo-soprano Stephanie Blythe (1970-) is valued for her “seamless range” (von Rhein, 2001) and “plush lower and middle register(s)” (Silverman, 2012). The *Santa Fe New Mexican* proclaimed mezzo-soprano Joyce DiDonato’s (1971-) performance a triumph in a 2013 production of Rossini’s *La Donna del lago*, noting that “this piece requires singers skilled in the strenuous demands of *bel canto* vocal style, artists who can negotiate rapid figuration by the furlong, passing seamlessly from their highest notes to their lowest and vice-versa (dauntingly difficult)….DiDonato did not sing a measure all evening that failed to proclaim mastery of this style” (Keller, 2013). While there are some specialized operatic roles that allow for obvious register transitions for stylistic or character purposes (e.g. the title role in Bizet’s *Carmen* and works by Kurt Weill), critics and audiences prefer to know that this is a singer’s stylistic choice rather than a vocal necessity. Perfect command of registration has never been an absolute requirement for a career in opera, but it is considered a great virtue and indication of technical prowess.

**Vocal Register Definitions: Voice Science**

In addition to his writings as a pedagogue, Garcia made contributions to voice science through his laryngoscopic investigations. His definition of the term “register,” which appeared in his 1840 treatise *École de Garcia*, has become ubiquitous in the pedagogic and scientific literature: “The word register means, a series of consecutive and homogeneous sounds produced by the same mechanism, and different essentially from
other sounds originating in mechanical means of a different kind (Garcia, 1924, p. 4).”

On first reading, the definition sounds clear and well delineated. It has, however, sparked numerous discussions due to its specification of a mechanical framework, rather than a foundation reliant on a combination of laryngeal and acoustic factors. In contrast, Ingo Titze describes the primo and secondo passaggi of the male and female voice as a “timbre transformation,” which he defines as an “abrupt quality change” originating from alterations in the acoustic spectrum, spectral slope, and glottal adduction (p. 183). R.H. Colton and Harry Hollien’s definition is similar to Titze’s with the addition of airflow, subglottal air pressure, and perceptual variables (1973). Hollien’s 1974 article, “On Vocal Registers,” recommended four criteria for creating future register-related definitions: “I maintain that, before the existence of a particular vocal register can be established, it must be operationally defined: 1) perceptually, 2) acoustically, 3) physiologically and 4) aerodynamically” (p. 2). Nathalie Henrich, while acknowledging the continuing debate in the singing community surrounding registration, blamed not the tremendous variety of terms describing these events for the continuing confusion, but the lack of specificity in their definitions (Henrich, 2006).

The terms used to describe register events are various indeed, spanning several languages and various historical interpretations applied without standardization. Table 1 is a list of terms commonly found in the voice science and pedagogy literature to describe the registers of the voice. This list does not include either the so-called “vocal fry” or “whistle registers” found at the extremes of the voice and is certainly not exhaustive. The frequent appearance of some terms in each of the columns illustrates the need for
operational definitions when discussing aspects of registration to avoid unnecessary misunderstandings and controversies. For example, some voice teachers would say that any part of the classical male singer’s range is sung in “chest” or “modal” voice. Others would identify the high range as sung in “head” voice. However, in the female voice, the chest voice may only be referred to when discussing the low register (especially in classical singing) and head voice would describe the register used in the middle and upper ranges. The scientific community’s use of the term “falsetto” to describe the female middle and upper registers is completely rejected by some voice teachers and some voice scientists, as is voice teachers’ identification of head voice (as it is applied in the high register of a male classical singer) as a separate register by some scientific researchers.

<table>
<thead>
<tr>
<th>Low Range</th>
<th>Middle Range</th>
<th>High Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest Voice</td>
<td>Chest Voice</td>
<td>Chest Voice</td>
</tr>
<tr>
<td>Modal Voice</td>
<td>Modal Voice</td>
<td>Modal Voice</td>
</tr>
<tr>
<td><em>Voce di petto</em></td>
<td><em>Voix mixte</em></td>
<td><em>Voce di testa</em></td>
</tr>
<tr>
<td>Mode 1</td>
<td>Mode 1</td>
<td>Mode 1</td>
</tr>
<tr>
<td>Mechanism I</td>
<td>Mechanism I</td>
<td>Mechanism I</td>
</tr>
<tr>
<td>Belt</td>
<td>Belt</td>
<td>Loft Voice</td>
</tr>
<tr>
<td>Heavy Mechanism</td>
<td>Heavy Mechanism</td>
<td>Heavy Mechanism</td>
</tr>
<tr>
<td></td>
<td>Mechanism 2</td>
<td>Mechanism 2</td>
</tr>
<tr>
<td></td>
<td>Mode 2</td>
<td>Mode 2</td>
</tr>
<tr>
<td></td>
<td>Light Mechanism</td>
<td>Light Mechanism</td>
</tr>
<tr>
<td></td>
<td>Head Voice</td>
<td>Head Voice</td>
</tr>
<tr>
<td></td>
<td>Middle Voice</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mixed Voice</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Registration Terms by Voice Range
Register Transitions: Singing Research

Voice researchers view register transitions from two main standpoints: as phenomena or learned behaviors.

Register Transitions as Phenomena. The untrained singer’s register transition is often characterized by the appearance of subharmonics and instabilities in fundamental frequency ($F_0$), amplitude, and resonance. Ron Baken (2005) theorizes that these instabilities and irregularities are evidence of the voice as a fundamentally chaotic system. Instead of viewing vocal imperfection as something with an “identifiable proximate cause” (p. 233), he argues that in a chaotic system “an infinitesimal change in some system parameter can cause a significant change in the system’s behavior” such as an abrupt change from “modal register to glottal fry because of some change in laryngeal status that we could never register.” These sudden and often perceptually obvious changes are “bifurcations” that are “characteristic, and in fact, diagnostic, of chaotic systems” (p. 235). It is the element of controlled chaos in the vocal system that Baken credits with enabling it to produce a wide spectrum of phonatory possibilities (2005). Biomechanical modeling, excised human laryngeal modeling, and human studies of the transition between the chest and falsetto registers have provided evidence of laryngeal bifurcations (Švec & Pešák 1994; Švec, Schutte & Miller, 1999; Tokuda, Horáček, Švec, & Herzel, 2007; Vilkman, Alku, & Laukkanen, 1995) in the form of “period doubling, period tripling, and irregular vibrations” (Švec et al., 1999, p. 1526). Especially notable was the brief simultaneous appearance of both register vibratory systems, which researchers describe as a form of hysteresis, defined as “a one-dimensional example of a
Register Transitions as Learned Behaviors. Register transition phenomena are modified and controlled by behaviors and learned techniques, usually acquired through voice training. In addition to preventing possible bifurcations, a singer must also perceptually attempt to match the loudness and tonal quality of the registers on either side of the transition. Donald Miller and Harm Schutte (2005) identify the “abrupt change of quality and/or loudness…in the female voice at the point in an ascending scale where the chest register yields to a lighter register” as “one of the most conspicuous, and thus problematical, register transitions” (p. 279). Not all singing genres avoid these transitions – register breaks are considered an expressive tool in the current pop music and country music styles. Yodelers intentionally exploit and exaggerate the register break for artistic effect. However, it is widely accepted by the scientific and artistic community that perceptually equalized registers and a smooth transition between them are principle goals of Western classical voice training (Sundberg & Kullberg, 1999).

Interestingly, the general acknowledgement of the difficulties surrounding register blending in the female voice has not led to many studies of register transitions in trained female singers. Despite a growing number of studies examining professional male singers, countertenors, sopranos, and untrained singers, a review of the literature revealed only a few studies that included professional mezzo-sopranos as subjects (Miller & Schutte, 2005; Roubeau, Chevrie-Muller, & Arabia-Guidet, 1987; Wendler & Seidner, 1982). The number of mezzo-soprano subjects in all three of these studies combined is only four. Additionally, because a definition of “professional” was not included in the
methodologies, questions remain about whether these singers had sufficient skill to be considered highly successful practitioners with the ability to demonstrate best practices. The information yielded from studying behaviors of successful singers is of interest to researchers, singers, and singing teachers.

**Need for the Study**

Singing pedagogy and research literature provide insufficient detail about singing strategies in the low to middle pitch range of the professional mezzo-soprano voice. In addition, the data must be collected in a way that is noninvasive and comfortable enough for the subject to reasonably simulate operatic performance singing conditions. The collected information would be of interest and of use to singing teachers who are seeking pedagogic solutions to this technical and developmental challenge.

While there is some general agreement that register transitions in this part of the female voice take place in the 200-550 Hz range (Andrade, 2012; Henrich, Roubeau, & Castellengo, 2003; Howard, 1995; Keidar, Hurtig, & Titze, 1987; Miller & Schutte, 2005; Roubeau et al., 1987; Sundberg & Kullberg, 1999; Titze & Spencer, 2001), this range is quite general (spanning a Major 6th) and lacks context for intra and inter-subject variability. Additionally, there is insufficient information about the interaction of the sung tasks and the transition locations with respect to phrase direction or vowel choice. Vowel selection, particularly in this pitch range where R1 locations for some vowels

---

1 The abbreviation R1 for “resonance frequency” replaces the formerly used F1 for “formant frequency,” reflecting recent multi-disciplinary consensus about applying differentiated terminology for these two previously combined concepts (Titze et al, 2015).
approach and exceed $F_0$, may influence registration strategies because of the interactions of pitch and vowel resonances. Since the selected tasks and instructions may significantly alter the subject’s behavior, especially a subject who has a high level of skill and is, therefore, able to perform the maneuver in many ways, care must be taken in the study design to pinpoint the behaviors that the researcher desires to observe. In this way, more specific evidence can be acquired concerning the function of the mezzo-soprano operatic voice in a performance situation with an orchestra.

The over four hundred-year-old art of operatic singing owes much of its longevity to its ability to conform to contemporary musical tastes. The success of a singer is at least partially dependent on the ability to respond to audience aesthetics through vocal and musical choices. Documenting characteristics of today’s successful singers through acoustic and physiological measures also adds to the historical record of the singing voice that extends beyond a studio or live performance recording.

**Purpose of the Study**

The purpose of this study was to determine if professional operatic mezzo-sopranos display evidence of a register transition between C4 (approximately 260 Hz) and G4 (approximately 391 Hz) through the examination of alterations in electroglottographic (EGG) waveform characteristics and vibrato rate and extent patterns. These measures were used to help determine potential transition and adduction strategies used by two professional mezzo-sopranos as they sang ascending and descending chromatic scales between C4 and G4.
Research Questions

For ascending/descending phrase conditions and four vowel conditions (/a/, /i/, /o/, /u/) between the pitches of C4 and G4,

1. Are there transitions identifiable from the EGG signal intensity and waveform characteristics,

2. Are there transitions identifiable from the vibrate rate (Hz) and extent (cents),

3. Are there differences between the transition pitch locations, and

4. Are there patterns of EGGW50 measurements relative to the transitions?

Limitations

Microphone Distance. Although care was taken to standardize the distance between the subject’s mouth and the microphone through the use of two positioners and precise instructions before each recording, no video recording exists to confirm the consistency of body position relative to the microphone.

Intensity Standardization. The subjects were instructed to sing as they would in a professional performance with orchestral accompaniment, which presumably would prompt them to sing at substantial volume. That loudness level was not, however, monitored during the recording process with the goal of creating intensity uniformity between tasks and repetitions.

Electroglottography. The limitations of electroglottography as a representation of some aspects of glottal movement and adduction will be discussed in the data analysis section of Part Three: Method
This chapter surveys empirical studies investigating aspects of voice registration, with special emphasis on the transition between modal and falsetto voice. The review is organized by the physiologic or acoustic variable observed in relation to voice registration (aerodynamics, vocal tract resonance, subglottal pressure) under the main headings of (a) Voice Source, (b) Subglottal Resonance, (c) Vocal Tract Resonance and Configuration, (d) Nonlinear Source-Filter Interaction, (e) Register Smoothing Considerations and Behaviors, and (f) Transition Pitch Locations.

Because the voice pedagogy literature primarily is based on perception and sensation rather than scientific research, those works will not be included in this review.

**Voice Source**

**Glottal aerodynamics.** Studies have utilized excised canine larynges and live human subjects to examine the relationship of subglottal pressure and glottal airflow to registration phenomena and transition behaviors.

**Canine larynges.** Alipour, Finnegan, and Scherer (2009) examined modal/falsetto register shifts in excised canine larynges (n=10) with subglottal pressure and glottal airflow adjustments, reporting abrupt changes in F0 and modes of vibration. Thus, increases in pressure alone, with constant adduction, may alter the register between
modal and falsetto. Similarly, Berry, Herzel, Titze, and Story (1995) increased and decreased subglottal pressure with excised canine larynges (n=5) and observed where bifurcations, instabilities, and register transitions occurred. They reported modal-like vibrations at higher subglottal pressures than falsetto-like vibrations.

**Untrained singers.** Švec and Pešák (1994) studied an untrained baritone (n=1) who sang individual pitches while varying expiratory airflow to induce vocal breaks, subharmonics, and bifurcations between the modal and falsetto registers. The speed of airflow was theorized to be a major factor in achieving a bistable laryngeal position. Murray, Xu, and Woodson (1997) found a strong interaction between register, adduction, and mean airflow rate among male and female untrained subjects (n=8) in addition to identifying posterior glottal gaps in subjects phonating in both register modes. Aerodynamics does not appear always to be a decisive factor in register control in other studies, especially those that employ untrained singers. Kitzing, Carlborg, and Löfqvist (1982) and Sundberg and Kullberg (1999) did not observe a significant change in subglottal pressure during modal/falsetto shifts of a non-singer male subject (n=1) or untrained female singers (n=11) respectively - suggesting that the use of subglottal pressure in phonatory control may be related to singing training.

**Trained/Professional singers: male.** Large, Iwata, and von Leden (1972) noted in a study of professional and semi-professional singers (n=5) that the modal register was associated with lower airflow, greater spectral energy in higher harmonics, and greater adduction than falsetto. Herbst, Hess, Müller, Švec, and Sundberg (2015) also observed lower airflow during modal voice phonations from a classically trained baritone (n=1),
along with higher subglottal pressure, Maximum Flow Declination Rate (MFDR) and Sound Pressure Level (SPL). Elevated subglottal pressure was linked to increased loudness, rising $F_0$, and higher contact quotient (CQ) in professional baritones and tenors in both modal and falsetto registers by Sundberg and Högset (2001). Countertenors were observed to sing with relatively low subglottal pressures compared to their baritone and tenor colleagues (n=13).

**Trained/professional singers: female.** Modal register was associated with higher subglottal pressures and lower normalized amplitude quotient (NAQ) values than head register in trained and professional female musical theater singers (n=7) studied by Björkner, Sundberg, Cleveland, and Stone (2006), suggesting a more adducted glottal condition in the modal register.

**Adduction/vocal fold vibratory regime.** Measuring patterns of vocal fold adduction and aspects of vocal fold movement provides information about the behavior of the voice source and properties of vocal fold vibratory systems while phonating in different registers.

**Levels of adduction in modal vs. falsetto.** Although inter-subject measures of adduction (e.g. CQ, MFDR) vary appreciably, studies reveal a general trend of greater adduction in modal register compared with falsetto in singers (Björkner et al., 2006; Echternach & Richter, 2010; Henrich Bernardoni, Smith, and Wolf, 2014; Herbst et al., 2015; Kochis-Jennings, Finnegan, Hoffman, and Jaiswal, 2012; Large et al., 1972; Miller & Schutte, 2005; Roubeau et al., 1987; Sundberg & Högset, 2001) and nonsingers (Askenfelt, Gauffin, Sundberg, and Kitzing, 1980; Echternach, Sundberg, Zander, and
Richter, 2011; Kitzing et al., 1982; Sundberg & Kullberg, 1999; Švec, Sundberg, and Hertegård, 2008). Howard (1995) reported contrasting data in his investigation of $CQ_{\text{EGG}}$ in trained and untrained adult female singers (n=26). CQ values were lower for pitches D4 and below and higher for pitches B4 and higher, with higher CQ values above B4 positively correlated to years of singing training. In the author’s opinion, this pattern may be explained partly by the subjects’ choral background and the British choral system’s discouragement of the female “chest tone” quality in favor of a “boy treble,” more aspirate aesthetic (R. Miller, 1997).

**Glottal closure: professional male singers.** Incomplete adduction often is observed in falsetto production, even in trained singers, but many of the subjects in the Sundberg and Högset (2001) study showed complete glottal closure in falsetto production at a variety of dynamic levels. The professional countertenors who participated in the Welch, Sergeant, and MacCurtain (1989) study with baritones and tenors (n=19) and used a falsetto production that produced an EGG waveform shape similar to the tenors and baritones’ modal production in that it displayed a longer closed phase and a “knee” in the opening phase. The tenors and baritones, however, did have EGG waveforms with reduced amplitude compared to the countertenors for their falsetto production, suggesting less vocal fold contact area in that register.

**Glottal closure: untrained singers.** A study of male untrained singers (n=18) by Echternach et al. (2010) that observed adduction patterns through EGG and rigid laryngoscopic high-speed digital imaging revealed a consistent posterior glottal gap during the falsetto voicing of 8 subjects and in the modal voice of four subjects.

16
Although open quotient as measured by EGG (OQ_{EGG}) was expectedly lower in modal voice compared with falsetto, there was little change between OQ_{EGG} in modal voice compared with values extracted from register transitions. Vilkman et al.’s (1995) study of amateur male and female singers with inverse filtering and EGG (n=5) suggested that the transition from falsetto to modal voice on lower frequencies was dependent on creating enough vocal fold contact area, subglottal pressure, and crispness in the contacting gesture to form a “critical mass.” Sundberg and Kullberg (1999) noticed that untrained female subjects employed a variety of glottal adduction strategies during the register transition and theorized that professional singers may increase adduction in order to camouflage transitions.

**Intensity, glottal adduction, and register.** Henrich, d’Alessandro, Doval, and Castellengo (2005) specifically examined the relationship between register, intensity, and OQ_{EGG} in classically trained male and female singers (n=18). Those data indicated that OQ_{EGG} was likely to decrease with intensity levels in the modal register and with F_0 in falsetto register. They also observed that OQ_{EGG} can predict register but can be inaccurate in overlapping value ranges.

**Cartilaginous adduction.** Herbst, Qiu, Schutte, and Švec (2011) specifically examined the role of laryngeal configuration as it relates to adduction and register in male and female singers and nonsingers (n=13) using videostroboscopy and videokymography (VKG) to determine whether subjects could be trained to independently control cartilaginous adduction with increased vocal fold bulging. During voicings with an adducted posterior glottis, CQ, as measured by VKG (CQ_{VKG}), values were higher in
modal voice than falsetto. However, during modal voicings with an abducted posterior glottis, CQ\textsubscript{VKG} values were often similar to falsetto voicings with an adducted posterior glottis. The authors caution that contact quotient should not be the sole measurement used to determine register.

**Intrinsic laryngeal musculature.** Intrinsic laryngeal muscle adjustments are responsible for frequency, intensity, and timbre control through aerodynamic interactions with pulmonary airflow (Sundberg, 1987). The levels of intrinsic laryngeal muscular activity are assessed through laryngeal electromyography (LEMG) using hook wire electrodes. Measuring laryngeal structure movement through radiography- or dynamic magnetic resonance imaging (dMRI) can provide indirect observations about muscular activation as it relates to registers, as well as valuable information about phonatory and prephonatory positioning.

**Register control.** Hirano, Vennard, and Ohala’s (1970) study of male and female professional singers with laryngeal electromyography (LEMG) for four subjects identified the vocalis as the primary register regulatory muscle, with increasing contribution from the cricothyroid (CT), interarytenoid (IA), and lateral cricoarytenoid (LCA) muscles as F\textsubscript{o} rises, with the exception of high falsetto, when other muscles and airflow modulation control pitch. Martin, Thumfart, Jolk, and Klingholz (1990) used LEMG to observe the role of the posterior cricoarytenoid (PCA) muscles, primarily a muscle of abduction, in a trained male singer (n=1). Results showed that activity in the PCA, vocalis, and CT muscles increased with F\textsubscript{o} and intensity in modal and falsetto registers. Register definitions may have been a confounding variable in this study if the
subject was singing in what some would identify as a low-intensity modal production at high pitches rather than falsetto, which typically is associated with an active CT muscle and deactivated TA.

**Register and $F_0$ interactions.** The occasional use of the terms “thyroarytenoid-dominant production” (TDP) to describe modal register and “cricothyroid-dominant production” (CDP) to describe the falsetto register in the voice pedagogic literature and community (Peckham, 2010; Phillips, Williams, Edwin, McPhearson, and Welch, 2012) is not supported by empirical research. As Hirano et al. (1970) stated, register, $F_0$ control, and intensity are inter-dependent variables. Evidence of a relationship between variable $F_0$ control systems and register/pitch range has been documented by a number of studies (H. Hollien, Brown Jr., & K. Hollien, 1971; H. Hollien & Moore, 1960; Lindstad & Södersten, 1988; Roubeau et al., 1987). Titze, Luschei, and Hirano (1989) proposed a theory based on data from male singers and nonsingers (n=4) of how the increase in TA muscle activity can either raise or lower $F_0$: when SPL and the amplitude of vocal fold vibration increase, the $F_0$ will probably rise. Conversely, decreasing vibrational amplitude will result in an $F_0$ drop. However, this theory is applied more directly to the modal register. The voice community would benefit from more studies that compare these two populations and falsetto-register pitch control systems.

**Register and $F_0$ interactions – trained singers.** Kochis-Jennings, Finnegan, Hoffman, Jaiswal, and Hull (2014) examined the degree of thyroarytenoid (TA) and CT activity in trained and untrained singers (n=7), which did not support the theory that the modal register is TA dominant and falsetto CT dominant. Instead, the data suggested that
TA and CT muscle activity is more pitch-level dependent. A sample of commercial singers, classically and non-classically trained (n=7), also displayed varied patterns of TA and CT activity in tasks performed in chest, chestmix, head, and headmix registers, with the classically trained subjects showing lower levels of TA muscle engagement than commercially-trained singers (Kochis-Jennings et al., 2012), especially at higher frequencies. This suggests that the type of voice training received may influence muscle activity usage in voice registers. An observed, but not statistically significant, effect of voice training was seen in Sonninen and Hume’s (1998) review of radiographic data from male and female singers and non-singers (n=67), possibly because of disparities in sample size (13 singers and 57 non-singers). The average vocal fold strain value, based on radiographically based measurements of vocal fold length, was lower (shorter vocal folds) for singers than for non-singers (longer vocal folds), suggesting somewhat different muscular balance strategies to achieve the same F₀. Sonninen, Hurme, and Laukkanen (1999) evaluated a trained dramatic soprano (n=1) with a similar methodology. This subject displayed decreased thyroarytenoid distance (corresponding to vocal fold length) at the highest pitches, again suggesting a different F₀ control mechanism in the high pitch register (in general, shorter vocal folds would indicate either more TA activity or less CT activity).

**Subglottal Resonance**

Theories surrounding the role of subglottal resonance, its role in involuntary register transitions, and its constructive or destructive interference potential have been discussed by Titze (1988; 2008) and Vennard (1967). Evidence of these effects was
found in studies of excised human larynges (Austin & Titze, 1997) and canine larynges (Berry et al., 1995). However, more recent studies involving human subjects conclude that subglottal resonance does not appear to be a significant factor in registration or on the location of register transitions (Spencer & Titze, 2001; Sundberg, R. Scherer, Hess, Müller, & Granqvist, 2013).

**Vocal Tract Resonance and Configuration**

*Vocal tract resonance.* The perceptual quality of vocal registers is dependent largely on its acoustic correlates. While modal voice is typically associated with high frequency spectral energy, falsetto is often characterized by the energy in the fundamental and lower harmonics, an acoustic variation that is a critical perceptual aspect of transition (Berry et al., 1995; Castellengo, Chuberre, & Henrich, 2004; Castellengo, Lamesch, & Henrich, 2007; Colton, 1972; Echternach, Traser, Markl, & Richter, 2011; Henrich et al., 2005; Kochis-Jennings et al., 2012; Large, 1968; 1973; Large et al., 1972; D. Miller & Schutte, 2005; Švec et al., 2008).

*Resonance strategies – trained singers.* Trained singers use a variety of resonance strategies in the modal and falsetto registers in an effort to create the expected quality of tone throughout the vocal range. The tendency of the first resonance (R1) to follow $2F_0$ in ascending phrases through the *passaggio* is one of the primary perceptual elements of the commercial music singing genre and belt technique (Schutte & D. Miller, 1993). For classical singing, R1 must fall below $2F_0$ in the lower middle pitch range to avoid a register violation. Using a resonance strategy that boosts R2 can be a very effective technique for amplifying fundamental frequencies in this range and smoothing
register transitions in the female voice (D. Miller, 2008). Henrich, Smith, and Wolfe (2011) also noted resonance-tuning relationships between R2 and 2F_o and R2 and 3F_o in the lower range of female classical singers (n=22), along with an approach based upon a combination of R1 amplification of the F_o with R2 amplification of 2F_o. Studies suggest that these strategies are somewhat subject- and voice classification-dependent, however. For example, the voice pedagogue, Richard Miller, and Harm Schutte’s 1984 study of voce di petto (modal), voce mista (mix), voce di testa (head voice), voce finta (feigned voice), and falsetto in a professional tenor (n=2) revealed few substantial spectral differences between these productions except for the intensity variance in 2F_o, but the reliance on a strong singer’s formant through all voicings was evident. D. Miller & Schutte (2005) discuss the register transition of a dramatic soprano subject who made no effort to camouflage her register transition, but nevertheless perceptually matched the registers on either side through maintaining spectral energy in the high harmonics, giving her tone a “metallic quality.” Large (1972) observed that the most perceptually similar sample pairs recorded by female classical singing students (n=9) in modal and falsetto voicings exhibited spectral energy in higher harmonics during both samples. Researchers have documented a variety of resonance patterns in trained and professional singers, even at the highest level of the profession (D. Miller, 2008; D. Miller & Schutte, 2004; 2005; Schutte, D. Miller, & Duijnste, 2005; Schutte and R. Miller, 1984).

**Vocal tract configuration.** Studies that employ dynamic magnetic resonance imaging (dMRI) and radiography are able to document the movements of glottal and supraglottal structures during voicing. Welch et al. (1989) observed through
radiographic-based measurements that professional tenors, basses, and baritones did not make obvious vocal tract adjustments between registers. Lindestad and Södersten’s (1988) comparison of vocal tract configuration in trained countertenor and baritone singers (n=5) also observed a lack of pharyngeal adjustment in the baritone subject but a tendency to narrow the lower pharynx with rising F₀ in the countertenors. The countertenor subjects (n=7) in Echternach et al. (2011b) were seen to narrow their pharynx in their modal-to-falsetto transitions, along with lip, jaw, tongue, uvula, and laryngeal position adjustments. They also found that classically trained sopranos did not make large vocal tract adjustments in their register transitions. However, significant vocal tract adjustments were observed in soprano subjects as the F₀ crossed the R1 location for the sung vowel. More investigations of female singers, especially mezzo-sopranos and altos, are needed.

Nonlinear Source-Filter Interaction

Deterministic chaos and nonlinear source-tract coupling. While intrinsic muscular control is doubtlessly important to excellent singing in any register or genre, Švec et al.’s (1999) experiment with an excised human larynx illustrates that pitch jumps, bifurcations and hysteresis can occur when small, gradual laryngeal muscular adjustments are made. During bistable phonation, modal and falsetto vibratory regimes co-exist without apparent muscular adjustments. Hatzikirou, Fitch, and Herzel’s (2006) experiment with coupling various length tubes to a two-mass model led to “subharmonic variations and deterministic chaos” if the nonlinear source-vocal tract coupling is high, even in the absence of vocal fold asymmetries. Nonlinear laryngeal-vocal tract coupling
is high when the epilaryngeal space is narrowed, as it often is in operatic singing. Similar phenomena have been observed with canine larynges (Alipour et al., 2009; Berry et al., 1995), a three mass model (Tokuda et al., 2007), and in human subjects (Echternach et al., 2010a; Švec & Pešák, 1994; Vilkman et al., 1995). This may explain why the dramatic soprano studied in Sonninen et al. (1999) widened her previously narrowed epilaryngeal space as she approached the passaggio from the lower range. Perhaps this adjustment helped her avoid sound irregularities caused by source-tract acoustic interactions in the transition. Nonlinearities may also be responsible for the elevated perturbation measures that characterize transitions, whether or not the transitions are judged to be perceptually smooth, in the form of jitter (frequency perturbation) (Large, 1968), shimmer (amplitude perturbation), and relative average perturbation (RAP) (Echternach et al., 2011a; Echternach & Richter, 2010; Echternach, Traser, & Richter, 2012), or vibrato disturbances (Large, 1968).

**Pitch jumps.** D. Miller, Švec, and Schutte (2002) noted the existence of subharmonics during transition as they examined the magnitude of pitch jumps in a measurement they termed the “characteristic leap interval” (CLI). In their study of experienced male and female singers (n=11), females exhibited a CLI of a perfect fifth or less while males leapt closer to an octave when controlled for subglottal pressure and intensity. Švec et al. (1999) observed similar gender-related CLI variations in three subjects and Bloothooft, van Wijck, and Pabon (2001) in untrained males (n=7).

**Register Smoothing Considerations and Behaviors**

**F₀ control systems.** An argument could be made that a major factor in the success
or failure of a register transition is \( F_0 \) control. If a singer does modulate \( F_0 \) in the modal register with a different method than in the falsetto register, this could complicate the transition maneuver as one system gives way to the other. Roubeau et al. (1987) described the process of the modal/falsetto register transition as starting with a change in laryngeal vibratory pattern (responding to a neural command) resulting in an EGG waveshape modification and signal intensity discontinuity, disruption in the \( F_0 \) control system (perhaps from changes in subglottal pressure and the laryngeal musculature), and an \( F_0 \) correction (or reestablishment of \( F_0 \) control) that may be a response to the singer’s aural feedback. Their subjects, male and female trained and untrained singers (n=18), displayed relatively quick register transitions characterized by decreased volume, abrupt frequency jumps (followed by a pitch correction), and recovery of volume. D. Miller et al. (2002) and Sonnin (1999) also observed the \( F_0 \) system recovery delay (often coinciding with subharmonics and/or other instabilities) during transition. Perhaps singers skilled at navigating the transition establish a system of \( F_0 \) control that requires no range- or register-related adjustments or can be adjusted throughout the range in small increments.

**Adduction and \( F_0 \) control.** Titze (1988) presented a theoretical muscular contraction and glottal adduction strategy for a smooth register transition by abducting the vocal processes as \( F_0 \) rises in modal voice to balance the increase in TA activity pre-transition, then progressively decreasing TA activity while increasing CT activity post-transition. This strategy may, in the author’s opinion, encourage a posterior glottal gap in the falsetto register, which some view as undesirable, especially in the operatic genre.
Intensity and resonance control also are important components of a smooth register transition. Henrich et al. (2005) observed a smoothing strategy in a countertenor subject who reduced vocal intensity and increased OQ before transition in order to match the characteristics of the falsetto register (perhaps with the above smoothing procedure described by Titze, 1988). Other subjects attempted to smooth their transitions with an intensity decrease and a reduction in the spectral energy of higher harmonics. Both techniques were judged to be successful perceptually, but perhaps not to a certain operatic standard that highly values the preservation of resonance balance and loudness. D. Miller and Schutte (2005) (n=3) reported a CQ decrease before a perceptually-smooth transition in a professional mezzo-soprano but did not judge that transition a success because the singer failed to counteract the natural intensity reduction as she entered the falsetto register. From an operatic standard, this transition would have caused an audibility issue in a performance situation with orchestral accompaniment. A discussion of resonance matching during register transitions can be found in the Vocal Tract Resonance section (above).

“Mixed” Voice. Colton and Hollien (1973) described registers as “multidimensional phenomena: each register exhibiting a specific and specifiable set of perceptual dimensions….This physiological state produces, acoustically, certain combinations of fundamental frequency, vocal intensity and spectrum” (p. 278). This also allows skilled singers to adopt the attributes of one register while singing in a different register for blending purposes. Castellengo et al. (2004) studied male and female professional singers (n=5) to ascertain if the voix mixte (mixed voice) was, in fact,
indicative of a separate vibratory mechanism, and whether or not the technique was capable of smoothing register transitions. They found that *voix mixte* was always related to either mechanism 1 (modal) or mechanism 2 (falsetto), but mimicked the sound of the other register through manipulation of vocal tract acoustics (particularly the singer’s formant) and sound intensity (approximately +5 dB in mechanism 2 to simulate mechanism 1; -10 dB in mechanism 1 to simulate mechanism 2). Castellengo et al. (2007) reported similar results with professional countertenor and soprano subjects (n=2). Unlike the two previous studies, Echternach et al.’s (2012) study of *voix mixte* and register transitions in professional tenors did not produce EGG or spectral data that would support classifying the mix definitively as either modal or falsetto register. They also did not observe other common registration phenomena such as abrupt CQ, frequency, or intensity disturbances, although perturbation measures did increase during ascending pitch glides. The subjects’ ability to navigate the register transition skillfully was ascribed to vocal training.

**Transition Pitch Locations**

Table 2 displays the reported transition pitch locations for subjects in previous studies.
<table>
<thead>
<tr>
<th>Study</th>
<th>Subjects</th>
<th>Task(s)</th>
<th>Transition Location (Hz or pitch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spencer and Titze (2001)</td>
<td>Males (n=2) Females (n=2)</td>
<td>Glissandos (ascending/descending)</td>
<td>Males: 341.5 (asc.) 339.5 (desc.) Females: 415 (asc.) 378.5 (desc.)</td>
</tr>
<tr>
<td>Roubeau et al. (1987)</td>
<td>Males (n=10) Females (n=8)</td>
<td>Glissandos (ascending/descending)</td>
<td>All subjects: 195 – 540 (most within 200-300 range)</td>
</tr>
<tr>
<td>Keidar et al. (1987)</td>
<td>Males: (n=1) Females: (n=1)</td>
<td>Sustained tones between G#3 and A#4; ascending/descending 5-note scales</td>
<td>Males: 320 Females: 353</td>
</tr>
<tr>
<td>Sonninen et al (1999)</td>
<td>Female (n=1)</td>
<td>Sustained tones between C#3 and G#6</td>
<td>Female: D#4 – F4</td>
</tr>
<tr>
<td>Švec et al. (2008)</td>
<td>Female (n=1)</td>
<td>Ascending scales and pitch glides</td>
<td>Female: 525</td>
</tr>
<tr>
<td>Björkner et al. (2006)</td>
<td>Females (n=7)</td>
<td>Sustained tones</td>
<td>Female: D4-C5 (294-523)</td>
</tr>
</tbody>
</table>

Table 2: Transition Pitch Locations by Subject and Task
Part Three: Method

Subjects

Two female subjects, ages 35 and 38, were identified and selected based on the following criteria: 1) identification as operatic mezzo-sopranos, 2) at least one performance of a major role with a Level 1 Opera Company (as designated by Opera America) or an orchestra with an annual operating budget of at least $20 million during the past five years, and 3) between the ages of 21 – 65 years old. Using the Chapman and Bunch (2000) taxonomy, Subject 1 would be classified as 3.4 (National: Concert/Oratorio/Recital) and Subject 2 as 2.1 (International: Opera Principal). Both subjects are native English speakers and reported that they were in good vocal health at the time of the experimental recording.

All proposed activities were reviewed and approved by The Ohio State University’s Institutional Review Board (IRB) under protocol #2013B0406. Subjects were read a script with a project description to obtain verbal consent and completed a brief questionnaire to confirm that they met the selection criteria.

Equipment

All data recordings took place in the Helen Swank Research and Teaching Lab at The Ohio State University. Subjects were recorded in a sound attenuated booth using an Earthworks M50 omnidirectional microphone (frequency
response: 10 Hz – 30k Hz ± 1 dB) placed 10 cm from the corner of the subject’s mouth opening out of the air stream. The EGG signal was recorded simultaneously with a Glottal Enterprises Two-Channel Electroglottograph EG2-PCX. Both recordings were ported through an Edirol UA-25 USB audio interface and recorded at 16K samples per second. Signals were stored in .wav format on a Dell Optiplex 9010 computer using the Sopran (Granqvist) software program. A schematic diagram of the data collection system can be found in Figure 1.

![Recording System Schema](image)

**Figure 1.** Recording System Schema

The recording level was calibrated before each task at 94dB with an Amprobe Microphone Calibrator using a reference tone and Sopran’s calibration function. The subjects were positioned 10 cm from the microphone through the use of two positioners placed on either side of the nose. The distance was measured between the subjects’ mouth opening and the microphone before commencing their performance of each task. Subjects were asked to stay as stationary as possible while performing the tasks.
The integrity of the EGG signal was monitored during recordings using VoceVista software (D. Miller). Subjects elected to use a neck strap to position the EGG electrodes. To avoid fastening the neck strap too tightly for comfortable singing, they added light pressure to the sensors with their fingers. If the signal was lost or disrupted during the recording of a task, the subject was asked to adjust the electrode locations and the recording was restarted.

Protocol

Tasks. Subjects sang ascending and descending chromatic scales between C4 and G4 (see Figure 2) on the six vowels /a/, /æ/, /e/, /i/, /o/, and /u/. An iPhone with a keyboard application was used to check starting and ending pitches as often as subjects desired. Subjects were allowed unlimited attempts at each task, indicating whether the phrase they just sang was acceptable by saying “yes” or “no.” The recording ended after the third acceptance of a task performance. The twelve combinations of vowel and phrase direction conditions were requested in random order.

Figure 2. Ascending and Descending Chromatic Scale Tasks
**Directions.** In order to observe the vocal production used during an operatic performance insomuch as the situation allowed, the subjects were given the following performing instructions.

1. Please stay as stationary as possible.
2. Please sing the phrase in one breath.
3. Sing as you would in an operatic performance in a 3,000-seat theater with a Mozart or Donizetti-sized orchestra.
4. Sing whatever vowel you would use if the score asked you to use these vowels in the text.
5. Please start in “head voice” when descending and end in “chest voice” and the reverse when ascending.
6. You may repeat the task as many times as you would like. Please let me know which three you would like to keep by saying “yes” after the ones you feel best reflect what you would do on stage.

Subjects were encouraged to sing as much as desired between task recordings, either in or out of the sound booth. The performance of repetitive tasks in a sound-attenuated environment is not conducive to professional-quality singing. Therefore, subjects often elected to sing other phrases in a different sound environment between recordings for orientation purposes. Recording sessions lasted approximately 45 minutes.

**Observational Study Design**

Because this was an observational study designed to elicit subject behavior that would mirror aspects of operatic performance, some variables were only indirectly
controlled through task instructions, including loudness, duration of the task performance, precise intonation in the chromatic sequence, vowel quality consistency, and uniformity of task performance between repetitions. While chromatic passages are not a ubiquitous component of day-to-day vocal exercises (as one subject remarked), they do appear fairly frequently in the bel canto repertoire, both in composer-notated form and in traditionally interpolated cadenzas.

Analysis

Perception. The identification of transition locations (pitches within the scale) was made by perceptual judgments of the author, who is a professional operatic mezzo-soprano with over ten years of voice teaching experience. The perceptual identification of register transitions was heard as quality changes, which were corroborated with other measures (see below).

EGG analysis factors. The limitations of EGG waveform analysis as a method of noninvasively evaluating glottal vibratory characteristics have been explored by Baken and Orlikoff (1987), Childers, Hicks, Moore, Eskenazi, and Lalwani (1990), Colton and Conture (1990), Herbst and Ternström, (2006), Kitzing (1990), Rothenberg (1992), Sapienza, et al. (1998), and Titze (1990) who described difficulties with obtaining valid measurements during the occurrence of vertical laryngeal movements, neck muscle contraction, and signal interference from extraglottic movements, excess adipose tissue, or mucus strands across the glottis. Female subjects present additional difficulties in obtaining valid waveforms because of their relatively smaller larynx structures and larger thyroid cartilage angle (compared to male subjects). The inability to determine from EGG
waveform aspects such as complete glottal closure or contact change along the superior/inferior or anterior/posterior glottal axes creates difficulties in interpreting the EGG waveform. It is for this reason that Herbst and Ternström (2006) suggest describing certain EGG measures of glottal contact as the “quasi closed quotient.” EGG waveforms that have unusual appearances must be considered carefully and eliminated from consideration if there are reasons to suspect the meaning of waveform measures.

Despite these difficulties, there is general agreement that a strong EGG signal without irregularity or distortion can yield sensitive intra-subject medial vocal fold contact patterns during registration maneuvers and relative changes in general vocal fold adduction (Baken and Orlikoff, 1987; Childers et al., 1990; Henrich et al., 2003; Herbst and Ternström, 2006; R. Scherer, Vail, & Rockwell, 1995) such as vertical phasing, medial surface bulging (Titze, 1990), and longitudinal contacting and de-contacting events (Herbst et al., 2014). The two subjects selected for this study sing with a stable laryngeal position and sufficient loudness to generate clear EGG signals, supporting the assumption that the recordings are valid and contain relevant information.

**EGG signal intensity (EGGSI).** The intensity of the EGG signal was obtained with Praat software’s (Boersma & Weenink, 2010) intensity module. Discontinuities viewed in the signal trace that had an inverse relationship with phrase direction (e.g. an abrupt signal intensity increase during a descending scale task, see Figure 3) were identified as register transitions (Askenfelt, et al., 1980; Echternach et al., 2012; Henrich et al., 2003; 2005; D. Miller et al., 2002, Roubeau et al., 1987; Vilkman et al., 1995). Those locations were confirmed through perceptual evaluation and qualitative EGG
waveform examination.

Figure 3. Example of an EGGSI Trace with Transition in Brackets (Descending Chromatic Scale). In the example illustrated here, the abrupt change in the EGG intensity curve indicates an abrupt increase in EGG waveform height as the singer transitioned from falsetto to modal register.

**EGGW50.** The EGG signal was analyzed with SIGPLOT, a custom, Matlab-based software program, to obtain EGGW50 values. EGGW50 is an estimate of glottal closure based on the equation in Figure 4, with A representing the waveform width at 50% of its peak-to-peak amplitude and B the period of the glottal cycle (R. Scherer et al., 1995; Fisher et al., 1996).

Figure 4. EGGW50 Definition
The recorded EGG signal for each task with an identified transition was separated into the following five locations for data analysis: (1) the first pitch of the phrase, (2) the pitch preceding transition, (3) the transition pitches, (4) the pitch following transition, and (5) the last pitch of the phrase. If no evidence of a transition was identified, the recording was separated into the following 3 locations for data analysis: (1) the first pitch of the phrase, (2) E-flat 4 (chosen as a general phrase midpoint), and (3) the last pitch of the phrase. Seven to twelve adjacent waveforms were selected from the approximate midpoint of each location to obtain an average EGGW50 value.

Figure 5 illustrates the relationship between EGGW50 measurements, EGGSI, and EGG waveforms. The EGGW50 scatterplot shows the values obtained for the five locations and for each of the three tokens sung during the task. The EGG waveforms to the right of the graph are extracted from the waveforms sampled to obtain the EGGW50 values marked by the arrows. The approximate locations of the EGGW50 data points in the EGGSI are also marked with arrows at the bottom of the figure. The increasing relative width of the EGG waveforms from the pre-transition falsetto voice register to the post-transition modal voice can be seen in the EGG waveforms. The increasing amplitude of the EGG waveforms is reflected in the upward direction of the EGGSI trace.
Figure 5. Subject 1: EGGW50 Scatterplot, EGG Waveforms, and EGGSI Trace for the /i/ Descending Task. The two arrows refer to the pre-transition (falsetto register) note and the last sung note (modal register), respectively. The EGGW50 value is less for the head register note than for the modal register note, suggesting more adduction for the modal note. The EGG waveform on the left refers to the pre-transition note, and the larger magnitude EGG waveform on the right to the last note. The greater magnitude results in the larger EGG intensity values shown in the lower trace.

**DEGG wavegrams.** The DEGG wavegram was developed by Herbst, Fitch, and Švec (2010) to provide a method for visualizing how pertinent waveform characteristics derived from the first mathematical derivative of the EGG signal (DEGG) change over time. The DEGG is a useful elaboration in that it highlights details about the rate of change in the EGG signal that are difficult to detect by visual examination. When clear single peaks in the DEGG are present, they are assumed to coincide with glottal opening and closing events (Henrich et al., 2004). A recent investigation with a canine larynx by Herbst et al. (2014) comparing EGG and High Speed Videoendoscopy with a frame rate of 27,000 frames/second concluded that even single, clear peaks in the DEGG did not
consistently correspond with the moments of glottal opening and closing. However, many vocal fold oscillation patterns with phase differences along the anterior-posterior axis display more progressive contacting and de-contacting events that may be more accurately described as intervals. Herbst et al. (2014) did determine from their data that DEGG peaks coincide with contacting and de-contacting intervals, possibly providing information about the presence of longitudinal phase differences in the glottal cycle.

DEGG wavegrams offer a visual representation of glottal closing and opening intervals that can assist with qualitative EGG analysis. The positive and negative peaks of the normalized DEGG waveform, representing glottal closing and opening intervals, are respectively represented with black and white intensity values. This information is recorded in a color-coded strip representing one glottal cycle, rotated 90 degrees, and linked together for a continuous visualization of opening and closing events. The EGG signal magnitude is converted into color intensity values (Herbst et al, 2010). For the purposes of this project, the DEGG wavegram provides a qualitative tool for evaluating contact quotient trends. A visual representation of the construction of the DEGG wavegram from the DEGG signal can be viewed in Fig. 6. In ascending scales, an abrupt narrowing of the distance between the closing and opening events is evidence of a transition (and thus reflects less adduction in the falsetto register). In descending scales, an abrupt widening of the distance between the closing and opening events is evidence of a transition (and thus reflects more adduction in the modal register).
Figure 6. DEGG Wavegram Construction Process (Herbst et al., 2010, Figure 4, pp. 3073)

**Spectrographic analysis.** Fast Fourier transform (FFT) spectral slices were taken from approximately the same locations identified for EGGW50 analyses (described above) through Praat’s spectrum module.

**Vibrato.** Vibrato rate (number of cycles per second) and extent (the maximum $F_o$ value minus the minimum $F_o$ value within the vibrato cycle, in cents) were analyzed for each task. An illustration of vibrato rate and extent parameters can be found in Figure 7. The $F_o$ readings at the maximum and minimum of each vibrato cycle were extracted from Praat’s pitch analysis module and recorded with the corresponding time marker into a
spreadsheet. Each cycle’s vibrato rate and extent were calculated from that data using the following formulas:

\[
\text{Rate (Hz)} = \frac{1}{\text{period(s)}}
\]

\[
\text{Extent (cents)} = 1200 \log_2 \left( \frac{a}{b} \right),
\]

where \( a \) and \( b \) are the maximum frequency and minimum frequency values in Hz within the vibrato cycle. Variations of rate and extent that occurred near or during a register transition defined by the EGG signal intensity alteration were considered to be evidence of transition phenomena (Large, 1968), which helped to further define the borders of the transition zone. Areas of enhanced variation were defined as a rate or extent maximum or minimum, or a disruption of a subject’s general pattern of cycle-to-cycle vibrato variability intra-phrase, intra-pitch, and/or during pitch changes (Prame, 1992; Sundberg, 1979).

![Diagram showing vibrato rate and extent](image)

Figure 7. Vibrato Rate and Extent (after Lee, Dong, & Chan, 2011). The vibrato rate is the inverse of the period. The extent is the peak-to-peak difference (in cents).
Part Four: Results and Discussion

Transition Pitch Locations

Subject 1. Subject 1 displayed perceptually identifiable, although occasionally subtle, register transitions in all tasks and repetitions. All transitions also were discernable by changes in the EGGSI trace. Increased vibrato rate and extent variation was observed in 79% of her samples (19 of 24). Transitions occurred in areas surrounding pitch changes between C#4 and F4 (see Figure 8) with the greatest number of transitions taking place between D#4 and E4. There are two visible, but not statistically significant, trends: (a) /a/ transitions tend to occur at higher pitches than /i/, and (b) transitions tend to occur at higher pitches in ascending tasks and at lower pitches in descending tasks.
Figure 8. Subject 1: Transition Pitch Location Frequency Graph. The number in each cell represents how often a register transition took place on that pitch. Each transition occurred around a pitch change, beginning at the end of one pitch and ending on the next consecutive pitch. Both pitches received one mark. For example, the graph indicates that four transitions occurred between D# and E, one transition between D and D#, and one transition between E and F for the vowel /a/. The color key indicates the pitch direction of the task corresponding to each transition.

**Subject 2.** Subject 2 displayed perceptually identifiable transitions in only 5 of her 24 samples (in contrast to the first subject who had identifiable transitions in all tokens). Three of those five transitions were for the /a/ ascending chromatic scale task repetitions. A transition was detected in one out of the three /o/ descending tasks and one out of the three /u/ descending tasks. Four out of the five transitions were perceptually identifiable. All five transitions were discernable by
changes in the EGGSI trace. Increased vibrato rate and extent variation was observed in 80% of the samples (4 out of the 5). Subject 2’s transitions for the /o/ and /u/ descending tasks began and ended on the same sustained pitch. The transitions on the /a/ tokens occurred during a pitch change, beginning on one pitch and ending on the next consecutive pitch (see Fig 9). The most common transition pitch location was F4.

![Figure 9. Subject 2: Transition Pitch Location Frequency Graph. The register transitions in the /o/ and /u/ tokens took place on one pitch. The register transitions on the vowel /a/ occurred around a pitch change.](image)

**Subharmonics**

Subharmonics were observed in one descending /a/ task from Subject 2 on E-flat4 with evidence of hysteresis visible in the intermittent adjacent alternation of two distinctive waveshapes (see Figure 10). This could indicate a subtle change of laryngeal vibratory mode or the presence of nonlinear source-vocal tract effects.
However, because there was no perceptual evidence of a transition visible in the EGGSI, this was not classified as a transition. No subharmonics were found in the analyzed tasks for Subject 1, but an example was found (see Figure 11) in a descending /æ/ task that will be analyzed in a later phase of this project.

Figure 10. Subject 2: Alternating Adjacent Waveshapes (/a/ Descending Task). The amplitude of the EGG waveform shows a minor two-cycle pattern.

Figure 11. Subject 1: Alternating Adjacent Waveshapes (/æ/ Descending Task). The amplitude of the EGG waveform shows a definitive two-cycle pattern.

Descending Chromatic Scale Task - /a/ Vowel

Combined graphs.
Subject 1. Figure 12 shows the identified register transition (falsetto to modal) for the second token of the /a/descending chromatic task for Subject 1. There is a vibrato rate maximum followed immediately by a minimum in the center of the shaded transition zone that corresponds with an EGGSI increase. These two parameters defined this transition zone. The vibrato extent did not show enhanced variation in this particular transition. The spectrogram with F\(_0\) trace from Praat indicates that the estimated R2 location migrates from 3F\(_0\) to 4F\(_0\) during the transition, indicating a possible acoustic transition element.
Subject 2. The EGGSI and DEGG wavegram shape is constant or descends with pitch in Subject 2’s descending /a/ task (see Figure 13), providing no evidence of a register transition. While there are some small perceptual changes in tone quality from pitch to pitch that correspond to vowel modifications, the phrase has no obvious tone quality transformations, instabilities, or irregularities associated with transition. A consistent increase in vibrato rate and decrease in extent at pitch changes (at the moments of the vertical lines) are similarly undisturbed. In contrast
with Subject 1’s R1/2F₀ strategy in the lower half of the phrase, Subject 2 uses a R1/3F₀ pairing.

Figure 13. Subject 2: /a/ Descending Chromatic Task (token 1)

**Spectral comparison: Subjects 1 and 2.** Corresponding FFT spectral slices (see Figure 14) were obtained from the midpoints of G4, Transition, and C4 for Subject 1 and G4, E-flat4, and C4 for Subject 2 of the task tokens illustrated in Figs 1 and 2. The highest-intensity harmonics (1F₀ and 2F₀) for Subject 1, amplified by R1, suggest a general lower-register R1 dominant strategy in this task. Contrastingly,
the spectral slices taken from Subject 2’s task show R2’s amplification of $3F_o$ and $4F_o$, suggesting an R2 strategy. Subject 2’s closer spacing of estimated R1 and R2 locations may help her boost the spectral energy in R2. Both subjects display an increasing spectral prominence in the singer’s formant frequency (or R3) range as they descend to C4.

Figure 14. FFT Spectral Slices (/a/ Descending Task)

Ascending Chromatic Scale Task - /a/ Vowel

Combined graphs.

**Subject 1.** Subject 1’s EGGSI has two areas of abrupt downward change preceding an upward change in preparation for the last note of the phrase, G4 (see Figure 15). Perceptual indications, two large peaks in the vibrato rate separated by a large relative decrease, and an extent minimum surrounded by increased variation
identify the first downward change as the most important transition point. Large decreases in vibrato extent coinciding with the second downward change in the EGGSI, subsequent EGGSI increase, and widening of the DEGG wavegram suggest the possibility of smaller post-transition laryngeal adjustments. S1 uses an acoustic strategy similar to the descending /a/ task, mostly preferring an R1/2F₀ and R1/1F₀ strategy.

Figure 15. Subject 1 - /a/ Ascending Chromatic Task (Token 3)
Subject 2. The three tokens of the /a/ ascending task all had identifiable register transitions. All other tasks had no transitions, except for one token each in the descending /o/ task and descending /u/ task. Subject 2’s EGGSI has a slight downward trend up to the beginning of the transition (see Figure 16). The intensity abruptly decreases and then regains its previous level during the transition interval, which coincides with an obvious narrowing of the DEGG wavegram. Because this is an ascending scale, the signal direction’s inverse relationship with pitch direction may indicate that the singer is making small laryngeal adjustments in preparation for the transition from almost the beginning of the phrase. Two vibrato rate maxima and an extent minimum occur within the transition boundaries as defined by the EGGSI trace. Subject 2 uses an acoustic strategy similar to her descending /a/ task, with an R1/3F₀ and R1/2 F₀ tracking closely paired with an R2/4F₀ to R2/3F₀ progression.
Spectral comparison: Subjects 1 and 2. The FFT spectral slices again illustrate Subject 1’s R1 dominant and Subject 2’s R2 dominant strategy preferences (see Figure 17.) for an /a/ ascending task. In contrast with the /a/ descending task, Subject 2’s C4 spectral slice does not display significant energy in the singer’s formant (R3) range. The attenuation of the singer’s formant in all three slices could be evidence of a transition smoothing strategy in preparation for the acoustic
expectations of the middle register. Subject 1’s spectral slices are similar to the /a/ descending task with the exception of C4, which also shows a comparative decrease in singer’s formant (R3) energy.

Figure 17. FFT Spectral Slices (/a/ Ascending Task)

**Descending Chromatic Scale Task - /i/ Vowel**

**Combined Graphs.**

**Subject 1.** The transition in this task is identified by an increase in EGGSI and a clear DEGG wavegram widening (see Figure 18). The DEGG wavegram shows a relatively short contact quotient before the transition followed by a longer contact quotient with a shorter open interval after the transition into modal register. The vibrato rate and extent graphs do not reflect an obvious increase in variation, but there is a minimum followed two cycles later by a peak that is among the highest in
the phrase. The vibrato pattern seen in this phrase, with clear vibrato rate maxima and extent minima near pitch transitions, is similar to the pattern seen in Subject 2’s descending /a/ task.

![Graph showing vibrato rate, vibrato extent, EGG intensity, and DEGG wavergram with a transition highlighted.](image)

Figure 18. Subject 1 /i/ Descending Chromatic Task (Token 2)

**Subject 2.** Neither the EGGSI nor the DEGG wavegram suggests a transition in this phrase (see Figure 19). There is a smooth, gradual decrease in contact quotient
from the beginning to the end of the phrase, revealing no evidence of a change in laryngeal vibratory regime. The vibrato rate and extent patterns are especially consistent (and remarkable) examples of the patterns observed in Subject 2’s descending /a/ and Subject 1’s descending /i/ task, with a faster, narrower (extent) vibrato cycle found at each pitch change. Also notable is the estimated location of R3 and its close proximity to R2 in the spectrogram.

Figure 19. Subject 2 /i/ Descending Chromatic Task (Token 2)
Spectral comparison: Subjects 1 and 2. In contrast to the /a/ vowel, both subjects display an R1/Fo dominance in the /i/ vowel tasks (see Figure 20), recognizing that R1 is low for the /i/ vowel (relatively close to Fo). The difference in acoustic strategies between the two singers is more evident at higher frequencies. Although Subject 1 has some spectral prominence in the R2 area, the majority of the high frequency energy is in the singer's formant (R3) region, whereas Subject 2 shows a preference for acoustic energy in the R2 area, especially in the G4 and transition locations. Subject 2 displays a more spectrally prominent singer's formant (R3) in the C4 slice, still with more energy in the R2 area.

Figure 20. FFT Spectral Slices (/i/ Descending Task)
Ascending Chromatic Scale Task - /i/ Vowel

Combined Graphs.

Subject 1. All of the graphs provide evidence of modal-to-falsetto transition within the shaded zone (see Figure 21). The vibrato rate graph does not display some of the extremes found in other tokens within the transition interval, but it does have several large adjacent fluctuations. The vibrato extent graph shows a minimum followed by a maximum, reflecting an area of variability that appears in the spectrogram’s F<sub>0</sub> trace. An abrupt change in laryngeal vibratory mode is reflected in the EGGSI decrease and abrupt DEGG wavegram narrowing, suggesting lower EGG amplitude and less adduction.
Subject 2. Both Subjects 1 and 2 present a different vibrato rate pattern relative to pitch changes in their ascending /i/ tasks compared to the descending /i/ tasks. The vibrato rate in the pitch transition does not increase as dramatically, but retains the decreased vibrato extent seen in the descending /i/ task tokens (see Figure 22). Rather than a quicker vibrato cycle with a decreased extent accomplishing the pitch change, Subject 2 prefers a more gradual excursion between
pitches in this token with a smaller range of vibrato variation. In some sections of the phrase, it seems to resemble a controlled glissando with vibrato. No register transition is observed in the EGGSI trace, DEGG wavegram or vibrato rate and extent graphs. As with the descending /i/ task token, the estimated R3 location is located at a lower frequency than in Subject 1’s /i/ descending token spectrogram and is in closer proximity to R2.

Figure 22. Subject 2 /i/ Ascending Chromatic Task (Token 3)
**Spectral comparison: Subjects 1 and 2.** Both subjects again show an R1 spectral dominance in all of the FFT spectral slices (see Figure 23) because of the low R1 for an /i/. Subject 1’s slices display a similar preference for energy in the singer's formant (R3-R4) range as seen in previous task tokens, especially in harmonics that coincide with the location of R4 (around 3550 Hz). Subject 2 also exhibits increased energy in R4-boosted harmonics, but again R2 spectral energy equals or exceeds that of R4. Subject 2’s narrowing of the space between R2 and R3 may acoustically fortify R2 with the joined resonance potential it offers to the harmonics in that frequency range.

Figure 23. FFT Spectral Slices (/i/ Ascending Task)
EGGW50 Analysis

**Analysis of Variance (ANOVA).**

*Subject 1.* An analysis of variance was performed to examine the interactions between sung phrase direction, vowel, and transition-related moments (pre-transition, transition, etc.). The ANOVA results (see Table 3) indicate that vowel is the only variable that does not have a significant effect on Subject 1’s EGGW50 values. However, vowel does have a significant interaction with phrase direction, the most significant being the /o/ vowel’s relationship with ascending phrase direction.

The location data point significantly interacts with EGGW50 values, indicating that Subject 1 may include preparatory and recovery laryngeal gestures around the transition that are suggested by the EGGW50 values. The location data points also interact significantly with phrase direction, which implies that Subject 1’s adductory strategy is different for ascending versus descending phrases.
<table>
<thead>
<tr>
<th>Variable</th>
<th>P-Value (.05 significance level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition-Related Moment</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.3923</td>
</tr>
<tr>
<td>Phrase Direction</td>
<td>0.00423</td>
</tr>
<tr>
<td>Vowel/Phrase Direction</td>
<td>0.0082</td>
</tr>
<tr>
<td>Transition-Related Moment/Phrase Direction</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Transition-Related Moment/Vowel</td>
<td>0.06001</td>
</tr>
</tbody>
</table>

Table 3. Subject 1: ANOVA Effects with P-Values

**Subject 2.** The only significant interaction with EGGW50 values in Subject 2’s dataset was between the vowels /i/ and /a/. This is evidence of a change in the adduction pattern between these two vowels, perhaps in response to relatively low R1 location for the vowel /i/ compared with the relatively high R1 location for the vowel /a/. The other main effects were not significant.

**Visible trends.**

**Subject 1 descending tasks.** The scatterplots in Figure 24 show the progression of EGGW50 values through the five location data points of the descending chromatic scale for Subject 1. Boxes have been drawn around the values obtained for those points from each of the three task tokens. The lines between the boxes represent an estimate of the general, progressive direction of the values. As
expected (because greater adduction is expected for modal compared to falsetto register), the EGGW50 values rise after Singer 1 transitions into modal voice. This increasing trend is both visible and statistically significant. Varying patterns are found from G4 through transition, demonstrating no statistically significant trends.

**Figure 24.** Subject 1: EGGW50 Values for Descending Tasks (Vowels /a/, /i/, /o/, and /u/)

**Subject 2 descending tasks.** The scatterplots in Figure 25 contain three location data points for task tokens without an identifiable transition and five location data points for task tokens with an identifiable transition. With the exception of the /i/ task, Subject 2’s EGGW50 values are within a range of less than 0.1. All of the /u/ task tokens are within a range of less than 0.05. These values also reflect the general shape of the EGG5I trace in these tasks, which either gently decreased or stayed fairly constant. The /i/ task patterns are both visibly and
statistically different than the other three vowels because of the relatively large decrease in EGGW50 between G4 and E-flat4.

Figure 25. Subject 2: EGGW50 Values for descending tasks (vowels /a/, /i/, /o/, and /u/)

**Subject 1 ascending tasks.** Subject 1 showed an increasing trend in EGGW50 values between C4 and pre-transition followed by a decrease to transition (see Figure 26). There is also a decreasing trend between pre-transition and post-transition, followed often by an increase from post-transition to G4. This is a statistically significant pattern identified in the ANOVA, which predicts an EGGW50 increase between the C4 and pre-transition, followed by a decrease at transition and post-transition. The vowel that did not conform to this general pattern was the /o/, which was found to be significantly different than the other vowels in the ascending task.
A comparison of first and last pitches demonstrates an inverse relationship between pitch direction and EGGW50 values in both ascending and descending tasks. However, the decrease in EGG50 values in the ascending task between the C4 and G4 location data points is noticeably smaller than the corresponding EGGW50 value increase between G4 and C4 data points in the decreasing tasks.

Figure 26. Subject 1: EGGW50 Values for Ascending Tasks (Vowels /a/, /i/, /o/, and /u/)

**Subject 2 ascending tasks.** As observed in the descending tasks, the /o/ and /u/ ascending EGGW50 location data point values are within a narrow range (see Figure 27). The ascending /i/ task shows a reversed pattern as compared with the descending /i/ task: fairly consistent values between C4 and E-flat 4 with a larger increase between E-flat 4 and G4. Subject 2’s /a/ ascending task provides the best opportunity for comparison with Subject 1 because both subjects have identified
transitions in all three task tokens. Both Subject 1 and Subject 2’s EGGW50 values decrease between pre-transition and post-transition and increase between post-transition and G4, suggesting an adductory control strategy involving reduced adduction taking place often within the transition when the phrase is ascending.

![Figure 27. Subject 2: EGGW50 Values for Ascending Tasks (Vowels /a/, /i/, /o/, and /u/)](image)

**Discussion**

The instruction given to the subjects to sing the chromatic scales as they would in an operatic performance with orchestra was designed to elicit sufficient intensity to be heard over an orchestra and to sing with professional performance quality. This may be very different from how the subjects would perform the task if they were asked to completely mask the register transition maneuver. Henrich (2006) describes a register smoothing technique that relies on an increase in open
quotient and decrease in loudness to navigate the camouflage transition. Performing the task in this manner may not always be possible in a performance situation when decreasing loudness may reduce the singer's audibility by the audience. In such a predicament, a singer might sacrifice a perceptually smooth transition for audibility in the hope that the relative distance from the audience and orchestral texture will mask any perceptual transition-related irregularities. This was likely the basis of D. Miller and Schutte’s (2005) preference for singers who preserved their intensity levels with little regard for a smooth transition over a mezzo subject who was able to successfully minimized her transition while sacrificing loudness.

As seen in Table 2, previous studies have reported transition pitch locations in subjects ranging from 195 – 540 Hz, which spans well over an octave. In this study, Subject 1’s register transitions occurred between C#4 and F4 (approximately 270 – 340 Hz), with the greatest number of transitions taking place between D#4 and E4 (approximately 300 – 320 Hz). These results are within the 195-540 Hz transition frequency range reported for the subjects in Roubeau et al. (1987), and somewhat lower than for the subjects studied by Björkner et al. (2006), Keidar et al. (1987), Sonninen et al. (1999), Spencer and Titze (2001), Švec et al. (2008), and Sundberg and Kullberg (1999). The register transition pitch location range for Subject 2 was E4 to F#4 (approximately 320 – 360 Hz), which aligned her results more closely to previous studies. The chosen singing task and the subject’s preferred singing genre are also likely influences on the transition pitch location. The literature search uncovered no previous studies that used a chromatic scale
task to determine transition pitch locations and few that have studied any professional mezzo-sopranos.

The difficulty level of these tasks, even for high-level professional subjects, must be appreciated. Quotes from the subjects during data collection included “This is hard!” “We never practice chromatic scales, but we should.” “This is just the worst part of the voice!” and “I’m not sure I know how to sing anymore.” While the last quote was certainly tongue-in-cheek, it does illustrate some of the major challenges faced by mezzo-sopranos. The mezzo-soprano repertoire requires frequent high-intensity singing in this pitch range. Not by coincidence, this is also a highly problematic pitch region for tenors. However, this range is located higher in the pitch range of the tenor voice than it is in a mezzo-soprano’s (and therefore inherently louder), making the higher intensity requirements a special challenge for mezzos. There is some truth also in the statement that chromatic scales are not often used in vocalizing, either for warming up the voice or for exercise (diatonic scales are generally favored in this context). The chromatic scale was chosen because it allows for the precise determination of transition location in the context of a musical phrase, rather than a glissando. Glissandos are an inherently different vocal maneuver for a singer and do not necessarily reflect the challenges of this register when singing operatic repertoire. A chromatic scale also avoids the combination of whole and half steps found in a major or minor scale, a factor that may influence where transition occurs. The role of different intervals in the singer’s choice of transition location would be an interesting study of its own.
Both singers had vowels they preferred over others in this pitch range. As the results indicate, Subject 1’s transitions occur at lower pitches when performing /i/ tasks compared with the /a/ tasks. Statistical analysis indicates that Subject 2 uses a different adductory pattern when she sings /i/ compared with /a/. Most female singers will anecdotally acknowledge that /i/ simply feels easier and more resonant in this pitch range than /a/. Although Henrich et al. (2005) and Tokuda et al. (2010) discuss the difficulties that can arise in the resonance interactions that occur when F₀ is in close proximity to R1 (including involuntary register breaks), this interaction can also be quite useful when singing /i/ in this range. So long as the singer can exploit the relative acoustic advantage that the R1 location provides to the /i/ around E4 or E-flat4, she can execute the transition at a lower pitch without perceptually decreasing resonance (quality) or loudness. Moving the transition slightly away from the R1 location may also avoid some of the less desirable nonlinear vocal tract-resonance interactions. Therefore, she gets the “best of both worlds.” This strategy is not available in this range on the /a/ because its R1 location is approximately an octave above the highest pitch sung in these tasks. This creates a particular challenge for the singer. Subject 1’s transitions for the /i/ tasks were perceptually smoother than for the /a/ tasks. Subject 2 had no identifiable transitions in the /i/ tasks, whereas she did with all /a/ ascending tasks. These results are not surprising when the relative acoustic advantage of the /i/ and disadvantage of the /a/ in this pitch range are considered.

Informal observations of the audio samples (without identifiers linked to the
subjects) among the author, the author’s document supervisor, and advisors from the communication sciences department at Bowling Green State University generated some discussion about the considerable amount of apparent vowel modification both singers use in this register, especially in the /u/ vowel tasks. The author and author’s document supervisor from the School of Music at The Ohio State University did not find them particularly noticeable, most likely because the emphasis in this range for a classical singer is audibility and tone quality rather than vowel purity. The advisory faculty from the communication sciences and disorders department at Bowling Green State University immediately noticed the vowel quality changes, perhaps because vowel quality is an important evaluative tool in both singing and speaking research. Perhaps this is anecdotal evidence of an aural perception hierarchy that aligns with professional discipline.

The ANOVA did not detect any significance in the relationship of vowel and EGGW50 values in Subject 1 or Subject 2 (with the exception of the /i/ vowel), indicating that the vowel did not have a relationship with adduction changes indicated by changes in EGGW50 values during the progress of the phrase. This may be because the frequent resonance adjustments the subjects make with respect to vowel quality and high frequency spectral energy that help them maintain consistent intensity and tone quality through this range (a goal of the classical singing aesthetic). It would be interesting to compare these results with untrained singer subjects to determine if the vowel effect diminishes with training.

The data for Subject 2 poses some intriguing questions. The absence of
abrupt change in EGGSI, mostly consistent EGGW50 values, and lack of alteration in the DEGG wavegram all indicate that she is not using modal voice as it is most commonly defined. Although the sound is not lacking in substance at lower pitches, the perceptual tone quality and adductory measurements suggest that she may be using what Kochis-Jennings et al. (2012) termed “mixed” voice or Castellengo et al. (2004) identified as “voix mixte.” Both productions are defined as having elements associated with both modal and falsetto (head) voice. In this case, a relatively low contact quotient (associated with falsetto or head voice) is combined with spectral energy in high harmonics (particularly in the R2 and singer’s formant locations). Even the two transitions identified by the EGGSI trace in one /o/ descending task token and one /u/ descending task token were dissimilar to Subject 1’s in that they occurred on a single note (rather than on either side of a pitch change) and showed no obvious adjustments to the laryngeal vibratory mode in the DEGG wavegrams or EGGW50 values. Since the evidence of transition is only observed in the EGGSI, it is possible that the EGGSI is providing information influenced by acoustics, nonlinearities, or another variable entirely. One of the two transitions in question was not perceptually located. Had the definition of transition used in this investigation been formulated to require agreement between the DEGG wavegram and EGGSI criteria, these two task tokens would not have met that standard.

An argument could be made that Subject 2’s preference for spectral energy in the R2 area would be an alternative resonance strategy for /a/ in this pitch range. Subject 2 used an R2 dominant strategy in both the ascending and descending /a/
tasks, but had identifiable transitions in all three tokens of the ascending /a/ task. The same resonance strategy led to a different result, suggesting that it must be accompanied by an equally ideal adduction strategy to avoid a noticeable transition. These transitions fit all of the criteria of a transition, including obvious changes in the DEGG wavegram. These may be the only examples of Subject 2’s modal voice in this dataset as it is usually defined, due to its larger EGGW50 values, higher EGGSI, and evidence of a vertical phase difference in the DEGG wavegram. Would Subject 2 revert to a more traditionally defined modal voice at a lower pitch than C4? Would she combine register characteristics into a different “mixed” voice? Apart from extending the singing task to a lower range, these questions would be best answered with the addition of nasendoscopy to further illuminate and define the laryngeal postural and vibratory characteristics associated with these register behaviors.

Subjects 1 and 2 provide a truly fascinating contrast, especially considering their high level of singing skills and not dissimilar physiques. They have also both been successful singing some of the same operatic repertoire, despite having some perceptually different tonal characteristics and divergent register strategies. Some of those tonal characteristics may perhaps be attributable to the differences in resonance and adduction strategies observed in this study.
Part Five: Conclusions

The results of this study support the assumption that Subjects 1 and 2 employ highly contrasting registration strategies in this pitch range relative to adduction and acoustics. Subject 1 displayed evidence of modal to falsetto and falsetto to modal register transitions in all vowels and all tasks. Subject 2 only displayed evidence of register transitions in /a/ ascending tasks and inconsistent evidence of a register transition in the /o/ and /u/ descending tasks. Transitions were defined by criteria related to (a) perceptual changes in sound quality, (b) changes in EGG signal intensity (increasing in descending tasks; decreasing in ascending tasks), and (c) increased variation of vibrato rate and extent. DEGG wavegrams added qualitative depth to transition observations. Increased variation of vibrato rate and extent related to register transition was discovered in 79% (19 of 24 task tokens) of Subject 1’s samples and 4 of 5 task tokens of Subject 2’s samples.

Subject 1’s transitions mostly occurred at higher pitches when singing /a/ tasks than /i/ tasks. There was a visible, but not statistically significant, trend for transitions to occur at higher pitches in ascending tasks than descending tasks. Subject 2 did not produce enough transitions to generate meaningful data in relation to transition location.
Based on analysis of FFT spectral slices relative to location data points, Subject 1 employed a resonance strategy dominated by R1 and, to a lesser extent, spectral energy in the singer’s formant range. Subject 2 employed a resonance strategy characterized by the spectral prominence of harmonics boosted by R2, which is occasionally assisted by the close proximity of either R1 or R3.

The contrasting EGGW50 value patterns observed suggest that Subject 1 may have a different adductory strategy than Subject 2. For Subject 1, EGGW50 values tend to increase for descending tasks and decrease for ascending tasks relative to the transition. The ANOVA for Subject 1 revealed significant interactions between transition-related locations and EGGW50 values, phrase direction and EGGW50 values, vowel and phrase direction, and transition-related locations and phrase direction. These interactions indicate that Subject 1 may also employ a different adductory strategy for descending tasks than for ascending tasks. The ANOVA for Subject 2 identified a significant difference in the EGGW50 value patterns between the /a/ and /i/ vowels.

The data indicate that Subject 1 has an audible, objectively identifiable transition in all of her 36 samples. The pre-transition, transition, and post-transition adduction behaviors occasionally draw parallels to the “smoothing technique” described by Henrich et al. (2005), characterized by an increased open quotient and decreased intensity in preparation for the transition from modal to falsetto (head) voice. Subject 2 does not have an audible, objectively identifiable transition in 29 of her 36 samples, which is perhaps attributable to the possible use of a register mix in
this range that combines a relatively low contact quotient with high frequency spectral characteristics.

In conclusion, this study provides evidence of divergent acoustic and adduction-related registration strategies in two professional operatic mezzo-sopranos. Much would be gained by gathering similar data from other professional operatic mezzo-soprano subjects and mezzo-soprano subjects at other experience levels. The extension of the subject population to other female voice types may also be illuminating.

Future studies may utilize the chromatic scale tasks used in this study with the inclusion of additional, complementary tasks. Tasks involving glissandos, larger intervals, isoparametric tones, or *messa di voce* may allow for observations of contrasting register behaviors in the same subject. To further contextualize any mixed voice strategies observed in subjects, a task involving pitches in the extremes of the pitch range may clarify register norms.

In addition to further study of the contacting and de-contacting intervals as provided by EGG and DEGG, inverse filter- and videokymographic-based adduction measurements would yield data that would refine our understanding of these singers’ register behaviors at the glottal level. The maximum flow declination rate (MFDR) in particular would add dimension to the EGG and DEGG data describing the glottal closing gesture. Videokymography would add definition to the information in the DEGG signal suggesting longitudinal phase differences in the vocal fold vibratory cycle, and identify any general inconsistencies related to issues of glottal closure.
The inclusion of videokymography, with its necessary accompaniment of laryngoscopy and imaging equipment represents a “luxury item” in that it requires specialized facilities and personnel to facilitate subject recordings. The unavailability of many professional operatic subjects because of their travelling and singing schedules may make the use of higher-cost imaging prohibitive and impractical. Encouraging singers of this caliber to participate in studies employing more portable, affordable instruments should be prioritized when more sophisticated imaging is inaccessible.
Bibliography


Mexican. Retrieved from
http://www.santafenewmexican.com/news/local_news/review-joyce-didonato-
dazzles-in-la-donna-del-lago/article_ce1169f8-8889-5c51-ae04-
21b626ea8eb3.html


Smith, Jr., M. L. (2011). *Adolphe Nourrit, Gilbert Duprez, and the high C: The influences of operatic plots, culture, language, theater design, and growth of orchestral* 86


