Vocal Vibrato Variability: Novel Analytical Tools & Diagnostic-Pedagogical Approaches in Diverse Genres

Theodora Ivanova Nestorova

Interdisciplinary Studies:
Computational Acoustic Modeling Laboratory - Music Technology Area
Music Performance and Body Laboratory - Music Education Area
Vocal Science & Pedagogy - Voice Area

Departments of Music Research and Music Performance, Schulich School of Music
McGill University
Montréal, Québec, Canada

July 2025

A thesis submitted to McGill University in partial
fulfillment of the requirements
of the degree of Doctor of Philosophy

Abstract

This dissertation presents a cross-disciplinary compendium of acoustic, perceptual, and biomechanic studies on naturally-occurring, time-varying vocal vibrato across several musical genres. Vibrato is a complex, multi-factorial feature tied to singing technique and artistry that is both volitionally and non-volitionally controlled by the vocalist.

A comprehensive literature review of existing research identifies the need for advanced analytical tools accounting for vibrato's complexity and time-variant nature. Existing vibrato analysis methods are based on calculating averages that do not consider ecologically valid time domain-based changes in singing. Normative vibrato standards are derived from Western classical singing studies, signaling a lack of diversity. This culturally biased gap overlooks other music styles and aesthetic traditions, leading to misrepresentation and mischaracterization of vibrato.

To address these challenges, a methodological study introduces a novel theoretical framework to classify vibrato patterns into half-extent time profile archetypes. Focusing on professional singing samples drawn from Opera, Musical Theatre, and Jazz genres, the study proposes novel parametric techniques for vibrato analysis, providing a systematic approach including Coefficient of Variation, 4-PL regression models, and new metrics like Vibrato Onset Time to differentiate vibrato types.

A secondary, extended analysis of interview data examines how vibrato variability influences timbre semantics, analyzing cognitive interpretations from vocalists and instrumentalists describing unaccompanied vocal stimuli sung by professional singers. Using qualitative content vocabulary and thematic analysis, this perceptual study reveals frequent mentions of vibrato and strong correlations between vibrato variability and perceived vocal function. Findings underscore the perceptual validity and salience of temporally variant vibrato and its relevance for future assessment methodologies.

Field recordings collected in the Balkan Mountains enable acoustic analysis of vibrato in Bulgarian Folk Singing. Results highlight distinctive vibrato patterns and their function within traditional performance practices, expanding the scope of vibrato research to an Eastern musical tradition known for its distinctive vocal styles. Highly complex, multiphasic vibrato contours and unique vibrato patterns are interpreted alongside effective teaching strategies of cultural-pedagogical significance.

As no consensus or codified pedagogy exists on addressing vibrato in teaching contexts, an integrative investigation synthesizes the practical application of scientific vibrato research. Direct (explicit) vs. indirect (implicit) instruction using variable vibrato as a more comprehensive and style-specific diagnostic tool for assessment and training is explored. Created exercises promoting vibrato versatility and efficiency are recommended.

Building on these foundations, vibrato variability, its potential as a diagnostic tool, and its parallel to tremor are investigated. A pilot study leverages machine learning to assess vibrato as a biomarker for functional and neurogenic voice disorders, with applications in clinical and pedagogical contexts. The study further explores vibrato as an indicator of vocal stress and proposes methods for integrating vibrato analysis into vocal health assessments.

By integrating perceptual research, computational modeling, ethnographic analysis, and pedagogical innovation, this dissertation advances the field of vibrato studies. These interdisciplinary, mixed methods studies demonstrate vibrato variability as perceptually important and propose a more genre-inclusive system of complex, time-varying vibrato acoustic parameters beyond averages. The implications of this research to evidence-based pedagogical and medical/clinical practices have the potential to transform and modernize vibrato training in singing teaching and vocal health rehabilitation in voice therapy.

Résumé

Cette thèse présente un recueil interdisciplinaire d'études acoustiques, perceptives et biomécaniques sur le vibrato vocal naturel et variable dans le temps dans plusieurs genres musicaux. Le vibrato est une caractéristique complexe et multifactorielle liée à la technique de chant et à l'art, qui est contrôlée par le chanteur de manière volontaire et non volontaire.

Une analyse documentaire complète des recherches existantes a mis en évidence la nécessité de disposer d'outils d'analyse avancés tenant compte de la complexité du vibrato et de sa nature variable dans le temps. Les méthodes d'analyse du vibrato existantes sont basées sur le calcul de moyennes qui ne prennent pas en compte les changements du chant dans le domaine temporel, valables d'un point de vue écologique. Les normes relatives au vibrato sont dérivées d'études sur le chant classique occidental, ce qui témoigne d'un manque de diversité. Cette lacune culturelle néglige d'autres styles de musique et traditions esthétiques, ce qui conduit à une représentation et une caractérisation erronées du vibrato.

Pour relever ces défis, une étude méthodologique introduit un nouveau cadre théorique permettant de classer les modèles de vibrato en archétypes de profils temporels de demi-extension. En se concentrant sur des échantillons de chants professionnels issus de l'opéra, du théâtre musical et du jazz, l'étude propose de nouvelles techniques paramétriques pour l'analyse du vibrato, en fournissant une approche systématique incluant le coefficient de variation, des modèles de régression 4-PL et de nouvelles mesures comme le temps d'apparition du vibrato pour différencier les types de vibrato.

Une étude perceptive transversale examine comment la variabilité du vibrato influence la sémantique du timbre, en analysant les interprétations cognitives de chanteurs et d'instrumentistes décrivant des stimuli vocaux non accompagnés chantés par des chanteurs professionnels. En utilisant un vocabulaire de contenu qualitatif et une analyse thématique, l'étude révèle des mentions fréquentes du vibrato et de fortes corrélations entre la variabilité du vibrato et la fonction vocale perçue. Les résultats soulignent la validité perceptuelle et l'importance de la variabilité temporelle du vibrato, ainsi que sa pertinence pour les futures méthodologies d'évaluation.

Les enregistrements de terrain réalisés dans les Balkans permettent une analyse acoustique du vibrato dans le chant folklorique bulgare. Les résultats mettent en évidence des modèles de vibrato distinctifs et leur fonction dans les pratiques d'interprétation traditionnelles, élargissant

ainsi la portée de la recherche sur le vibrato à une tradition musicale orientale connue pour ses styles vocaux distinctifs. Des contours de vibrato très complexes, multiphasiques et des modèles de vibrato uniques sont interprétés parallèlement à des stratégies d'enseignement efficaces d'une grande importance culturelle et pédagogique.

Comme il n'existe pas de consensus ou de pédagogie codifiée pour aborder le vibrato dans les contextes d'enseignement, une enquête intégrative synthétise l'application pratique de la recherche scientifique sur le vibrato. L'enseignement direct (explicite) et indirect (implicite) utilisant le vibrato variable comme outil de diagnostic plus complet et spécifique au style pour l'évaluation et la formation est exploré. Des exercices créés favorisant la polyvalence et l'efficacité du vibrato sont recommandés.

Sur ces bases, la variabilité du vibrato, son potentiel en tant qu'outil de diagnostic et son parallèle avec le tremblement sont étudiés. Une étude pilote s'appuie sur l'apprentissage automatique pour évaluer le vibrato en tant que biomarqueur des troubles vocaux fonctionnels et neurogènes, avec des applications dans des contextes cliniques et pédagogiques. L'étude explore également le vibrato en tant qu'indicateur du stress vocal et propose des méthodes pour intégrer l'analyse du vibrato dans les évaluations de la santé vocale.

En intégrant la recherche perceptive, la modélisation informatique, l'analyse ethnographique et l'innovation pédagogique, cette thèse fait progresser le domaine des études sur le vibrato. Ces études interdisciplinaires et à méthodes mixtes démontrent que la variabilité du vibrato est perceptible et proposent un système de paramètres acoustiques du vibrato complexes et variables dans le temps, au-delà des moyennes, qui englobe davantage de genres. Les implications de cette recherche pour les pratiques pédagogiques et médicales/cliniques fondées sur des preuves ont le potentiel de transformer et de moderniser la formation au vibrato dans l'enseignement du chant et la réhabilitation de la santé vocale dans la thérapie vocale.

Acknowledgements

I owe the completion of this dissertation– and the wild, beautiful journey it represents (examining vibrato "in the wild")– to many people whose support, guidance, and collaboration shaped every step of the way.

First and foremost, I extend my deepest thanks to my supervisors, Gary Scavone and John Mac Master, for their steadfast support and expert guidance. Gary, thank you for encouraging me to explore the acoustics of voice with rigor and curiosity and fervor of my own path. John, your encouragement, support, and belief in my pedagogical and artistic voice have been deeply formative. And to my doctoral advisor, Isabelle Cossette– thank you for your vision, mentorship, and boundless genuine care. You all saw the value and potential in me charting my own interdisciplinary path before I had the vision to see where it would take me and the language to describe it. Developing academically while Working with you all has been both a privilege and a joy.

To my longtime mentors– many of whom I am lucky to call collaborators on the very projects presented here– thank you for shaping not just my research but who I am as a scholar, performer, teacher, and practitioner: Johan Sundberg, Ingo Titze, Ron Scherer, Dominique Labelle, Luc Mongeau, Stephen McAdams, Malte Kob, Ian Howell, Dana Varga, Martin Vácha, Josh Glasner, John Nix, Yvonne Gonzales Redman, Nick Perna, Allen Henderson, Matt Hoch, Ken Bozeman, Joanne Bozeman, Amelia Rollings Bigler, Lynn Maxfield, Kittie Verdolini Abbott, Eric Hunter, Pasquale Bottalico, Aaron Johnson, Jenevora Williams, Kari Ragan, Filipa Lã, Brian Gill, Marcel Montrey, Lindsey Reymore, Kit Soden, Jay Marchand Knight, Maude Desjardins, Michael McMahon, Ana Sokolović, Martin Spencer, Christian Herbst, Bodo Maass, Vasileios Chatziioannou, Matt Edwards, Maria Gradeshlieva, Galia Petrova-Kirkova, Margarita Ilieva, Velizara Karayankova, Stephen Hargreaves, Andrea Creech, Liliana Araújo, Fabrice Marandola, Martha de Francisco, Pilar Bertuzzi, D.D. Michael, Marci Rosenberg, Deanna McBroom, Jeanie LoVetri, Aaron Williamon, Martin Dubé, Stéphane Lemelin, Nancy Czemmel, Hélène Drouin, Carole Haber, Lorraine Manz, and many, many others who are not listed here (but to whom I am no less grateful)– thank you, truly. Your guidance, support, intellectual, artistic, pedagogical, and clinical generosity, collaboration, and friendship have meant the world and shaped my journey. To the multi-disciplinary international organizations of which I am a member and/or board member, the Pan American Vocology Association (PAVA), National Association of Teachers of Singing (NATS), National Center for Voice and Speech (NCVS), The Voice Foundation (TVF), American Speech-Language-Hearing Association (ASHA), Performing Arts Medicine Association (PAMA), Vocal Health Education, Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Analysis, Creation, and Teaching of ORchestration Project (ACTOR Project), McGill Music Graduate Students' Society (MGSS), Fulbright Alumni Association; thank you for believing in me and empowering me to grow and shine from a young investigator to an early career professional. From paper and conference awards to educational resources to individualized mentorship, it means the world to me that you have believed and continue to believe in me. You remind me that the way is dreaming to believing to achieving, that the sky is truly the limit, and that continually innovating my work with my natural inclination of thinking outside the box and doing things uniquely makes a difference to someone, somewhere.

A heartfelt thank you to Françoise Chagnon and your Vocalex Clinique ORL / ENT Clinic, LLC. (formerly AXiO Performing Arts & ENT Clinic); your mentorship and our collaborative projects and initiatives in the realm of voice medicine, clinical health, and science has been truly meaningful and formative for me. My immersive internship training experience with you brought my work to life. Thank you, too, to the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) for resources, inspiration, and a home base for so many of my ideas. Thank you to the School of Communication Sciences and Disorders for the academic grounding, and to the Faculty of Engineering for the resource support. I also acknowledge the Applied Performance Sciences Hub with thanks for nourishing the scientist-musician in me. My unquenchable thirst and pursuit for knowledge and skills has been supported by all of you.

To my dear colleagues at the Computational Acoustic Modeling Laboratory (CAML) and Music Performance and Body Laboratory (MPBL) (and also Music Perception and Cognition Laboratory (MPCL)); Alberto Acquilino, Miranda Jackson, Champ Darabundit, Ninad Puranik, Xinmeng Luan, Diako Kaboodi, Takuto Yudasaka, Mark Rau, Song Wang, Margaret Hopkins, Esteban Maestre, Douglas Beaton, Caroline Holden, Jacqueline Woodley, Mariane Generale, Erivan Duarte, Ziyue Piao, Mathieu Boucher, and many more (including in voice studios and classes): thank you for being my labmates, peers, collaborators, and LaTeX lifesavers. Your kindness and insight made the research process not only possible but joyful.

To my wonderful students, both at McGill University (espeically in my curated, inaugural Voice Science, Health, & Disorders course [MUPG 474-001], and in the Vocal Techniques course [MUCT 235-001]), in Canada, in the U.S., in Europe, and in my private voice studio around the globe: working with you has been one of the most fulfilling parts of my academic life. Your questions, voices, and curiosity continually inspire me, and remind me why I teach.

From my years as an undergraduate student of Voice at Oberlin College & Conservatory, to my time as a Fulbright Study/Research Grant scholar at the Universität für Musik und darstellende

Kunst Wien (University of Music and Performing Arts Vienna) in Austria, to my master's degree in Vocal Pedagogy at New England Conservatory, to my M.B.A. in Arts Entrepreneurship from the Global Leaders Institute for Innovation, to this Ph.D. at McGill University in Interdisciplinary Studies– gratitude is at the center of what I feel when I look back on my academic and personal journey.

To my current and new institution, Viterbo University: thank you for welcoming me as I take the next leap in my career to teach as an Assistant Professor of Speech-Language Pathology in this new Communication Disorders and Sciences department, while building new cross-disciplinary initiatives with the Conservatory for the Performing Arts. I am so grateful to continue this journey of life-long learning, growing, and giving back to the communities through the power of the communication.

To my partner-in-life-and-music, my beloved fiancé Alex: thank you for supporting me unconditionally, walking this eclectic and non-traditional path beside me with so much love, humor, and music. Our Pizzicanto Soprano-Cello Duo has been a constant reminder that artistry and partnership can be one and the same. I can't wait for our next adventures; both personally and professionally!

To my parents and my brother; Мама, Тати, и Аци: I would not be where I am without you. You have always nurtured the driven and determined trailblazer in me, sacrificed the world so that I could pursue my varied interests, and encouraged me to reach for the stars and shine. Thanks for allowing me to be the "curious philosopher" who is always voicing, communicating, and expressing, from the very moment I came into the world. Thank you for your unconditional love and unwavering support through the many years, cities, countries, miles, and chapters. Though I was born in Manchester, U.K., and grew up across both coasts of the U.S., my heart has always belonged to Europe and our homeland culture– I carry my 100% Bulgarian roots with pride and gratitude. Being a third-culture, multi-national, multi-lingual kid taught me how to listen deeply, communicate across borders, and feel at home in many worlds– a fitting metaphor for interdisciplinary research, too.

With deep gratitude to my homeland, Bulgaria– land of song, spirit, and ancient wisdom. To my grandmother, my Баба, who is unwavering and steadfast, and to my grandfather, Дядо, and grandparents, who are watching over me from above: I only hope to make you proud. To my great-grandmother, Атанаска Тодорова / Atanaska Todorova, and great-uncle Манол Тодоров / Manol Todorov: when I sing, you sing through me with your voices. I offer my heartfelt thanks to my cultural heritage and legacy of all those who came before me and all of my family and friends who have supported me on this path of learning, music, and service to others. As a logophile and linguaphile, my voice, my music, and my pursuit of knowledge are bridges across cultures, hearts, and generations. I dedicate this work to the belief that learning never ends, and that through both

art and science, we can heal, connect, and inspire.

Figure 1: A modern performance opportunity for the author of this thesis, to sing with her great-grandmother, one of the first Bulgarian folk singers ever recorded on the radio, Атанаска Тодорова / Atanaska Todorova, who gave the author her voice, spoken, sung, and written, and the ultimate impetus to write this thesis.

And finally: Знанието е сила.

Contributions

This thesis and the research it describes is original work by the candidate, except for commonly understood and accepted concepts, or where explicit reference to the work of others is made. Multiple references throughout the thesis chapters are made to conference-only presentations, which are not included in the list below. The dissertation is formatted as a traditional monograph thesis comprised of eight chapters and four appendices. It includes restructured and reintegrated content from the following journal publications:

1. **Included in Chapter 3:** Nestorova, T., Nestorov, I., and Scavone, G. (2025). Analysis and modeling of time-varying vibrato contours and complex patterns. *The Journal of the Acoustical Society of America*. [Manuscript under review].

2. **Included in Chapter 4:** Nestorova, T., Reymore, L., Marchand Knight, J., and Soden, K. (2025). Speaking of vocal vibrato: Timbre perceptions and semantic descriptions with pedagogical and clinical implications. *Voice and Speech Review*. [Manuscript accepted; expected publication in late 2025].

3. **Included in Chapter 5:** Nestorova, T., Nestorov, I., Nestorova, D., and Gradeshlieva, M. (2025). Bulgarian folk singing: The art and the science. *The Journal of Voice*. [Manuscript under review].

4. **Included in Chapter 6:** Nestorova, T. (2026, in press-a). To teach or not to teach vibrato?: Implicit versus explicit instruction in the vocal studio; Part I: Vibrato pedagogy in past to present pedagogical trends. *Journal of Singing*. [Manuscript accepted; expected publication in early 2026].

5. **Included in Chapter 6:** Nestorova, T. (2026, in press-b). To teach or not to teach vibrato?: Implicit versus explicit instruction in the vocal studio; Part II: Practical and applicable tools, tips, and tricks. *Journal of Singing*. [Manuscript accepted; expected publication in early 2026].

6. **Included in Chapter 7:** Nestorova, T., Nestorov, I., and Chagnon, F. (2025). Diagnosability

of vibrato for voice (re)habilitation: Clinical implications of vibrato variability. *Logopedics Phoniatrics Vocology.* [Manuscript submitted for publication].

For each study and corresponding manuscript and chapter in this thesis, candidate conceptualized, designed, and performed all data collection and analysis, including, but not limited to: recruiting study participants, creating experimental protocol and procedures, taking measurements, conducting quantitative acoustic analyses, numerical and computational modeling, qualitative perceptual analyses, and statistical calculations. The candidate wrote and prepared all manuscripts and presentation materials for the conference and journal references cited above (as well as many others). Preliminary script coding and original programming was conducted by I. Nestorov, but the candidate, T. Nestorova, as principal author, was responsible for generating research questions, objectives, gathering/collecting data, designing experiments, recruiting and running participants, analyzing data, interpreting results, and writing the manuscripts listed above.

A detailed account of the individual contributions of each author to these publications is provided at the beginning of each corresponding chapter.

Supervisors Dr. Gary Scavone, Professor John Mac Master, and advisors Dr. Isabelle Cossette and Dr. Françoise Chagnon reviewed, provided feedback, and suggested edits for the majority of manuscripts. They also provided guidance and work space for laboratory research.

Contents

List of Figures

List of Tables

List of Acronyms and Symbols

The list below includes the most frequently employed acronyms and symbols in this thesis. While other acronyms are singularly defined, the ones in this list are included multiple times:

- Research question (RQ)
- Mean (M or $mean$)
- Standard deviation (SD)
- Confidence interval (CI)
- Fundamental frequency (f_o)
- Hertz (Hz)
- Decibels (dB)
- Amplitude modulation (AM)
- Frequency modulation (FM)
- Spectral Envelope modulation (SEM)
- Sound Pressure Level (SPL)
- Fast Fourier Transform (FFT)
- First vocal tract resonance / formant (R1)
- Second vocal tract resonance / formant (R2)
- Third vocal tract resonance / formant (R3)
- First harmonic (H1)
- Second harmonic (H2)

- Third harmonic (H3)
- Linear Predictive Coding (LPC)
- Signal-to-Noise Ratio (SNR)
- Rate of Frequency Change (RFC)
- Just-noticeable difference (JND)
- Half-Extent (HE)
- Half-of-Extent (HoE)
- Relative Half-Extent (rHE)
- Akaike Information Criterion (AIC)
- Bayes Information Criterion (BIC)
- Second (sec.), seconds (secs.)
- Millisecond(s) (ms)
- Term frequency-inverse document frequency (TF-IDF)
- Generalized linear mixed model (GLMM)
- Likelihood ratio test (LRT)
- Intraclass correlation coefficient (ICC)
- Western art music (WAM)
- World music (WM)
- Western classical singing (WCS)
- Musical Theater / Musical Theatre (MT)
- Jazz (J)
- Contemporary Commercial Music (CCM)
- Bulgarian folk music (BFM)
- Bulgarian folk singing (BFS)

- Vocal fold (VF)
- Thyroarytenoid (TA) muscle
- Cricothyroid (CT) muscle
- Pulse (M0), modal (M1), falsetto (M2), and whistle (M3) vocal registers
- Otorhinolaryngologist (ENT)
- Speech-Language Pathologist (SLP)
- Muscle Tension Dysphonia (MTD)
- Temporomandibular joints (TMJs)
- Temporomandibular joint dysfunction or disorders (TMD)
- Primary Muscle Tension Dysphonia (pMTD or MTD-1)
- Painful Temporomandibular Disorder (pTMD)
- Critically Appraised Topic (CAT)
- Heart rate variability (HRV)
- Essential tremor (ET)
- Multiple sclerosis (MS)
- Amyotrophic lateral sclerosis (ALS)
- Graphical user interface (GUI)

Other Notes

In Chapter 4, vocalist participants or listeners refer to the listener participants interviewed for the perceptual study. Singers refer to those recorded as source recorded stimuli samples / excerpts for the perceptual study.

Throughout this thesis, several terms are used interchangeably (and should be considered synonyms for the purposes of this thesis), including, but not limited to:

- Genre / Style
- Time-varying / Temporally-variant
- Half-extent time profile(s) / Vibrato archetype(s)

Chapter 1

Introduction and Motivation

The functional use and artistic training of the vocal mechanism is a challenging and multifactorial process. An integral part of this process in singing includes vibrato, a oscillation in pitch and intensity of sound which is ubiquitously known across musical genres and cultures. Vibrato is a natural feature used for artistic means in both instrumental and vocal music. Mastery of vocal technique and creative expression, particularly in relation to vibrato, represents one of the most nuanced and lifelong pursuits for singers across genres. Vibrato has long been regarded as a hallmark of vocal artistry, influencing both technical precision and emotional expressivity. Its variability not only serves as an indicator of stylistic authenticity and genre conventions but also may reflect the health, training, and adaptability of the voice. However, despite its centrality to vocal performance, the analytical, diagnostic, and pedagogical approaches surrounding vibrato remain steeped in normative tradition and subjectivity. This dissertation addresses this gap by proposing innovative analytical tools integrated into diagnostic-pedagogical frameworks to explore vibrato variability across diverse musical genres.

The study of the singing voice, and vibrato in particular, is of profound anthropological importance for several reasons. First, the voice, as a primary instrument of human communication, transcends cultural and linguistic boundaries. Evolutionarily, vibrato has been considered a natural state of the voice; present in animalistic utterances and heightened speech patterns. Viewed as a universal extension of emotional arousal, pitch vibrato "occurs in sustained speech; and in very hearty laughter of the adult… It is present in the vigorous crying of the newborn infant, in the bark of the dog, in the cooing of the dove" (Seashore, 1936). Vibrato, an inherent attribute of the voice's organic nature, is often added to the pitch and/or intensity production of instruments seeking to replicate the sound of the singing voice.

Understanding vibrato enables a deeper exploration of how physiological, acoustic, and stylistic factors intersect, offering insights into vocal health, performance optimization, and artistic identity. Second, vibrato may serve as a diagnostic biomarker for vocal function and pathology, making its study essential for interdisciplinary applications in both pedagogical and clinical settings. Third, the exploration of natural vibrato variability can illuminate style-specific nuances, as vibrato is broadly known as a stylistic tool used differently across various genres, traditions, and contexts. Moreover, vibrato study fosters culturally-informed and evidence-based pedagogical practices that respect the diverse needs of vocalists.

Unlike fields such as athletics, where technology and science have become integral to understanding and optimizing performance, vocal pedagogy has only recently begun to incorporate and apply systematic, quantitative methodologies from voice analysis. Yet, singers are identified as vocal athletes, utilizing the vocal subsystems with refined coordination. Singers have historically been guided by qualitative interpretations rooted in stylistic dogma and overgeneralized perspectives rather than empirical evidence. This reliance on anecdotal and subjective perspectives can hinder the development of science-informed and individualized teaching strategies.

Vibrato, much like other systems engaged in vocalization, is largely idiosyncratic. It therefore may illuminate biomechanical function and use of the vocal mechanism as a whole. As a voice performer, pedagogue, and clinician myself, I have encountered several critical incidents related to vibrato that have inevitably led to my research interests in the realm of vibrato variability. Deconstructing the cause and effect of coordination issues for various singing skills, including vibrato, has continually raised more and more curiosity for me, my own teachers, my colleagues, and my own students. For most, vibrato seems to be a signal to the "inner workings" of their voice.

Through both quantitative and qualitative methods, the research presented in this thesis is intended to harness acoustic and psychoacoustic analysis modeling alongside ethnographic and pedagogical approaches. Quantitative analyses will provide objective insights into vibrato's acoustic properties and informed inferences into vibrato's physiological underpinnings. Qualitative studies will uncover the lived experiences and perceptions of singers, their pedagogues, and their audiences. Together, these methods foster a more comprehensive and inclusive understanding of vibrato variability and its implications.

The development of novel diagnostic-pedagogical tools for vibrato variability in this dissertation has the potential to revolutionize vocal pedagogy and performance practices, contributing a genre-sensitive pedagogical paradigm to the broader fields of voice science, music education, and vocology.

1.1 Background on Vocal Vibrato

Vibrato involves many different kinds of fluctuations over the duration of a sustained tone, contributing to the perceived pitch, loudness, and timbre (Miller, 1986). Acoustically, vibrato is defined as a "vibrating quality of musical sounds, corresponding to simultaneous modulations of amplitude (AM), frequency (FM) and/or spectral envelope (SEM)" (Verfaille and Guastavino, 2005). These quasi-periodic aspects of vibrato often co-exist when produced by most instruments; there is a correlation between amplitude modulation and frequency modulation (McAdams and Rodet, 1988). It is believed that this frequency-amplitude connection contributes to vibrato's influence in timbre perception (Henry, 2021). While singing vibrato is usually understood in terms of oscillations in pitch, these pitch fluctuations are dependent on spectral (or frequency composition) and intensity (or perceived loudness/volume) changes because the spectral response of the vocal instrument is a strong function of its fundamental frequency (or perceived pitch).

In the singing voice, vibrato is not only an acoustic and perceptual phenomenon, but also a physiological and aerodynamic one, governed neurologically. Production of vibrato involves cortically coordinated (Lester-Smith et al., 2022) simultaneous and synchronous oscillations of various tissues, airflow, and acoustic properties of the vocal tract (Nandamudi and Scherer, 2019; Rothenberg et al., 1988; Dejonckere et al., 1995; Hsiao et al., 1994). Nix (2014) defines vocal vibrato as follows:

> Vibrato is a periodic oscillation of the fundamental frequency (perceived as pitch) and all its harmonics, amplitude (perceived as apparent intensity or volume), timbre (a result of harmonics sweeping through vowel formants), subglottic pressure, closed quotient (the percentage of each vibration cycle the vocal folds are in contact), and formant frequencies.

The mechanism of vocal vibrato has been proposed as a reflex-resonance model with a long latency by Titze et al. (2002). This reflex-resonance model suggests that the plausible physiology of vibrato production stems from cortical commands controlling phonation and average fundamental frequency of the sung sound, f_o, being modulated by an oscillatory system. The centrally controlled oscillators then act upon the laryngeal cricothyroid (CT) and thyroarytenoid (TA) muscles and involve agonist-antagonist muscles that change vocal fold length, as displayed in Figs. 1.1 and 1.2. Detecting these vocal fold length adjustments, laryngeal sensory receptors initiate a motor response opposing the modulation. This proprioceptive response repeats and maintains the simultaneous oscillation of vocal fold length correlated with f_o. While accepted to be prototypical in scope, this reflex-resonance model of vibrato may not fully capture aeroacoustic properties, proprioceptive significance, or differences in the model's behavior during time-varying vibrato changes. Auditory feedback in the form of a control loop with both feedforward and feedback mechanisms has been identified as the regulating agent of vibrato (Leydon et al., 2003; Deutsch and Clarkson, 1959; Ackermann and Altenmüller, 2021). Though still subject to future inquiry, physiologic tremors from the kinesthetic somatosensory reflex may interact with the auditory f_o reflex through the reflexive brainstem network when producing vibrato (Jones and Keough, 2008). While experienced professional singers may be able to cultivate vibrato by increasing the gain, scaling, and timing in the reflex loop, real-time volitional regulation of vibrato excursion amplitude is compositely adjusted from targeted speed (rate) (King and Horii, 1993). Historically, rate has been posited to be generally consistent within a singer and singers have been found to be able to voluntarily control the amplitude or excursion of vibrato fluctuation (Titze et al., 1994).

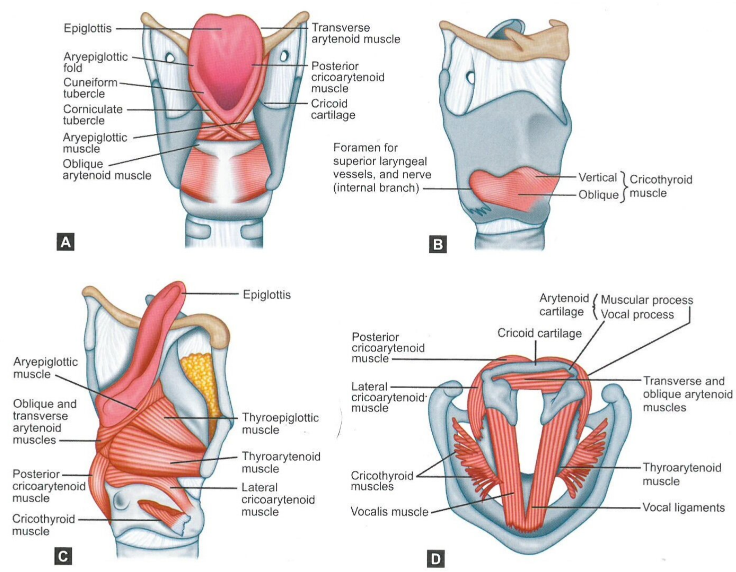


Figure 1.1: The anatomical and physiological structure of the larynx, with labeled laryngeal cartilages and intrinsic musculature. Reprinted with permission from Sataloff et al. (2014).

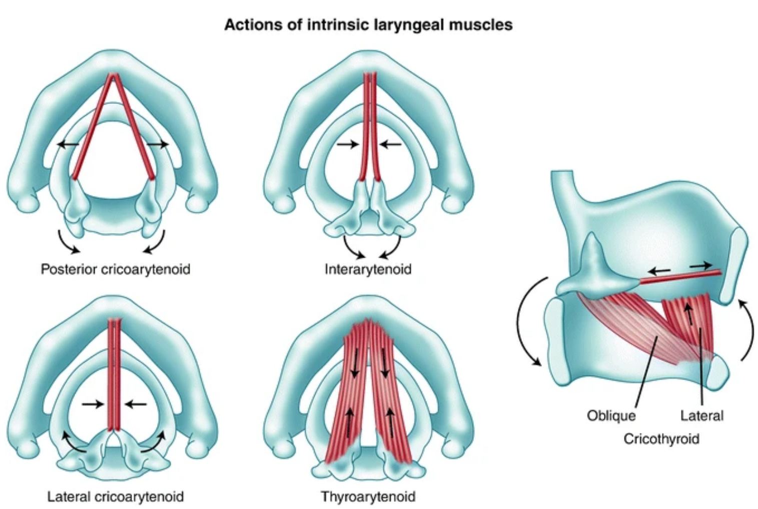


Figure 1.2: Actions of intrinsic muscles of larynx, including cricothyroid and thyroarytenoid muscles and agonistic-antagonistic movements that change vocal fold length and are active in vocal vibrato production. Reprinted with permission from Wadie et al. (2013).

Because vibrato involves complex coordination of neurological, biomechanical, acoustical, and perceptual elements, it is viewed as a by-product of vocal technique and musical style. Sundberg (1987) furthers this by stating that "it is important to stress that different types of vibrato exist and they probably differ from a physical point of view." Moreover, differing production techniques dictated by musical genres and singing styles have shown potential as extremely impactful factors distinguishing vibrato features (Nestorova et al., 2023). As vibrato can be both volitional (voluntary) and non-volitional (involuntary), less explored territory is the implication of vibrato variability for efficient voice functioning, both in training and therapeutic settings.

This research is centered on addressing three key gaps in the existing literature:

1. Investigating the interaction of vibrato and different musical genres/styles,
2. Understanding naturally-occurring quasi-periodicities in vibrato and why extent may be more volitionally controlled, and
3. More comprehensively evaluating vibrato as indicative of underlying voice function.

This provides rationale for examining time-varying aspects of vibrato, proposing a new system of vibrato metrics which considers the regularity, variability, and shape of the full tone in more genres over time.

1.2 Acoustic Parameters for Quantifying Vocal Vibrato

The properties of vibrato have been quantified most commonly in the frequency realm using averages, as vibrato is typically quasi-periodic and quasi-sinusoidal. Vibrato rate refers to the number of fluctuations of the fundamental frequency, f_o, above and below a mean value per second. Average modern vibrato rates of operatic singing lie between 4.5 and 6.5 Hertz, with some literature indicating 5-7 Hertz (Ferrante, 2011; Titze et al., 1994). Vibrato extent describes how far above and below the mean f_o that the fundamental frequency oscillates each cycle. Extent can be expressed as either full or half, the former a measure of peak to trough and the latter a measure of peak to a calculated mean f_o (see Fig. 1.3 for a simplified visualization of rate and half-extent). The unit of measurement for vibrato extent is commonly reported in cents or percentages of an octave, where there are 1200 cents per octave or 100 cents per equally-tempered semitone. To convert to cents, it is necessary to use a reference frequency for the transformation (Herbst et al., 2017). See Chapters 2 and 3 for more details and Appendix C for the exact transformation calculations and data pre-processing.

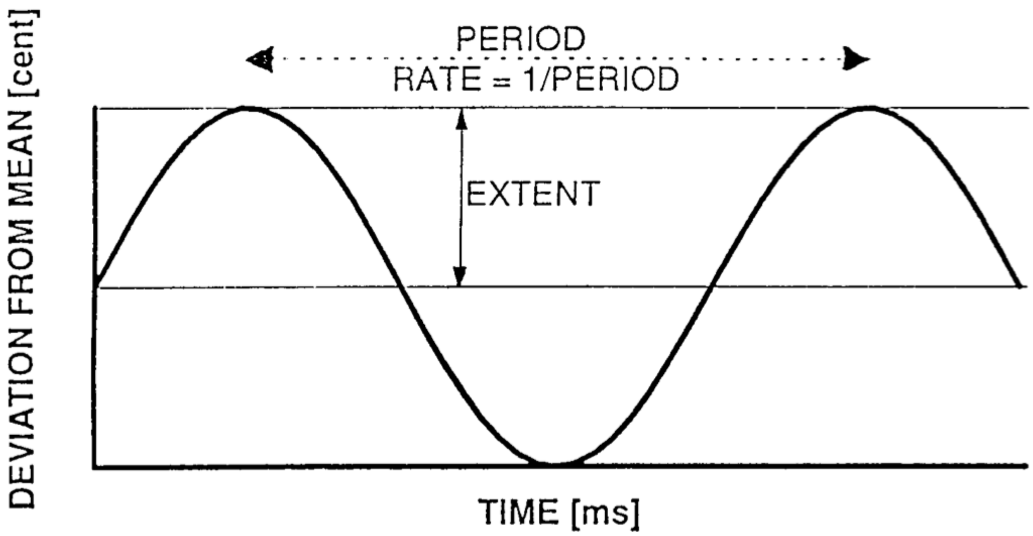


Figure 1.3: Graphic of a simple sine wave, as an ideal sinusoid, representing the fundamental frequency contour of vibrato and showing period, rate, and half-extent (Sundberg, 1994a).

Historically in singing voice science and pedagogy, the generally accepted range for full vibrato extent is approximately ±50-100 cents depending on loudness (Dejonckere et al., 1995; Titze et al., 2002; Miller, 1986; Shipp et al., 1979; Guzman et al., 2012; Hakes et al., 1987; Sundberg, 1994a). In *Singing: The Mechanism and the Technic*, Vennard (1968) stated that full extent may reach a whole tone. However, there is literature that calls into question and reframes the normative measures as a matter of taste or preference for an "acceptable vibrato rate or extent" (Sundberg, 1975). This matter of "preference" also points to the need for further study of historical audio recording technology potentially altering listener perceptions of vibrato (as extent measurements recorded with historical audio recording technology have been shown to be slightly overestimated from the original voice output signals) (Glasner and Johnson, 2020). Vibrato extent and rate are generally related; narrower extents usually correspond to faster rates and vice versa (Wooding and Nix, 2016; Sundberg, 1975).

Vibrato rate and extent have also been shown to differ as a function of different influence factors: pitch and vocal loudness (Nawka et al., 1993), emotion (Guzman et al., 2012; Sundberg, 1998; Holmes, 2013), gender (Shipp et al., 1980; Reddy and Subramanian, 2014), voice type (Müller et al., 2021), vocal training (Bottalico et al., 2017; Brown et al., 2000; Mitchell and Kenny, 2010; Mürbe et al., 2007), warm-up (Moorcroft and Kenny, 2012), vowel (Shipp et al., 1979; Nix et al., 2016), mouth opening (Patinka and Nix, 2020), and room acoustic environment (Bottalico et al., 2022). Different musical genres and singing styles are arguably impactful factors distinguishing vibrato rate and extent across various contexts, emerging as a nascent field of vibrato research in recent years (Nestorova et al., 2023).

Perceptually, vibrato's co-occurring features have been shown to drive timbre (tone quality) and pitch or intonation perception (Daffern et al., 2012; Reddy and Subramanian, 2015; Geringer et al., 2010, 2015; Almeida et al., 2021; Loni and Subbaraman, 2019; Duvvuru, 2012). Within certain extent boundaries, vibrato has been shown to play a significant role in assigning pitch to a tone. Specifically, the wider the vibrato extent is on a certain tone, the further is the deviation of the mean perceived pitch (Daffern et al., 2012; Reddy and Subramanian, 2015). In addition, vibrato rate and extent work inversely to affect perceived presence of vibrato versus non-vibrato (Järveläinen, 2002). Vibrato rate affects where the extent threshold occurs for the perception of a non-vibrato sung tone, existing on a "sliding scale" reflecting extremes (i.e., slow rates and wide extents were perceived as non-vibrato, whereas fast rates and narrow extents were perceived as vibrato) (Wooding and Nix, 2016). The vibrato extent threshold for rating a tone as non-vibrato has been demonstrated as approximately 27 cents peak-to-peak (though some listeners were found to discriminate vibrato with extent as low as 12 cents). Listeners' ability to perceive differences in vibrato characteristics depends on proportional changes between vibrato rate and extent, with these perceptual distinctions potentially varying over time (Vatti et al., 2014). At identical vibrato rates but differing vibrato extents, expert listeners demonstrated high rating error and variability (Glasner and Nix, 2023). Given this, vibrato extent seems to play a significant role in overall perception of vibrato and its temporally-variant changes. Perception of vibrato is outlined in further detail in Chapter 4.

It is important to acknowledge that the aforementioned vibrato rate and extent standards are derived from studies mostly involving both singers and instrumentalists of the predominantly represented Western classical music genre, where the overwhelmingly accepted performance aesthetic includes highly periodic, uniform, consistent, and persistent vibrato (Vennard, 1968). Though recent years have revealed more genre diversity in singing voice science studies (Nestorova, 2022), most vocal and instrumental vibrato studies still reference Western classical-based measurements for normative comparisons. Table B.1 within Appendix B exhibits this phenomenon. Sundberg (1994a) acknowledges this in his seminal text on vibrato: "as the vibrato occurring in Western operatic singing has been much more successful than other kinds of vibrato in attracting researchers' attention, I will focus on this type of vibrato mainly." This remark from 1994 remains relevant even today, 30 years later, as music researchers continue to historically analyze vibrato with tools presuming a Western classical aesthetic.

Comparing average rate and extent standards may not fully characterize non-uniform vibrato found in non-classical styles and several subgenres within the Western art music tradition. To date, there is a lack of research studies involving both singers and instrumentalists volitionally varying their natural vibrato, a stylistic device typical of many musical genres (Nestorova, 2021).

As Nix et al. (2016) reveal:

> Not many studies have been done on the topic of singers intentionally altering their typical vibrato. Also yet to be explored is the relationship between training for the Western classical tradition versus other styles and vibrato characteristics.

Furthermore, employing and reporting average vibrato rate and extent metrics is only fully representative if the vibrato is consistent, persistent, and omnipresent. In reality, vibrato sung by human subjects of any voice type may stray from this particular ideal aesthetic. When analyzing realistic, non-ideal vibrato, such as the illustrative sample shown in Fig. 1.4, regularity, variability, and time-varying contour of the vibrato also become critical analytical components. Yet, these parameters have been underrepresented in the acoustic analysis of vibrato Laukkanen et al. (1992). Sundberg (1994b) claims:

> There is nothing to suggest that regularity and waveform are of lesser perceptual relevance than rate and extent. Still, only rate and extent have been analysed extensively in the past. Therefore, awaiting that regularity and waveform will be studied in future investigations, we will henceforth focus on rate and extent.

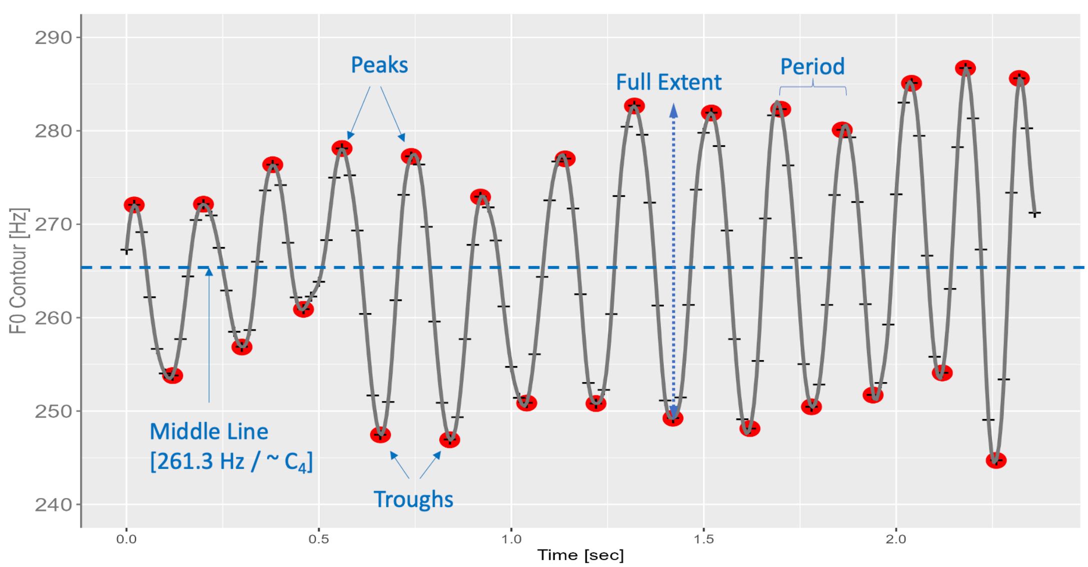


Figure 1.4: A digitally, synthetically generated vibrato sample on an isolated pitch (around middle C (C4, ≈ 261.3 Hz)), illustrating a realistically appearing vibrato typical in human vocal performance. The naturally-occurring fundamental frequency (f_o) contour demonstrates deviations from the theoretical ideal, particularly varying extent over time, with rudimentary labeled parameters. (gray dashes: values of the f_o vibrato contour; gray curve: smoothed vibrato contour approximation; red dots: peaks and troughs; broken blue line: short-term average f_o.)

1.2.1 Focus on Vibrato Extent

Early vibrato analytics suggest that vibrato rate tends to be fairly similar across singers and genres and generally consistent within a singer (Sundberg, 1975). This may be related to the theorized process of vibrato production outlined above. However, there are discordant views on the volitional regulation of vibrato with respect to changes in rate (King and Horii, 1993). Further confounding, extent has been shown to vary most across musical genre, used as a voluntary, controllable, stylistic function of vocalism in different contexts (Carter et al., 2010). The normative range of vibrato rate (5-7 Hz) is much narrower than that of extent (30-200+ cents in some reported cases), leaving a wider range for variability in vibrato extent rather than rate. For these reasons, and based on prior research, this study focuses solely on vibrato extent, the metric shown to be most voluntarily controllable and genre-representative in singing.

Discussion of Non-Vibrato

Attenuation of vibrato or non-vibrato (known colloquially as "straight tone") in singing remains fiercely debated and highly contentious among performers and teachers. A common pedagogically held belief denounces "deliberate alteration" of natural vibrato as potentially "injurious to the vocal health and natural progress of young voice students" when used over extended periods (American Academy of Teachers of Singing, 1994). Yet, the majority of genres beyond Western classical (even several of its subgenres), such as Contemporary Commercial Music (CCM) and Non-Western or World music, employ both volitional control over vibrato and non-vibrato stylistically (LoVetri and Weekly, 2003; Bartlett and Naismith, 2020). This makes vibrato versatility a lucrative and necessary skill for singers, as the music industry demands cross-training as more normative and expected. As vibrato itself is an indication of muscle balance, to physiologically achieve non-vibrato, Titze et al. (2002) believes that the dominance of one vocal fold length changing muscle, the cricothyroid (CT), over the other, the respective antagonist muscle, the thyroarytenoid (TA), damps out the extent of the frequency swings, resulting in what is perceived as a more "straight" tone. This biomechanical process is further evidence of the significance of vibrato extent. Counter to the prevalent pedagogical misconception above, Titze emphasizes that non-vibrato is not necessarily a by-product of either a healthy or unhealthy technique, explaining that "it may actually be a good exercise to learn how to disengage your vibrato ... the extent is in the singer's control" (Olson, 2008).

1.3 Dissertation Summary

1.3.1 Objectives

The objectives of this research are to:

1. Perceptually investigate variable vibrato in various voices and singing styles
2. Create a genre-inclusive system of complex, time-varying vibrato acoustic parameters beyond averages
3. Examine the practical application of variable vibrato diagnosis, analysis, and training toward technical efficiency, artistic effectiveness, and vibrato versatility in singing

An ultimate goal is to apply and implement vibrato variability analysis (including the perceptual, acoustic, and ethnographic tools developed above) into vocal pedagogy. Future steps are recommended through a final integrative review, examining potential implications for voice functionality and vibrato diagnosability, plus the integration of vibrato variability in evidence-based medical/clinical practice.

1.3.2 Research Questions and Hypotheses

The research questions underlying the objectives of this interdisciplinary work include:

1. How may complex, non-uniform vibrato shapes and temporally-variant patterns be quantitatively modeled both intra- (within) and inter- (between/across) genre?
2. Are the proposed methods of time-varying vibrato classification and empirical analysis viable and generalizable across performances in Western and Eastern musical cultures?
3. What are vocalists' perspectives of voice function, pedagogues' instructional approaches, and the health-related implications related to vibrato variability?

The hypotheses based on these questions are that (1) diverse Western and Eastern musical genres introduce more complex, time-varying vibrato patterns that require sophisticated quantitative models for accurate representation; (2) comprehensive, systematic analysis of variable vocal vibrato may improve the understanding of voice function and efficiency in singing training and performance, and (3) this may lead to diagnostic advancements in singing pedagogy for effective teaching and learning of vibrato as a versatile artistic technique.

1.3.3 Chapter Summaries

The following descriptions summarize each of the subsequent thesis chapters.

Chapter 2 highlights and investigates the state of the literature, gaps/limitations, and current analytical approaches. This narrative (traditional) review explores relevant questions that emerge from the literature as guiding principles for research on vocal vibrato variability. Several related themes and topics are examined, including:

- Quasi-Periodic, periodic, and aperiodic components connected to vibrato
- Instrumental studies on vibrato, perception of vibrato as it pertains to pitch
- Frequency tracking of modulations: algorithms and software platforms
- Theoretical and computational approaches to vibrato parameter extraction
- Recent advances in parametric modeling and analysis in vibrato research.

Chapter 3 focuses on vibrato within the Opera, Musical Theater, and Jazz genres, synthesizing prior research and theoretical developments. Summarizing the vibrato characteristics across these Western music genres and their artistic and technical implications, a subsequent methodological study is introduced that defines a theoretical framework to differentiate and classify simple vibrato archetypes, characterized by consistent modulation, from complex vibrato archetypes, encompassing more time-varying, intricate patterns. The novel parametric methods are examined in their utility for complex vibrato analysis, incorporating advanced analytical techniques tailored to temporally-variant vibrato in these musical styles.

Chapter 4 details a secondary, extended analysis exploring how variations in vibrato influence the semantic perception of timbre in singers of three genres. It aims to explore the perceptual impact and salience of vibrato variability across these conditions, utilizing a qualitative approach to understand cognitive experiences of both instrumentalist and vocalist listeners and their subjective interpretations of vibrato. The results, derived from content and thematic analysis, reveal significant interaction and correlation between vibrato and timbre perception. Vibrato is found to be perceptually salient across both trained vocalists and instrumentalists, and significantly impacts perceptions of vocal function and production, affecting the timbre semantics of the singing voice. These insights have implications for refining quantitative acoustic analyses, offering a bridge between perceptual and measurable vibrato parameters.

Chapter 5 delves into the vibrato practices within Bulgarian folk music, a tradition known for its distinctive vocal styles. A cultural and historical overview of Bulgarian folk singing is provided, as well as a scoping literature review, setting the context for this field research study.

An exploratory pilot study involving in-the-field acoustic recordings gathered on a research trip to the Bulgarian Balkan Mountains is outlined, with results demonstrating several notable highly complex characteristics and functions of vibrato in this genre. Unique vibrato patterns and their cultural significance are highlighted, contributing to a deeper understanding of vibrato's role in folk traditions and teaching strategies in diverse pedagogical traditions.

Chapter 6 centers on the quasi-philosophical question of whether "to teach or not to teach vibrato," as there exists no consensus on addressing vibrato in the voice studio and therapy clinic. To this day, there exists no codified pedagogical methodology for vocal vibrato. Direct (explicit) versus indirect (implicit) instruction is perhaps the most unresolved topic related to artistic, aesthetic targets of vibrato. As vibrato production physiology and acoustics are nebulous, it is evident that singing teachers of all styles need objective, tangible, and updated information and resources in order to appropriately and practically address vibrato with their students. This integrative literature review synthesizes relevant research on vibrato pedagogy in singing and offers targeted pedagogical tools to transform and modernize vibrato training. Applied exercises that promote vibrato versatility and efficient voice function as integral and feasible facets of singing in any style and context are recommended.

Chapter 7 explores the implications of this work, including the potential of vibrato analysis as a more genre-inclusive and comprehensive diagnostic tool in voice health and rehabilitation. Convergences and divergences of functional and neurogenic voice disorders are defined, examining vibrato variability as a distinguishing marker. Following this, preliminary observations from a currently ongoing study employing machine learning and AI models to analyze vibrato variability as a biomarker for vocal health is described. Applications of vibrato analysis in both clinical and pedagogical settings are explored, highlighting its potential as a stress indicator and its implications for ongoing research in voice therapy and pedagogy (including neurodegenerative diseases, aerodynamic analyses, and organic/phonotraumatic, neurological, psychogenic, and functional voice, swallowing, and upper airway disorders).

Chapter 8, as a final chapter, summarizes the key findings of the thesis, presenting concluding remarks and outlining future directions for this research.

Chapter 2

Literature Review on Vibrato Variability

Previous research demonstrated that current methods of measuring vibrato rate and extent are not consistent across platforms and software programs (Nestorova and Glasner, 2021). Such discrepancies in the literature may result in unreliable and non-generalizable conclusions.

There are several critical factors relating to vibrato variability that must be further explored before temporally-variant vibrato itself is investigated. This chapter provides a narrative (traditional) review of the literature on vocal vibrato, highlighting key themes and examining both contemporary and foundational studies. Topics explored include the quasi-periodic, periodic, and aperiodic components connected to vibrato, as well as instrumental studies on vibrato and its perception. The perception of vibrato in relation to pitch is discussed alongside efforts to define the threshold between unskilled or irregular pitch production and intended variable vibrato. The chapter also reviews advances in frequency tracking of modulations through various software platforms and algorithms, as well as approaches to calculating and reporting vibrato extent. Theoretical and computational methods for vibrato parameter extraction and recent advances in parametric modeling are explored, culminating in a comparative table of reported vibrato characteristics across Western and Eastern musical genres. Through this review, emerging questions and existing gaps are identified, setting the stage for the studies investigating aspects of vibrato variability discussed in subsequent chapters.

2.1 Vibrato Periodicity

It is generally accepted that perfectly periodic signals do not exist in nature. Ontologically, periodicity or a repeating, cyclical behavior is inherently tied to vibrato. Even in the definitions for vibrato at the beginning of this thesis, periodicity is qualified as "quasi-periodic" and "quasi-sinusoidal." Related terms such as fluctuation, modulation, tremor, variability, and perturbation are often compounded with periodicity and even used interchangeably in some literature. It is critical, however, to decorticate them scientifically, as vibrato itself contains some but not all of these components, both periodic and aperiodic. Thus, vibrato is said to be quasi-periodic.

Modulation quantifies the systematic change of a cyclic parameter (usually frequency or amplitude) of a signal (Boyanov and Hadjitodorov, 1997). In the context of the voice, vibrato is often conceptualized as a modulation of the fundamental frequency (f_o), though this modulation is not necessarily periodic or sinusoidal. While modulation is a useful framework for describing vibrato, it may not fully capture all aperiodicities present in the variable vibrato.

Fluctuation suggests a longer-term, periodic deviation forming a pattern (Titze, 1994b). Like perturbation (see definition below), it reflects instability in the system. However, fluctuation is usually of greater magnitude and may imply inherent lack of control, whereas perturbation includes the inference that it will eventually return to normal (Titze, 1994b). Due to its presumed periodicity, vibrato is quite often referred to as a repeating fluctuation of frequency and amplitude, but this comparison may prove fallible when naturally aperiodic elements come into play. Moreover, the term "fluctuation" lacks specificity in distinguishing between controlled, expressive modulations and irregular, unintended variations, making it less precise for describing the nuanced variability of vibrato.

Tremor is a low-frequency, periodic fluctuation in amplitude or frequency (or both) with an origin that is usually neurologic. Physiologic tremors in the body (potentially caused by heart rate) and various brain waves (Gamma, Beta, Alpha, Theta, and Delta) have fluctuation rates between 0-15 Hz (Aronson and Bless, 1992). Flutter, a type of tremor, has been connected to the adductor-abductor control system in vocal phonation. Wow is another named tremor type. These essential tremors of the nervous system have a mid-range rate (4-8 Hz) that is mechanically stabilized through a reflex-resonance loop and auditory feedforward and feedback for regulation. Due to this similarity in the systematic process, tremors have been posited to be linked to vibrato in the singing voice (Hakes et al., 1987; Lester, 2014). Trained singers are distinguished by their ability to cultivate these tremors, achieving their desired vibrato through volitional processes (Lester-Smith et al., 2021).

Variability, as a property, is the amount of variation that occurs or is observed in any given input, including, in this case, an audio signal. Although variability may be quantitatively determined through statistical calculations as variance, variability is also an important qualitative descriptor of dispersion, or how spread out a distribution of a random variable is (Miksza and Elpus, 2018). In relation to audio signals, variability often becomes synonymous with the degree of regularity of the period, or periodicity. This restrictive notion may lead to faulty implications that periodic or quasi-periodic vibrato itself does not include any variability (such as momentary aperiodicities within the overall periodic signal of vibrato).

While variability may be juxtaposed with regularity over time, regularity when referring to

Western vibrato is intended to denote the temporal similarity of frequency excursions (Nestorova, 2021). Waveform refers to the time domain and is a visual representation of a physical or electrical signal that, in the case of audio, charts the sound pressure variation of a signal over time (Howell, 2025). To study vibrato's temporal intensity changes, analysis of the waveform is suggested. The vibrato waveform has been typically less studied than the vibrato parameters of the frequency realm (Sundberg, 1994a).

Beyond these, other rarely reported acoustic features associated with vibrato include jitter and shimmer, measures of perturbations in the frequency and amplitude, respectively, and onset/offset modulation, describing features of the first and last several cycles of the vibrato.

Perturbation is a measure of short-term aperiodicity in the signal. Jitter and shimmer have been utilized in speech studies, more so than singing, as measures of the uniformity of perturbations in an ideal sinusoidal wave (Rasch, 1984). As it bears important clinical implications for disordered voices, it has been a central focus of acoustic voice analysis and from speech science made its way into singing voice science. This has led to its flawed application in vibrato analysis. Mean jitter, defined as the average absolute difference in f_o period between adjacent pitch pulses, is the most common metric used to quantify frequency perturbation (Titze, 1994b). Since there is a strong correlation between mean jitter and the average f_o period, jitter is often expressed as a percentage of the average f_o period. However, studies have cautioned that this simple method may not adequately normalize for differences in average f_o, as those seen in vibrato (Horii, 1989). The discrepancy in computational methods of jitter (and those of its companion in the amplitude realm, shimmer) may have contributed to the limited agreement across laboratories on a host of issues related to perturbation measures (Hillenbrand, 2011).

Therefore, there is a lack of consensus in the voice science community on a mathematical definition for jitter or shimmer; many differing ones are in use (Titze, 1994b). Recent studies have further highlighted the arbitrary nature and inconsistent reporting of jitter and shimmer (Brown, 2007; Rasch, 1984; Herbst et al., 2017). While jitter quantifies pitch perturbations by analyzing variations between consecutive glottal cycles, it is unreliable in highly aperiodic voices. Alternative metrics derived from jitter include the Relative Average Perturbation (RAP) which offers a smoothed assessment of pitch variation, and the Amplitude Perturbation Quotient 3 (APQ3) which evaluates amplitude variations over neighboring cycles (Baker et al., 2023). Despite these, recent studies have recommended employing jitter and shimmer solely as generic descriptors of frequency and amplitude variability, rather than attempting to connect them to vibrato or ascribe metrics to them. Instead, scientists advocate for the use of more precise metrics for vibrato and more standard terminology of signal processing and statistical analysis to quantify error measurements.

Moreover, jitter and shimmer are generally short-term, often imperceptible phenomena and therefore may not adequately characterize perceived long-term vibrato variability in singing. Ongoing, minor perturbations in f_o often occur in phonation regardless of stylistic intention or musical genre, due to normal human irregularity. Typically, perturbations are such that they do not indefinitely alter the qualitative appearance of a visual or temporal pattern; they are small irregularities that are for the most part overlooked (Titze, 1994b). Thus, jitter and shimmer typically do not impact the overall percept of pitch, since overall pitch perception is a composite phenomenon, accommodating short-term instability of the maximum and minimum frequency (Brown, 2007).

In singing voice science applied to pedagogy, there is a potentially unfortunate misunderstanding that can arise from using the term shimmer. Many singing teachers use shimmer to describe a spinning, aesthetically pleasing, flexible voice, perhaps even extending to timbral metaphors such as "bell-like vocal quality" in the context of a voice training studio (Doscher, 1994). In its scientific definition, shimmer as a random short-term amplitude perturbation is usually perceived as a crackling, buzzing, or rough sound (Titze, 1994b). Therefore, it is essential to distinguish field jargon (even in adjunct and applied realms such as voice science and vocal pedagogy) in order to communicate the context in which these terms are used.

Given the phenomenological convergences and divergences in the terms above, inconsistent term usage to describe vibrato variability could very well be psychophysically misleading. The aforementioned periodic and aperiodic components influence both the acoustic analysis and psychoacoustic perception of vibrato; it is reasonable to imagine that vibrato and pitch perception lies in the periodicity (or lack thereof) of the audio signal. Regardless of whether a sound is perceived as having vibrato or not ("non-vibrato" or "straight tone"), oscillation to a degree of 1–2 Hz is always present. There are always slight aperiodic perturbations in frequency or intensity (perhaps stemming from essential neurophysiological tremors in the muscle contractions involved in vibrato) that occur in all singing, thereby contributing to the fluctuations and modulations in vibrato production (Titze, 2014). In addition, in processing audio recording data, the sampling rate of the original signal must be sufficiently high to ensure that no aliasing occurs and f_o extraction must be performed in such as way as to avoid introducing artifacts into the vibrato analysis (see more of this in Chapter 3).

Such refined motor control of both periodic and aperiodic components of vibrato extends beyond what has been quantified in the current literature surrounding stability and consistency of voice modulations, as most often, averaged short-term perturbations of voicing (including jitter and shimmer) are transferred from speaking voice research to singing vibrato (Manfredi et al., 2015; Baker et al., 2023). Recent investigations have examined incorporation of frequency variation tracking, including relative fundamental frequency (RFF) (Stepp et al., 2011) and even vibrato in clinical and ambulatory monitoring measures of vocal hyperfunction in voice disorders such as Muscle Tension Dysphonia (Willis and Mehta, 2024).

2.2 Instrumental Studies on Vibrato

Similar to vocal vibrato, instrumental studies on vibrato are overwhelmingly steeped in the Western classical tradition. It must also be acknowledged that in the voice (and in stringed instruments), frequency vibrato dominates. In stringed instruments, rhythmic variations in intonation are synchronized with the average frequency extent, meaning that producing a vibrato with greater frequency extent deviation may require increased resistance between the string and the bow to maintain stability and control. In this regard, the stringed instrument is similar to the voice. Rather than frequency vibrato, wind and brass instruments predominantly produce amplitude vibrato by stimulating periodic variations in air pressure through thoracoabdominal and even laryngeal movements (Cossette et al., 2010), or lip pressure variations in clarinets and saxophones (where the vocal tract has a distinct influence, shown through pressure and impedance measurements (Scavone et al., 2008)), which vary the reed tip opening. Since frequency and amplitude have a direct relationship, the pitch is dependent on the intensity of the air pressure, so correspondingly slight variations in frequency arrive synchronously with variations in intensity (Gärtner, 1981).

In the voice, frequency vibrato is accompanied by synchronous variations of amplitude due to altering the relation of intensity between the source harmonics and the vocal tract resonances during the vibratory cycle (Rothenberg et al., 1988). The waveshape of the glottal airflow pulse may have a significant effect on perceived loudness. Amplitude vibrato therefore has previously been attributed to variations in subglottal pressure (Sundberg, 1999), though more recent studies on airflow vibrato have indicated a variety of biomechanical interactions as potential driving forces of amplitude vibrato (Nandamudi and Scherer, 2019).

Studies involving vibrato on string instruments have explored pitch production more deeply than studies involving the voice. Furthermore, pedagogical and empirical literature on string vibrato extensively delineate the location of the pitch center and direction of vibration. Unfortunately, no such research, pedagogical, or empirical studies involving vocal vibrato have explored either pitch center or vibration direction. Even Titze (2014) says: "studies similar to the viola vibrato study should be done for vocal vibrato."

Geringer et al. (2010) conducted a novel study design investigating pitch perception differences of vibrato and non-vibrato cello and violin tones across university music students. Both the stimulus and response were actively paired, with the goal of discerning differences in perception between string players and other instrumentalists. Both groups of music student listeners perceived the pitch of vibrato tones very near the mean frequency of the vibrato for cello and violin tones.

However, string players demonstrated significantly less deviation in tuning judgments than non-string players for both violin and cello tones. Results corroborate the longstanding pedagogical practice that string performers vibrate both above and below the intended pitch (Geringer et al., 2010).

Conversely, a study conducted by Geringer et al. (2015) investigated whether there were perceived differences between timbres in vibrato and non-vibrato performances of mistuned intervals in unaccompanied melodies performed by trumpet, violin, and voice. Participants were university music students who listened to tones with vibrato and non-vibrato conditions, either in tune, sharp 25 cents, or flat 25 cents relative to equal temperament. Across participants, all three stimuli were perceived as more out of tune when there was no vibrato compared to vibrato. In performances without vibrato, the violin was judged as more out of tune than the voice and the trumpet across all three tuning conditions. With vibrato, the violin was judged as least in tune for intervals mistuned in the flat direction, the trumpet was heard as least in tune for intervals mistuned sharp, and the voice was judged least in tune when intervals were in tune (Geringer et al., 2015). The results from this study substantiate the common claim that vibrato may mask intonation inaccuracies in instruments, but not necessarily the singing voice. A potential implication of these pitch perception findings is that when listening to the voice, there seems to be two expectations; (1) that all sounds are more in tune with vibrato, and (2) that vocal vibrato should perhaps be a bit sharp or flat relative to equal temperament.

2.3 Perception of Pitch During Vibrato

Epistemologically, vibrato and pitch are intimately linked (Lester-Smith et al., 2021). In vocal vibrato production, lower-level cortical controls facilitate reflexive resonance, while higher-level commands adjust modulation based on aesthetic intent. If these neural activations are too widely spaced or occur too slowly, the resulting pitch instability is perceived as a sagging frequency, commonly referred to in vocal pedagogy as “wobble.” Conversely, when activations are too frequent or rapid, the pitch increases excessively, leading to what is described as “flutter.” Both phenomena are governed by the auditory-motor system’s negative feedback control loop in sustaining vibrato-like oscillations (Lester-Smith et al., 2022). Especially in singers, the pitch-shift reflex is triggered with a low latency timing (approximately 100 milliseconds) and the frequency and magnitude of these reflexive responses may contribute to the rate and extent of vocal vibrato (Leydon et al., 2003).

Similarly, in the perception of pitch, the brain processes the input from the ear by computing the average of the undulating frequency. Perceived pitch corresponds closely to this average geometric rather than the arithmetic mean that determines the numeric pitch from the acoustic frequency. In most vibrato found in Western art music, the difference between these two means is very slight (Shonle and Horan, 1980; Sundberg, 1975). According to Seashore (1938/1967), the auditory processing cortices of the brain operate with interpretation when listening to singing, allowing for and filling in pitch information during creative departures from acoustic frequency. However, their sensitivity has limits when perceiving small, deliberate, and meaningful deviations in pitch during vibrato (Seashore, 1932).

It is often erroneously assumed that the perceived pitch of a vibrato tone corresponds to the upper frequency excursion of the undulation. However, it has been definitively demonstrated that the perceived pitch in the presence of vibrato corresponds to the average frequency (Sundberg, 1975; Shonle and Horan, 1980). As corroborated by the aforementioned instrumental literature, it has also been assumed that vibrato is useful in musical practice because it reduces the demands on accuracy of the fundamental frequency (Winckel, 1953). This might lead to the conclusion that the pitch (or pitch interval) of a vibrato tone is less accurately perceived than the pitch in a non-vibrato tone.

Sundberg (1994b) tested this, finding that vibrato reduces pitch-perception accuracy slightly, but only for low frequencies. Substantiated by a following study, Sundberg et al. (1996) found a wide variety in the judgments of “in tune” mean fundamental frequencies within a band of about ± 7 cents. It is important to note, however, that the difference limen (or Just Noticeable Difference, JND) for pitch varies as a function of the frequency band and spectral complexity of the stimulus, as well as the source stimulus itself (speech or music) (Moore, 1973, 1995, 2013; Gockel et al., 2004). Furthermore, several sources, including research conducted on non-Western music, have found the JND threshold for trained listeners to be between 2-6 cents, with results higher or lower depending on the experimental method (Marandola, 2004a; Demany and Semal, 2002; Zarate et al., 2012; Edmonds and Howard, 2025).

Sundberg’s various studies involving vibrato verify the common performance reality that deliberate deviations from the center fundamental frequency are employed in singing, based on context, genre, style, and intention. Surveyed singers adhered to certain principles in their deviations from the equally tempered tuning of f_o, including singing high tones sharp in particular acoustic spaces or to sharpen and flatten scale tones situated on the dominant and subdominant in Western art music (Sundberg, 1999).

As with acoustic metrics derived from vibrato, duration of the sung tone directly impacts pitch perception and vibrato. An element separating professional from nonprofessional singers was that their long tones were observed to adjust in average frequency in various ways over the course of the tone (Bjørklund, 1961). Trained Western art music singers have been found to alter their vibrato to adhere to the timing, both in meter, tempo, and rhythm of the musical line during pitch transitions and interval shifts (Myers and Michel, 1987), though this warrants future research on

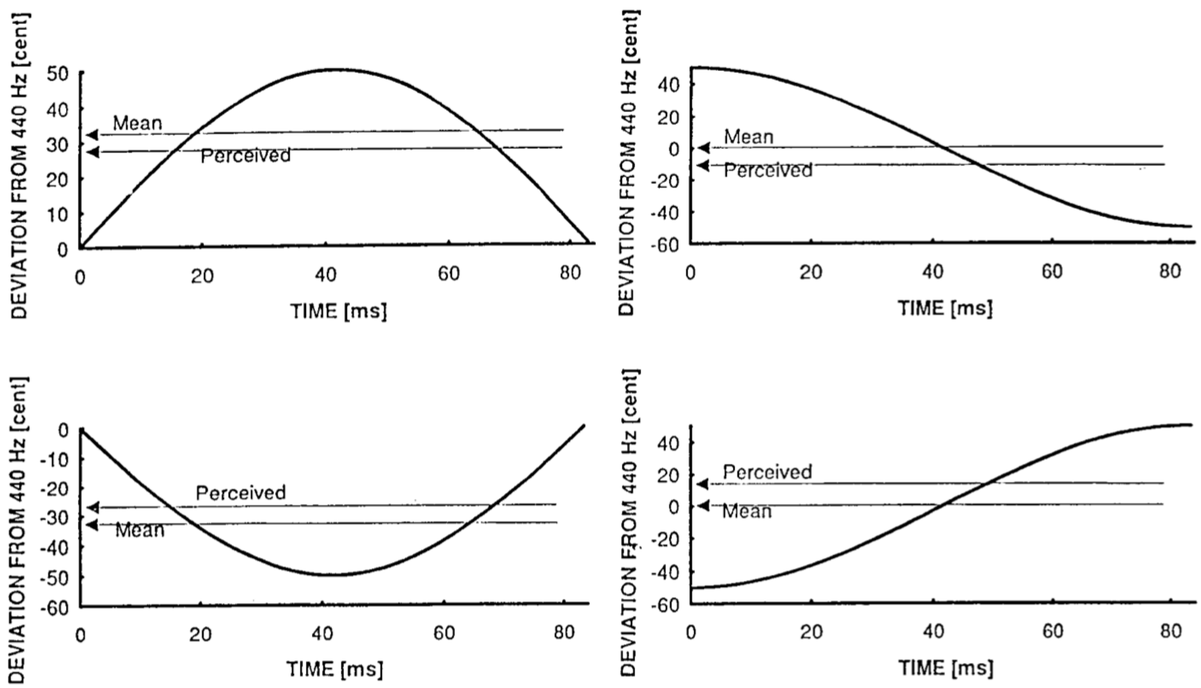


Figure 2.1: Fundamental frequencies corresponding to the pitches perceived from short tones with a fundamental frequency changing according to a half sine wave curve. The linear mean and the frequency corresponding to the perceived pitch are shown. Data from d'Alessandro and Castellengo (1991), reprinted in (Sundberg, 1994a).

a broader population (Jonck et al., 2024). One of earliest studies done on fundamental frequency correction or adjustment by Seashore in the 1930s found that long tones began slightly below the fundamental frequency (about 90 cents on the average) and were corrected by the singer during the initial 200 ms of the tone (Seashore, 1932). While these adjustments may be purely connected to frequency correction, another possible interpretation is that pitch, like vibrato, is used as a means of musical expression.

With tones shorter in duration, the relationship is slightly less linear and a bit more complicated. d'Alessandro and Castellengo (1991) found that within one vibrato cycle presented alone, the rising half was perceived as 15 cents higher than the mean while the falling half was perceived as 11 cents below the mean, yet the ending of the vibrato cycle was more significant to pitch perception than was the beginning (d'Alessandro and Castellengo, 1991). This may be attributed to proven concepts of tonal memory and recall in music perception and cognition (Krumhansl, 1991). However, caution is warranted when interpreting d'Alessandro and Castellengo (1991)'s findings, as perceiving pitch within a single vibrato cycle is challenging unless the rate is extremely slow. Moreover, even if such judgments are made at slow vibrato rates, it remains uncertain whether they generalize to typical vibrato speeds.

d'Alessandro and Castellengo (1991) also studied synthetic stimuli, in which vibrato perception may be ambiguous above a threshold of duration, called herein "threshold of fusion." This phenomenon can be defined as the fractional number of cycles where a stimulus is integrated into one tone. Above this threshold, a glissando or two distinct, consecutive tones are perceived. This threshold was estimated to appear at approximately the 2/3 mark of a vibrato cycle beginning at zero phase. One notable discovery was that participants were still able to assign a pitch to the stimuli above the threshold when the tone was long enough. See Fig. 2.1 for a depiction of this.

In addition to enhancing pitch perception, vibrato can affect listeners' ability to identify the number of voices in an audio sample. Erickson (2016) demonstrated that vibrato, combined with formant patterns, influenced untrained listeners' capacity to differentiate between two overlapping voices and detect intonation variations. These results are consistent with research indicating that smaller vibrato extents improve listeners' accuracy in perceiving pitch differences.

2.3.1 Intended vs. Irregular Pitch Production

Pitch accuracy (including regular and skilled pitch production) has been shown to increase with training (Bottalico et al., 2017). The parameters that might define a perceptual threshold between an irregular or unskilled pitch production versus the intended use of vibrato lie in whether the frequency change is periodic or aperiodic. As described above, there may be random, short-term aperiodic changes rather than longer-term changes in periodicity, implying some pattern regularity. When listening to music, information is extracted by recognizing patterns formed by sequences of sound. All of these sequences have a beginning and an end; the smallest parts of the whole units are called motifs or *Gestalts*. In musical performance practice, not disrupting the continuity within a section or motif of patterns and marking boundaries by creating a discontinuity is a learned, cultivated skill. Likewise, in sound perception, any discontinuity that disrupts the perception of a timbral or pitch *Gestalt* signals a new pattern (Sundberg, 1994b).

Disregarding the phenomenon of absolute pitch, the perception of pitch changes is fundamentally different in listening to speech versus singing. In speech, the quantity of pitch change is perceived in a continuous fashion (Burns and Ward, 1978; Balaguer-Ballester et al., 2009; Toscano et al., 2010). The amount, magnitude, and manner of pitch change is noted and interpreted in speech prosody, whereas in music, pitch changes are perceived categorically, categorized into a limited number of musical intervals. In Western art music, melodies are constructed as sequences of visits to fixed pitch stations called scale tones, and the pitch changes thereby produced are the musical intervals (such as third, fifth, octave, etc.) which constitute the categories (Krumhansl, 1991; Sundberg, 1994b; Burns, 1999; Bidelman et al., 2011). However, it is crucial to note that not all musical scales function with a set number of pitches, particularly those found in non-Western genres, World music styles, and polyphonic traditions (Marandola, 2004b). Tone length difference is another instance of categorical perception in musical listening (Clarke, 1999). Humans perceive tone lengths in terms of note values, while the actual temporal durations of the corresponding tones may vary within very wide limits, depending on the musical tempo, meter, rhythm, interpretation, etc.

Therefore, pitch discontinuity in vibrato may depend on the periodicity randomness or pattern of the fundamental frequency contour, perhaps even of the full acoustic signal's radiated harmonics. Perception of pitch discontinuity as irregular pitch production versus intended variable vibrato depends on the musical categorization of the pitch changes themselves. Irregular pitch production and perception would likely include several discontinuities of timbre (and perhaps loudness) indicating an unintended boundary, whereas pitch discontinuity in intended variable vibrato indicates a musically intended boundary. An intended variable vibrato often does not discontinue the pitch abruptly, but rather changes the phases of the periodicity of the vibrato in a musically patterned, rather than randomly erratic, manner. Pitch discontinuities may be due to a lack of control of the periodicity of the vibrato; hence, it is not considered an intentional vibrato but more a sound that contains unintentional aperiodicities. Therefore, vibrato perception categorization (irregular pitch versus intended variable production) is dependent on changes in frequency periodicity with musical *Gestalts* of discontinuity/continuity and pattern recognition.

Finally, a variational range that may affect the perceptual threshold of irregular pitch perception versus intended variable vibrato is a novel concept termed rate of frequency change (RFC) by Vatti et al. (2014). Wooding and Nix (2016) found that extent had to decrease by 11 cents per 1 Hertz increase in rate in order to still be perceived as non-vibrato. This dependency of the extent threshold on the vibrato rate is a salient factor in perceiving frequency changes and particularly their impact on pitch. The rate of frequency change (RFC) is a product of rate multiplied by extent with the unit in cents per second. As RFC varies with a change in frequency vibrato rate or extent, it can be viewed as an amalgamated perceptual threshold in assigning a single pitch to a vibrato sample, rather than multiple pitches (or irregularity in pitch control) to a vibrato sample.

RFC may explain listeners' ability to recognize or discern differences in vibrato characteristics which may be dependent on the proportional, time-varying relationship between vibrato rate and extent. As long as the RFC remains constant, there is a sliding scale range of values in vibrato rate and extent that are auditorily permissible for pitch categorization. RFC is likely the determining culprit in vibrant "singleness in pitch," (Bozeman, 2021) a concept that has existed since the dawn of vibrato studies in the 1930s, with Ramsdell (1936) conducting a study on the limits of rate and extent for perception of pitch fusion (Sundberg, 1999).

The perceptual distinction between intentional vibrato and irregular pitch production depends critically on the listener's recognition of periodic patterns within musical *gestalts*, where timbral continuity determines categorical pitch perception. RFC thus constitutes a psychoacoustic threshold differentiating unified pitch from fragmented frequency components, with significant implications for understanding the neural mechanisms of musical pitch categorization.

2.4 Frequency Tracking of Modulations: Algorithms & Software Platforms

There are a variety of frequency tracking algorithms incorporated in various voice analysis programs. YinFFT, pYIN, and YAAPT are pitch-tracking algorithms with distinct strengths, tailored for various applications. Conventional YIN deterministically estimates pitch by minimizing a difference function in the time domain to identify periodicity (de Cheveigné and Kawahara, 2002). YinFFT, an adaptation of the Conventional YIN algorithm, integrates the Fast Fourier Transform (FFT) for efficiency, making it suitable for real-time tasks like live performance analysis. It estimates fundamental frequency using autocorrelation in the spectral but not temporal domain, and thus may struggle with noisy or rapidly changing signals due to its deterministic nature. pYIN builds on Conventional YIN by incorporating a hidden Markov model (HMM) for probabilistic pitch tracking, smoothing pitch contours and robustly handling voiced/unvoiced decisions, offering greater accuracy in complex or noisy contexts. It outputs a probabilistic pitch track with confidence measures, making it ideal for offline analysis (Mauch and Dixon, 2014). YAAPT (Yet Another Algorithm for Pitch Tracking) uses a hybrid normalized cross-correlation approach, combining autocorrelation and cepstral [1] methods, along with spectral and time-domain processing, to achieve high robustness and precision, particularly in noisy environments or where voiced/unvoiced decisions are challenging (Zahorian and Hu, 2008). While YinFFT excels in speed, pYIN provides detailed probabilistic modeling, and YAAPT balances robustness and precision, the choice among these algorithms depends on the application's needs, such as real-time processing, signal complexity, or noise levels.

Praat, Tony, and Sonic Visualizer are accessible, open-source software programs that integrate a variety of pitch tracking algorithms. For successful pitch extraction, optimal parameter setting is recommended, regardless of chosen pitch tracking algorithm (Evanini and Lai).

Praat's default spectral-based pitch detection lag domain autocorrelation method (Boersma, 1993; Makov et al., 2019) extracts the fundamental frequency on the annotated audio samples. Using the raw autocorrelation method proves to be successful for isolated signals, but somewhat unsuccessful for signals with either competing background noise or accompaniment mimicking

[1]Cepstral analysis is a variation on speech spectral analysis; the nominal goal is a voiced speech signal representation, the cepstrum, in which the excitation information is separated from the dynamics of the speech system impulse response. Therefore, the speech system dynamics and the excitation signal can be studied, coded, and manipulated separately. Applications of the cepstrum have included pitch detection and "homomorphic" vocoding (Chen, 2004).

the melody of the singing voice. Even when using the manually provided annotations for minimum and maximum f_o for each note, the f_o extraction may not always prove accurate for a sung sample exhibiting low signal-to-noise ratios (SNR). It is important to note that while there is an FFT involved, the window type in duration matters, and the pass filtering needs more refinement, there is an inverse FFT to generate parts of the comparisons. While autocorrelation is the default setting, raw and filtered cross-correlation can be manually configured within Praat. Cross-correlation also has the advantage of requiring less computation (Samad et al., 2000). Previous studies have shown that the cross-correlation method yields more accurate pitch detection than the auto-correlation method when used with linear predictive coding (LPC) techniques to track the f_o contour. For that reason, when analyzing time-varying vibrato samples that are not already isolated from background noise or music, a custom approach favoring the cross-correlation method for the extraction of the f_o is recommended when using Praat (Nestorova, 2022).

Both Tony (Mauch et al., 2015) and Sonic-Visualizer (García et al., 2014) utilize the pYIN algorithm, which provides an easy way to transform the parameters and can then be used in a Python script for extraction of a larger number of audio files. CREPE (Kim et al., 2018), a deep convolutional neural network that operates directly on the time-domain waveform to perform pitch extraction, may potentially have the most user-friendly interface with equivalent results to pYIN.

2.5 Theoretical & Computational Approaches to Vibrato Parameter Extraction

For theoretical vibrato parameter extraction, onset and offset frequency modulation of a tone plays a significant role in the intersection of intonation and vibrato. Prame (1994) studied this phenomenon in depth, analyzing Schubert's "Ave Maria" from ten professional singers' CD recordings. A major finding was that each singer's vibrato rate typically increased at the end (offset) of each tone, while no typical structure could be found in the beginning (onset) of a tone; the onsets varied more in their vibrato characteristics. The rate accelerated during the last three vibrato cycles on average, regardless of whether or not the tone was followed by a pitch change. This discovery was made not only in singing but also corroborated in string performances and its origin posited as organic neuro-physiological muscular activity and behavior of musicians. This hypothesis is yet to be studied thoroughly.

Many conventional approaches to vibrato parameter extraction (including Discrete Fourier Transform (DFT)) result in single averaged values for rate and extent over long-term durations. This is the most common case when using pitch detection algorithms (YIN, pYIN, etc...) that are the basis of pitch tracking and extraction functions (auto- or cross-correlation) on commercially available platforms, such as Praat, MATLAB, Python, etc...

Furthermore, d'Alessandro and Castellengo (1991) observed that the frequency changes between Western art music scale tones tended to be timed such that the changes happened in phase with the vibrato (as a prolongation of the frequency change due to the vibrato). Both of the above findings have led to the hypotheses that vibrato rate changes in accordance with pitch changes and that speeding up vibrato rate is more physiologically conducive than slowing down. This idea has been so ubiquitous that famed vocal pedagogue Vennard (1968) stated that a "good singer sets the tempo in multiples of his vibrato; if the tone cannot be shortened/prolonged, a good singer slows or speeds his vibrato to conform to the duration of the tone. "

As a practical consequence of the above vibrato onset/offset findings, Prame (1994) recommended that the last three vibrato cycles should be excluded in the calculations of the average vibrato rate for an artist. As the vibrato rate tends to increase during the last three vibrato cycles of a tone, the average vibrato rate of a tone will depend on its duration, resulting in shorter tones showing higher average rates than longer tones. This cautionary practice has been adopted in the majority of vibrato studies following Prame (1994), such that excluding the first three vibrato cycles (to equalize the approach) has also become normative in vibrato analysis.

Consequently, Manfredi et al. (2015) excluded cycles with an average extent of less than 17 cents, due to defining boundaries based on prior studies from Ferrante (2011) and Anand et al. (2012) with normative extent values constituting vibrato per gender. It could be argued that this practice disregards significant stylistic characteristics of many performed genres with non-normative, aperiodic vibrato features, especially at the onset and offset of a tone. While average rate and extent, onset/offset modulations, and even perhaps jitter/shimmer and other non-linear metrics may be characteristics of some vibrato, most are averages and thus not particularly useful in contexts where there is naturally-occurring, time-varying, genre-dependent vibrato (Nestorova, 2021).

Manfredi et al. (2015) presented one of the first attempts to parametrically model vibrato regularity, jitter, and shimmer as an objective and precise way of assessing vibrato stability. From the acoustic waveform, vibrato jitter evaluates variations in the timing of vibrato cycles, while vibrato shimmer measures fluctuations in amplitude. Vibrato jitter may capture the short-term frequency changes and vibrato shimmer may detect short-term amplitude changes within vibrato cycles, especially in the case use of comparing Western Opera and Jazz. However, vibrato jitter and vibrato shimmer still do not offer deep insights into the long-term, across vibrato cycles source of variability and sometimes emphasize variations that are not perceptually significant leading to overinflated and misrepresented regularity measures. Moreover, these perturbation measures do not allow for comparison across samples (e.g., between singing styles that contain even greater stylistic dissimilarity such as Western and Eastern musical traditions).

Prame (1994, 1997)'s method performs a short-term vibrato analysis, providing the results for the vibrato parameters as a function of time. Pang et al. (2020) proposed a method based on sinusoidal modeling of the f_o trajectory, simultaneously estimating the intonation, vibrato rate, and vibrato extent. According to experimental results, the proposed method is robust to the irregularities in the f_o trajectory. In addition, the proposed method was compared to those of Prame's and produced different results for intonation and vibrato extent when the f_o trajectory was not perfectly sinusoidal (Pang et al., 2020). This is particularly promising for analyzing vibrato intonation as well as rate and extent, while also considering the variability (short- and long-term) of the vibrato signal.

Vibrato can also be extracted and parametrized in less traditional methods (Driedger et al., 2016; Lee et al., 2011; Lee and Dong, 2011). Both periodic and quasi-periodic low-frequency components are present in steady-state vibrato portions of sustained musical sounds. In order to build a hierarchical multi-level representation of sound for transformation with musical purposes, these components should be isolated and separated from the amplitude and frequency envelopes. Using Spectral Modeling Synthesis (SMS) and other adjunct tools applied to the frequency envelopes in each frame allows for re-construction of an envelope that has been cleaned of those quasi-periodic low-frequency components (Herrera and Bonada, 1999).

Similarly, various theories assume implicitly or explicitly that each spectral component of a sound signal introduces a series of subharmonics (harmonics below the fundamental frequency). The spectral-compression method for pitch determination can be viewed as a direct implementation of this principle, using an algorithm described in terms of this non-linear subharmonic summation system theory. Hermes (1988) argues that favorable performance of the subharmonic-summation algorithm may stem from it corresponding more closely with current pitch-perception theories than does the harmonic sieve.

Yang et al. (2016) provided computational and scalable solutions for the automatic extraction and analysis of performed vibratos and portamenti (gliding motion connection from one musical tone to the next). By automatically detecting vibratos and estimating their parameters, a novel method based on the Filter Diagonalisation Method (FDM) was proposed, remaining robust over short time frames and allowing frame sizes to be set at values small enough to accurately identify local vibrato characteristics and pinpoint vibrato boundaries. The vibrato and portamento detection and analysis methods were implemented into AVA, an interactive graphical user interface (GUI) tool implemented in MATLAB for automated detection, analysis, and visualization of vibrato and portamento. A cross cultural analysis of vibrato and portamento differences between erhu and violin performance styles, and between typical male or female roles in Beijing opera singing was performed using this program, with meaningful results (Yang et al., 2016). Interestingly, Yang et al. (2016) used a Logistic Model to describe portamento parameters after extraction. As will be exemplified in Chapter 3 of this thesis, a related approach to analyze variable vocal vibrato samples will be applied.

For analysis over the course of an entire song, Yang et al. (2016)'s AVA interactive graphical user interface would be most beneficial, as it may perform vibrato and portamento analyses simultaneously to estimate and parametrize the intonation changes both within a single note and across the entire song. Eventually, applying both vibrato and portamento models to vibrato intonation, rate, and extent over the course of an entire song might be visualized concatenated together with the models created in this thesis research, illustrated in Chapter 3. To evaluate vibrato intonation, rate, and extent over the course of a single note, the use of Pang et al. (2020)'s method based on sinusoidal modeling of the f_o trajectory would be most appropriate. However, if assessing variability and aperiodic components of regularity is also of interest parametrically, then the creation and application of the parametric model proposed in Chapter 3 of this thesis research would be beneficial.

It must be noted that regularity and complexity, which are subcomponents of variability, are not sufficiently addressed or captured by any of the above in the time domain-dependent realm. Additionally, the ease of use of the aforementioned vibrato parameter extraction approaches may be restricted and challenging outside of specialized research settings due to the lack of accessibility in practical application settings. Therefore, while these tools are valuable to adapt to specific voice analysis purposes, they should be applied thoughtfully and selectively according to proper contexts. Working toward a ubiquitous parameter extraction tool for variability that also includes a user-friendly, real-time visual biofeedback aspect is thus needed and the subject of this thesis research.

2.5.1 Recent Advances in Parametric Modeling & Analysis in Vibrato Research

In recent years, a number of new models and parameters connected to acoustic evaluation of vibrato have arisen in the signal processing and music audio computing fields. Many of these involve some measure of variability, regularity, and complexity, but are still somewhat separated from the perceptual aspects of vibrato.

More recent studies have posited the use of non-linear metrics, such as entropy and recurrence analysis, as single quantifiers of vibrato regularity (Acosta Martínez and Daffern, 2023b). A salient finding from this study was that sample entropy (SampleEn)'s implementation (though parameter dependent) outperformed L (line length or mean prediction time), which was slightly more difficult to obtain, more sensitive to embedding parameters, and provided normality and additional recurrence plot information (Acosta Martínez and Daffern, 2023a). Yet, the authors themselves state that "although the individual values from each particular metric or singer are compelling to observe, the relationships or trends across the data were not easy to infer" (Acosta Martínez and Daffern, 2023a). Therefore, a parametric model that enables both within- and across-sample analysis and comparison of vibrato regularity is needed. Furthermore, a deeper understanding of how the vibrato regularity changes across commonly found time-varying samples of vocal vibrato offer the aforementioned inferences that are not possible with solely individual values. Finally, these studies have only sourced from professional singers' commercial recordings, both variables that may skew truly representative results of regularity. Time-varying, unstable vibrato is typical and indeed expected in samples that are drawn *in vivo* from a broader demographic of human subjects. Variability is commonly encountered in vibrato, existing in singers in training and therapy, in whom the many stakeholders have an interest.

Two imperative points for further investigation are raised by the authors of the aforementioned study, stating first: "analysis in the present study is based on purely acoustic and objective measures, and the perceptual significance of the results is still unknown" and second: "Given the renowned status of singers such as Callas and Sutherland, it could be speculated that high complexity can be psychoacoustically pleasant and characteristic of particular singers" (Acosta Martínez and Daffern, 2023a).

To the first point, vocal vibrato has been historically studied either solely in the acoustic or perceptual realm without examining the psychoacoustic connection that is most relevant for voice practitioners in education and habilitation. Given that vibrato is a perceptual phenomenon, it is critical that vibrato parameters analyzed acoustically are validated perceptually. Prior vibrato studies examining perceptual characteristics open pertinent questions on regularity emerging from vocal pedagogues and practitioners (Wooding and Nix, 2016; Glasner and Nix, 2023). It is clear that parametric models that encapsulate both the acoustic and perceptual realm are necessary, especially since data gathered from vibrato tones are often non-stationary and short, as outlined in the aforementioned study.

To the second point, complexity in the aforementioned study does not appear to be conceptually defined in full. This may be confounding, since complexity in non-linear systems may be conceptually misaligned with that of perceptual and acoustic complexity. When examining a broader data set of vibrato samples from live human subject studies that range from unprofessional to professional singers, vibrato changes become even more dependent on the time domain, and complexity takes on another meaning entirely (Nestorova, 2022). Excising or divesting vibrato from the domain of time as the most crucial variable may cause a misrepresentation of complexity if solely comparing complexity within and across two genres of homogeneously grouped professional singers. Changes in voicing over time even within individual utterances are a natural phenomenon often observed in samples derived from humans in real-world contexts, such as those in vocal training or therapy. Thus, it is critical that researchers investigate vibrato in the population most often encountered: singers.

Moreover, while these recently emerged metrics in signal processing and music audio computing are certainly advanced and interesting, they may have several limitations, including more complex computational and analytical demands that are not easily accessible or intuitive to stakeholders in vibrato analysis. However, it is crucial that these tools are practical for a wide range of stakeholders. Furthermore, these studies consistently highlight several limitations, such as reliance on computationally advanced software and the need for specialized technical expertise, making them difficult to implement and interpret in reality.

2.6 Calculating, Reporting, & Analyzing Vibrato Half-Extent

Acoustic analysis of frequency vibrato using conventional voice analysis software often produces differing results for identical audio samples, observed especially in vibrato extent, or frequency excursion. Sundberg (1994a), in reference to full extent, writes that "often the extent is preceded by plus and minus signs so as to prevent confusion with an amplitude measure," though vibrato studies both pre- and post-dating this statement do not always adhere to this practice (Nestorova and Glasner, 2021).

As a parameter of vibrato, extent is difficult to compare across studies, due to its temporally-variant nature outlined above. First, one must understand how extent is being calculated, as there are different approaches used by various researchers. Second, it is important to accommodate the minor, temporary, naturally-occurring perturbations in the f_o oscillation that occur in phonation regardless of genre, while remaining aware of their potential to confound the extent calculation.

There has been a lack of consensus in using and interpreting half- versus full extent in prior voice science literature. Historically, researchers have analyzed vibrato using full extents. A recent trend is the analysis of vibrato using half-extent (HE), however, at least two variations of the HE formulas are used (Herbst et al., 2017; Glasner and Johnson, 2020). These two variations have not yet been comprehensively compared and their utility has not been determined. It can be argued that HE is easier to interpret in the context of vibrato analysis, as it represents the magnitudes of the f_o excursion above and below an "intended pitch," which can be constant or time-varying, depending on the intention of the singer. In contrast, the measurement of the full extent disregards the notion of "intended pitch." Vibrato HE proves to be an adaptable and comprehensive, yet still practical, accessible, and understandable measure, able to characterize observed variation in the f_o contour.

Recent literature, including Herbst et al. (2017) and Glasner and Johnson (2020), has utilized and applied two slightly divergent formulae to derive average vibrato extent from the original frequency vibrato contour. However, neither paper offers an elaboration on their differing approaches. In addition, neither author addresses the specific properties, interpretability, and applicability of each measure. This is a focus area of the methodology outlined in Chapter 3 of this thesis.

The process for average vibrato HE calculation is outlined in the existing literature as follows. First, the f_o contour is extracted from the originally recorded signal, then the data is converted into cents via the following equation reported in Herbst et al. (2017) and also applied in Glasner and Johnson (2020), both of which used a reference point relative to the musical pitch, middle C (C4):

$$c[i] = 1200 \frac{\log\left(\frac{f_o[i]}{f_{o\_C4}}\right)}{\log(2)}, \tag{2.1}$$

where $f_o[i]$ is the fundamental frequency estimate at time index i, and f_{o_C4} is the fundamental frequency of the musical pitch C4, calculated as $440 \times 2^{(-9/12)} \approx 261.63\,\text{Hz}$.

Then, the average of all vibrato extents is calculated in two different ways resulting in two different metrics of HE. The approach of Herbst et al. (2017) is referred to as Relative Half-Extent (rHe) in this thesis, while that of Glasner and Johnson (2020) is referred to as Half-of-Extent, as described below. For more detailed descriptions, applications/case uses, and interpretations of both rHE and HoE, see Chapter 3.

1. **Average Relative Half-Extent (rHE)—:** From n samples in time of the f_o contour, given by $c[i]$ and expressed in cents from middle C, the average musical pitch ($\bar{c}$) is computed as:

$$\bar{c} = \frac{1}{n}\sum_{i=0}^{n-1} c[i]. \tag{2.2}$$

   The average absolute deviation from the mean musical pitch $\overline{\Delta c}$, representing vibrato half-extent, is then computed as:

$$\overline{\Delta c} = \frac{1}{n}\sum_{i=0}^{n-1} \left|\bar{c} - c[i]\right|. \tag{2.3}$$

2. **Average Half-of-Extent (HoE):** From the f_o contour given in cents, an average musical pitch $\bar{c}$ is computed. Glasner and Johnson (2020) use Praat's "Get pitch" command to do this, which is likely equivalent to Eq. 2.2. Also from the f_o contour, they identify local maxima (peaks) and minima (troughs). Then, the average half extent is calculated as:

$$\overline{\Delta c} = \frac{1}{n}\sum_{i=0}^{n-1} \frac{|(\bar{c} - c_i[f_{o\_min})] - (\bar{c} - c_i[f_{o\_max}])|}{2} \tag{2.4}$$

where $c_i[f_{o\_min}]$ is the indexed minimum f_o, or trough, of the deviation of a vibrato cycle expressed in cents, and $c_i[f_{o\_max}]$ is the opposite peak of the same vibrato sample. The vibrato extent for each participant is determined by averaging the calculations from individual samples within each recording condition.

Equation 2.4 can be rewritten as follows:

$$\overline{\Delta c} = \frac{1}{n}\sum_{i=0}^{n-1}\frac{|(\bar{c}-\text{trough}_i)-(\bar{c}-\text{peak}_i)|}{2} = \frac{1}{n}\sum_{i=0}^{n-1}\frac{|\bar{c}-\bar{c}+\text{peak}_i-\text{trough}_i|}{2} = \frac{1}{n}\sum_{i=0}^{n-1}\frac{|\text{peak}_i-\text{trough}_i|}{2}, \tag{2.5}$$

from which is it seen that HoE is equivalent to the use of half of the full extent.

2.7 Previous Relevant Work and Dataset

In prior research performed for a Master's thesis project, Nestorova et al. (2024b) conducted a pilot study on the state, gaps, and methods of vocal vibrato analysis. A two-stage quantitative acoustic and perceptual study explored temporal vibrato variability and its connection to genre and perception, with the objectives of (1) examining vibrato half-extent variability as an acoustic genre indicator, (2) investigating the perceptual salience of vibrato variability for genre classification, and (3) identifying new parameters and quantitative models for vibrato half-extent time profiles.

First, a voice study protocol with 15 professional Opera, Musical Theater, and Jazz mezzo-sopranos and sopranos was conducted, using the Coefficient of Variation (CV)[2], linear mixed effects analysis, and logistic regression to analyze vibrato half-extent variability over time. The results showed genre-specific variability. Second, a proof-of-concept perceptual survey with seven vocal pedagogue judges confirmed that the vibrato variability that was most representative per genre was a significant predictor of accurate classification. Experts found Musical Theater the most difficult to classify and Opera the easiest, with Jazz in between.

Results demonstrated that CV varied greatly across genres, with Opera singers showing the lowest CV and a regular, uniform vibrato pattern, while Musical Theater and Jazz singers exhibited more complex patterns with higher CV. Acoustic analysis showed that vibrato in song repertoire tasks better differentiated genres, while perceptual results suggested the opposite. A

[2]The CV is a dimensionless statistical measure of variability, calculated as the ratio of the standard deviation to the mean, expressed as a percentage, which allows for comparison of dispersion across datasets with different units or magnitudes. Coefficient of Variation is sometimes used interchangeably with Coefficient of Variance.

linear explanatory model of the perceptual study confirmed that vibrato from Opera and Jazz singers was more accurately recognized than from Musical Theater singers.

A key strength of this study was the validation of the acoustic and perceptual stages, addressing the psychoacoustic connection that is particularly relevant for voice practitioners. Future perceptual surveys should include randomly selected vibrato samples from the entire dataset to enhance genre classification validity. The study concluded that vibrato half-extent variability is an important factor in genre classification, highlighting the acoustic and perceptual salience of vibrato in Opera, Musical Theater, and Jazz, with implications for non-Western singing styles. Future study iterations are needed to expand the dataset and include a more diverse demographic for broader applicability.

At the end of this prior research, several complex patterns in the vibrato f_o contour were noted, consisting of two or more vibrato phases with differing extents. It was established that a four-parameter logistic regression model may more successfully fit the half-extent time profiles of such complex patterns (most often observed in Musical Theater and Jazz samples). However, conclusions indicated that further systematic research is warranted, since in-depth understanding, interpretation, and implementation of this approach was outside of the scope of the prior research effort. A new methodology study and comprehensive analysis of this proposed approach using both HE measures is offered in Chapter 3 of this thesis.

Chapter 3

Acoustic Modeling & Analysis of Time-Varying Vibrato in Western Musical Genres

Author Contributions

This chapter includes restructured and reintegrated content from the following research article: Nestorova, T., Nestorov, I., and Scavone, G. (2025). Analysis and modeling of time-varying vibrato. *The Journal of the Acoustical Society of America*. [Manuscript under review].

Author-specific contributions: *Theodora Nestorova:* conceptualization of the study, data acquisition and pre-processing, post-processed acoustic analysis and interpretation of results, development of interpretation and conclusions, writing, review, and editing of the manuscript. *Ivan Aleksandrov:* coding of preliminary script, development and programming of original numerical and computational modeling for data and results processing, review and editing of the manuscript draft. *Gary Scavone:* supervision of the overall study project, results processing, and review and editing of the manuscript draft.

Chapter-specific contributions and differentiation from original journal publication: In this chapter, the article manuscript above is contained solely within Sec. 3.3 and its subsections. The other sections included in this thesis have neither been submitted nor published in any article.

3.1 Introduction

In this chapter, the results of two related investigations are presented: (i) a methodology study defining various types of vibrato contours and the results of the classification of vibrato samples according to quantitative and qualitative criteria; and (ii) the results of the variability analysis of the vibrato samples with two different vibrato metrics. The overall discussion and conclusions, as well as ideas about future developments and implementations of the proposed methodology are presented in the final section.

As noted in the pertinent sections below, the detailed data analysis procedures as well as the annotated scripts of the original source code material used to develop these models and parameters are made available by the author of this thesis at https://github.com/theodora-nestorova/vocalvibratovariability

3.2 Sources of Variation in Vibrato Contour

The major sources of vibrato contour variability in the frequency domain are: (1) the variability of the vibrato extent; (2) the variability of the vibrato rate; and (3) the temporal change of the intended fundamental frequency (or heard pitch), henceforth referred to as "pitch drift" over the duration of the vibrato sample. A fourth potential factor contributing to vibrato contour data variability is the presence of spurious or erroneous estimated values of f_o caused by background noise during an audio recording or inaccuracies in the f_o estimation algorithm. When conducting vibrato variability analysis, it would be ideal to define and derive more specific and comprehensive variance measures that disentangle the contributions of all sources of vibrato variability.

There are quite sophisticated ways of describing f_o variability, based on complex mathematics and stochastic processes, including non-linear models outlined in Sec. 2.5.1. However, such methodologies may not be easily interpretable and practical to all stakeholders for activities such as characterization, conceptual understanding, diagnostics, and correction. In multifactorial cases such as vocal vibrato, variance metrics that first characterize or isolate one or a few of the variability factors are preferable. For this and aforementioned reasons, including the vocalist's ability to volitionally control vibrato extent and vibrato extent's role as a key stylistic distinguishing factor in genre-specific vocal performance, this thesis research focuses on extent, the first source of vibrato variability outlined above.

Before calculating vibrato half-extent (HE) in either of its forms (Half-of-Extent (HoE) or Relative Half-Extent (rHE)), the f_o of the vibrato needs to be converted to cents and maximum and minimum values (peaks and troughs of the vibrato) need to be determined. To convert the f_o values from Hertz to cents, Eq. C.1 from Herbst et al. (2017) is used.

3.2.1 Vibrato Half-Extent Time Series Calculations

In order to allow the evaluation of time-varying properties of vibrato, methods for computing vibrato half-extent time series are described. These approaches are based on the average HoE and rHE calculations previously defined in Sec. 2.6, though revised to produce results that vary

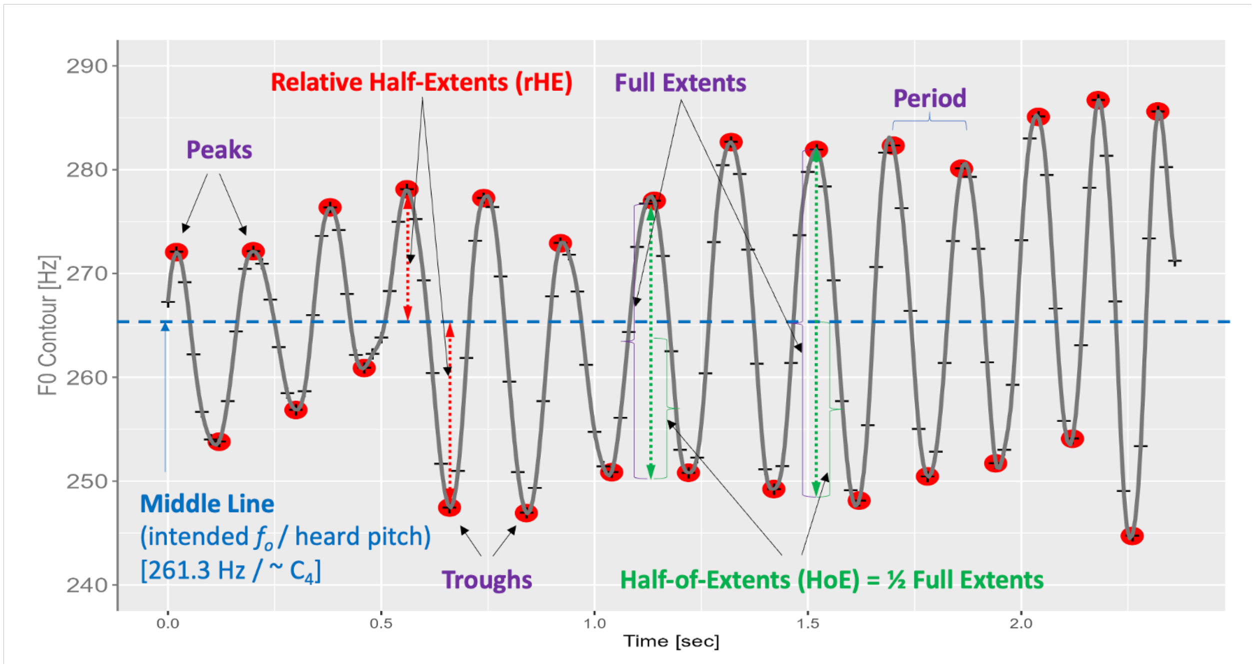


Figure 3.1: A digitally, synthetically generated sample of a singer's naturally-occurring fundamental frequency (f_o) contour, demonstrating a realistically appearing vibrato of varying extent over time, with labeled half-extent parameters (HoE & rHE) (gray dashes: values of the vibrato contour f_o; gray curve: smooth vibrato contour approximation; red dots: peaks and troughs; broken blue line: average f_o over the sample.

over time. An example time-varying f_o contour is illustrated in Fig. 3.1. The two ways to calculate the vibrato HE shown in this example are detailed below:

1. **Half-of-Extent (HoE):** The Half-of-Extent (HoE) metric is representative of the overall variability of full vibrato extents. As the HoE name belies, it corresponds directly to $1/2$ of the full extent (the difference between successive peaks and troughs of f_o). HoE does not rely on an estimated intended pitch middle line. Instead, halving the full extent effectively introduces an actual pitch middle line, which shifts dynamically at each vibrato peak and trough, passing through the midpoint of each extent. As a result, the HoE values correspond to the distances between each peak-trough pair and this shifting actual pitch middle line. Because the HoE is based on these dynamically shifting midpoints which may be asymmetrical rather than a fixed reference, it largely, though not entirely, minimizes the influence of naturally occurring pitch drift in vibrato production.

   To calculate the Time Series Vibrato HoE: This transformation procedure, adapted from Glasner and Johnson (2020) and modified to account for temporal variance, provides a more dynamic representation of vibrato variability. A simplification of the process to calculate HoE is described as follows and displayed in Fig. 3.2.

- Convert the quantified f_o contour into cents (via formula in Appendix C, adapted from Herbst et al. (2017))
- Identify peaks (p_1, p_3, p_5, and p_7) and troughs (p_2, p_4, p_6, p_8).
- For each successive peak-trough pair, calculate full extent: $e_i = |p_{i+1} - p_i|$
- Calculate half-of-extent: $h_i = 0.5 * e_i$

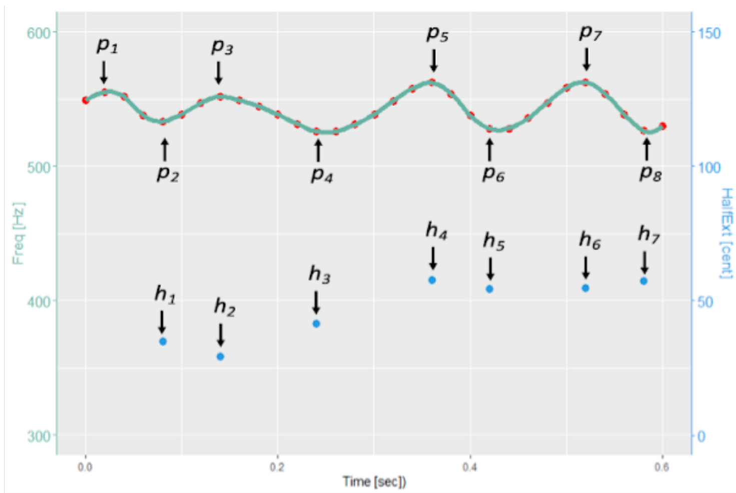


Figure 3.2: Illustration of the half-of-extent calculation from the original f_o contour. (Green curve: f_o contour; Red dots: measured values of f_v contour; p_i points: peaks (p_1, p_3, p_5, and p_7) and troughs (p_2, p_4, p_6, p_8); Blue h_i points: calculated half-extents in cents.)

2. **Relative Half-Extent (rHE):** The Relative Half-Extent (rHE) metric provides a broader representation of vibrato variability by incorporating both the extents and fluctuations of the intended pitch middle line. Unlike the Half-of-Extent (HoE) metric, which relies on dynamically shifting midpoints, rHE is derived with respect to a middle pitch line— an estimated intended pitch over a window of samples (see broken blue line labeled in Fig. 3.1)— calculated for a range of vibrato samples with individual peak and trough of the f_o vibrato contour. This metric represents the absolute deviation from the mean between each peak or trough and the intended pitch middle line.

   Because rHE reflects both the variability of the vibrato extent and fluctuations in the actual middle line, which may drift in pitch over time, it captures a more comprehensive picture of vibrato behavior. This calculation assumes a relatively stable pitch middle line, typically aligning with the intended pitch, though this assumption introduces additional variability factors. As a result, rHE values are expected to exhibit greater variation than HoE, as they account not only for vibrato extent but also for the pitch drift of the middle line.

To calculate the Time Series Vibrato rHE: This calculation requires individual peaks and individual troughs and an intended pitch middle line. This transformation procedure, adapted and modified from Herbst et al. (2017), calculates the intended pitch middle line using only the identified peaks and troughs within the vibrato signal, rather than all points in the f_o contour.

The half-extent values (calculated either as HoE or rHE) derived from the f_o vibrato then form the dataset from which CV is calculated and interpreted. It should be noted that, for an "ideal" vibrato, defined as having a perfectly sinusoidal f_o, the two HoE and rHE measures should converge. To visualize the above differences in rHe as compared to HoE, see Fig. 3.3.

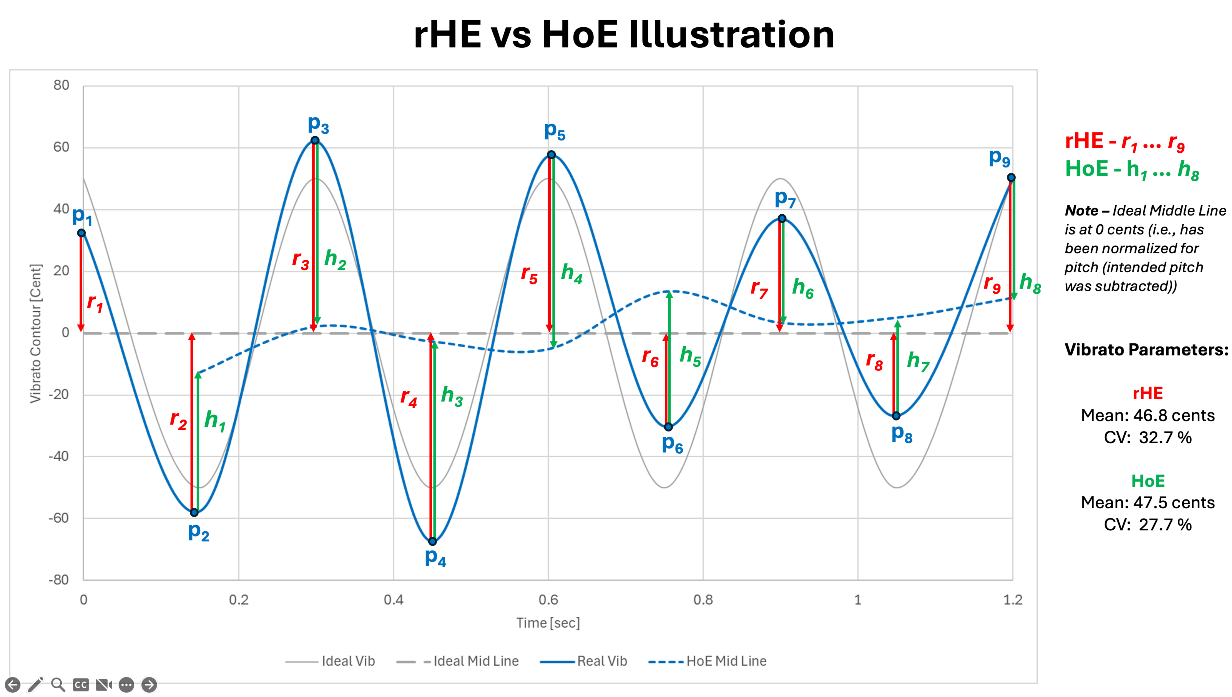


Figure 3.3: A labeled and defined illustration of rHe as compared to HoE over the course of a time-varying vibrato sample.

3.3 Methodology Study: Modeling & Defining Time-Varying Vibrato Contours

Historically employed average vibrato extent parameters and more recently tested variability metrics may not fully characterize naturally-occurring, time-varying vibrato if the shape of the temporal contour is ignored. It becomes clear that a comprehensive analytical system of vibrato metrics considering the regularity, variability, and shape of entire vibrato sample patterns in more genres over time is essential.

In the prior acoustic and perceptual research studies described in Sec. 2.7, the Coefficient of Variation (CV) of vibrato half-extent was proposed as a measure of vibrato variability independent of pitch (Nestorova et al., 2024b). Preliminary comparisons of vibrato half-extents across genres, without consideration of the vibrato patterns, indicated that average vibrato of Opera singers is less variable (having lower CVs) compared to Musical Theater or Jazz singers. Upon more in-depth acoustic signal evaluations, it was observed that some singers exhibit more complex vibrato patterns than the commonly found regular, uniform vibrato samples. Apparent in the fundamental frequency contour, these complex vibrato shapes consisted of two or more vibrato phases (vibrato with markedly different half-extent), with transitions of various duration and steepness between them. One of the most common vibrato patterns observed was a transition from no vibrato (henceforth referred to as "non-vibrato" and colloquially as "straight tone") to a uniformly regular vibrato.

Analyzing the variability of such complex vibrato contours while ignoring the complexity of the shape leads to an overestimation of the CVs. Several vibrato samples exhibited this in Nestorova et al. (2024b), as the CV calculations of variability across all half-extents without distinguishing vibrato phases or transitions led to inflation of the variability estimates. This was shown to be problematic, leading to potential misinterpretation, particularly in genres where complex vibrato patterns are inherent stylistic features. A main conclusion of this prior research was the need for an in-depth, systematic investigation to establish a definition and classification framework of time-varying vibrato contours. Therefore, this thesis chapter includes a new methodology study that details, interprets, and applies this approach.

The objectives and aims of this methodology study are as follows:

1. To define differing shapes and patterns based on time-varying vibrato contours.
2. To propose a method of identification and modeling for quantitative analysis of complex vibrato half-extent time profiles.
3. To apply this approach to selected ecologically valid vocal samples and interpret the results.

3.3.1 Experimental Data Sourcing

The proposed methods were developed and illustrated on live singer vibrato samples obtained from a prior data set that was validated both acoustically and perceptually (Nestorova et al., 2024b). Acoustic recordings were collected from 15 professional mezzo-soprano and soprano singers (five each from the Opera, Musical Theater, and Jazz genres), aged 20–30, with at least ten years of training and professional experience. Each singer was recruited by professional referral and self-reported to fulfill the inclusion criteria. The singer participants performed two genre-neutral songs, “Summertime” and “Amazing Grace,” and *messa di voce* (crescendo-decrescendo) exercises on the vowel [*a*] across several notes and were instructed to sing the tasks, maintaining the notated pitch, rhythm, and tempo, while in the authentic style for each genre. The singers were permitted to re-record the tasks as many times as they wished, as long as each recording was done in a single take, without interruptions. Furthermore, the singers were instructed to sing in their ”best, most natural Musical Theater / Operatic / Jazz voice” in order to capture the most authentic, raw, and representative performances. Despite COVID-19 pandemic-related limitations that prevented fully supervised or uniform recording conditions, each singer used a high-fidelity external microphone—– either an Earthworks M23 omnidirectional microphone or an Audix TM1 condenser microphone—– producing high-quality audio signals suitable for frequency-based analysis of vibrato.

See Appendix A for the detailed study protocol, or refer to Nestorova et al. (2024b) for the detailed information on the source of vibrato sample recordings for the current methodology study included in this thesis.

Sixteen pitch samples from each singer’s recordings were analyzed spectrographically using the software Praat (Boersma and Weenink, 2025). Three representative individuals, one per genre, were selected with the common criteria of having the widest variety of complex profiles in each group, to evaluate and illustrate the concepts introduced in this paper: S1 (Musical Theater), S2 (Jazz), and S3 (Opera).

The fundamental frequency (f_o) for each sample was isolated and extracted using bandpass filtering and pitch detection algorithms within Praat, ensuring full onset and offset retention to preserve natural vibrato variability. The f_o contour data was subsequently converted into cents, and using a custom code written in R (R Core Team, 2021), the peaks and troughs of the sinusoid and their half-extents (as HoE for this section) were calculated over the duration of the samples. This code was then used to fit the half-extent time profile of the f_o vibrato contour to produce the tabular and graphical output detailed below.

The detailed algorithm for the fundamental frequency (f_o) contour preparation/pre-processing is described in Appendix C.

3.3.2 Statistical Analysis of the Half-Extent Time Profiles

To determine the complexity of the vibrato contour, the time profiles of the half-extent (calculated as HoE) are fit with a linear regression (function *lm()*) or a 4-parameter logistic regression (functions *L.4* or *LL.4* from R package *drc*) in R (R Core Team, 2021) and RStudio (RStudio Team, 2023). The model discrimination is determined by the value of the Akaike Information Criterion

(AIC) and the Bayes Information Criterion (BIC) (Mohammed et al., 2015). The model selection favors the approximation with the lower AIC and/or BIC. If the linear regression provides a better fit, the vibrato is classified as “simple” or “monophasic”. If the logistic regression provides a better fit, the vibrato is classified as “complex” or “biphasic”.

There are a number of mathematical functions that describe S-shaped curves and the detection of their transition phases (Sebaugh and McCray, 2003); one of the most frequently used being the logistic regression curve, widely applied to analyze bioassay data or to model dose-response relationships in pharmacology (frequently referred to as “The Four-Parameter Curve”) (Mager et al., 2003). For the purposes of this investigation, we used the 4-parameter form of the logistic function (defined in the R package *drc* as *LL.4*) (Knezevic et al., 2007; Ritz et al., 2015) as follows:

$$f(x) = c + \frac{d - c}{1 + \exp(b(\log(x) - \log(e)))} \tag{3.1}$$

where y is the dependent variable (half-extent), x is the independent variable (time), and b, c, d, e are the parameters estimated by the regression.

For the definitions and classifications of vibrato complexity archetypes, refer to the subsections under Section 3.3.3 below.

To quantitatively characterize complex vibrato samples using these mathematical and statistical functions, the transition phase of the curve can be defined as the profile between approximately 20% and 80% of the transition between the lower and upper asymptote, where the logistic curve can be assumed as approximately linear. These thresholds are somewhat arbitrary, but coincide approximately with the linear ascent of the logistic curve. The statistical summaries as well as the start and end time of the transition phase are calculated for each vibrato contour that has been identified as “complex.” The formulae for the approximations and model discrimination are listed in Appendix D and the detailed data analysis procedures as well as the annotated scripts of the original source code material used to develop these models and parameters are made available by the author of this thesis at `https://github.com/theodora-nestorova/vocalvibratovariability`

3.3.3 Methods & Definitions for Vibrato Contour Archetyping

Upon more detailed time-domain vibrato signal evaluation, it was observed that the vibrato contour shapes of some samples were simple (or monophasic), while others were more complex (or multiphasic). Determination of the vibrato archetypes took a mixed methods approach, based both qualitatively on the original f_o contour and quantitatively on the threshold of half-extent perception of discernible vibrato versus non-vibrato (defined in prior voice science literature as approximately 13.5 cents on average) (Wooding and Nix, 2016). The iterative process of vibrato archetyping is outlined in greater detail in Sec. 3.3.6 below. The following subsections describe how simple and complex vibrato shapes and patterns can be defined, identified, and classified as vibrato half-extent time profiles, with examples from vibrato samples drawn from the voice study data set:

Simple / Monophasic Contour Patterns:

The simple or monophasic types of vibrato contour patterns are characterized by half-extent time profiles in which no evidence of distinct phases or transitions is observed. In this case, the half-extent time profiles are best approximated by linear regression lines. Examples of representative simple / monophasic vibrato contours extracted from the voice study data set are presented in Figs. 3.4 and 3.5.

1. **Uniform Vibrato:** In uniform vibrato, the half-extent is approximately constant with time; the frequency excursion of the vibrato contour does not significantly or perceptibly change over time. This subtype is best described by a nearly horizontal linear regression line with a slope of approximately 0 (see example in Fig. 3.4). The linear fit (broken black line) of the half-extent time profile of this specific uniform vibrato sample (blue dot time-series on the plot) has a slope of 0.97 cents/second. The intercept of the linear fit is 49.5 cents.

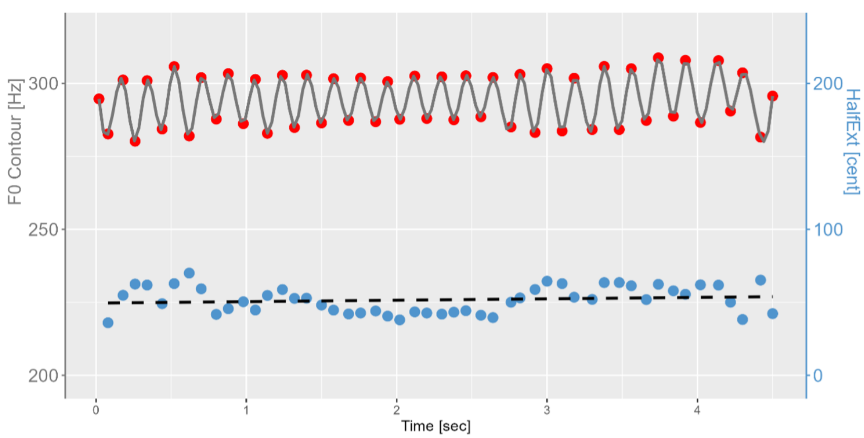


Figure 3.4: Example of simple, monophasic, uniform vibrato profile from S2 (ID12_16). (*Upper half of plot:* gray curve displaying original, smoothed fundamental frequency (f_o) contour connecting the red dots (peaks and troughs of the f_o). *Lower half of plot:* blue dot series representing the half-extent vibrato values calculated from the peaks and troughs with broken black line representing the linear regression fit through the half-extent vibrato time series).

2. **Uniformly Increasing Vibrato:** The half-extent of this vibrato archetype increases uniformly (gradually so) over time. This subtype is best described by a linear regression line

with a positive slope – an example is given in Fig. 3.5. The half-extent time profile of this specific increasing vibrato sample (blue dot time-series on the plot) is best fit by a linear regression (broken black line) which has a positive slope of 27.7 cents/second. The intercept of the linear fit is -1.85 cents, which is close to 0. The interpretation of this linear fit is that the vibrato starts with almost no half-extent (a non-vibrato initial onset at time = 0 seconds) and gradually, proportionally increases the half-extent to approximately 100 cents at the end of the vibrato (time = approximately 3.2 seconds).

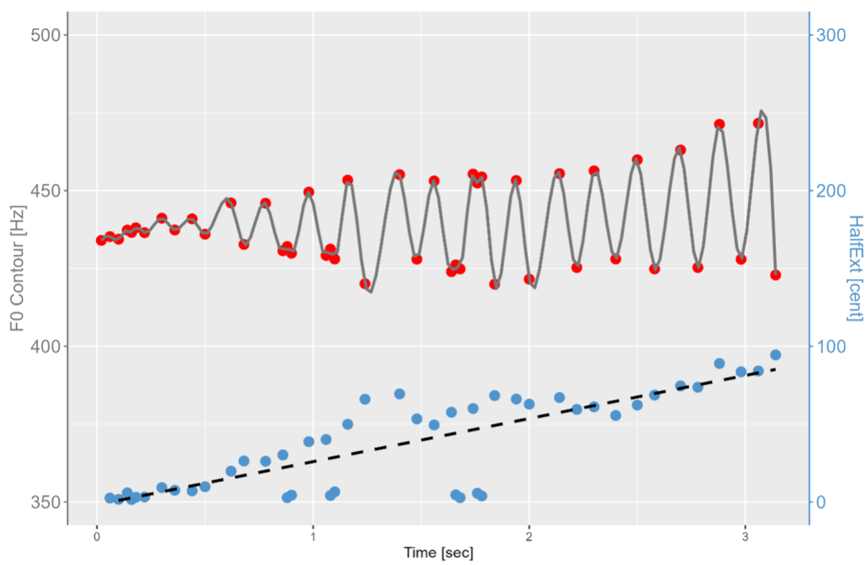


Figure 3.5: Example of simple, monophasic, increasing vibrato contour from S3 (ID13_14). (Orientation, lines, and dot colors are as described in Fig. 3.4).

3. **Uniformly Decreasing Vibrato:** The half-extent of this vibrato archetype decreases uniformly (gradually so) over time. This subtype is best described by a linear regression line with a negative slope. From the entire data set of the voice study, this vibrato type appeared to be the most uncommon.

Complex / Multiphasic Contour Patterns

The complex or multiphasic vibrato contour patterns are characterized by half-extent time profiles in which more than one distinct vibrato phase can be identified; each of these phases can be considered as being made up of simple (monophasic) vibrato with transitions between them. The transitions themselves can also be considered simple increasing or decreasing patterns. Therefore, a complex vibrato will include at least two, possibly more, simple vibrato shapes, which can be isolated and analyzed as separate phases. In many cases, the properties of the transitions between these phases may also be of interest.

Examples of complex vibrato contours drawn from the voice study data set are presented in Figs. 3.6 and 3.7.

1. **Step-Up Vibrato:** The half-extent is initially low (often non-vibrato) in the first phase, transitioning to a higher half-extent vibrato second phase (Fig. 3.6). An S-shaped 4-parameter logistic (4PL) curve best fits the half-extent time data. The step-up shape emerges as the most common complex vibrato shape.

   For the example shown, the half-extent time profile starts with a low half-extent phase (light blue dot time-series on the plot); the logistic regression yields an estimate of the half-extent for this phase at 12.6 cents (borderline flat tone, within the aforementioned non-vibrato perception threshold) and lasts up to 1.12 second. The short transition phase that follows (yellow dot time-series on the plot) spans between 1.12 and 1.32 sec and includes 3 half-extents, starting at 21.7 cents and rising to 49.2 cents. The high half-extent phase (blue dot time-series on the plot) starts at 1.32 sec and is characterized by a half-extent of 58.4 cents.

   A detailed interpretation of the logistic approximation in terms of the vibrato properties is given in the text below.

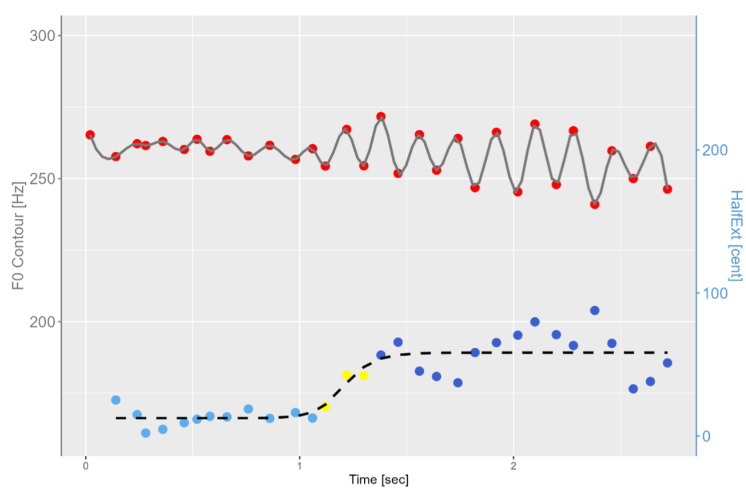


Figure 3.6: Example of complex, step-up vibrato contour from S2 (ID12_6). (Orientation, lines, and dot colors are as described in Fig. 3.4, except broken black line here represents the 4-parameter logistic regression fit through the half-extent vibrato time series; light blue dots denote Phase 1, dark blue dots denote Phase 2, and yellow dots denote the transition phase in between).

2. **Step-Down Vibrato:** Though much less common, this shape starts with an initial higher half-extent vibrato phase, transitioning to a low half-extent (Fig. 3.7).

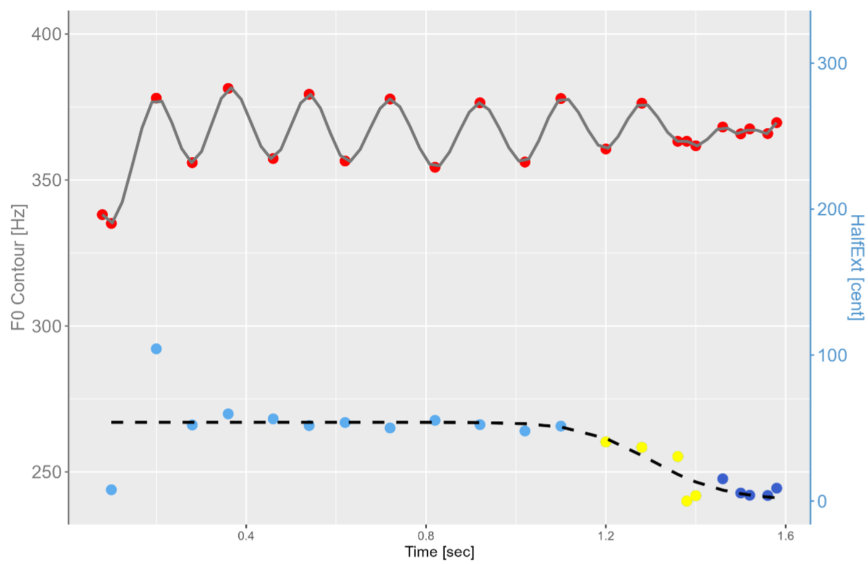


Figure 3.7: Example of complex, step-down vibrato contour from S2 (ID12_13). (Orientation, lines, and dot colors are as described in Fig. 3.6).

3. **More complex (multiphasic) vibrato patterns:** In contrast to the other complex patterns outlined above, patterns with more than one transition may exist. Preliminarily, it may be useful to analyze these by decomposing them into an interpolated series of biphasic vibrato tokens (such as those in Figs. 3.6 and 3.7). However, more refined approaches are needed to fully represent such complex, multiphasic vibrato half-extent time profiles. (See Fig. 3.10 for such an example.)

3.3.4 Resulting Characterization of Complex Vibrato Contour Patterns

As stated above, the definition of "complex vibrato patterns" inherently specifies an S-shaped form of the curves to fit to the half-extent time profile data.

This logistic function approximation is illustrated in greater detail in Fig. 3.8. As a result of the fit, the parameters *b, c, d, e* of the logistic curve in Equation 3.1 are calculated. Each of those parameters is related to a particular feature of the curve as annotated in the figure. Further explanation of the logistic curve parameters and their interpretations is given below:

- *b* characterizes the direction and the slope of the transition between the two asymptotes, corresponding to a large ($> |2-3|$) units) estimate indicating a fast transition. The sign of the transition denotes the direction: from high to low half-extent (positive sign, +) or from low to high (negative sign, -).

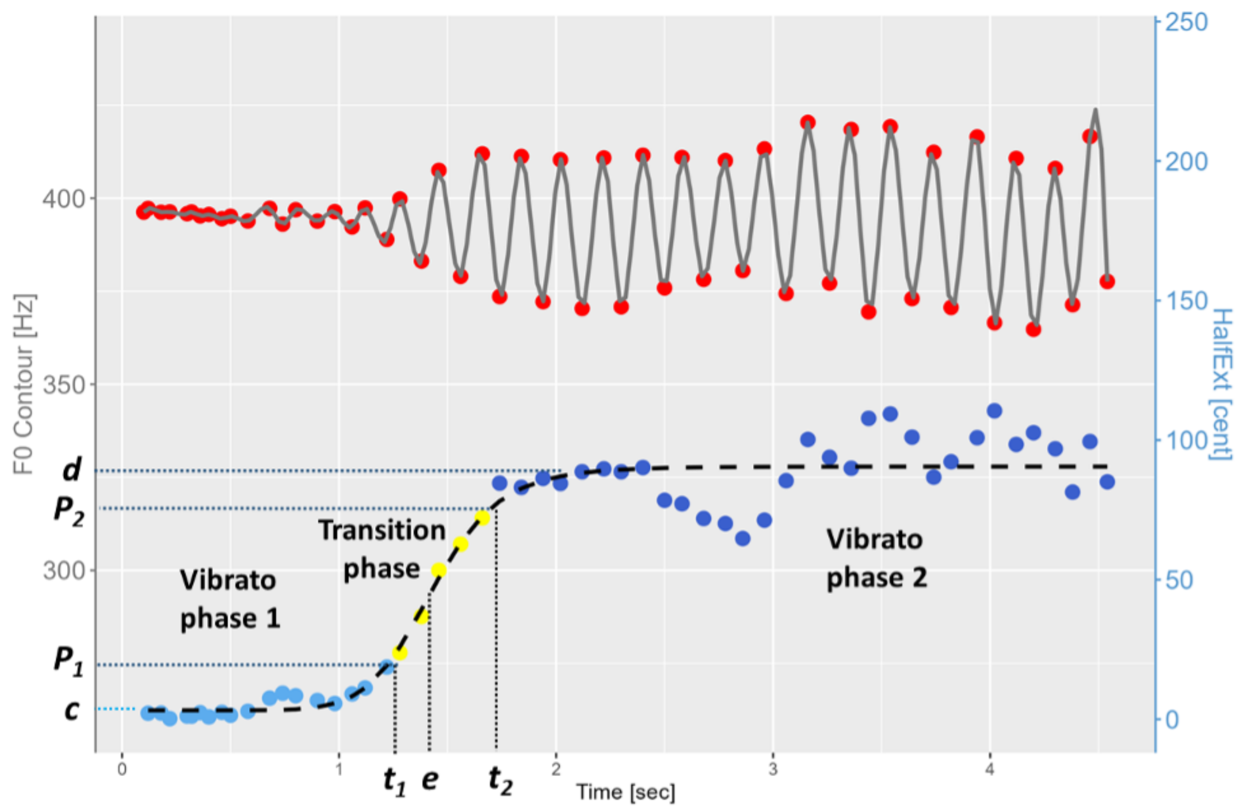


Figure 3.8: Logistic function approximation of a complex vibrato contour from S3 (ID13_4). (Orientation, lines, and dot colors are as described in Fig. 3.6. The logistic parameters *c*, *d*, and *e* are annotated on the X- and Y-axes).

- *c* is the lower[1] asymptote, corresponding to the half-extent [in cents] of Phase 1 of the complex vibrato.
- *d* is the upper[2] asymptote, corresponding to the half-extent [in cents] of Phase 2 of the complex vibrato.
- *e* is the middle time value, corresponding to the approximate time [in seconds] when half of the transition is completed.

Parameters of the logistic approximation in Fig. 3.8 and their interpretations are as follows:

- *b*: b = -9.28; Transition phase from low half-extent to high half-extent is quite short, lasting for $t_b - t_c$ = 0.43 seconds.
- *c*: Lower asymptote = 3.20 cents; Half-extent of Phase 1 is very low, corresponding to a non-vibrato onset.

[1]For inverted S-shaped complex vibrato, the Lower and Upper asymptotes are transposed.
[2]Ibid.

- *d*: Upper asymptote = 90.0 cents; Half-extent of Phase 1 is high, indicating a great vibrato excursion.

- *e*: Middle time value = 1.43 sec; Time in which half of the transition from 3.20 cents to 90.5 cents is 1.43 seconds.

- P_2, t_2: P_2 = 73.0 cents, t_2 = 1.66 seconds.; Phase 2 starts at 1.66 seconds.; any half-extents lower than 20.7 cents belong to the transition phase.

- P_1, t_1: P_1 = 20.7 cents, t_1 = 1.23 seconds.; Phase 1 ends at time 1.23 seconds.; any half-extents higher than 20.7 cents belong to the transition phase.

The estimation of the logistic curve parameters permits the separation of the three vibrato phases within the complex vibrato pattern: (i) Phase 1 – low half-extent vibrato; (ii) Transition Phase – transition from low to high half-extent vibrato; and (iii) Phase 2 – high half-extent vibrato. In concordance with the definition from the previous section, each of the phases is a simple vibrato pattern by itself: Phase 1 and 2 are uniform vibrato patterns, while the transition represents a uniformly increasing one.

The transition phase is defined as the segment of the half-extent time profile between 20% and 80% of the change between the lower and upper asymptote (the difference d-c). These thresholds are somewhat arbitrary, but coincide approximately with the linear ascent of the logistic curve. The times (t_1 and t_2) and the half-extent values ($P1$ and $P2$) that separate the phases in the example from Fig. 3.8 are given below the figure. For the exact formulae used, refer to Appendix D and for the As noted in the pertinent sections below, the detailed data analysis procedures as well as the annotated scripts of the original source code material used to develop these models and parameters by the author of this thesis, refer to `https://github.com/theodora-nestorova/vocalvibratovariability`. It can be seen that in this case that Phase 1 vibrato contains 19 half-extent observations, Phase 2 vibrato contains 31 half-extent observations, while the transition phase vibrato contains only 5 half-extent observations. It should also be noted that the half-extent of Phase 1 is very low, therefore, this phase can be classified as a non-vibrato segment.

3.3.5 Identification of Complex Vibrato Contour Patterns Using Statistical Criteria

As previously indicated in Nestorova et al. (2024b), the qualitative analysis and quantitative characterization of vibrato contours without considering the complexity of the vibrato patterns may lead to biased results. When analyzing complex vibrato contours, it is essential to segment the different vibrato phases from the transitions and analyze each of these fragments both separately and as a whole.

In such an approach, a critical question is how to assign a given vibrato contour to either the simple or the complex type. Based on the definitions given above, this should be done based on the Goodness of the Regression Fit (GoF) to the half-extent time data of either a straight line (indicating simple or monophasic vibrato) or a more complex curve – e.g. S-shaped curve (for a complex or multiphasic vibrato). In other words, the classification of a vibrato to one of the two classes defined herein should be based on both qualitative criteria (the presence of several phases in the vibrato contour, established visually and paired with the aforementioned perceptually validated vibrato versus non-vibrato extent threshold) and quantitative criteria (through formal statistical evaluation of regression fits with different mathematical models executing multiple computational fits to the data by straight line versus an S-shaped curve). Subsequently, model discrimination during the archetyping process of the vibrato half-extent time profiles should involve calculations to select the best approximation.

The use of the Akaike Information Criterion (AIC) and the Bayes Information Criterion (BIC) are proposed–– two widely used criteria for model selection (discrimination) reference–– to categorize the complexity of a vibrato contour. Both the AIC and BIC criteria account for the GoF of the model to the data, the sample size, and contain a penalty for the complexity (number of the parameters) of the model. As such they are routinely used to select the most parsimonious model from several concurrent models, given the data approximated. The model with the lower AIC and/or BIC values is favored in this process. For the exact AIC and BIC formulae used, refer to Appendix D.

Table 3.1 illustrates the model selection and vibrato complexity determination of the samples presented in Figs. 3.4-3.8. The first two vibrato samples (from Figs. 3.4 and 3.5 are determined by AIC and BIC as “simple vibrato,” whereas the last three (from Figs. 3.6-3.8 are classified as “complex vibrato” according to this proposed definitions and algorithm.

Sample	**Linear Fit**		**Logistic Fit**		**GoF Selection**	**Vibrato Type**
	AIC	***BIC***	***AIC***	***BIC***		
S2 (ID12_16)	***366.5***	***372.2***	*371.3*	*380.9*	Linear	Simple
S3 (ID13_14)	***398.4***	***403.9***	*402.2*	*411.3*	Linear	Simple
S2 (ID12_6)	*262.2*	*266.5*	***250.8***	***258.0***	Logistic	Multiphasic
S2 (ID12_13)	*197.8*	*201.1*	***193.4***	***198.9***	Logistic	Multiphasic
S3 (ID13_4)	*470.9*	*476.9*	***397.3***	***407.3***	Logistic	Multiphasic

Table 3.1: Model selection and vibrato complexity determination of the samples presented in Figs. 3.4-3.8. ***Note:*** *The bolded numbers indicate where the preferred (lower) AIC and BIC values are.*

3.3.6 Methodology Study: Results & Discussion

Over the course of quantitative examinations of the vocal vibrato characteristics in the samples from this voice study data set, f_o contours with different half-extent time profiles were encountered, requiring diverse analytical approaches. This finding suggests that any comparative evaluation of vibrato should consider existing shapes and patterns of the time profiles to prevent bias in the metric results. Consequently, several common half-extent time profile types were identified and thus separated and defined as "simple" or "monophasic" versus "complex" or "multiphasic" half-extent time profiles. Furthermore, in this research, novel applications of methods were proposed to discriminate between vibrato types by isolating the different vibrato phases within a complex vibrato sample.

Vibrato Extent Threshold Determination

As mentioned in Sec. 3.3.3 above, it is recommended that researchers approach the vibrato archetyping process more holistically and iteratively. Designating vibrato contour patterns based on their half-extent time profile shapes both qualitatively and quantitatively is suggested. Such mixed methods allow for both an individual examination and more accurate representation of each singer's vibrato samples rather than setting explicit limits or thresholds in favor of more stringent categorization criteria.

A relatively common complex vibrato pattern appeared to be the S-shaped contour in which a low half-extent non-vibrato phase is followed by a higher half-extent vibrato phase with a transition phase between the two, often representing a simply "uniformly increasing" vibrato. In the majority of cases, the lower phase had an estimated half-extent of several cents (5-10 cents), essentially representing a non-vibrato segment. Previous perceptual studies concerning the extent threshold separating vibrato from non-vibrato indicated that listeners were able to discriminate vibrato with full extent as low as 12 cents (6 cents half-extent), although the average vibrato extent threshold for rating a tone as non-vibrato was found to be approximately 27 cents peak to peak (13.5 cents half-extent) (Wooding and Nix, 2016). This S-shaped pattern was encountered in most of the complex vibrato samples in this data set and aligns with prior perceptual validation, notably occurring in the Jazz and Musical Theater genres.

From the proposed definitions and examples, it becomes immediately clear that complex vibrato patterns often consist of a series of adjoined simple vibrato tokens that can be separated and analyzed both individually and collectively. Therefore, the simplest "multiphasic" vibrato should consist of at least two distinct phases. Such a case is shown in Fig. 3.9 (S3, ID13_15), a vibrato sample from an Opera singer's extracted repertoire recording. In this case, the algorithm favored a logistic curve approximation (top panel) by AIC/BIC, with a very short Phase 1 (two half-extents) and almost non-existent transition (one half-extent). This preference is likely due to the first two half-extent points, which seem to constitute a separate vibrato phase. However, this vibrato is also very close to a simple "increasing" pattern (bottom panel). It is therefore recommended that the final decision as to the classification of this sample should be left to the discretion of the researchers themselves.

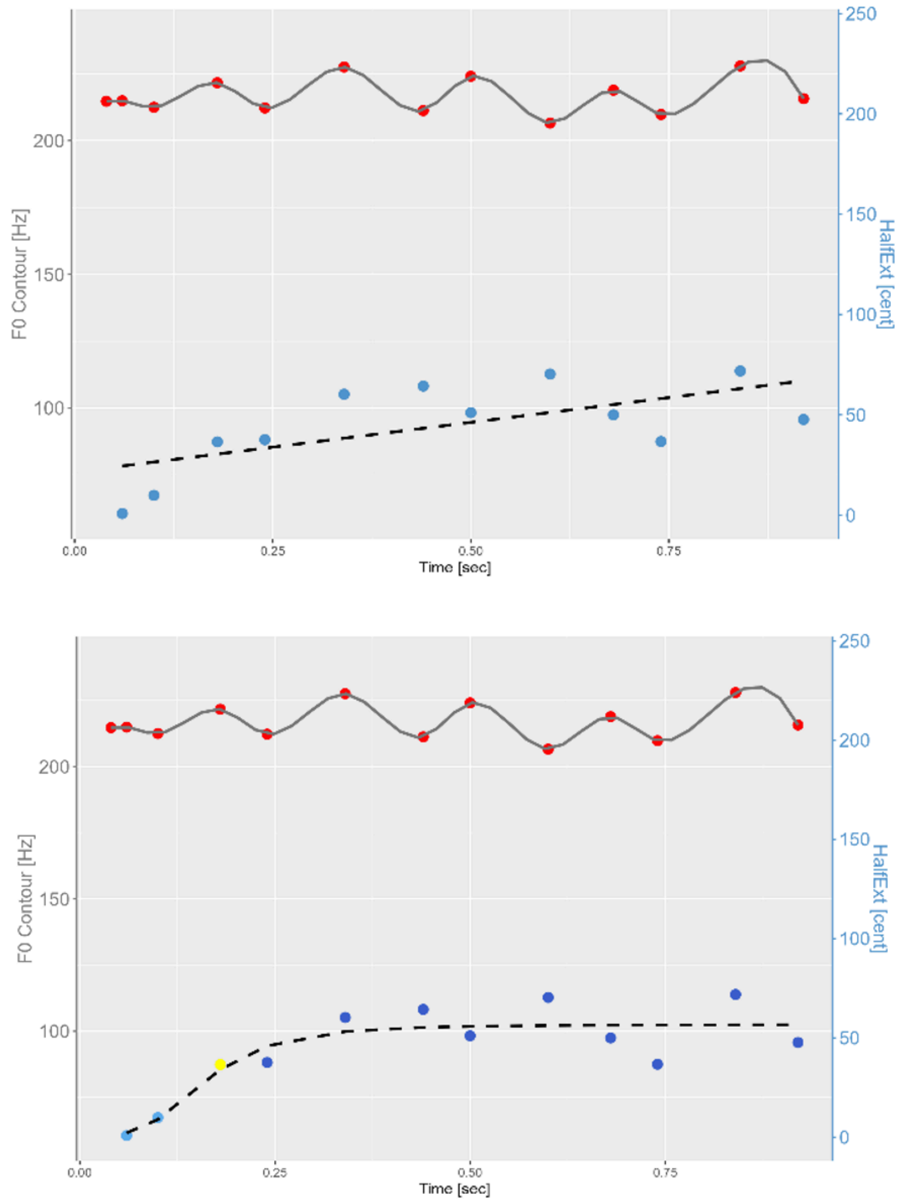


Figure 3.9: Example of the simplest "multiphasic" vibrato consisting of two phases from S3 (ID13_15). (*Top panel:* approximation of the biphasic vibrato by a straight line linear regression with AIC = 106.9 & BIC = 108.3. *Bottom panel:* approximation of the biphasic vibrato by a logistic curve regression with AIC = 99.1 & BIC = 101.5).

Infrequently, very complex vibrato patterns may emerge; an example of such is given in Fig. 3.10. Initially, the quantitative algorithmic fit by AIC/BIC criteria identified the first biphasic and a phase transition between Phases 1 and 2. However, upon closer visual inspection, it was determined that four distinct phases are present in this vibrato sample. Based on this secondary qualitative examination, the algorithm was manually modified to detect more than three phases. A more detailed characterization including four phases and their transitions were isolated and obtained by several superimposed and concatenated logistic regression approximations applied to each pair of successive phases. Occasionally, even more complex, multiphasic vibrato samples emerge; best practices for approaching analysis of these cases is currently being conducted by the author.

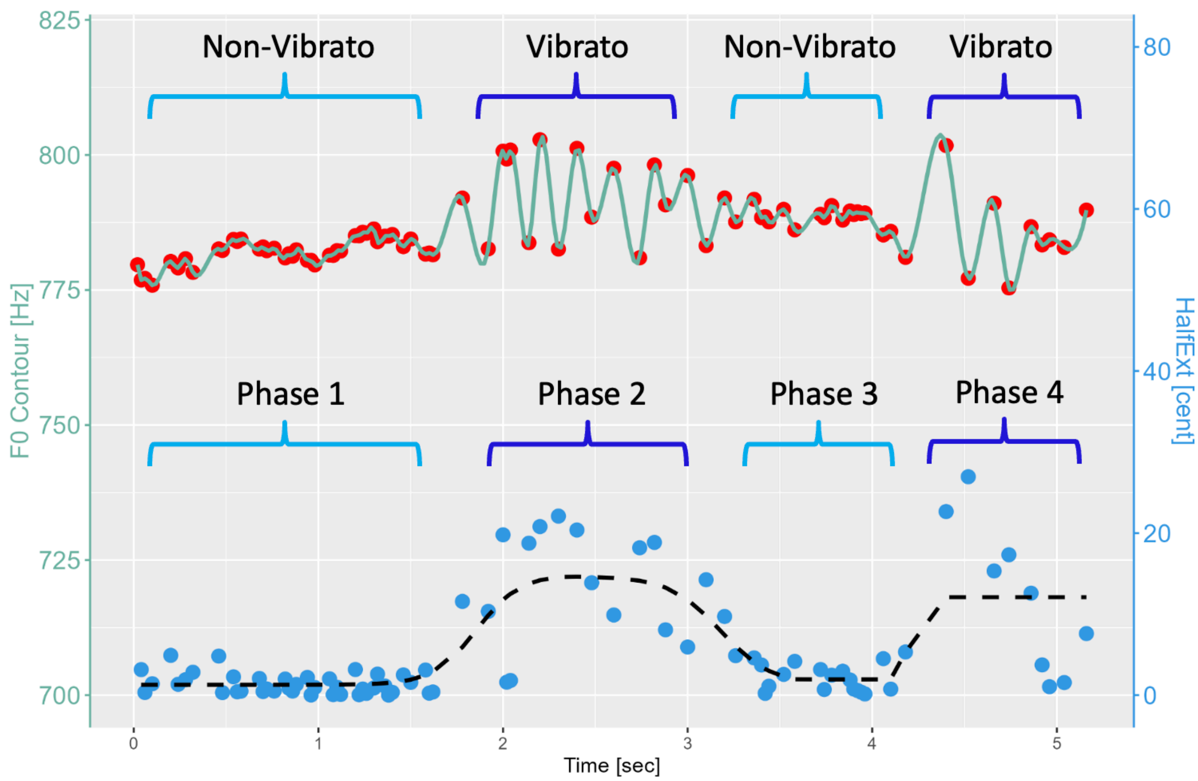


Figure 3.10: Example of the complex "quadriphasic" vibrato from S2 (ID12_12).

From a biomechanical standpoint, the "quadriphasic" vibrato in Fig. 3.10 may signal instability in the voice (see more on the implications of this to vocal dysfunctions, voice disorders, and other abnormalities/irregularities/pathologies in Chapter 7). However, the cause and underlying intention of this instability is not easily discernible solely from the vibrato pattern itself. Vibrato production is a complicated, multi-factorial process which is both volitional and non-volitional for singers, and may be due to several concomitant factors. Therefore, the classification of such a vibrato ultimately needs to be made using the subjective judgment of an expert, and not solely relying on the AIC/BIC which can be quite sensitive.

Steady-State vs. Full-Sample Vibrato Extraction

The extraction of the vibrato f_o contour from authentic singing excerpts within Praat may present some technical issues for the subsequent analysis, including isolation of the vibrato sample from the adjoining vocal phrases at the beginning or end of the time-series. Such cases require more scrutiny from the analyst as they might produce somewhat misleading outcomes of the AIC/BIC based classification scheme proposed.

This study differs in its ecological approach of including all cycles within the full vibrato sample, rather than the steady-state, as has been perpetuated as common practice in vocal vibrato analyses. Prame (1994) observed that singers' vibrato rates typically increase by approximately 15% at the end of each tone, while no consistent pattern is evident at the onset. This finding has led subsequent vibrato studies to exclude both the initial and final vibrato cycles, focusing instead on the steady-state segment for accurate analysis of vibrato rate and extent.

This research reconsiders inclusive analysis of the full samples of vibrato, selectively including rather than excluding and avoiding all onset and offset artifacts. Vibrato-specific information is lost when cutting off a pre-determined, maybe arbitrary (if the samples are of various lengths) number of seconds of onset and offset of the whole sample. As the vibrato rate tends to increase during the last three vibrato cycles of a tone, the average vibrato rate of a tone will depend on its duration, resulting in shorter tones showing higher average rates than longer tones (Prame, 1994). While there is justification in excluding the onset and so as not to skew average measures, the methodology proposed in this paper is focused on vibrato extent and not averages. Another justification for including all cycles from the vibrato proper is the limited duration of some vibrato samples that can be isolated and analyzed. Several vibrato samples drawn from authentic live singing may only include a limited number of cycles that fall at the boundary of the critical minimum of data points for a meaningful statistical analysis. As detailed variance measures and representations are the focus of this methodology study, we believe it is important to use researcher discretion and reconsider including the full representative sample, in the interest of most accurately characterizing the natural variability found in the above vibrato pattern representations.

Discretization / Sampling Frequency Considerations

Another technical problem encountered during data processing affected the determination of the peaks and troughs of the f_o contour, which is critical to this analysis. The discretization of the f_o contours was originally done with a default time step of 0.02 seconds, corresponding to a sample rate of 50 Hz, within Praat. This default step should be sufficient in the context of the Nyquist theorem, given the typical vibrato frequency of 5-7 Hz. Occasionally, however, spurious values appear in the f_o contour due to several possible sources, including artifacts of the frequency estimation algorithm and/or noise in the recording.

Such spurious values present in the f_o contour are illustrated in Fig. 3.11, displaying the effect discussed. In the second half of the top plot, a significant number of extra half-extent peak and trough values are identified because of the presence of high-frequency noise in the f_o contour. As one consequence, the variability of the half-extent in the second half of the profile is falsely overestimated. As a second, related consequence, the AIC/BIC preference for this profile is falsely assigned to the logistic regression, misclassifying this vibrato as "complex".

For the purpose of investigating this effect, the same vibrato was analyzed using increased discretization steps of 0.03 and 0.05 seconds (middle and bottom panels of Fig. 3.10). With a discretization step of 0.03 seconds, the spurious values are almost fully removed and there remains only one incidence in the whole profile. The AIC/BIC preference in this case is correctly assigned to "Linear" regression. With a discretization step of 0.05 seconds, there are no spurious values and the AIC/BIC preference is again correctly assigned to "Linear" regression. While it seems beneficial to increase the discretization step within Praat in order to limit the spurious values, another effect was noted, as illustrated by the results of the Linear fits of the three profiles, described above. With the increase of the discretization step from 0.02 seconds to 0.03 seconds to 0.05 seconds, the linear regression intercept decreases from 65.7 cents to 61.9 cents to 53.5 cents, while the slope also decreases from -11.9 cents/second to -4.75 cents/second to -3.49 cents/second. It is clear that a discretization step of 0.05 seconds, although still sufficient from a sample rate perspective, misses some of the "true" peaks and troughs of the f_o contour, leading to potential meaningful underestimation of the real half-extents and bias in all subsequent results. Discretization steps of 0.02 or 0.03 seconds fare much better in that respect. It is important to note, however, that the increase of the step does not impact the classification of the vibrato. Therefore, when only classification of the type of vibrato is pursued, the above considerations are of lesser consequence. Following this evaluation, it may be concluded that this vibrato with a half-extent of around 60 cents can be classified as uniform.

A few options exist to deal with these spurious values, including (1) using a more refined peak picking algorithm that ignores spurious values between peaks and troughs; or (2) filtering the f_o contour before peak picking. For example, a simple two-point moving average performed on the sample in Fig. 3.11 produced results similar to those of the manual approach (trial and error optimization of the discretization step detailed above), but the two-point moving average was more effective, reliable, and successful in smoothing the f_o contour before peak and trough finding. These recommended solutions warrant further investigation across a broader data set of

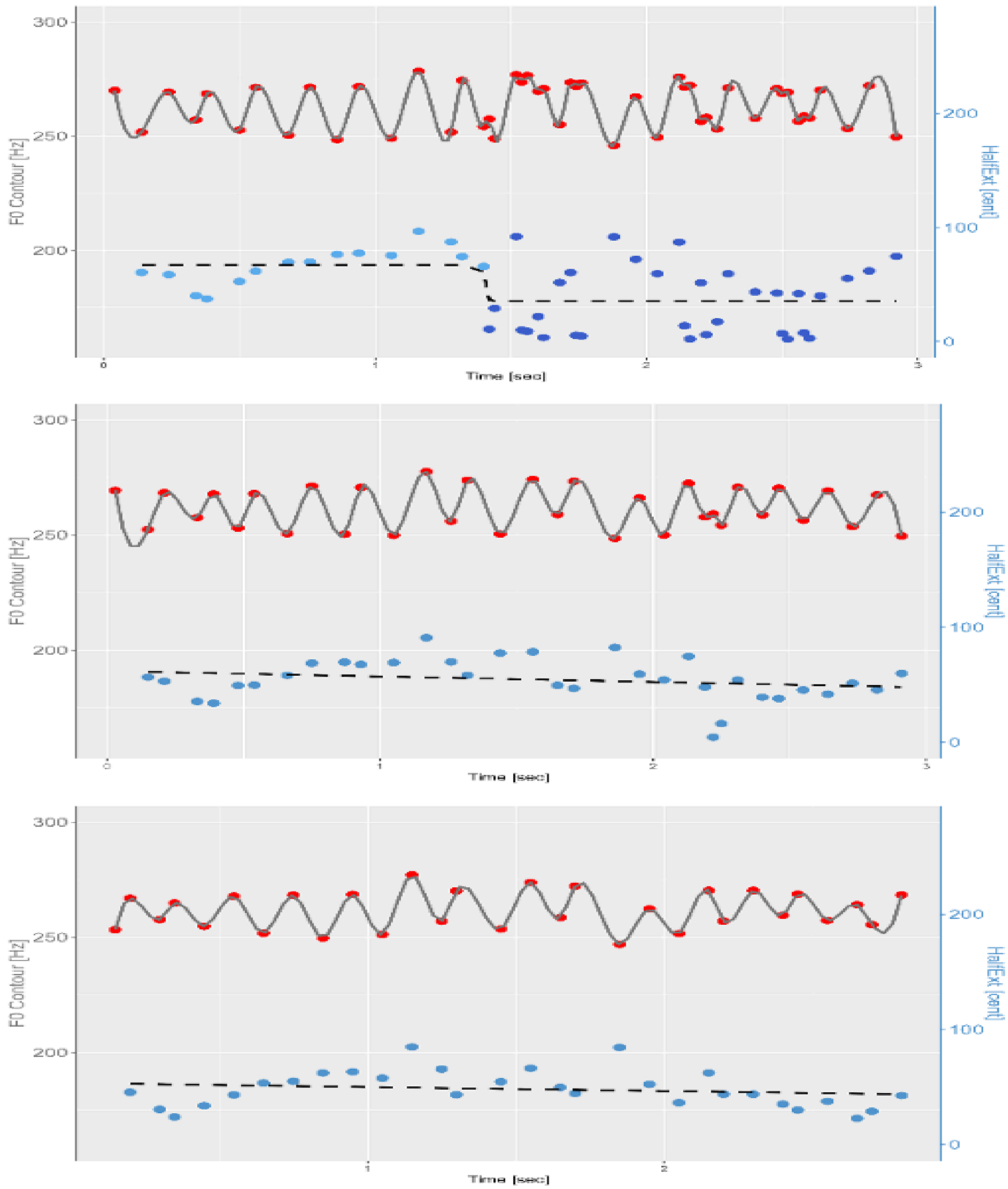


Figure 3.11: Example of the effect of spurious values in the f_o contour and parameters of S1 (ID2_6). (*Top panel:* Initial algorithmic fit with discretization step: 0.02 seconds; *Middle panel:* Secondarily refined algorithmic fit with discretization step: 0.03 seconds; *Bottom panel:* Tertiarily refined algorithmic fit with discretization step: 0.05 seconds. Note that the continuous curve through the red dots is a spline approximation, which rarely but occasionally surpasses them.

time series data.

3.3.7 Methodology Study: Conclusions & Future Research

The objectives of this research were met by defining different shapes and patterns of time-varying vibrato contours, developing modeling methodology for identification and quantitative analysis of simple (monophasic) versus complex (multiphasic) vibrato half-extent time profiles, and applying this methodology to ecologically valid samples from a prior voice study's data set and interpreting the results.

The archetype definitions outlined in this methodology study are as follows: Simple / Monophasic Vibrato Patterns (including Uniform, Uniformly Increasing, and Uniformly Decreasing profiles), and Complex / Multiphasic Vibrato Patterns (including Step-Up, Step-Down, and More Complex / Multiphasic profiles). Based on the participants in this voice study's data set, genres may not differ based on single measures of variability (such as CV), but rather on prevalence of vibrato pattern (generally more complex in Musical Theater and Jazz, and predominantly uniform in Opera, though exceptions certainly exist).

This methodology study emphasizes the need to decompose and classify complex, multiphasic vibrato samples before conducting comparative analyses. Vibrato complexity, vibrato archetype prevalence, and properties of the transitions between vibrato phases may become characteristics of interest on their own. These characteristics are the subject of current and continuing research by the author.

The differentiation of simple (uniform) and complex (multiphasic) vibratos and the methodology of evaluating the variability of complex vibrato contours have not yet been proposed and discussed in the literature. It requires a comprehensive understanding and specialized methodology that may potentially present some challenges related to data quality, approach limitations, and results interpretation. Therefore, it is critical to realize that applying any measure of vibrato half-extent variability (such as Coefficient of Variation of the half-extents) (Nestorova et al., 2024b) without accounting for the complexity in the temporal profiles can bias the results and interpretations.

This study contributes to vibrato analysis by incorporating a data set validated through perceptual evaluation, reinforcing its psychoacoustic relevance. Future research aims to integrate these models into open-source platforms for voice practitioners. It is crucial that these tools be accessible, practical, and interpretable for a wide range of stakeholders, including vocal pedagogues, therapists, and performers. Further investigation with a broader data set of singers, genres, and evaluators is also necessary to refine and validate these approaches. The proposed time-varying models for analyzing vibrato variability previewed in this paper may be a useful, pragmatic, and progressive addition addition to the current systems of vibrato evaluation (notably half-extent variability metrics and adding time-varying models as viable parameters).

3.4 Interpretation & Evaluation of Vibrato Contour Variability using the HoE vs. rHE Metrics

As discussed in Sec. 3.2.1, two distinct time-series metrics— Half-of-Extent (HoE) and Relative Half-Extent (rHE)— are utilized to interpret time-varying vibrato contours. This section compares these metrics to elucidate their differences and respective benefits. Both HoE and rHE are applied to the same datasets to validate and refine previous results, with the Discussion providing a detailed analysis of their comparative advantages.

To test the applicability of the methodological framework outlined in Sec. 3.3, the vibrato samples from recordings of three singers representing three genres in a cross-section sampling of the Nestorova et al. (2024b) voice study were analyzed. The purpose of the following sections is to apply the methodology to a subset of ecologically valid vibrato samples and analyze overall trends in the results. Detailed results of all samples included in this three-individual representative sampling analysis and interpretations below will be made available by the author of this thesis upon request.

3.4.1 Comparison Results: Simple vs. Complex Vibrato Incidence

Three individual singer subjects representing each genre (Opera, Musical Theater, Jazz) with samples including the widest variety of complex vibrato were selected as a cross-section from all total vibrato samples (N = 240) and singer subjects (N = 16) of the prior voice study outlined and mentioned in Sec. 2.7 (Nestorova et al., 2024b).

HoE Metric Analysis

The results of the vibrato type classification using HoE for these individuals by genre are given in Table 3.2. The total number of simple (Uniform, Uniformly Increasing, or Uniformly Decreasing) vs. complex (Biphasic Step-Up, Biphasic Step-Down, and Multiphasic) vibrato contours by individual singer are listed in the last two rows of the table and represented on the bar plot in Fig. 3.12. It is important to note that some complex vibrato samples listed below potentially include more than one transition phase (belonging to the More Complex / Multiphasic archetype).

[3] Complex vibrato with potentially more than one transition phase

[4] Ibid.

[5] Ibid.

[6] Ibid.

[7] Ibid.

Sample #	Sample Name	S1: Musical Theater	S2: Jazz	S3: Opera
1	Time	Complex[3]	Complex	Complex
2	Easy	Simple	Complex	Complex[4]
3	Jumping	Complex	Simple	Complex
4	High	Complex	Complex	Complex
5	Rich	Simple	Simple	Complex
6	Looking	Simple	Complex	Complex
7	A3	Complex	Complex	Simple
8	C4	Complex	Complex[5]	Complex
9	A4	Complex	Complex	Complex
10	C5	Complex	Complex	Complex
11	E5	Complex	Simple	Complex[6]
12	G5	Simple	Complex[7]	Complex
13	Grace	Simple	Complex	Complex
14	Me	Complex	Simple	Simple
15	Found	Simple	Simple	Simple
16	See	Simple	Simple	Complex
Number of Complex		9	11	14
Number of Simple		7	5	2

Table 3.2: Type of simple vs. complex vibrato half-extent time profiles per sample using HoE in representative sampling of 3 singers across 3 genres.

Complex vibrato samples dominate versus simple types across all three singer subjects and their respective genres, with S3 possessing only two simple vibrato contours. This underlines further the need and importance of creating and using a methodology to analyze vibratos by accounting for the complexity. The preponderance of complex vibratos by genre is apparent in this sampling of three singers, but this hypothesis needs to be tested on a larger sample size of singers per genre and compared with varied metric approaches (i.e., with both HoE, as in Table 3.2, and rHE, as in Table 3.3) to draw more generalizable conclusions.

The distribution of simple and complex vibrato types across the singers is illustrated graphically on Fig. 3.12. It is important to note that the trend from singer subjects and samples in the prior study outlined in Sec. 3.3 indicated a greater homogeneity of simple vibrato types in Opera than in Jazz or Musical Theater, making S3 with the wide variety of complex vibrato types different in this regard.

When applying the AIC and BIC criteria for vibrato type classification, it should be noted that in a limited number of samples, the AIC criterion prefers the logistic function fit, while the BIC criterion prefers the linear function fit. This is the case in two samples from S1 and one sample from S2. Upon inspection, it was observed that this is due to the presence of a very small number of onset/offset modulation artifacts in the half-extent time profiles, usually at the start or at the

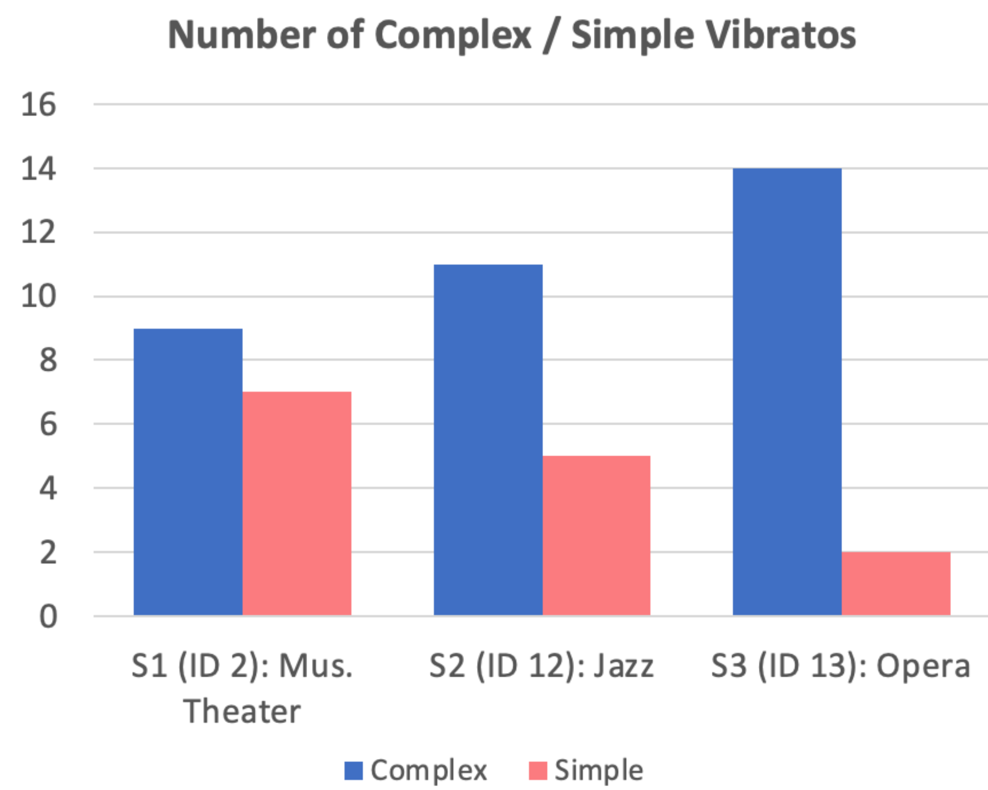


Figure 3.12: Number of complex and simple vibratos by individual singer, calculated with HoE.

end of the sample. Due to the higher discriminating power of the BIC criterion, a simple vibrato type preference was concluded for these samples.

The most frequently encountered case of multiphasic vibrato consisted of two phases connected by a transition phase. A number of samples include multiple phases and transitions with a pattern of several interpolated biphasic step-up / step-down profiles. This was the case in one sample from S1, two samples from S2, and one sample from S3 (see Table 3.2). This observation demonstrates that either more complex types of the vibrato half-extent time profiles exist, or the increased complexity is a consequence of inherent irregularity of the singer's vibrato, reflected in the half-extent contour itself. One way to handle this situation requires more nuance and approach modification by decomposing each sample into several biphasic segments, and sequentially applying the 4-PL logistic regression to each of the segments (see Fig. 3.10 as an example).

rHE Metric Analysis

When each vibrato sample was analyzed with the rHe metric instead of the original HoE metric, the vibrato archetype classification for these individual singers' results remained almost identical to those given in Table 3.2. However, there were three classifications that switched from simple to complex and two classifications that switched from complex to simple when analyzed with rHE compared to HoE. See Table 3.3 for the samples that experienced changes in vibrato archetypes from HoE to rHE metrics by individual singer and genre.

#	Sample Name	S1 (ID2): MT	S2 (ID12): Jazz	S3 (ID13): Opera
5	Rich	Simple $\rightarrow$ *Complex*	Simple $\rightarrow$ *Complex*	Complex
15	Found	Simple $\rightarrow$ *Complex*	Simple	Complex $\rightarrow$ *Simple*
16	See	Simple	Simple	Complex $\rightarrow$ *Simple*

Table 3.3: Changes in vibrato archetypes from HoE to rHE metrics per singer per genre; appearing only in samples 5, 15, and 16.

As demonstrated in Tables 3.2 and 3.3, the classification of vibrato archetype remains largely consistent using either HoE or rHE, however, in certain instances, the selection criteria based on AIC and BIC resulted in a marginal shift from HoE to rHE without an apparent underlying justification. This variability can be attributed to the increased fluctuation observed in half-extent values derived from the rHE measure. The heightened variation in rHE introduces additional complexity, making it more challenging to distinguish existing patterns in the fundamental frequency f_o vibrato contour. In those cases, the AIC/BIC suggested inference was considered artifactual and the HoE-based classification was mainly retained.

3.4.2 Summarized Results: Complex Vibrato Profiles

HoE Metric Analysis

Most of the complex vibrato half-extent time profiles consist of a Phase 1 with a half-extent that is approximately or less than 10 cents (essentially a non-vibrato tone) transitioning to a Phase 2 with much higher vibrato; the latter is henceforth referred to as "vibrato proper." Three Phase 1 vibrato samples of S3 have half-extents greater than 20 cents, which can be qualified as vibrato proper. Since analyzing the variability of a non-vibrato tone is not the subject of current interest, Phase 1 CV results are not interpreted here; the results shown refer to the variability of the proper vibrato profiles from Phase 2. The summarized results from the vibrato variability analysis of the complex vibrato half-extent time profiles and their CV by individual and by phase are listed in Table 3.4. The same data are illustrated graphically in Fig. 3.13.

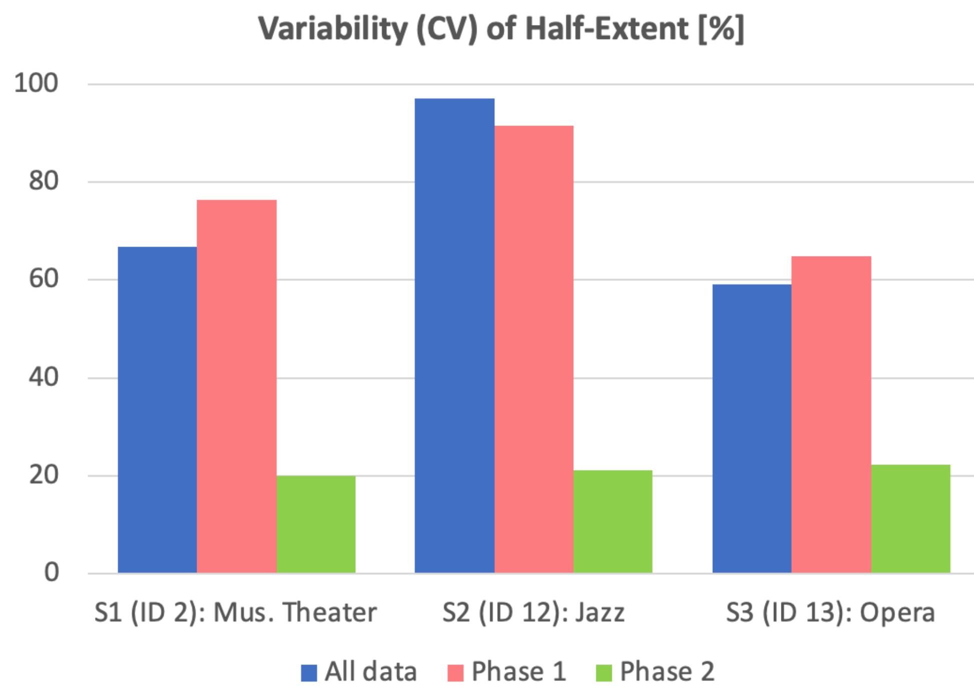


Figure 3.13: Complex Vibrato: Variability (CV in [%]) calculated with HoE) by phase and collectively in each individual singer.

Data from		Half-Extent[8] [cent]	CV[9] [%]
S1 (ID2): Musical Theater	All data per sample	55.9	66.8
	Phase 1	10.9	76.3
	Phase 2	79.0	20.0
S2 (ID12): Jazz	All data per sample	20.1	97.1
	Phase 1	5.77	91.5
	Phase 2	46.8	21.1
S3 (ID13): Opera	All data per sample	46.3	59.1
	Phase 1	11.3	64.9
	Phase 2	63.6	22.3

Table 3.4: Complex Vibrato Profiles: Half-extents & variability (CV) calculated with HoE by phase and collectively in each individual singer.

It can be seen that the CV estimates of Phase 2 (vibrato proper) are significantly lower than the CV estimate over the whole complex sample (Phase 1 + Transition + Phase 2). For S2, the CV decreases by a factor more than 3 from 66.8% to 20.0%; for S2, the difference is almost 5-fold from 97.1% to 21.1%, and for S3, it is 3-fold from 59.1 to 22.3%.

[8] Averaged over the average Half-Extent of each sample within an individual.
[9] Averaged over the average CV of each sample within an individual.

These results are an important validation of the proposed approach as they demonstrate the significant overestimation when variability is measured without consideration of the vibrato complexity of the half-extent time profile. Such quantitative variability estimates (i.e., using CV) are truly representative only when the vibrato proper phase is analyzed, rather than taking an average aggregate CV of multiple phases of varying complexities within one vibrato sample.

Table 3.4 also shows that the mean half-extents are consistently lower when the whole sample is analyzed, compared to the results for the vibrato proper from Phase 2. For S1, the half-extent increases from 55.9 cents to 79.0 cents; for S2, the increase is from 20.1 cents to 46.8 cents; and for S3, the increase is from 46.3 cents to 63.6 cents.

This is another important result, indicating again that ignoring the qualitative complexity of vibrato half-extent time profile inevitably leads to underestimation of the quantitative vibrato half-extent, as the statistical summary averages together half-extent from the two very different vibrato phases.

A third important observation from these results is that the CV of the Phase 2 vibrato for all three singers is quite similar— approximately 20-22% —evidence that most samples' vibrato proper was fairly regular. This is also further indication that the true quantitative variability of vibrato may be invariant with respect to genre; a hypothesis that also needs to be tested on a larger sample size of singers per genre.

rHE Metric Analysis

Results using rHE are provided in Table 3.5. As with the HoE metric use, when using rHE, Phase 1 half-extents for all three singers are consistently low, around or below 20 cents, approaching what could be classified as a borderline non-vibrato tone. The abnormally high variability observed in Phase 1 suggests that these fluctuations may primarily reflect background noise artifacts rather than meaningful artistically-relevant variation.

The CV of Phase 2 of the vibrato samples across all three singers remains relatively consistent, ranging from approximately 29% to 37%. This consistency raises the same question as with the HoE metric of whether vibrato variability is invariant across musical genres, which warrants further investigation.

For Phase 2, half-extents appear to vary by individual, yet the overall variability of the vibrato proper remains relatively low. This raises the possibility that these vibrato proper segments may be considered stable, warranting further exploration into individual differences in vibrato consistency. See Table 3.5 for the numeric details.

Similarly to the use of HoE above to analyze each sample, when using rHE, the CV for the half-extents calculated across all samples for each singer was substantially higher than when calculated specifically for Phase 2 (the vibrato proper) in each individual. This underscores the necessity of employing a methodological approach that accounts for vibrato complexity to mitigate potential biases in variability estimates.

Data from	Half-Extent[10] [cent]	CV[11] [%]
S1 (ID2): Musical Theater		
All data per sample	57.1	66.2
Phase 1	15.5	81.5
Phase 2	79.1	28.6
S2 (ID12): Jazz		
All data per sample	22.4	92.6
Phase 1	8.08	76.1
Phase 2	47.2	37.1
S3 (ID13): Opera		
All data per sample	48.2	61.3
Phase 1	19.2	56.7
Phase 2	64.3	33.4

Table 3.5: Complex Vibrato Profiles: Half-extents & variability (CV) calculated with rHE by phase and collectively in each individual singer.

The comparison between the HoE and rHE metrics reveals notable differences in variability estimation. Specifically, Phase 1 half-extent estimates increased from below 10% in HoE to a range of 8%–19% in rHE. Similarly, the CV for Phase 2 increased from approximately 20% under HoE to 29%–37% with rHE. These differences can be attributed to the greater variability captured by rHE, which seems to more comprehensively account for variability over the course of the vibrato samples. Refer to Fig. 3.14 for a more side-by-side visualization.

3.4.3 Summarized Results: Slopes of Simple Vibrato Profiles

HoE Metric Analysis

The summarized results of the simple vibrato samples' half-extents, their CVs, and the slopes of the linear regressions describing the simple vibrato types by individual and by phase are listed in Table 3.6. The simple vibrato half-extent time profiles of S1 and S2 are predominantly uniform, with slopes close to zero (vibrato half-extents do not change over time within the sample). This is as illustrated by Figs. 3.4 and 3.5.

Both simple vibrato profiles of S3 are of the uniformly increasing half-extent archetype, yielding large slopes in their linear approximations. Due to the low number of simple profiles and predominantly uniformly increasing archetype, S3's simple vibrato results are not interpreted.

[10] Averaged over the average Half-Extent of each sample within an individual.
[11] Averaged over the average CV of each sample within an individual.

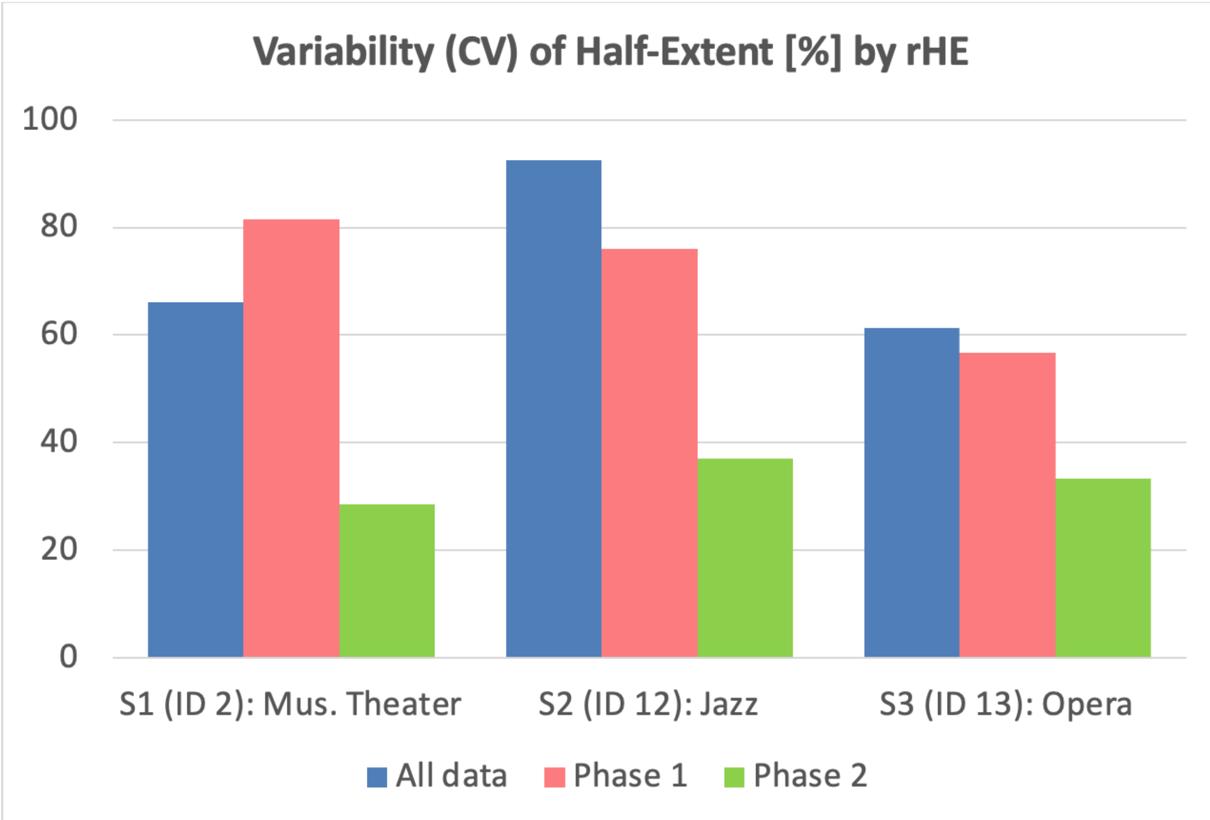


Figure 3.14: Complex Vibrato Profiles: Variability (CV in [%]) calculated with rHE by phase and collectively in each individual singer.

While this may bear potential functional voice use implications, the overall trend does not conform to the other singers in the same genre within the prior voice study and does not accord to documented technical and artistic conventions distinguishing the Operatic genre.

Data from	Half-Extent[12] [cent]	CV[13] [%]	Slope[14] [cent/sec]
S1 (ID2): Musical Theater	71.3	29	-0.151
S2 (ID12): Jazz	49.6	20	3.01
S3 (ID13): Opera	30.7	68.1	16.5

Table 3.6: Simple Vibrato Profiles: Half-extents, variability (CV), and slope calculated with HoE in each individual singer.

An important observation here is that the simple vibrato half-extent estimates are similar to the vibrato proper estimates of the complex profiles. For S1, these are 71.3 cents vs. 79.0 cents respectively, while for S2, these are 49.6 cents vs. 46.8 cents respectively. The same is true for the

[12] Averaged over the average Half-Extent of each sample within an individual.
[13] Averaged over the average CV of each sample within an individual.
[14] Estimated from the linear regressions of each sample within an individual.

CV estimates. For S1, these are 29.0% vs 20.0% respectively, and for S2, these are 20.0% vs 21.1% respectively. The consistency of the half-extents and their variability estimates between uniform simple vibratos and Phase 2 of the complex vibratos both within and across singers is remarkable and will be discussed further in the Discussion section.

rHE Metric Analysis

As shown in Table 3.7, vibrato patterns of S1 and S2 predominantly match the uniform archetype, with stable half-extents that remain consistent over time. As a result, the slopes of their linear approximations are minimal, reflecting a lack of significant variation. In contrast, the vibrato patterns of S3 demonstrate an increasing vibrato half-extent time profile, yielding substantially larger slope values in the linear approximations. This raises the question of whether such vibrato patterns have artistic or perceptual implications that warrant further exploration.

Data from	**Half-Extent**[15] **[cent]**	**CV**[16] [%]	**Slope**[17] **[cent/sec]**
S1 (ID2): Musical Theater	72.0	33.4	4.34
S2 (ID12): Jazz	48.5	30.2	4.39
S3 (ID13): Opera[18]	67.5	48.4	14.5

Table 3.7: Simple Vibrato Profiles: Half-extents, variability (CV), and slope calculated with rHE in each individual singer.

When comparing HoE and rHE (excluding S3), half-extent and slope estimates remain largely similar. However, the CV for half-extent increases from approximately 20% under HoE to over 30% with rHE. This increase is attributable to the greater variability captured by rHE, which provides a more detailed representation of vibrato variability.

3.4.4 Summarized Results: Transition Phases of Complex Vibrato Profiles

HoE Metric Analysis

Summarized results of the parameters of the transition phases of the complex vibrato half-extent time profiles by individual and by phase are listed in Table 3.8. These include the estimated asymptotic Phase 1 half-extent, asymptotic Phase 2 half-extent, the difference between the two asymptotes (referred to as Delta Half-Extent), and the estimated duration of the transition phase (referred to as Vibrato Onset Time, having drawn inspiration from Voice Onset Time (VOT), an acoustic parameter that quantifies the duration between the release of a plosive stop consonant's

[15] Averaged over the average Half-Extent of each sample within an individual.
[16] Averaged over the average CV of each sample within an individual.
[17] Estimated from the linear regressions of each sample within an individual.
[18] S3/ID13 has only 2 Simple Vibratos.

oral constriction, marked by a noise burst, and the initiation of voicing for the following sound, indicated by the onset of periodicity (Kuhlewind, 2014; Lisker and Abramson, 1964; Klatt, 1975)). These parameters are calculated from the logistic regression fits. Thus the Phase 1 and Phase 2 half-extent asymptote values are approximations of the values, presented in Table 3.6. The similarity of the asymptotes to the data in Table 3.6 indicates that the logistic functions describe well the data. Additionally, both Delta Half-Extent and Vibrato Onset Time are novel parameters with features that deserve more attention, and are part of the future work of this research.

Data from	Phase 1 Half-Extent[19] [cent]	Phase 2 Half-Extent[20] [cent]	Delta Half-Extent[21] [cent]	Vibrato Onset Time[22] [sec.]
S1 (ID2): Musical Theater	9.18	80.2	71	0.233
S2 (ID12): Jazz	5.04	47.9	42.9	0.487
S3 (ID13): Opera	8.93	65.4	56.5	0.49

Table 3.8: Complex Vibrato Profiles: Transition phase parameters from the logistic function fits, change in half-extent, and Vibrato Onset Time, calculated with HoE and averaged by individual singer.

Several conclusions can be drawn from the results reflected in Table 3.8; across the singers, the Phase 1 asymptotes are fairly similar. However, the Phase 2 asymptotes, the Delta Half-Extent, and the Vibrato Onset Times differ across these individual singers and hence have the potential to differentiate across genres, though this hypothesis warrants further testing. Interestingly, the Vibrato Onset Times do not seem to be correlated to the Delta Half-Extents within the individuals. However, this conclusion is considered preliminary due to the small sample size and the presence of several Onset Time outliers in many of the vibrato samples.

rHE Metric Analysis

According to Table 3.9, the analysis of transition phases of the complex vibrato half-extent time profiles by individual and phase indicates that Vibrato Onset Times consist of the estimated asymptotic Phase 1 half-extent, asymptotic Phase 2 half-extent, the difference between the two asymptotes (referred to as Delta Half-Extent), and the estimated duration of the transition phase (referred to as Vibrato Onset Time). All calculated by using rHE were comparable to the HoE calculations (shown in Table 3.8) across samples, with no evident correlation between Vibrato Onset Time and the half-extent of the Phase 1 to Phase 2 transition, as measured by Delta Half-Extent.

[19] Averaged over the Half-Extent logistic regression fit of each sample within an individual singer.
[20] Ibid.
[21] Differential between the phase 1 and phase 2 half-extent averaged over samples within an individual singer.
[22] Duration of the transition phase averaged over each sample within an individual singer.

Data from	Phase 1 Half-Extent[23] [cent]	Phase 2 Half-Extent[24] [cent]	Delta Half-Extent[25] [cent]	Vibrato Onset Time[26] [sec.]
S1 (ID2): Musical Theater	12.6	81.3	68.7	0.439
S2 (ID12): Jazz	6.76	48.4	41.6	0.554
S3 (ID13): Opera[27]	14.6	65.9	51.3	0.513

Table 3.9: Complex Vibrato Profiles: Transition phase parameters from the logistic function fits, change in half-extent, and Vibrato Onset Time, calculated with rHE and averaged by individual singer.

When comparing HoE and rHE (excluding S3's sample, ID 13_16, due to spurious value outliers), the results show minimal differences, suggesting that the choice of metric does not substantially impact the overall estimation of vibrato characteristics in this context.

3.4.5 Discussion of Interpretative Results

HoE Metric Analysis

The results of the vibrato variability analysis based on HoE lead to a number of conclusions, as follows:

1. The proposed AIC and BIC statistical criteria are successful in discriminating between simple and complex vibrato archetypes of time-varying vibrato. This justifies the approach in utilizing such models, though their sensitivity to outliers at the onset and offset of the samples must be acknowledged (and accounted for with researcher discretion).

2. The comparisons of the CV estimates derived from the whole vibrato samples, and the ones from Phase 2 segments illustrate the benefit of considering vibrato type when evaluating vibrato variability. The potential overestimation of variability and underestimation of half-extent in the former approach have been convincingly displayed. Such a bias in the estimates can be misleading.

3. The trends in typical vibrato half-extent variability indicate that the metric of variability (CV) remains reliably around 20% for vibrato proper phases, while the qualitative shape of the half-extent time profile for each vibrato sample itself varies within singer and genre (though patterns may emerge with a greater sample size).

[23] Averaged over the Half-Extent logistic regression fit of each sample within an individual.
[24] Ibid.
[25] Differential between the Phase 1 and Phase 2 Half-Extent averaged over samples within an individual.
[26] Duration of the Transition Phase averaged over each sample within an individual.
[27] ID13_16 has been excluded as an outlier, and therefore not included here.

During data processing, spurious values occasionally appeared in the f_o contours, likely due to artifacts in the frequency estimation algorithm or from background noise. These anomalies, though limited in number, potentially inflated variability estimates. To mitigate this, the time step for discretizing the f_o contours was increased from the default 0.02 seconds to 0.03 seconds. This adjustment effectively reduced spurious values without compromising the algorithm's ability to detect f_o peaks and troughs, thereby maintaining the integrity of the half-extent estimates. Therefore, filtering spurious values while preserving valid half-extent estimates remains a complex signal processing challenge that warrants further investigation.

A number of quantitative and qualitative questions arise following this evaluation, conducted in several stages and from several stakeholder perspectives. Amongst others not outlined here, these questions are currently under investigation and include: (1) Where specifically do tendencies of between-genre variability potentially exist— is it the predominance of vibrato type, the half-extent of the vibrato, or the parameters of the transition phase (Delta Duration, Onset Time)?; (2) Where specifically in the vibrato segment phases do similarities across CV (%) most exist?; (3) What is the most perceptually and proprioceptively salient aspect in the transitions of complex, multiphasic vibrato?

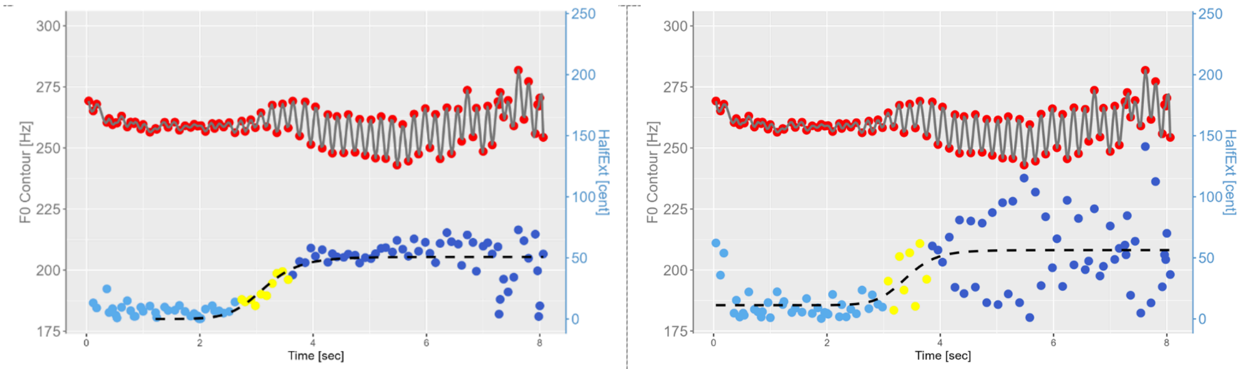
rHE Metric Analysis

As anticipated, the variability in vibrato samples calculated by using the rHE metric is higher than those calculated by the HoE metric, as rHE accounts for both the intrinsic variability of the vibrato extent and the pitch drift of the middle line. This effect is illustrated in one of Singer 3's vibrato sample in Fig. 3.15, where the observed variation in rHE values (right panel) is substantially broader compared to the HoE values (left panel). While the mean rHE estimates remain relatively stable, the CVs for rHE—– particularly in Phase 2 (the vibrato proper)—–are significantly higher. This increase in variability is attributed to the noticeable drift of the pitch middle line, which is especially pronounced at the onset of Phase 1 and persists throughout Phase 2.

For complex vibratos, the Phase 2 (vibrato proper) and all data half-extent estimates remain largely unchanged regardless of the metric, as indicated in the first columns of Table 3.5. In contrast, Phase 1 half-extent estimates exhibit greater variability, ranging from under 10% when using the HoE metric to approximately 8%–19% with the rHE metric.

Across all samples from the three singers, the estimated Phase 2 coefficient of variation (CV) increases from approximately 20% using the HoE measure to 29%–37% with the rHE measure (second columns of Tables 3.5 and 3.7). These results indicate that the contribution of Middle Line drift to vibrato variability is substantial, leading to a 50%–90% inflation of the CV.

A similar effect is observed in simple vibratos, where the half-extent estimates remain largely unchanged across both measures (first columns of Table 3.7), yet CV estimates are consistently



S3 (ID 13_8)	HoE [Cent]	CV [%]
All	31.7	75.0
Phase 1	5.03	73.1
Phase 2	50.1	31.7

S3 (ID 13_8)	rHE [Cent]	CV [%]
All	36.1	90.7
Phase 1	11.9	111.0
Phase 2	55.8	58.0

Figure 3.15: rHE (right panel) and HoE (left panel) values derived from the same f_o contour (red dots) in the same vibrato sample taken from a repertoire excerpt from S3 (ID13_8). (Orientation, lines, and dot colors are as described in Fig. 3.6).

higher with rHE. The parameters defining the Transition Phase are also consistent across the rHE and HoE measures (Table 3.9).

The key conclusion that vibrato complexity must be accounted for when evaluating vibrato variability holds regardless of the metric used. Notably, Phase 2 (vibrato proper) CV values are consistently lower than all data CV values across both rHE and HoE measures (see Figs. 3.13 and 3.14). Furthermore, the consistency of Phase 2 estimates—– both in mean half-extent and CV—– for complex vibratos aligns with the corresponding estimates for simple vibratos, regardless of the variability measure employed (Tables 3.4 and 3.5). The low slopes of linear regression fits further support the predominance of uniform vibrato in the simple vibrato class.

The classification of vibrato type is minimally affected by the choice of half-extent measure. This is attributed to the increased variation in rHE-derived half-extent values, which makes patterns in the f_o contour more difficult to discern. As a result, in some cases, AIC/BIC-based classification marginally shifted from HoE to rHE without a clear justification.

However, it is important to emphasize that vibrato classification— simple vs. complex— is determined by the presence of segments with notably different extents across the f_o contour, independent of the pitch middle line drift. Consequently, while rHE provides a more comprehensive characterization of vibrato variability by incorporating additional sources of variation, HoE remains a more reliable metric for vibrato type classification. This distinction highlights the importance of considering both metrics in vibrato analysis.

Justification in Comparing rHE vs. HoE

The above in-depth examination of half-extent and its HoE versus rHE calculations not only distinguishes and examines both metrics of half-extent but also delineates their respective properties, advantages, limitations, and practical applications. By systematically evaluating these aspects, these findings contribute to a more comprehensive understanding of the metrics and their implications in vibrato analysis.

Calculation of half-extent as HoE compared to rHE yields differing results for the CV variability. The definition of half-extent as HoE implies that its CV reflects the variability of the vibrato extents, but somewhat ignores the temporal changes (pitch middle line drift) of the actual pitch over the course of the vibrato. Hence, the CV of the HoE is assumed to solely describe the variability of the vibrato extent. On the other hand, rHE is more a comprehensive estimate of variability, as it includes a consideration of the pitch middle line, which may drift. Therefore, the HoE is perhaps better suited for vibrato time profile classification because the definitions of the vibrato archetypes do not rely on any pitch related assumptions. As a result, it can be recommended that the rHE vs. HoE usage are thus: if one seeks only classification of vibrato time profile archetypes, use HoE; if one is interested in more comprehensive quantitative characteristics of the variability, use rHE.

Furthermore, the difference between an rHE CV estimate and a HoE derived CV estimate of the same vibrato sample would reflect the impact of the actual pitch drift on vibrato variability. Importantly, the HoE metric does not rely on any assumption regarding the variability of the actual pitch middle line, as there is no guarantee that half-extents will be symmetrical around a mean value. Instead, this method bypasses the necessity of a uniform frequency basis in reference to vibrato, since the upper extent (peak to midline) may differ from the lower extent (trough to midline). The isolation of various sources of variability in frequency vibrato can be a very informative action allowing for adaptable variability calculations, as is the case with vibrato half-extent HoE or rHE metrics.

3.5 Overall Conclusions: Vibrato Variability Analysis

Traditional approaches that rely on averages when evaluating and reporting vibrato characteristics overlook the naturally-occurring variation over the full duration of the vibrato contour; the analysis of the latter can provide more meaningful insights. Furthermore, the utility of such an average-based approach is inherently limited, as it is applicable only to time invariant and regular, consistent, and persistent vibrato contours. When applied to vibrato samples that exhibit significant variation over time, averaging methods may yield misleading outcomes.

The major determinants of vibrato variability in the frequency domain include:

1. The variability of vibrato extent.

2. The variability of vibrato rate.

3. The temporal drift of the middle line (pitch or tonal center) over the course of the vibrato.

The presence of spurious perturbations in the signal, which may arise from technological factors or data processing artifacts, may also impact the vibrato variability estimates.

In this thesis chapter, a systematic evaluation of vibrato extent variability in recorded singing voice samples was undertaken, utilizing half-extent data. The CV was selected as a straightforward, accessible, and understandable measure of variability. Two half-extent metrics, previously applied in prior literature but not fully characterized, were employed and defined:

- *Half-of-Extent (HoE)*: Defined as half of the full vibrato extent at each peak-trough pair, centering values around an actual pitch middle line. As this definition disregards middle line drift, HoE accounts only for the variability of vibrato extent.

- *Relative Half-Extent (rHE)*: Defined as the distance between each individual peak and each individual trough of the vibrato's f_o contour and an intended pitch middle line, calculated as the average of all peaks and troughs. This definition incorporates both vibrato extent variability and variability arising from pitch middle line drift.

The vibrato variability analysis was conducted on samples from three professionally trained singers, representing Jazz, Musical Theater, and Opera, each performing 16 identical excerpts. The findings revealed that the majority of f_o vibrato contours exhibited complex half-extent time profiles, consisting of segments with notably different half-extents. This observation underscores the importance of vibrato complexity as a crucial characteristic of vibrato variability. Furthermore, applying variability metrics directly to complex vibrato contours introduces a significant upward bias by conflating different vibrato segments (phases). These insights highlight the necessity of segmenting vibrato phases before applying and interpreting vibrato variability measures, including CV.

As a result of this systematic and comprehensive analysis of vibrato extent variability, several key conclusions were drawn, and the following methodological advancements were proposed:

- Vibrato f_o contours can be classified into simple and complex types based on their temporal extent patterns. Proposed classifications include uniform, increasing/decreasing, and S-shaped complex vibratos.

- A novel classification method for vibrato f_o contours was introduced, utilizing linear and logistic regression models to distinguish between vibrato types. Model selection was based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), combined with qualitative visual inspection of the f_o time profiles. While these statistical criteria demonstrated strong sensitivity and power for vibrato type differentiation, expert discretion remains necessary, particularly for short vibratos or datasets with outliers.

- The proposed method allows for segmentation of vibrato phases, thereby mitigating the risk of overestimating vibrato variability when analyzing complex vibrato contours.

- Variability analysis confirmed that the CV of Phase 2 (sustained vibrato) is significantly lower than the CV calculated across the entire vibrato sample. Additionally, the CV values for Phase 2 closely matched those of uniform, simple vibratos in the dataset, further validating the necessity of vibrato segmentation.

- As expected, the CV derived from rHE was consistently higher than that from HoE, as rHE accounts for both extent variability and Middle Line drift, whereas HoE reflects only extent variability. Nevertheless, conclusions derived from rHE results were equally applicable to HoE, and vice versa.

- Simultaneous evaluation of both half-extent metrics enabled differentiation of the contributing factors to vibrato variability. Specifically, the pitch middle line drift was shown to increase vibrato extent variability by 50%–90%.

- The two half-extent metrics were further distinguished:

  – rHE provides a more comprehensive measure of vibrato variability by incorporating multiple contributing factors.

  – HoE demonstrates greater sensitivity for vibrato classification, as vibrato types are defined based on extent differentiation rather than the pitch middle line fluctuations.

- Consistent trends emerged in variability analysis, with CV ranges beginning to take shape for trained, professional Western classical singers:

  – CV (HoE): 20%–25%

  – CV (rHE): 29%–38%

  These ranges, when validated with larger sample sizes, may serve as normative references for this singer population.

The findings from this extensive evaluation advance the understanding and characterization of vibrato variability. The proposed methodologies provide new avenues for vibrato research, with potential applications in pedagogy, performance analysis, and perceptual studies. Future work should focus on expanding the dataset across diverse voice types and musical genres, refining classification models, and integrating perceptual analyses to assess how vibrato variability impacts listener perception and artistic interpretation.

3.6 Future Work

Vibrato is inherently tied to artistic and stylistic effects and expressive intentions in singing. This work serves to resolve the existing reductive limitation in analyzing all vibrato, regardless of genre, based on steady-state averages derived from a consistent and persistent Western classically-oriented scale. Expanding and diversifying the vibrato "canon" in our scientific and artistic understanding would allow for a description of singing variation across styles in a meaningful manner. This may bring about valuable teaching applications for more genre-specific vibrato training and rehabilitation for singers in the voice clinic by contributing more ecologically valid and genre-inclusive normative thresholds. Vibrato as a diagnostic by-product of the vocal mechanism and its interaction with voice function and technique is the subject of current research based on this original thesis study. This ongoing vocology work beyond the thesis ultimately aims to advance biomechanical understanding and auditory-perceptual-acoustic analysis of vibrato variability through a more comprehensive, multidimensional therapeutic and habilitative lens.

In working toward a comprehensive system of vibrato characteristics over time beyond averages of rate and extent, a continued aim is a corpus analysis of vibrato patterns that emerge as characteristic of genre. It would also be imperative to include genre contexts outside the Western singing traditions in this vibrato variability corpus analysis.

This research focused on frequency excursions in vibrato by genre over time, but a current investigation is the examination of interactions between the amplitude and frequency excursions in vibrato by genre over time. From this and prior studies, it is hypothesized that frequency excursions (of the half-extent in cents) may correspond with intensity or sound pressure level excursions (in dB), especially in *messa di voce* (*crescendo-decrescendo*) tasks sung by singers of certain genres. This warrants further research across contexts with a larger data set. In addition, future planned research involves the applications of Vibrato Onset Time. Furthermore, current work is being done to refine understanding of production physiology of vibrato and its related phenomena (such as tremor) in the body. Finally, it is important to note that while these models were developed on vocal vibrato, they are well-positioned to be transferable to instrumental vibrato. This research was conducted with that idea in mind and not intended to be isolated to solely vocal or singing vibrato.

Creating more accurate and genre-inclusive parameters and models of vibrato characteristics contributes to the future aim of describing vocal variation in a meaningful manner. While this study focused on analyzing variations in extent, there is certainly room for future work that also examines ways to characterize time-varying vibrato rate. Eventually, vocal vibrato could be conceptualized on a gradation scale or continuum as index of variability. Discovering novel ways to quantify and represent vibrato variation serves to close the existing gap and reductive limitation in analyzing all vibrato, regardless of genre, based on steady-state averages derived from a consistent and persistent Western classically-oriented scale. This may raise valuable teaching implications for more genre-specific vibrato training in the voice studio. Special attention should be placed on practical applications of this methodology, as the ultimate aim is to increase the understanding of health, function, production, and perception of vibrato with numerous potential benefits to enhance voice therapy and training.

Chapter 4

Timbre Perception of Temporally-Variant Vibrato Across Styles

Author Contributions

This chapter includes restructured and reintegrated content from the following research article: Nestorova, T., Reymore, L., Marchand Knight, J., Soden, K. (2025). Speaking of vocal vibrato: Timbre perceptions and semantic descriptions with pedagogical and clinical implications. *Voice and Speech Review*. [Manuscript accepted; expected publication in late 2025].

Author-specific contributions: *Theodora Nestorova:* conceptualization of the study, participant recruitment and interviewing, post-processed semantic analysis, interpretation of results, development of interpretation and conclusions, writing, reviewing, and editing of the manuscript. *Lindsey Reymore:* participant recruitment and interviewing, data acquisition and pre-processing, post-processed semantic analysis, coding of preliminary script, reviewing and editing of the manuscript draft. *J. Marchand Knight*: participant recruitment and interviewing, data post-processing analysis, reviewing and editing of the manuscript draft. *Kit Soden*: participant recruitment and interviewing, data post-processing analysis, reviewing and editing of the manuscript draft. *Bennett Smith:* creation of custom computer interface for study. *Stephen McAdams:* supervision of the overall study project.

Chapter-specific contributions and differentiation from original journal publication: In this chapter, all figures have been updated in a different color scheme than the submitted manuscript, as have edits for grammar and flow been conducted to optimally incorporate the article contents into this thesis chapter.

4.1 Background: Secondary, Extended Analysis

Perception of vibrato is inherently tied to both its acoustic properties and the cultural and experiential frameworks of the listener. As a perceptual phenomenon, it is crucial that vibrato analysis begins from its ecologically valid rendering; the lived experience of how it sounds. Vibrato interacts with the perception of vocal timbre, frequently referred to as the "color" or "quality" of a voice. Timbre is a multidimensional auditory attribute that allows individuals to differentiate between sound sources, even when those sources produce identical pitch and loudness (McAdams, 1989). Vibrato interacts with other acoustic properties to create complex auditory cues that listeners use to assess timbral aspects that often evoke particular emotional or stylistic connotations. These cues, highly affected by vibrato, are expressed through specific linguistic descriptions that play significant roles in the discussion and training of the voice. Investigating how vibrato affects timbre perception is crucial for the understanding of vibrato pedagogy and performance, as well as the diagnostic discernment of the voice. Insights into cognitive processes underlying voice perception will provide an ecologically valid basis to the development of analytical and pedagogical tools for vibrato skill building.

Prior perceptual studies done on vibrato indicate the power of vibrato's simultaneous features to influence the perception of pitch and intonation quality (Daffern et al., 2012; Reddy and Subramanian, 2015; Geringer et al., 2010, 2015; Almeida et al., 2021; Loni and Subbaraman, 2019; Duvvuru, 2012). Research indicates that within specific boundaries of vibrato extent, vibrato plays a crucial role in determining the perceived pitch of a tone. Notably, as the vibrato extent increases, the perceived pitch deviates further from the mean frequency (Daffern et al., 2012; Reddy and Subramanian, 2015). Additionally, vibrato rate and extent interact to define the threshold of perception of a vibrato versus non-vibrato tone. According to Wooding and Nix (2016), the perception of a tone without vibrato is influenced more by changes in vibrato rate than extent, although both factors contribute to distinguishing between vibrato and non-vibrato tones (Glasner and Nix, 2022). Further work by Wooding and Nix (2016) and Glasner and Nix (2023) revealed that there are certain limits of vibrato rate and extent that operate within an interdependent threshold yet jointly impact voice professionals' judgments of non-vibrato tones.

Furthermore, Wooding and Nix (2016) found that both vibrato rate and vibrato extent work together to affect the threshold of perception of vibrato versus non-vibrato tone. Most notably, they discovered that a singer can sing with a wider vibrato extent and still be perceived as singing non-vibrato if their vibrato rate is very slow. In their study, Wooding and Nix (2016) established the threshold below which most listeners rated a tone as non-vibrato was 27 cents peak-to-peak. Musical training level may also have an impact regarding perception of pitch as connected to vibrato; results from an earlier study demonstrated that vibrato does negatively affect pitch matching of uncertain singers, yet the presence of vibrato did not create an immediately perceptible out-of-tune model across singers (especially certain ones) (Yarbrough et al., 1992).

Vibrato characteristics also affect expert listeners' perception of pitch. For example, smaller vibrato extent has been linked to more accurate pitch assignment by experts when comparing operatic singers and early music performers (Daffern et al., 2012). More broadly, vibrato appears to assist in pitch perception, with the fundamental frequency of a vibrato note perceived as the geometric mean of its oscillating frequencies (Shonle and Horan, 1980). Furthermore, vibrato extent and rate influence pitch perception, with the interaction between the two contributing to the perception of pitch separation or fusion (d'Alessandro and Castellengo, 1991).

Research suggests that the ability to perceive vibrato characteristics may depend on proportional changes between vibrato rate and extent, as well as the temporal variance of those changes. Another crucial perceptual element distinguishing unintended irregular pitch production from intended variable vibrato is the regularity of the periodic cycle and its frequency excursion (extent). Sundberg (1994a) posited that regularity is considered a sign of the singer's vocal skill:"the more skilled the singer, the more regular the undulations." Moreover, in well-trained singers the regularity is generally not influenced appreciably even if the singer's sound is masked by noise (Schultz-Coulon, 1978).

In addition to perceptual thresholds, some studies have explored preferences among professional musicians. Choral singers, for example, tend to prefer narrower vibrato extents (Nix, 2013). Likewise, the use of listeners of varied musical backgrounds in vibrato perception studies often yields different results. For example, singer-judges (experts) demonstrate greater sensitivity to vibrato characteristics than non-expert judges (Reddy and Subramanian, 2015), yet the positive correlation between vibrato measurements and their detection by expert judges remains moderate (Howes et al., 2004). It has been thus recommended that the choice of listener type group should align with the study's objectives (Glasner and Nix, 2023). Studies with pedagogical or clinical applications may benefit from including trained listeners or relying on expert evaluations to ensure the findings are applicable in practical contexts. This provided significant impetus for this secondary, extended analysis study, whose goals were to explore timbre semantics related to vibrato in a dual demographic of trained musicians as expert listeners.

While significant progress has been made in understanding instrumental timbre (Reymore and Huron, 2020; Reymore et al., 2023; Zacharakis et al., 2015; Saitis et al., 2012, 2014; Traube, 2004; Hérold, 2012), the distinct and variable nature of vocal timbre— shaped by anatomical, stylistic, and cultural factors— demands a specialized and comprehensive investigation. The exploration of how listeners perceive and describe timbre in the singing voice remains an undervalued but critical area of inquiry within the broader field of music perception research. Moreover, there is a dearth of studies examining the difference in music perception and cognition, particularly focused on timbre semantics and their implications for pedagogical training. This research aims to fill this gap.

This thesis chapter presents a pilot study as a secondary, extended analysis of interview data originally collected in Reymore et al. (2025), focusing on listener perception of vibrato, a topic that was not examined in the original publication. Since this secondary, extended analysis is encapsulated within the overarching research project, the aim was to investigate noticeable differences in vocabulary and descriptive strategies between trained vocalists and trained instrumentalists on timbre perceptions of the singing voice. Convergences and divergences in how these listeners perceived and semantically expressed perceptions of vibrato were considered, shaped by their musical backgrounds as vocalists or instrumentalists. The potential for vibrato's concurrent features to affect timbre perception were a particular focus and implications for voice training and rehabilitation were examined. The underlying aim of this study is to contribute to a deeper understanding of cognitive sound processing of the voice, with particular emphasis on vibrato as a defining and dynamic element of the singing voice.

The main research questions underpinning the purpose of this pilot, multi-strategy investigation are as follows:

1. **(RQ1) Perceptual Salience of Vibrato:** In open-ended interviews where participants describe singers' voices in detail, how prominently is vibrato mentioned? Specifically, how frequently does vibrato emerge as a topic, and what types of remarks are made? This question involves analyzing vibrato-specific comments and comparing descriptive strategies among listeners.

2. **(RQ2) Differences between Vocalists vs. Instrumentalists**: Do trained vocalists and instrumentalists differ in their lexical vocabulary and frequency of vibrato-related comments? If differences exist, what might account for them? This question explores whether the unique characteristics of vocal timbre, potentially linked to specialized neural processing, influence perceptions among these groups.

3. **(RQ3) Genre-Specific Characterization:** Can listeners characterize singing within different genres using specific descriptors? Are there notable differences between genres in the semantic characterization of vocal timbre and vibrato? This question seeks to determine if and how genre influences the description and perception of vocal timbre and vibrato.

4. **(RQ4) Association Between Acoustic Variability and Perceptual Comments:** Is there any correlation between quantitative measures of vibrato variability (e.g., vibrato half-extent variability) and the frequency or nature of vibrato-related comments? What are the implications of such connections? This question aims to link acoustic measurements with

perceptual data to better understand how time-varying vibrato's acoustic properties influence listener perceptions.

In this thesis chapter, results pertaining to vibrato will be revealed and interpreted.

4.2 Methods

Recordings used as stimuli came from the voice study described in Sec. 2.7. Due to the extenuating circumstances in the recording conditions detailed in that section, all recordings, though resulting in high-quality audio signals, were furthermore post-processed with normalization of acoustic levels, conducted by a professional recording engineer. This ensured that the intensity levels remained uniform across all recorded stimuli. Three recordings were subsequently eliminated from the data set because they could not be sufficiently normalized to match the general sound characteristics of the other recordings. Therefore, there were 12 total singers sourcing the stimuli for this pilot study, with an equal number of 4 in each genre group (Opera, Musical Theater, and Jazz).

4.2.1 Listener Participants

This perceptual study was approved by the McGill University Research Ethics Board (REB# 201-1114), and participants were paid $10 CAD for taking part in the listening and interview study. The recruitment target was an approximately equal number of participants from each listener type group, vocalists and instrumentalists. Determining the appropriate sample size for interview-based studies is highly context-dependent (Baker et al., 2013). According to Morse (2000), considerations such as study scope, topic complexity, data quality, study design, and the use of shadowed data (not applicable here) should inform sample size decisions. Many researchers emphasize that thematic saturation is key for deciding sample size, with Green and Thorogood (2009) suggesting that for a focused research question, new insights typically plateau after 20 interviews per analytically relevant group.

Factoring in these considerations, participant numbers were balanced between two experiments, concluding that fewer listeners would be needed in this secondary, extended analysis focused on vibrato due to its narrower scope and controlled design. Consequently, the study included 30 participants (n = 30), split evenly between the listener type groups, including 15 vocalists and 15 instrumentalists.

Eligibility criteria for listener participants required vocalists to have at least two years of post-secondary formal vocal training, with no more than two years of post-secondary or five total years of instrumental training (excluding required keyboard skills classes). Instrumentalists needed equivalent formal training on their instrument, with the same limits for vocal training. Formal training was defined as consistent participation in private lessons and/or large ensemble rehearsals (e.g., choir, band, orchestra). The participants were recruited from McGill University's Schulich School of Music (19 female, 8 male, 1 non-binary, 1 unreported), and had an average age of 26.13 years ($SD = 5.54$, range 20–42). Of those, 76% were native English-speakers and 24% identified their primary languages as either French, Mandarin, Spanish, or Persian/Farsi. A visualization of the listener participant demographic charts is included in Fig. 4.1.

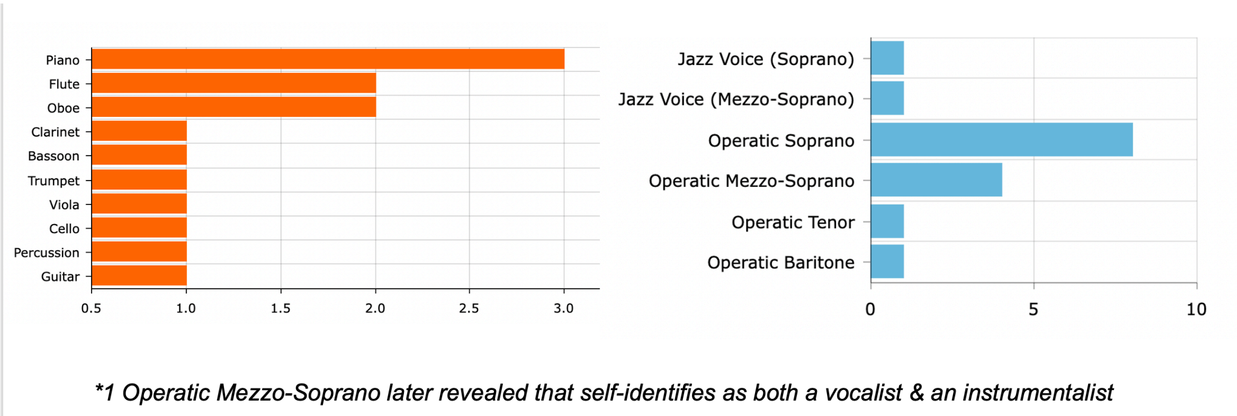


Figure 4.1: Side-by-side charts displaying listener participant demographics for each type group. On the left, the breakdown of instrumentalist listeners by primary instrument played. On the right, the breakdown of specific voice types within the vocalist listeners.

4.2.2 Procedure

Interviews were conducted in person at McGill University's Schulich School of Music in the Music Perception and Cognition Lab. Using over-ear, noise-canceling headphones connected to a custom computer interface designed by Bennett Smith, participants listened to 12 different unaccompanied singer excerpts from the Vibrato Corpus in a random order. Excerpts represented four individual singers for each of three genres (Opera, Musical Theater, and Jazz).

For a given singer, listeners heard two excerpts: one of two possible phrases from "Summertime" and one of two possible phrases from "Amazing Grace," each approximately 15 seconds long (see Table 4.1). The complete set of possible stimuli included a total of 48 musical phrases; each participant heard a subset of 24 of these phrases. Excerpts were randomized so that participants heard either the first half or second half of the longer musical phrase for each song, but randomization was organized across all listeners so that all four subphrases were heard equally often over the course of the interviews. To see the notated musical excerpts, see Figs. 4.2 and 4.3.

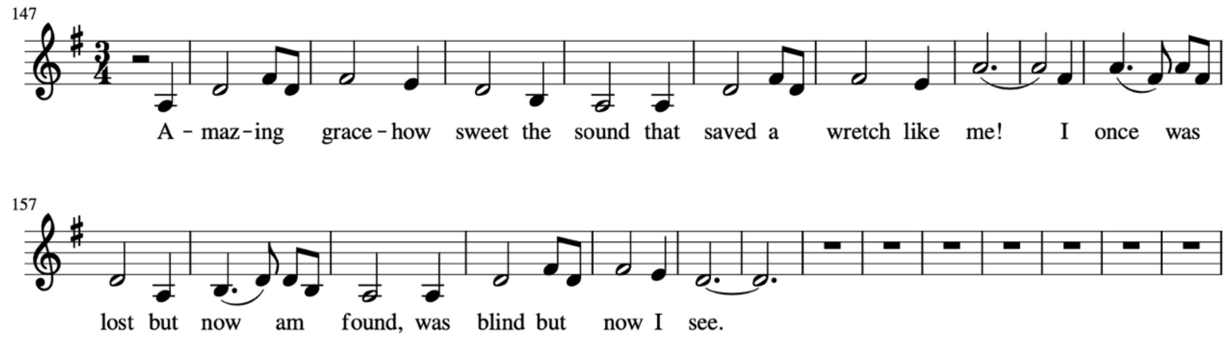


Figure 4.2: Notated musical excerpt of sung recording stimulus "Amazing Grace," including phrases used as a subset for the perceptual study.

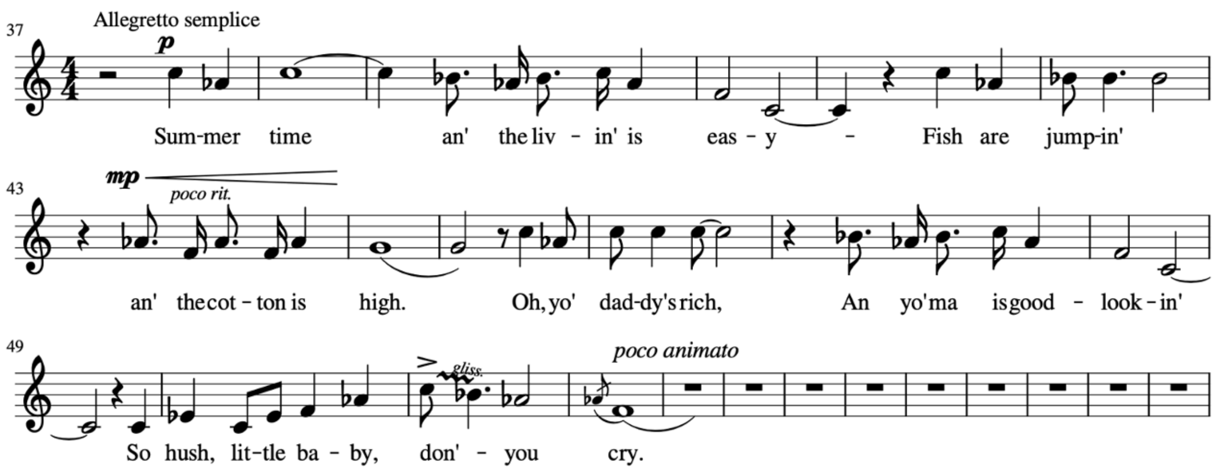


Figure 4.3: Notated musical excerpt of sung recording stimulus "Summertime," including phrases used as a subset for the perceptual study.

Amazing Grace	**Summertime**
Amazing grace, how sweet the sound / that saved a wretch like me	Summertime, and the livin' is easy / fish are jumpin', and the cotton is high
I once was lost, but now am found / was blind but now can see	Your daddy's rich, and your ma is good lookin' / so hush, little baby, don't you cry

Table 4.1: Phrases from the Stimuli Recordings (Singer Excerpts) in the Vibrato Corpus used to source the perceptual study.

Listeners were asked to describe the timbre of each singer's voice in as much detail as possible, using adjectives and descriptive words to express their timbral experiences of each heard excerpt. Prompting was not provided by the interviewer, nor was there any mention specific to vibrato in the instructions. Listeners could replay excerpts as many times as requested (all listeners were played each excerpt at least twice) but were encouraged to provide initial observations immediately after the first listening to capture their most genuine reaction. Following the second listening of the pair of phrases, listeners were encouraged to continue their description by the interviewer. The interviewers limited their verbal responses to the listeners to clarifying questions, if requested by the listeners. The listeners only removed their headphones while conversing with the interviewer. Listeners then completed a brief demographic survey, including information about musical background. Interviews typically took around 60 minutes, with an approximate range of 45–75 minutes.

4.2.3 Multi-Strategy Approach to Mixed Methods Analysis

Interview recordings were transcribed both manually and automatically, using Microsoft Office 365's transcribe feature on Word OneDrive, converting speech to a text transcript with each speaker individually separated and reviewed for accuracy. Observations were cataloged in several separate offline spreadsheets, with individual concepts extracted and categorized into phrases for further analysis.

Concept extraction was conducted in multiple steps. Initially, the $voice - description$ phase distilled observations by removing non-essential words but retaining the full meaning of phrases, such as refining "a little bit airy towards the end of phrases, I feel" to "airy towards the end of phrases." Subsequently, the descriptors phase further condensed this into key terms, e.g., "airy." This phase retained only unique terms and omitted repetition within a single participant's description of one singer. These steps ensured the data set's clarity and reproducibility while documenting every decision to provide a transparent reference for future researchers.

To standardize terms further, a $base - term$ phase was created manually through consensus of the full research team after automated stemming and lemmatization in R Studio (RStudio Team, 2023) proved unreliable. Similar terms were consolidated (e.g., "breath" and "breathy") but not conceptually merged if distinct meanings were implied (e.g., "forced" and "forceful"). The data set was cleaned by removing typos, negations, and extraneous characters.

For the purposes of this secondary, extended analysis, all non-adjectival descriptions of vibrato, such as mentions of its presence, amount, or changes, as well as all adjectival descriptors relating to vibrato like "wide" were retained and included by the author of this thesis. The finalized data set for this study comprised 4,088 observations, including 1,497 unique terms, and was analyzed for frequency and thematic patterns using R Studio (RStudio Team, 2023) and NVivo (Lumivero, 2025).

4.3 Results

Listeners employed diverse strategies to describe vocal timbre, consistent with prior research on timbre perception (Reymore and Huron, 2020; Porcello, 2004). These strategies included adjectives, metaphors, similes, comparisons to other timbres, and speculative narratives about singers' attributes, such as their personality or physical appearance. Common approaches included cross-modal descriptions, references to specific parts of excerpts, and comparisons across singers or genres, with frequent mentions of style and inferred characteristics like confidence or experience. The most frequently used terms were tabulated, then run through several iterations of a vocal timbre semantics model. This model was constructed by combining the pile sort method from Reymore and Huron (2020), the instrumental timbre semantics model, with elements of Grounded Theory used by Saitis and Weinzierl (2019). Term frequency-inverse document frequency (TF-IDF) was applied as a measure of word relevancy, analyzing descriptive terms across singers' genres in the stimuli, comparing language used by instrumentalist and vocalist participant, identifying notable terms and gender-based differences. Following this, genre timbre trait profiles were created.

4.3.1 Definition of Variables

For this study, the most pertinent results connected to vibrato, in addition to preliminary descriptive statistics and selective thematic content analysis, were extracted. For the purposes of this investigation, the comparison of semantic descriptors with the acoustic metrics and parametric models of vibrato variability outlined in Chapter 3 allowed for multidimensional analysis. This enabled a more comprehensive exploration of objective acoustic differences in the stimuli recordings and their relationship to subjective listener perceptions. Furthermore, this research's ultimate goal of developing measurable acoustic tools may be inspired and directly aided by the perceptual features that emerge from this secondary, extended analysis.

4.3.2 Quantitative Results: Content Analysis of Vibrato Mentions

The overall quantitative results of the descriptive statistics are clear and consistent, with no significant difference between the groups, a small effect size, and overlapping confidence intervals. From all remarks in the perceptual study's data set comprising of 4,088 total observations, including 1,497 unique terms, a statistical analysis table in Python revealed 240 total mentions of vibrato.

Vibrato seemed to be a critical feature in timbral descriptors regardless of group. The dispersion of vibrato mentions proved ubiquitous with high prevalence; in fact, vibrato was a predominantly

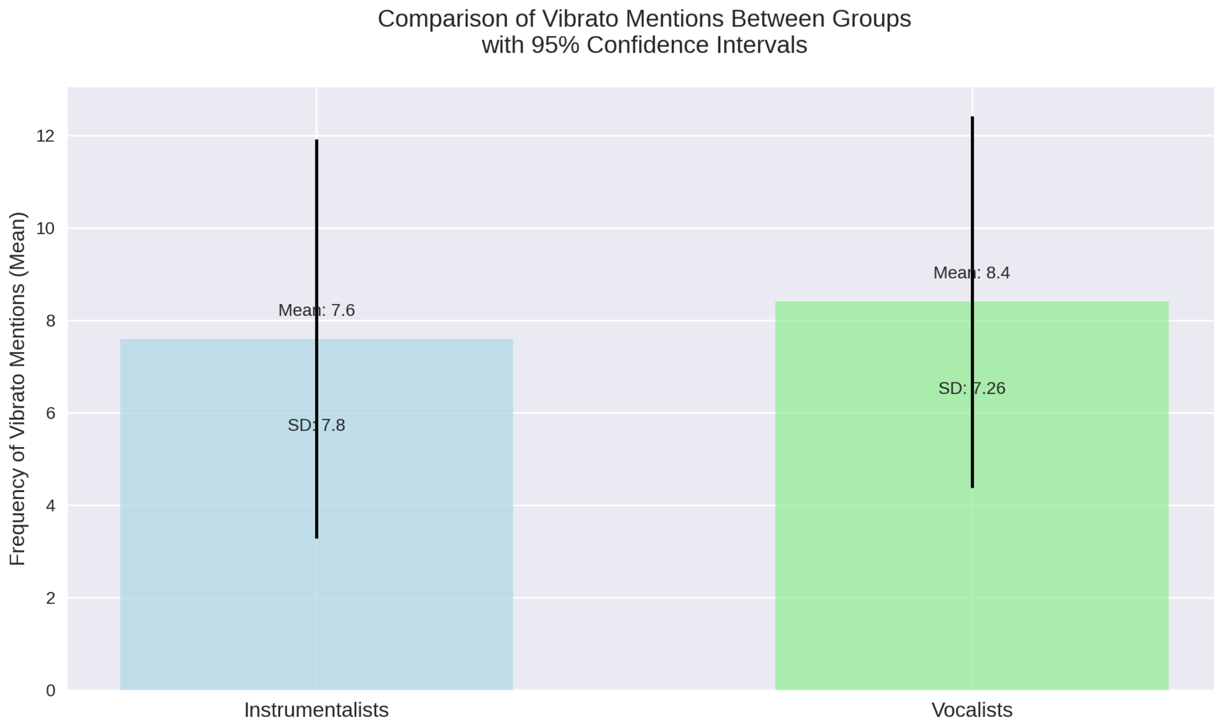


Figure 4.4: Comparison of vibrato mentions between instrumentalists (*n* = 114) and vocalists (*n* = 126) in a bar graph showing mean frequencies and 95% confidence intervals.

mentioned characteristic across groups. From all listeners, 94% (all except 2) mentioned vibrato at least once, with approximately 30% of listeners mentioning vibrato 10+ times. Some listeners exceeded 20 vibrato mentions. These results clearly demonstrate the strong role that vibrato plays and its strikingly pervasive perceptual relevancy.

Of the 240 total vibrato instances, instrumentalists had 114 total mentions ($mean = 7.6, SD = 7.8$). Vocalists had 126 total mentions ($mean = 8.4, SD = 7.26$). As dual visualizations of this, see Figs. 4.4 and 4.5. Instrumentalists and vocalists showed no statistically significant difference in their frequency of mentions, as indicated by a t-test ($t(28) = 0.291, p = 0.773$). The 95% confidence intervals for the means overlapped substantially, with instrumentalists ranging from 3.28 to 11.92 and vocalists from 4.38 to 12.42. The effect size was small ($Cohen's - D = 0.106$), suggesting minimal practical difference between the groups, if one exists. These results indicate that both groups discussed vibrato with similar frequency, despite individual variability.

When stratifying results by singer (from their recorded stimuli excerpts), listeners mentioned vibrato in their timbral descriptors as one of the top 10 remarks in 91.6% of the recorded singer excerpts (all except 1, in which vibrato was in the top 20 remarks). Interestingly, in both Musical Theater and Opera singer excerpts, there appears to be a notable spread of vibrato mentions across individual singer excerpts, with some showing significantly higher vibrato mentions (Singers 7

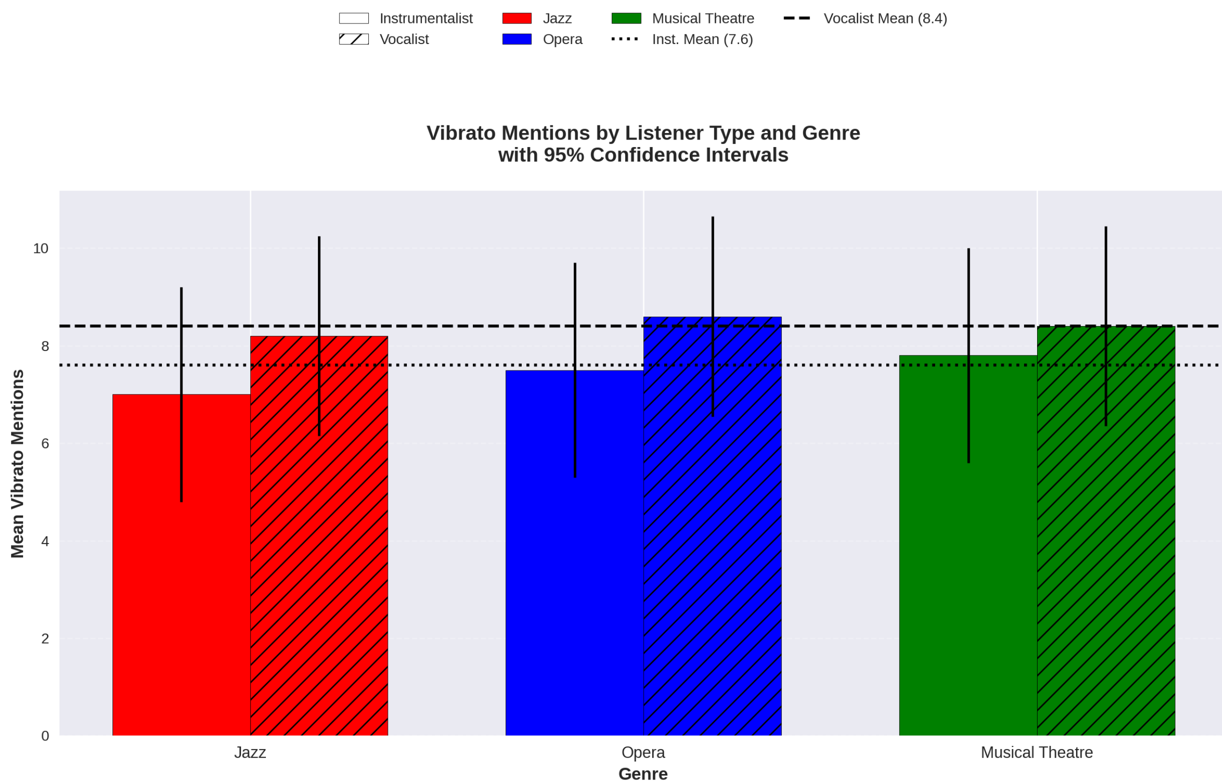


Figure 4.5: The same comparison of vibrato mentions, stratified by genre, between instrumentalists and vocalists. Corresponding colors and gradients outlined in the legend above the the histogram.

and 9 with 22 mentions each) compared to others with a significantly lesser number of vibrato mentions. This suggests that the perceptual impact of vibrato varies across different singers, different excerpts, and different phrases that listeners heard in the perceptual study. See Table 4.2 for a full list of the number of vibrato mentions extracted from the listeners' timbral descriptors, organized by singer and genre.

When stratifying results by genre, Opera shows the highest average number of vibrato mentions, followed by Musical Theater and then Jazz, as reported by all listener participants. To further characterize the data distribution, violin plots were generated and overlaid with individual data points and red lines representing the group mean and its 95% confidence interval (CI). For instance, for the Jazz category the mean number of vibrato mentions was $M = 10.75$ (95% CI [11.44, 20.06]. Similar analyses for Musical Theater yielded $M = 11.25$ (95% CI [-1.17, 23.67] and for Opera $M = 13.00$ (95% CI [1.31, 24.69]). The wide confidence intervals indicate considerable within-group variability relative to the group means. Figure 4.6 illustrates these violin plot distributions alongside the confidence intervals, providing a detailed view of the spread of the data and uncertainty around each estimate.

SINGER #	GENRE	*N* of VIBRATO MENTIONS
1	Jazz	18
2	Jazz	13
3	Jazz	6
4	Jazz	6
5	Musical Theater	5
6	Musical Theater	6
7	Musical Theater	22
8	Musical Theater	12
9	Opera	22
10	Opera	13
11	Opera	13
12	Opera	4

Table 4.2: Number of Vibrato Mentions in Timbral Descriptions per Singer and Genre.

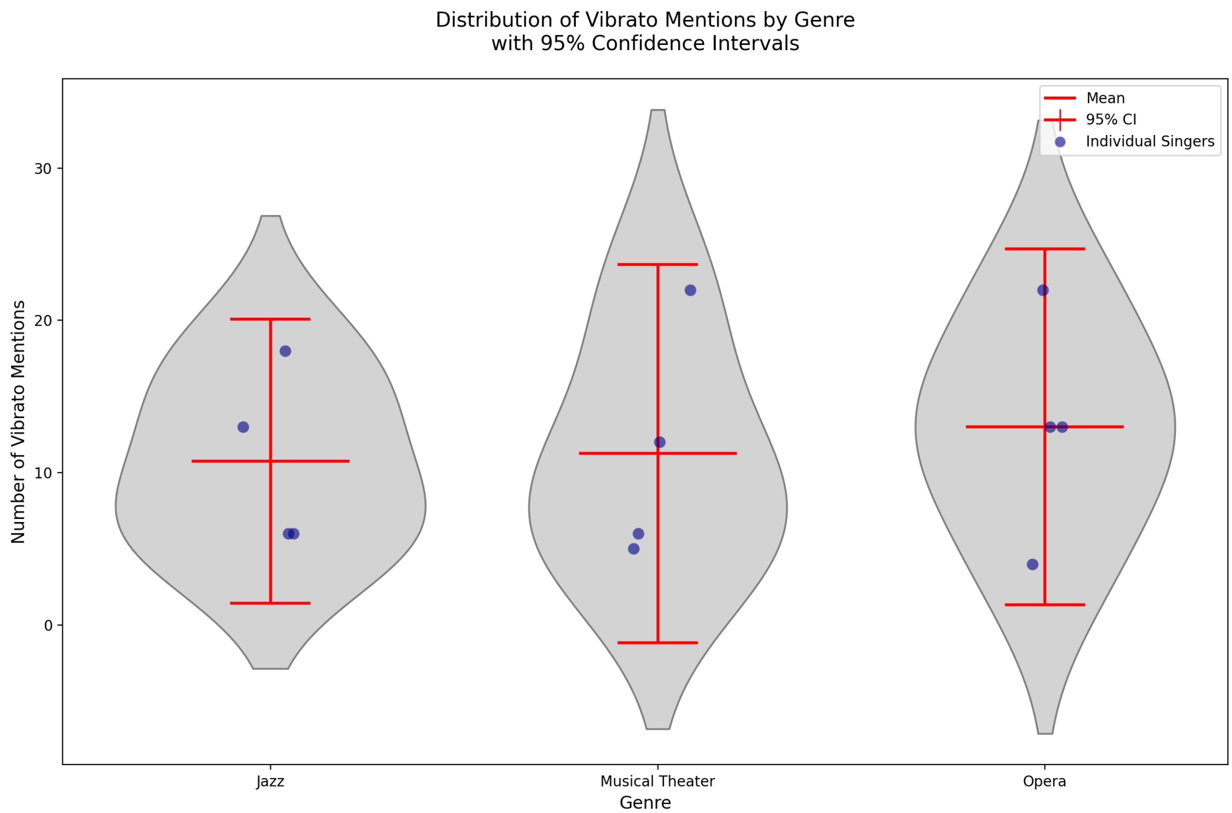


Figure 4.6: Violin plot showing the distribution of vibrato mentions across three musical genres: Jazz, Musical Theater, and Opera. Individual data points (blue dots) represent the number of mentions for each singer, while the red error bars represent the estimated mean and confidence interval ranges, illustrating variability within each genre. Wider sections of the violin plot indicate a higher density of data points, highlighting differences in vibrato usage patterns across genres.

A generalized linear mixed model (GLMM) with random intercepts for singers examined genre effects on vibrato mentions, with Jazz as the reference category. The likelihood ratio test ($\chi^2 = 0.000, p = 1.000$) suggested no significant random intercept variance. However, estimated variance components (random and residual intercept: 18.646) yielded an intraclass correlation coefficient of 0.500, indicating that 50% of variance in vibrato mentions was attributable to between-singer differences. Fixed effects revealed predicted means of 10.75 for Jazz ($SE = 3.05$) ($p < 0.001$), with non-significant increases for Musical Theater ($+0.50$), ($SE = 3.74, p = 0.894$) and Opera ($+2.25$) ($SE = 3.74, p = 0.547$). The Omnibus F-Test for genre effects was non-significant ($p = 1.000$), and pairwise comparisons showed no significant differences between genres (all $p > 0.05$). These findings suggest that singer-specific variability, rather than genre, primarily influences vibrato mentions. Despite Opera showing the highest predicted mean, the small sample size (*n* = 12 singers; 4 per genre) limited statistical power to detect significant genre random and fixed effects. The substantial overlap in the distributions of vibrato mentions across genres, as well as the non-significant omnibus test and fixed effects, indicate that genre differences are not a strong predictor of vibrato perception in this dataset.

These results suggest that when accounting for singer-specific variability, musical genre does not significantly predict the number of vibrato mentions. The substantial ICC value of 0.500 indicates that singer-level variation plays a more important role in vibrato mentions than genre differences. Despite the visible differences in mean vibrato mentions across genres (with Opera showing the highest predicted mean), none of the pairwise comparisons reached statistical significance. For a visualization of this, see Fig. 4.7.

Instead of implementing a chi-squared goodness-of-fit test, which may confound variation across singers with genre, the distribution of vibrato mentions across the twelve singers was analyzed using a more refined generalized linear mixed model (GLMM), fitted to examine the relationship between genre and vibrato mentions, with singer included as a random intercept to account for potential individual-level variation. The significance of the random effect was assessed using a likelihood ratio test (LRT) comparing the full mixed-effects model to a reduced model without the random intercept.

Outlined above, the GLMM was estimated to assess whether musical genre influenced the number of vibrato mentions while accounting for variability across singers by including a random intercept for each singer. However, the model exhibited convergence issues, as evidenced by the LRT test statistic of exactly 0.000 ($p = 1.000$) and identical random intercept and residual variances (both 18.646). These patterns strongly suggest computational problems in model fitting rather than substantive findings.

Given these convergence issues, the reported intraclass correlation coefficient (ICC) of 0.500 and the variance component estimates should be interpreted with extreme caution. The conver-

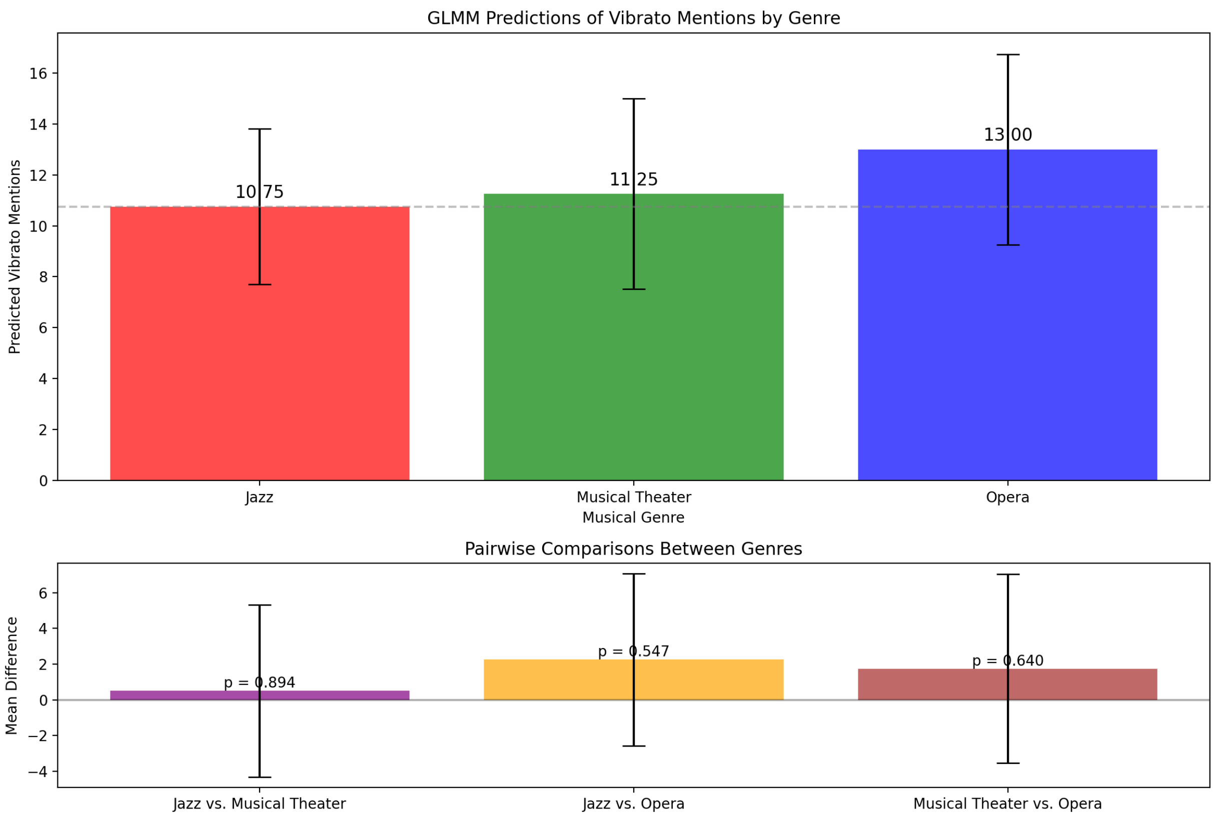


Figure 4.7: GLMM predictions of Vibrato Mentions by Genre with Pairwise Comparisons between Genres, consistent with the statistical estimates detailed above.

gence problems likely stem from the small sample size (n = 12 singers, with only 4 per genre) of this pilot study, which provides insufficient information for reliable and generalizeable estimation of both fixed and random effects.

Alternative modeling approaches, such as using different optimization algorithms, specifying different covariance structures, or considering Bayesian estimation with weakly informative priors, might help address these convergence issues. Additionally, the fixed effects estimates for genre differences should be interpreted cautiously given the model fitting problems.

The fixed effects portion of the model showed that, relative to Jazz singers ($intercept = 10.750$ ($SE = 3.053, p < 0.001$), Musical Theater singers exhibited a non-significant increase of 0.500 mentions ($SE = 3.740, p = 0.894$; 95% CI [-6.829, 7.829], while Opera singers showed a non-significant increase of 2.250 mentions ($SE = 3.740, p = 0.547$; 95% CI [-5.079 to 9.579]).

It is crucial to note that these results should be interpreted with considerable caution due to the study's limited sample size (n = 12 singers). The non-significant LRT result should not necessarily be interpreted as evidence for the absence of singer-level variation, as the small sample size substantially limits the statistical power to detect random effects. This limitation is particularly relevant given the meaningful ICC value of 0.500, which suggests practically important singer-level variation despite the lack of statistical significance. With singer included as a random intercept accounting for potential individual-level variation, there is not enough robust evidence to claim that there is meaningful variation between singers in the baseline rate of vibrato mentions. For the visualization of these results, see Fig. 4.8.

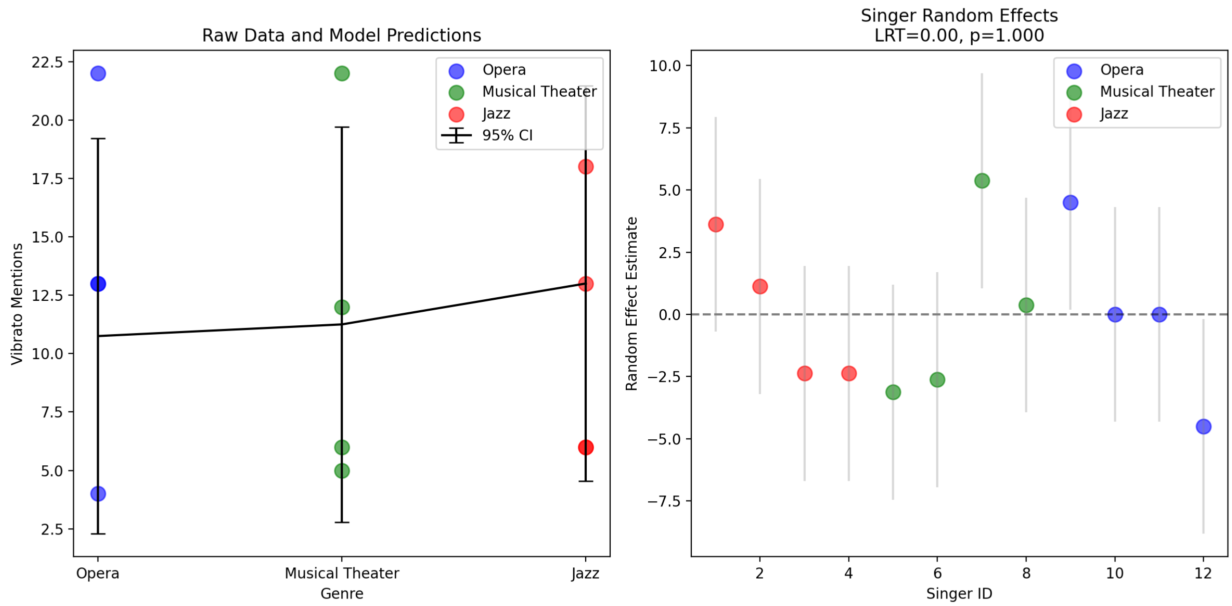


Figure 4.8: The GLMM overlayed with the LRT test findings investigating the relationship between genre and vibrato mentions, with singer included as a random intercept illustrated through a two-panel visualization: the left panel displays the raw data and model predictions across genres, while the right panel shows the estimated random effects for individual singers with their associated uncertainty. The substantial error bars in both panels visually reinforce the uncertainty in these estimates due to the limited sample size.

To further examine the homogeneity of variances across genres, pairwise F-tests were conducted, comparing the variance in vibrato mentions between each pair of genres. On the other hand, this analysis revealed substantial differences in variance across genres. Musical Theater showed the highest variance (60.92), followed by Opera (54.00), and then Jazz (34.25). However, F-tests comparing these variances revealed no statistically significant differences between any pair of genres: Jazz vs. Musical Theater ($F(3,3) = 0.562$, $p = 0.648$), Jazz vs. Opera ($F(3,3) = 0.634$, $p = 0.717$), and Musical Theater vs. Opera ($F(3,3) = 1.128$, $p = 0.923$). Figure 4.9 provides a visual representation of the variance differences across genres.

While the apparent differences in variance of vibrato mentions between groups were large, the small sample size may affect their detectability. Musical Theater showed nearly twice the variance

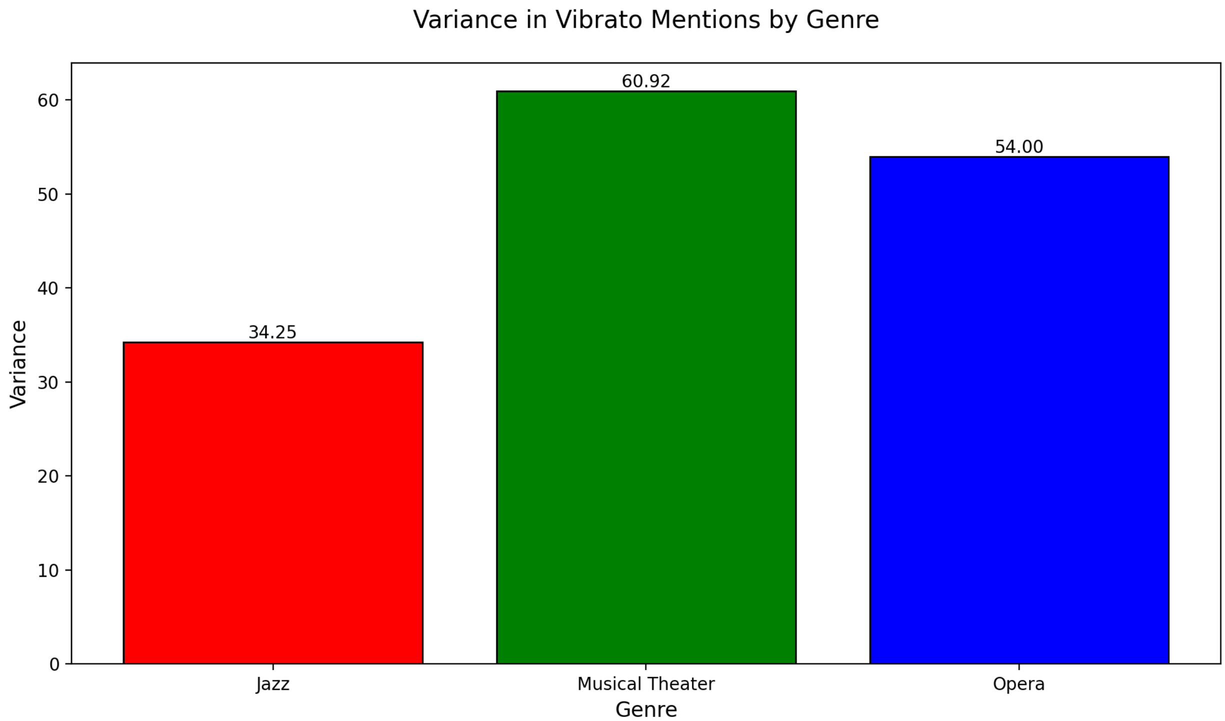
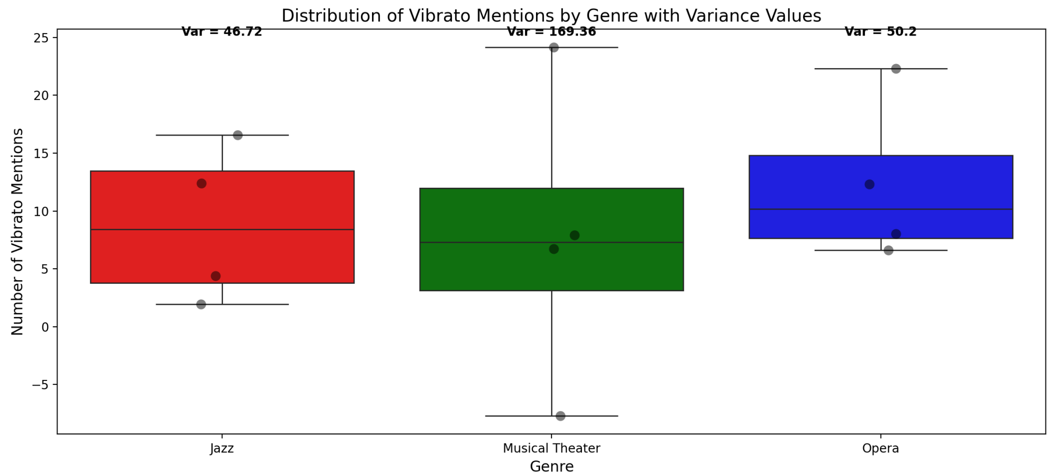


Figure 4.9: Histogram displaying the variance in vibrato mentions by genre, as calculated by an F-test.

of Jazz. However, Levene's test yielded a p-value of 0.7244, indicating insufficient evidence to reject the null hypothesis of equal variances across the three genre groups. The pairwise F-tests similarly show no statistically significant differences between any pair of genres (all p-values $>$ 0.05). The highest F-statistic (3.6251) was observed between Jazz and Musical Theater, but even this comparison failed to reach statistical significance ($p = 0.3183$). These results support the conclusion that while apparent differences in variance were substantial, the small sample size (*n*= 4 per genre group of 3 total genres) limited the ability to detect statistically significant differences in variance. Future studies with larger sample sizes are warranted to more definitively assess potential heterogeneity in variance across genres.

This finding supports the original conclusion that while variance patterns exist, the limited sample size prevents definitive statistical confirmation. The convergence issues in the GLMM remain a valid concern, as they stem from the small number of singers rather than the number of listeners. This distribution of vibrato mentions incorporating variance values can be observed in Fig. 4.10

To visualize the connections from the above study results to the original interview data set, all remarks specific to vibrato extracted from the timbral descriptors are categorized, grouped, and numbered by mention frequency and prominence in greater detail in the Mind Map in Fig. 4.11.



Comparison	F-statistic	p-value
Jazz vs Musical Theater	3.6251	0.3183
Jazz vs Opera	1.0744	0.9543
Musical Theater vs Opera	3.3740	0.3446
Levene's Test (all groups)	0.3342	0.7244

Figure 4.10: The distribution of vibrato mentions by genre with integrated variance values.

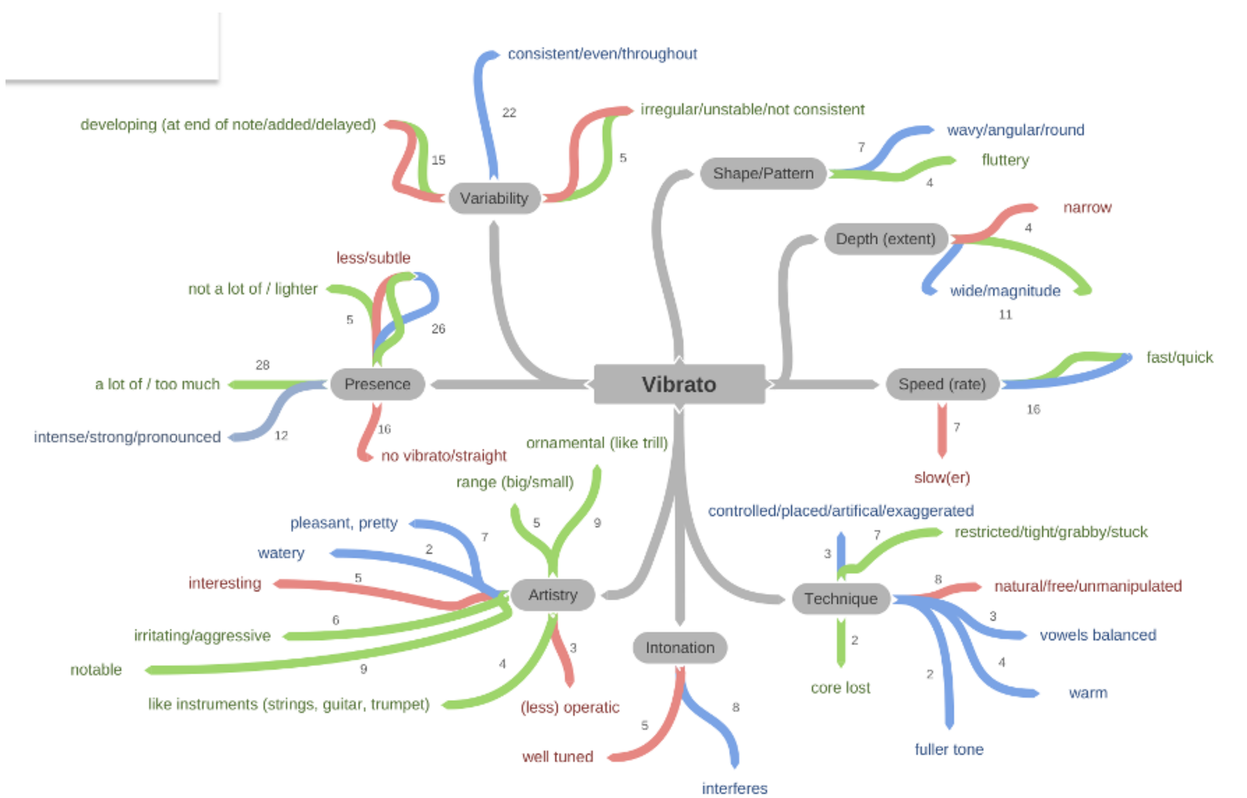


Figure 4.11: Web diagram/mind map of listeners' vibrato remarks, categorized by theme, grouped by type, numbered by frequency, and color coded by correlated excerpt genre. (Opera = Blue; Musical Theater = Green; Jazz = Red.)

These quantitative content analysis results confirm the perceptual power and salience of vibrato, regardless of listener participant group (vocalist vs. instrumentalist), across all singers and genres used as stimuli recordings within this data set. To delve deeper into the specific characteristics that are most prominent regarding vibrato, a qualitative analysis of the timbral semantic descriptors ensued.

4.3.3 Qualitative Results: Thematic-Semantic Analyses of Vibrato Remarks

Qualitative analysis of this perceptual study employed both thematic (to first systematically identify and interpret patterns within the data) and semantic analysis (a subfield of Natural Language Processing (NLP)), to interpret the contextual meanings of words and phrases), facilitating a comprehensive understanding of listener participants' experiences surrounding vibrato perception.

As the stimuli and excerpts in the perceptual study were alike in musical and lyrical content, the distinctions in timbral descriptors were subtler and more nuanced. The overall timbre semantic model included 64 descriptor categories, organized into 4 overarching groups: vocal registration, production, valence and arousal, and sound qualities, and 10 thematic groups, accounting for remarks grouped together semantically but not easily represented by a recurring term or terms (Reymore et al., 2025). For the purposes of this study, vibrato was considered within the four overarching groups and descriptor categories. While vibrato itself had the potential to occupy its own group, its categorical dimensions overlapped several of the already established categories. The most connected group to vibrato primarily included production (regarding the physiological and biomechanical production of the singing voice) and secondarily involved valence (the intrinsic positivity or negativity of an emotion) and arousal (the level of physiological activation or intensity associated with an emotion).

Listeners (both vocalists and instrumentalists) associated valence and arousal terms to describe expression and artistry when referring to the perceived vibrato. This perceptual study's findings confirmed what existing literature (outlined in Chapter 2) has posited as a perceived "sound of arousal" in the voice (Wallmark, 2022). Listener remarks regarding emotional arousal of the singers in the recorded stimuli excerpts connected to vibrato and included terms such as "happy," "sad," "irritated," and "anxious." In close proximity to these terms came timbral commentary, with many being cross-modal, often linked to sensory domains beyond sound. The valence and arousal category also included statements regarding the genre or type of vibrato used. These extended to qualifiers of vibrato, which correlated to the acoustic features of the vibrato in the recorded singer stimuli samples. It is known that acoustic resonance properties like spectral centroid, inharmonicity, and auditory roughness, guide emotional responses and influence the meaning attributed to specific timbral qualities. An under-explored acoustic characteristic that displayed significant impact on emotional and expressive language used in timbral descriptors in this study was vibrato and its variability.

Embodied Cognition, Functional Disorders, Muscle Tension Dysphonia, and Connection to Vocal Production Timbre Semantics

Heidemann (2016)'s system for vocal timbre description emphasizes embodied cognition, proposing that our understanding of vocal performance is deeply tied to the physical sensations of producing sound. These findings align with this premise, particularly within the production category of this perceptual study, where listeners frequently referenced physiological and acoustic voice production mechanisms and bodily states. These results suggest that such a framework linking timbral descriptions to physical sensation and sympathetic mirroring is already reflected in the natural lexicon of musicians. Since there was no significant difference across vocalist versus instrumentalist listeners in production descriptors, the level of complexity and shared terminology understandably was more homogenous within group rather than across the two listener groups. Vibrato was used in proximity or in reference to several common production terms, demonstrating its multifactorial nature as encompassing multiple aspects of vocal production. Remarks on vibrato encapsulated intersecting comments about vocal function and/or technique, including but not limited to airflow, subglottal pressure, vocal fold activity, and vocal tract shaping. This underscored the complexity of embodied cognition in timbral descriptors, especially those related to vibrato.

The production category primarily linked to vibrato also highlighted the role of perceived effort or ease, with descriptors like "pressed/pushed" and "free" indicating physical strain or the lack thereof. Prevailing production terms tied to vibrato included "controlled," "tense," and "natural/authentic." Particular terms such as "breathy," "flexible/agile," and "healthy" could have evolutionary implications for assessing a singer's health or state, in accordance with Wallmark (2022). Terms regarding physical health or stress, such as "sick/tired" or "phlegmy" appeared in 49 base terms and in close proximity to vibrato mentions, particularly when describing non-Operatic styles. This is another indication of the strong interaction of vibrato and recorded singers' genre influencing listeners' vocabulary. These embodied cognitive and production-based descriptor terms used alongside vibrato reflect an interplay between physical sensation, timbre perception, as listeners seemed to intuitively link sound qualities to bodily states and mental/emotional cues.

Analysis of syntactic proximity within the production category also exhibited importance to vibrato mentions, highlighting the potential of cognitive-somatic mapping occurring when perceiving timbre. This may confirm Wallmark (2022)'s claim that listening to timbre promotes sensorimotor engagement. Based on many of the vocalists' timbral descriptors and word choice, hearing vibrato in the stimuli samples seemed to activate an imagined physiological and motor response stemming from the action of sound production, ultimately evoking a profound embodied response.

Within the vocal production category, the excerpts with the highest number of remarks from both vocalist and instrumentalist listeners included speculative comments regarding some singers' potential compensatory tension conditions or voice-related disorders. Commonly used descriptive terms were related to specific maladaptive biomechanical functionality and production physiology, with phrases such as "overly tight," "forced by jaw," "muscle tension," and "imbalanced." Noting the erratic changes in variability in the vibrato over time in these excerpts, several listeners ascribed these qualities to potential paralaryngeal or functional voice disorders, such as Muscle Tension Dysphonia (MTD). Semantically, such descriptors align with common etiological and symptomatic characteristics of MTD, including "excessive perilaryngeal musculoskeletal activity" during phonation (Oates and Winkworth, 2008). Therefore, for several expert listeners, vibrato variability was qualitatively used as a signifier for MTD. As this pilot study is the first claim in the current literature of a connection of vibrato with hypo- or hyperfunctional paralaryngeal disorders such as MTD, more quantitative and qualitative research is certainly needed to further elucidate this hypothesized relationship. This theorized connection raises both clinical and pedagogical implications for the expansion of perceptual evaluation of voice disorders.

For more on the connection of embodied cognition, voice function, and vibrato, see Chapter 7.

4.3.4 Mixed Methods Results: Acoustic Model Correlates to Vibrato Remarks

The mixed methods results of this study revealed nuanced relationships between acoustic features of vibrato and listeners' subjective perceptions, as well as emergent thematic patterns. Quantitative analysis of the total 240 mentions of vibrato revealed that approximately 43% (102 mentions) were related to vibrato variability over time, categorized under the theme of "variability/complexity/shape/pattern." Of these 102 mentions, 52% (53 mentions) explicitly connected variability to vocal function and voice production, highlighting the perceptual salience of vibrato as an indicator of vocal technique. Thus, the selective thematic content analysis conducted with discrete qualitative categories exemplified the connection that vibrato has to timbral descriptor categories in voice perception.

Trends showed that, as in Table 4.2, several of the singers' stimuli samples garnered the most vibrato remarks, especially concerning variability and complexity. More detailed analysis offered by the thematic content using discrete qualitative categories method elevated variability and complexity as the most commonly shared subcategory, replacing the aforementioned category predominantly connected to vibrato, voice production.

As objective validation for the qualitative analyses conducted on this data set, a paired acoustic analysis verified the particularly visual remarks (i.e., describing shapes and patterns of the vibrato contours themselves) that many listeners verbally imagined. The quantitative acoustic metrics and qualitative parameters of variability (extracted and described in detail as Coefficient of Variation and logistic curves in Chapter 3 of this thesis) were highest in the same stimuli samples that had the greatest concentration of vibrato variability and complexity-related remarks. This corroboration reinforced the centrality of vibrato and its time-varying aspects to listeners' impressions of voice timbre, acoustic analysis, and the novel application of parametric models described in detail in Chapter 3.

The power of vibrato variability and its time-varying shapes and patterns is apparent in the listener descriptive remarks, across listener types, genre, and singers. To explore the most prominently mentioned adjectives related to vibrato within the remarks, see the illustrated Word Cloud in Fig. 4.12.

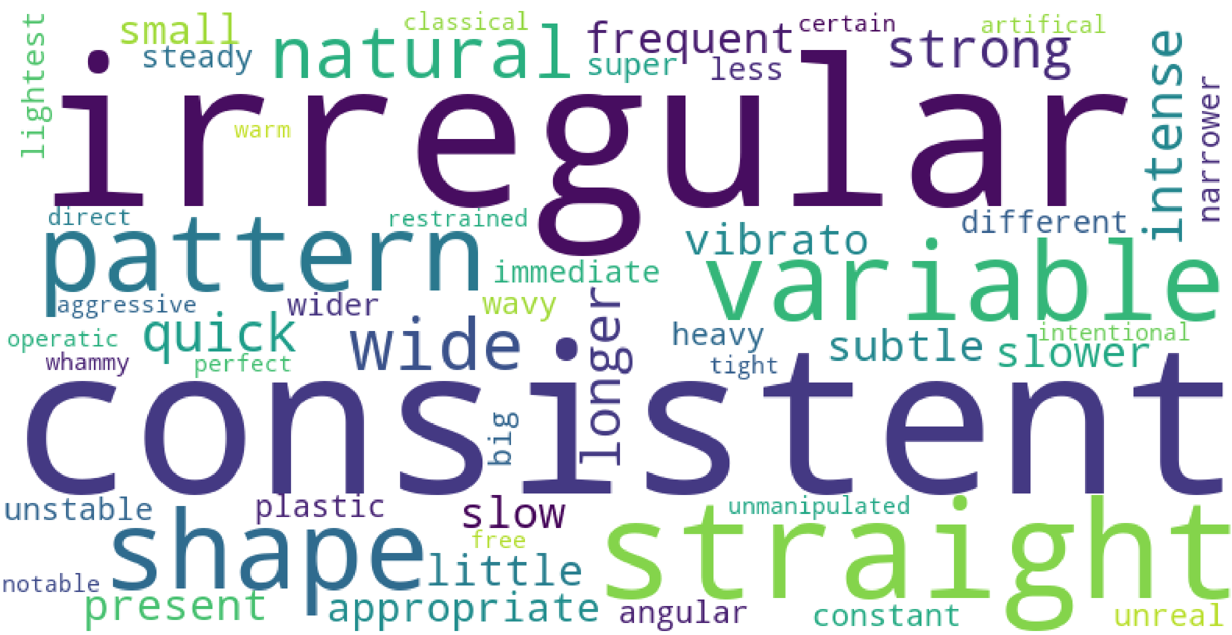


Figure 4.12: Word Cloud visualizing the most frequently mentioned adjectives related to vibrato and their prominence.

This study provides a multidimensional analysis of vibrato perception, integrating acoustic metrics, parametric models, and listener-based semantic descriptors. The results reveal that vibrato is a major perceptual feature across all listener groups, with no statistically significant difference in the frequency of vibrato mentions between trained vocalists and instrumentalists. While musical genre influenced vibrato mentions, statistical analyses indicated that singer-specific variability played a more substantial role than broad stylistic classifications. The findings underscore vibrato's intricate relationship with vocal timbre perception, spanning production, valence/arousal, and sound quality dimensions. Listener remarks frequently linked vibrato to physiological and acoustic mechanisms, reinforcing an embodied cognition framework and demonstrating an intuitive mapping between vibrato and vocal function.

Variability of vibrato, in particular, emerged as one of the most salient aspects of singing voice perception. A key perceptual cue across listener groups (both vocalists and instrumentalists), vibrato variability was closely tied to complexity and expressivity, with strong correlations between vibrato-related remarks and acoustic metrics validating these qualitative insights. Despite the study's small sample size and associated statistical limitations, the results support the feasibility of developing measurable acoustic tools informed by perceptual data. Ultimately, this study affirms vibrato's central role in shaping timbral perception, bridging cognitive, somatic, and affective dimensions in voice evaluation.

4.4 Discussion

The results of this study demonstrate that vibrato variability is both perceptually salient and a timbre-defining feature for trained vocalists and instrumentalists alike. Participants consistently mentioned vibrato when describing vocal characteristics, and their semantic accounts underscored its role in defining vocal timbre, influencing perceptions of pitch, intonation, and emotional expression. Vibrato was most predominantly tied to voice production, vocal function, and technique, as well as emotional valence and arousal categories of timbral descriptions. Particularly interesting were several listeners' novel remarks attributing vibrato variability as a qualitative signifier for paralaryngeal and functional voice disorders in singers, such as Muscle Tension Dysphonia (MTD). Additionally, the variability and complexity of the vibrato itself correlated to highest metric and parametric variability calculated and modeled quantitatively, and this specific aspect emerged as a pervading factor shaping listeners' perceptions of the singers' voices. The mixed methods approach— integrating quantitative acoustic measures of vibrato variability with qualitative content, thematic, and semantic analyses— revealed that the acoustic properties of vibrato are closely linked to listeners' subjective interpretations, with vibrato variability emerging as a key factor that shapes these perceptual judgments.

In addressing the research questions, the study found that:

1. **(RQ1) Perceptual Salience of Vibrato**: Vibrato was frequently and prominently noted in open-ended interviews, with participants providing numerous and varied descriptors that emphasized its timbre-defining role.

2. **(RQ2) Differences Between Vocalists and Instrumentalists**: Trained vocalists and instrumentalists differed significantly in their lexical choices and frequency of vibrato-related comments, suggesting that either vocalists' specialized neural processing of vocal timbre or instrumentalists' exposure over time enhances perception.

3. **(RQ3) Genre-Specific Characterization:** Listeners used distinct, genre-specific descriptors when discussing vocal timbre, highlighting the influence of musical style on vibrato perception.

4. **(RQ4) Association Between Acoustic Variability and Perceptual Comments:** A robust correlation was observed between quantitative measures of vibrato variability (e.g., vibrato half-extent variability) and the qualitative nature of vibrato comments, indicating that measurable acoustic fluctuations play a pivotal role in shaping listeners' interpretations. Implications for vibrato variability's qualitative and potential quantitative connection to functional voice disorders such as MTD were hypothesized.

4.5 Limitations and Future Work

As this was a secondary, extended analysis, there were several inherent limitations, including small sample size and purposefully constrained variables.

One of the listeners (classified in the study as an Operatic Mezzo-Soprano) revealed after the completion of the study that they personally self-identified as both a vocalist and instrumentalist, though they accorded with the inclusion criteria. This self-identification may have skewed the weight of the listener type groups, though this listener participant's results in terms of vibrato mentions and remarks did not immensely differ from the other participants' mentions and remarks. Finally, the homogeneity of several of the vocalist participants' responses may be attributed to the possible influence of the fact that the majority of vocalist participants shared the same studio voice teacher (voice professor). It is commonly accepted that semantic development is co-constructed within private musical instruction, as a shared lexicon is required for mutual understanding and progress. However, this study's intriguing findings warrant further investigation on a larger scale and different context.

The use of a visual sort and rate method for perceptual surveys suggested by Granqvist, and subsequently used with success in voice-related studies, might be a recommended approach for future iterations of the perceptual stage of this study (Glasner and Nix, 2023; Granqvist, 2003). Future work may also include working toward an operational conceptual framework for organizing descriptors, as the study revealed that many categories are not fully definitive, with some terms able to fit into multiple groups. The results of this pilot perceptual study provide a compelling lens for understanding semantic use across voice productions in different styles and contexts, adding more nuanced understanding of vibrato perception and highlighting the intricate relationship between timbre, musical genre, emotional interpretation, and social identity. Further investigations should delve deeper into these aspects, examining their wider impact on music perception and cognitive processing, especially applied to specific aspects of vocal production.

4.6 Conclusions

This perceptual study revealed that vibrato variability is a notable factor impacting singing voice timbre perception across both trained vocalists and instrumentalists, notably influencing how vocal function, production, and expression are interpreted. By incorporating comparative acoustic analysis of vibrato variability, this perceptual study investigated the extent to which objective acoustic properties influenced subjective perceptual judgments of vibrato and vice versa. This integration of acoustic measures and perceptual data provided deeper insights into the relationship between measurable features of vocal stimuli and their perceptual salience across both vocalist and instrumentalist trained listeners. Conducting this cross-sectional perceptual study as a follow-up to the acoustic analysis and variability modeling in Chapter 3 allowed for a more ecologically valid, perceptually verified approach that remains true to the natural context in which vibrato appears. The findings of this perceptual study help pave the way for more refined acoustic, pedagogical, and ethnographic analyses and interpretation of vocal vibrato variability in the chapters of this thesis that follow.

Chapter 5

Vibrato in Bulgarian Folk Singing

Author Contributions

This chapter includes restructured and reintegrated content from the following research article: Nestorova, T., Nestorov, I., Nestorova, D., and Gradeshlieva, M. (2025). Bulgarian folk singing: The art and the science. *Journal of Voice*. [Manuscript submitted for publication].

Author-specific contributions: *Theodora Nestorova:* conceptualization of the study, data acquisition and pre-processing, post-processed acoustic analysis and interpretation of results, development of interpretation and conclusions, writing, review, and editing of the manuscript. *Ivan Aleksandrov:* coding of preliminary script, development and programming of original numerical and computational modeling for data and results processing, review and editing of the manuscript draft.

Chapter-specific contributions and differentiation from original journal publication: In this chapter, Sec. 5.1 is significantly expanded from its scope in the original article manuscript. Figures 5.1-5.2 are also additions to this chapter and do not appear in the original article manuscript. Finally, editing for grammar and flow were conducted to optimally incorporate the article contents into this thesis chapter.

5.1 Background on Bulgarian folk music

Vocal vibrato in non-Western musical genres remains an under-researched topic. Published research on the Bulgarian folk singing style, in particular, is lacking. Moreover, studies on vibrato in Bulgarian folk singing (BFS) are practically non-existent. This chapter, therefore, describes the results and conclusions of an exploratory, field research pilot study on vibrato in Bulgarian folk singing, which will be henceforth referred to as BFS.

BFS, a part of the broader Bulgarian folk music framework, has roots tracing back over a millennium yet has been preserved and developed through generations to the modern day. The origins of BFS can be linked to the musical traditions of the Thracians, Slavs, and Proto-Bulgarians, dating back as early as the 7th century CE, when these cultures merged to form the foundations of Bulgarian identity (Rice, 1994). Over the centuries, BFS evolved through oral transmission and linguistic evolution, shaped by Byzantine liturgical influences, Ottoman-era cultural exchanges, and regional stylistic diversification (Kirilov, 2015). As the modern Cyrillic alphabet evolved from the Glagolitic script (developed in Bulgaria in the 9th century by Saints Cyril and Methodius to transcribe Old Church Slavonic) (Crampton, 2005), it is likely that BFS similarly evolved through several phonetic and orthographic iterations.

A defining feature of Bulgarian folk music is its complex rhythmic structure, characterized by asymmetrical, uneven meters such as 5/8, 7/8, 9/8, and 11/8, differing from conventional Western metric frameworks. These intricate rhythmic patterns, often described as “crooked” rhythms, are deeply embedded in both the vocal and dance traditions, necessitating an intuitive physical understanding of meter and phrasing. Additionally, Bulgarian folk music is known for its rich harmonic and timbral textures, including drone-based polyphony, parallel harmonizations in seconds and thirds, and the distinctive use of tight, close-voiced dissonances (Rice, 1994). These unique sonorities contribute to the resonant choral tradition particularly made famous by the Mystery of the Bulgarian Voices (*Le Mystère des Voix Bulgares*) world-renowned women’s choir that characterizes BFS throughout the world. BFS has made a profound global impact, with its striking vocal style and rich musical heritage gaining international recognition; most notably, the song “Излел е Дельо хайдутин” / “Izlel e Delyo Haydutin,” performed by Valya Balkanska, was selected to be included on the Golden Record of Voyager 1 and 2 in 1977, sending the powerful sound of Bulgarian folk music into space as a representation of Earth’s cultural legacy. Preservation and purity of this extraordinary Bulgarian folk music tradition and heritage are efforts by ethnomusicologists such as Prof. Manol Todorov, as well as their students (Levy, 2009), several of whom were participants interviewed for this study.

Original audio and video recordings gathered and collected as a part of this study are made available by the author of this thesis at:

https://github.com/theodora-nestorova/vocalvibratovariability

5.1.1 Vocal Technique and Production in BFS

BFS is characterized by a vocal production that may be perceptually described as bright, speech-like, and powerful. Unlike Western classical singing, which emphasizes timbral *chiaroscuro*[1] bal-

[1] “An Italian term meaning “bright–dark” or “clear–dark” borrowed from art history to describe an ideal resonance balance between low and high frequency components in Western classical singing. It is usually accomplished (in middle and lower ranges) by some balance in power between the first formant and the singer’s formant cluster, or possibly between the first and second formants” (Hoch et al., 2022).

ance and even mixing of the "head" and "chest" registers[2], BFS requires a mainly "chest-dominant" registration or "belt-like" resonance[3] (Henrich et al., 2007).

Acoustically, the BFS technique emphasizes an overtone-rich, resonant sound that singers often perceive as forward-centered, paralleling the default mode 1 (M1) register[4] in speech (Messner, 1980).

In the BFS vocal technique, two distinct vocal qualities, *тежка* (henceforth referred to in its Romanized / Latinized version, *tezhka*, meaning "heavy") and *лека* (henceforth referred to in its Romanized / Latinized version, *leka*, meaning "light"), are utilized. These are central to the vocal tradition, with their usage widespread across various schools and training institutions in Bulgaria, albeit with minor regional variations (Messner, 1980). The *tezhka* technique is characterized by a loud, projected sound, typically employed in slower, more mournful songs or ballads. It is generated through sustained subglottal pressure and predominantly thyroarytenoid (TA)-dominant coordination to produce a low-frequency, high-intensity phonatory output. This vocal modality conveys emotional intensity, particularly in expressions of sorrow or lamentation, with singers frequently utilizing extended resonance[5] to enhance the expressive power of the voice. In contrast, the *leka* technique is marked by a softer, more lyrical timbre, uniquely adapted to the Bulgarian folk idiom. It involves the use of lower subglottal pressure and predominantly cricothyroid (CT)-dominant coordination, resulting in a higher-frequency, lower-intensity phonatory output with a more diffuse tonal quality. The *leka* technique is typically employed in more melodic or delicate musical phrases, contributing to a tonal quality that emphasizes lightness and lyrical

[2] In singing, "the term register has been used to describe perceptually distinct regions of vocal quality that can be maintained over some ranges of pitch and loudness" (Titze, 1994a) and these registers are often colloquially referred to as "head" and "chest" voice from the original Italian terms *voce di testa* and *voce di petto*, corresponding to light and heavy mechanisms, respectively. Registration must be distinguished in its acoustic and laryngeal meanings; laryngeal registers are "delineated by differences in the vibratory pattern of the vocal folds. These are primarily due to the relative participation of the thryoarytenoid (TA—shortening, thickening) and cricothyroid (CT—stretching, thinning) muscles, and include pulse (M0), modal (M1), falsetto (M2), and whistle (M3)." (Hoch et al., 2022). Acoustic registers, on the other hand, correspond to "the timbral transitions and range segments that occur as a result of harmonic/resonance interactions. All acoustic register transitions involve interactions of the lower numbered source harmonics with the first and/or second resonance of the vocal tract. See also: close timbre, open timbre, *voce chiusa*, and *voce aperta*, as well as 'hoot' timbre and 'whoop' timbre (the last two terms cited from Donald Miller and Ken Bozeman, respectively)" (Hoch et al., 2022).

[3] "A type of singing with a longer closed phase of the vocal folds, higher subglottic pressure, greater thyroarytenoid (TA) activity, and a weaker fundamental frequency (f_o). A chest voice–dominant vocal quality used in many styles of Musical Theater and CCM singing" (Hoch et al., 2022).

[4] Vocal registration is a multifaceted phenomenon that can be assessed through proprioception, psychoacoustic perception, laryngeal mechanisms, vocal tract effects, and individual didactic systems, with apparent discrepancies in classification often arising from the differing frameworks within which they are defined. In reality, there are several simultaneous vibrating laryngeal structures, vocal tract influences, and categorical (distinct vs. "mixed") competing factors, as well as non-linear dynamics, that contribute to the common designation of vocal registers (Herbst, 2020). However, it is beyond the scope of this chapter to delve into the complex nuances surrounding the discussion on registration in singing, for which there is still no generally accepted clear definition (Sundberg, 1987).

[5] Extended resonance, while not formally defined in singing voice science, is commonly associated with techniques that enhance and prolong vocal sound through optimized use of vocal tract shaping.

beauty (Henrich et al., 2007). The dynamic interplay between *tezhka* and *leka* is a fundamental aspect of BFS, with these vocal techniques providing a diverse emotional palette.

Resonance strategies employed in BFS styles, particularly in the *tezhka* and *leka* techniques outlined above, involve specific adjustments to the vocal tract that enhance the timbre and projection of the voice. These adjustments often tune the first tract resonance close to the second harmonic, which is a defining characteristic of BFS (Henrich et al., 2007). This acoustic understanding is essential for both performers and educators, as it informs the methods used in teaching these traditional singing styles.

5.1.2 Acoustic Characteristics of BFS

Spectral analysis of BFS reveals a significant presence of high-frequency energy, with strong second and third harmonics that enhance the perceived brightness. Unlike Western classical singing, which often aims for a singer's formant cluster [6] around 2.5–3.5 kHz, the BFS voice projects through a different spectral balance, emphasizing even higher frequencies between 3-5 kHz (Doskov et al., 1995). This acoustic strategy allows the voice in Bulgarian folk music ensembles to carry over outdoor settings and typically dense, powerful instrumental textures with which the voice may need to compete, a necessity for traditional performance contexts (Rice, 1994).

In both *tezhka* and *leka* vocal qualities, a seminal single-subject case study by Henrich et al. (2007) found that the Bulgarian folk singer consistently tunes the first vocal tract resonance (R1) close to the second harmonic (H2) for most vowels. This resonance tuning significantly enhances the power output at H2, a defining characteristic of the BFS timbre. By amplifying energy in this harmonic, the technique contributes to the distinct richness and intensity of the vocal sound. Additionally, since this frequency range aligns with peak human auditory sensitivity, it enhances the perceived loudness of the voice, facilitating the remarkable natural projection associated with BFS (Robinson and Dadson, 1956).

This resonance tuning strategy seems to be applied consistently across the generally limited pitch range (F4–B4) characteristic of Bulgarian folk songs. Unlike the resonance tuning typically observed in Western classical sopranos, where R1-H2 tuning is selective, Bulgarian singers apply this formant tracking across most vowels (Henrich et al., 2007). However, some studies using linear prediction coding coefficients (LPC) and cepstrum methods found adjustment of the first formant to the second harmonic of the fundamental frequency only for the vowel /i/ but not for

[6] A grouping of frequency peaks in the sound's radiated spectrum created by clustering the third, fourth and fifth vocal tract resonances together within a narrower frequency range, approximately around 2300–3500Hz. This clustering provides amplification of the harmonics within this range. It is thought that this is a result of narrowing the epilaryngeal tube exit relative to laryngo–pharyngeal openness. This allows the voice to project through the masking effect of an orchestra (or other accompaniment) on spectral content below ca. 500Hz. (Hoch et al., 2022).

vowels /a/, /ɛ/, and /o/ within a song from the Bulgarian Rhodope region (Kirkov and Zieliński, 2019). The latter suggests that this resonance strategy may have developed in response to both tessitura and linguistically related factors, allowing for optimal vocal power and a distinguished spectral balance.

The role of R1 tuning in BFS may be influenced by the phonetic structure of the Bulgarian language, where second resonance (R2) distinctions play a greater role than R1 in vowel differentiation. The six-vowel system consists of three fronting-based pairs, with height distinctions primarily controlled by R1. However, in some dialects— particularly in Eastern Bulgarian informal speech— these height-based distinctions are reduced or even neutralized (Vutov, 2002). Given that vowel height contrast is of relatively low linguistic significance, tuning R1 to H2 in singing introduces minimal perceptual loss while yielding a substantial gain in radiated power. Ultimately, the prominence of H2 in BFS is a direct consequence of vocal tract tuning, which contributes to the recognizable timbre and high-intensity sound of the tradition.

Despite the perceptual differences between the *tezhka* and *leka* techniques, this resonance tuning strategy remains largely consistent, except in the case of the vowel /u/ (Henrich et al., 2007). Instead, these differences are likely attributable to glottal behavior, particularly in glottal closure speed and closed phase duration. Faster glottal closure is known to increase sound pressure level and high-frequency spectral energy, which may explain why *tezhka* appears more powerful and resonant, whereas *leka* exhibits the opposite qualities, resembling the Western classical "head" voice. However, further research conducted specifically with electroglottography (EGG) is needed to quantify the precise role of glottal dynamics and contact quotients[7] that underlie these vocal qualities.

For a more in-depth understanding of vocal physiology production patterns, evaluation of the acoustic properties of BFS should be paired with other methodologies, including biomechanic and aerodynamic instrumental studies.

5.1.3 Pedagogical Considerations in Teaching BFS

Given the distinct physiological and acoustical demands of BFS, pedagogical approaches in existing literature vary from source to source. As Bulgarian folklore, both narrative and sung, is commonly passed down through oral and aural transmission, the instruction of BFS parallels and draws from this tradition. Traditional teaching methods in BFS historically rely heavily on imitation, with students learning through direct exposure to and mimicry of master singers. This approach fosters an embodied understanding of the style, emphasizing cultural authenticity and

[7] "The percentage of vocal fold contact in a given vibratory cycle, used to estimate laryngeal register and closed quotient (i.e. the percentage of time in one vibratory cycle that the glottis is closed). While electroglottography (EGG) displays contact quotient reliably, inferring closed quotient from this may be less reliable" (Hoch et al., 2022).

kinesthetic memory (Petrova-Kirkova, 2019).

The pedagogy of BFS is deeply intertwined with the country's rich folk dance traditions, reflecting an inherently somatic method to vocal instruction. Unlike Western music genres, which often prioritize isolated vocal technique as separate from dance training, BFS instruction is effectively inseparable from physical movement. Singers are commonly vocally trained simultaneous to their dance training and are taught to vocally and physically coordinate the two domains with traditional rhythmic patterns, step sequences, and gestural motifs. This integration not only reinforces musical phrasing, metric complexity, and breath management but also serves as a mnemonic device for vocal execution. Research has shown that gesture techniques enhance vocal technique by improving motor and sound sensations among singers (Liao and Davidson, 2016), aligning with contemporary active learning pedagogies that emphasize physical and emotional engagement. Many Bulgarian folk ensembles feature singers who actively dance while performing, demonstrating the inseparability of vocalization and movement in this tradition. The synergistic relationship between movement and music not only enhances technical proficiency but also ensures the authentic transmission of regional stylistic elements, preserving the deeply communal and participatory nature of Bulgarian folk traditions.

Philosophers have long-believed that music evolved as accompaniment to bodily movement and signal calls used by primitive peoples for the purpose of communication (Rice, 1994). In the context of BFS, singing is seen an essential component of communal and ritualistic life. Historically, for the Bulgarian population, the ability to project a loud, penetrating voice was not only a musical necessity but also a cultural imperative, with singing playing a vital role in agricultural and seasonal ceremonies believed to influence natural cycles (Crampton, 2005).

Socio-ethnographic Specificities: Gender, Geographic Regions, and Ornamentation

A crucial element of BFS is the use of close harmonies and microtonal inflections, which are integral to the regional styles of different folk ensembles. The precise tuning of these harmonies, often involving intervals smaller than a Western semitone, requires both auditory acuity and fine auditory-motor control. Research in cognitive phonetics suggests that Bulgarian speakers (and likely by extension, singers) develop an enhanced ability to perceive and produce microtonal variations internal to the linguistic structure of Bulgarian itself (Vutov, 2002). This contributes to the unique intonational characteristics of the BFS genre.

The diverse geographic regions of Bulgaria have significantly shaped the country's folk singing traditions, with each region exhibiting distinct vocal timbres, melodic structures, and ornamentation styles influenced significantly by geolinguistic dialects (Antonova-Vasileva, 2019). As shown in Fig. 5.1, Bulgaria is typically divided into several ethnographic regions, each with its own unique musical characteristics, including:

- *Шоплук* (henceforth referred to as *Shopluk*, the Sofia region),
- *Пирин* (henceforth referred to as *Pirin*, the southwest region),
- *Родопи* (henceforth referred to as *Rhodope*, the south-central region),
- *Тракия* (henceforth referred to as *Thrace*, the south region),
- *Добруджа* (henceforth referred to as *Dobrudzha*, the northeast region), and
- The Northern and Northwestern regions.

In Shopluk, the singing style is often described as sharp, high-pitched, and nasal, with strong rhythmic drive and use of diaphonic singing, where one voice sustains a drone while another sings a moving melody above it. The Pirin region, influenced by neighboring Macedonia and Greece, features intricate melodic ornamentation and asymmetric rhythms, often performed in tight harmonies that emphasize close intervals. The Rhodope region, by contrast, is known for its slow, melancholic melodies with a wide vocal range and rich, open-throated resonance, reflecting the pastoral traditions of the region's mountainous landscape. Thracian folk singing is characterized by a hard glottal onset attack, fast tempos, dynamic vocal leaps, and a strong chest-dominant resonance. On the other hand, Dobrudzha's musical traditions incorporate ornamentation-rich melodies with an expansive vocal range, mirroring the region's flat and open landscapes. The Northern and Northwestern regions are known for their call-and-response singing forms and powerful, projected vocal delivery, shaped by the region's agrarian lifestyle (Borisova, 2013).

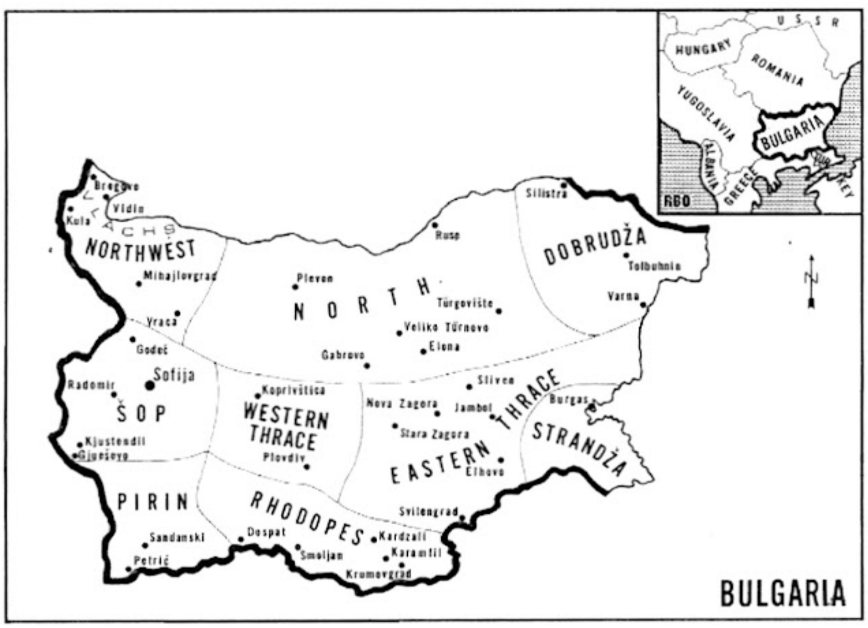


Figure 5.1: Map of folkoric regions of Bulgaria (Kozakand, 2012).

These regional styles are further influenced by Bulgaria's three major dialect groups (Girvin, 2016)— Eastern, Western, and Rup (spoken in the Rhodope Mountain region)— which impact vowel placement, phrasing, and intonational nuances in folk singing. For example, Western dialects, spoken in Shopluk, feature shorter vowels and sharper consonants, contributing to the bright, forward-placed vocal quality characteristic of the region's singing style. Meanwhile, Eastern dialects, dominant in Thrace and Dobrudzha, exhibit softer consonantal articulation and more open vowel production, aligning with the more fluid and expansive vocal lines of these musical traditions. The Rup dialect, spoken in the Rhodope Mountains, is marked by archaic phonetic features and influences the resonant, drawn-out phrasing found in the region's folk singing. Together, these geographic and linguistic factors contribute to the rich diversity of Bulgarian folk vocal traditions, demonstrating how regional identity is deeply embedded in the sonic and linguistic fabric of the country's musical heritage (Kremenliev, 1956b). See Fig. 5.2 for a dialectal map of the Bulgarian regions. Special focus on the phonological, linguistic dialect, and musical ornamentation in BFS will be given in Sec. 5.3.4 to follow.

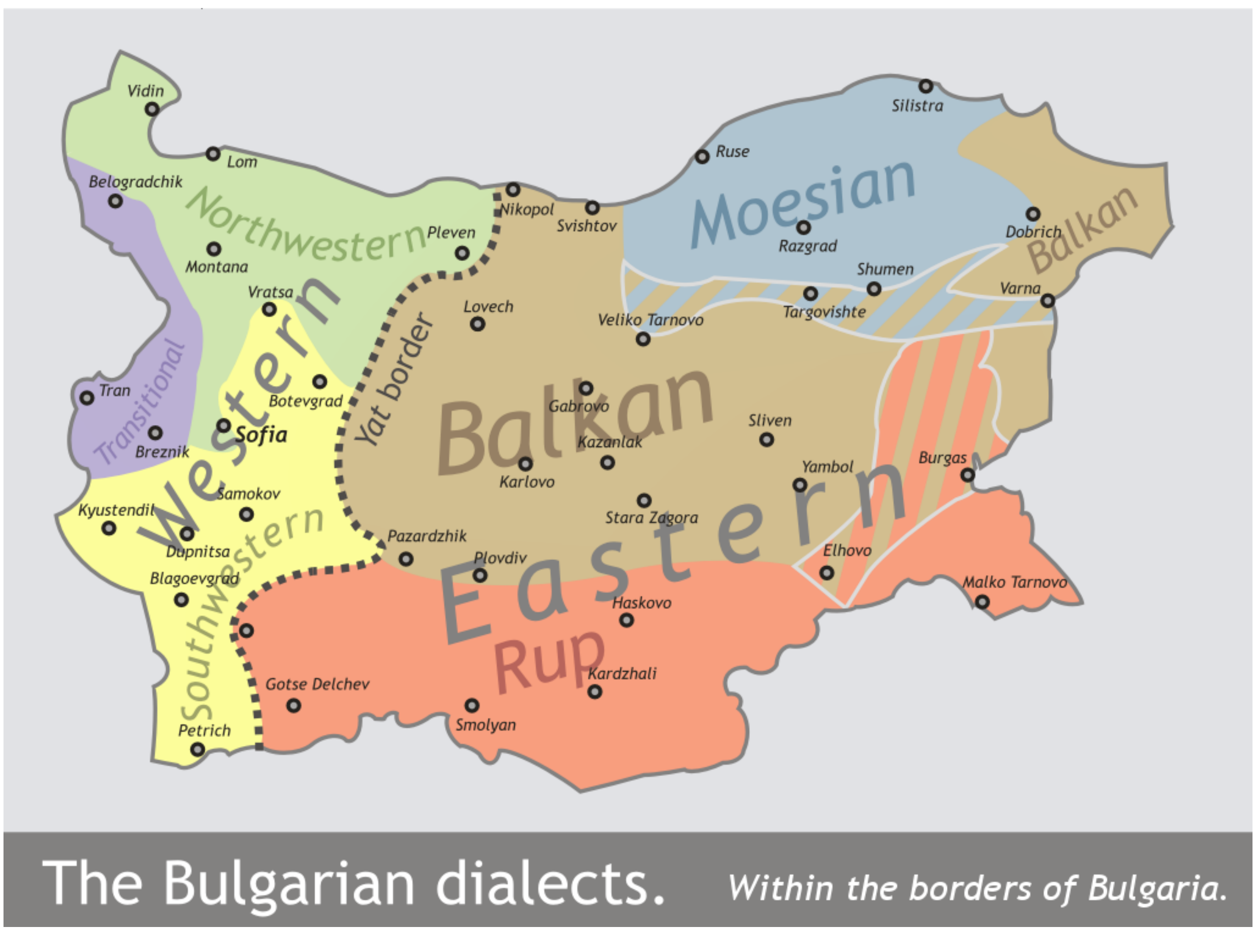


Figure 5.2: Map of dialectal regions of Bulgaria (Stoikov, 1962).

The predominance of women in BFS, both in teaching and performance, is deeply rooted in the country's sociocultural norms and traditions. Historically, Bulgarian village life was structured around communal labor, where women played a central role in agricultural work, domestic tasks, and social rituals. Singing was an integral part of daily life, accompanying activities such as harvesting, weaving, and childcare, reinforcing a strong oral tradition passed down through generations. Women were also the primary participants in ritual and seasonal celebrations, such as *lazaruvane* (a springtime rite of passage for young girls) and wedding ceremonies, where their vocal performances were believed to hold symbolic and even magical significance in ensuring fertility, prosperity, and social cohesion. Additionally, the close-knit nature of women's social circles fostered the preservation and transmission of folk repertoire, leading to their dominance as cultural bearers of traditional singing styles. During the 20th century, as Bulgarian folk music gained national and international recognition, women continued to be at the forefront, forming the backbone of professional folk ensembles such as the Philip Kutev National Folklore Ensemble and the renowned *Le Mystère des Voix Bulgares* (The Mystery of the Bulgarian Voices) choir. In the pedagogical sphere, women have remained key figures in folk music education, ensuring the continuity of vocal traditions through formal instruction and inter-generational mentorship. This historical and cultural lineage positions women not only as the primary voices of BFS but also as its preservers, educators, and innovators (Kremenliev, 1956a; Ivanova et al., 2021; Stankova, 2019; Silverman, 2011).

The socio-political impact on BFS and its pedagogy must be acknowledged, as during the Socialist-Communist era of 1944-1989, BFS was transformed from primarily oral transmission to written notation, *обработки* / *obrabotki* (arranged folk music), as the state sought to collectively own, standardize, and professionalize traditional music for mass appeal and Western-appearing ideological purposes. Folk ensembles such as the Philip Kutev National Folklore Ensemble were established at that time to create polished, harmonized, and orchestrated versions of regional folk songs, elevating them to a nationalized form that aligned with socialist ideals while simultaneously erasing much of the raw, village-based authenticity of earlier traditions (Willson, 2017). The socio-political impact on BFS is further expanded in Sec. 5.3.4 below.

5.2 Scoping Literature Review on Vibrato in BFS

There is a significant lack of published research on vibrato and ornamentation in BFS, in both native and non-native sources. For a scoping review conducted on this topic, a search strategy with specific screening procedures was used to retrieve all publications relevant to vibrato in BFS using the following keywords (multi-lingually; in English, Bulgarian, French, German, and Italian): BFS, Bulgarian Song, Bulgarian folk music, Bulgarian Folk Vibrato. Keywords were combined with the Boolean operator "OR," and the initial search was conducted in seven databases (Web of Science, Scopus, PubMed, JSTOR, RILM, ProQuest Central, EBSCOHost, and a Bulgarian-language research database through Saints Cyril and Methodius National Library). A significant portion of publicly available information on BFS emerges from non-traditional digital sources such as lessons, master classes, and video/audio recorded workshops offered by trained voice professionals. Therefore, the eligibility criteria of sources was broadened to include peer-reviewed papers published in academic journals as well as approved doctoral dissertations, and audio or video recorded tutorials by experts and professionals. No limitations were applied in terms of publication year. It is evident that there is a global knowledge gap and therefore, a comprehensive, focused, systematic, and first-hand investigation of vibrato in the BFS genre is warranted.

While the majority of literature sources on BFS are restricted to Bulgarian-language publications, a non-Bulgarian native source, Henrich et al. (2007), mentions vibrato as a part of one sentence, namely "...vibrato, which is used sparingly in this style of singing." However, several Bulgarian native sources seem to disagree with this opinion, which may be attributed to its premise as a single case pilot study; numerous Bulgarian sources hint that vibrato is an inherent part of artistic expression in BFS (Borisova, 2012; Vekilova and Minkova, 2010; Kukushev, 2008). Multiple scholarly sources suggest that various types of vibrato exist in both solo and choral women's folk singing, likely characteristic of the different geographic regions of the country.

Some sources hint at a "specific" type of vibrato in BFS (Meleki, 2021), which signals intensity and emotionality in the voice (Bosheva, 2013). Emergent from the literature is a clear connection of vibrato (and its various types) to the regional dialectal and ornamental practices that, as previously mentioned, vary geographically in BFS. Various sources indicate differing, potentially wider "Strandzha vibrato," the narrower "Dobrudzha vibrato," or the "Rhodope uncharacteristic vibrato" (Kaludova-Stanilova, 2022; Kehaiov, 2018).

Certain historic Bulgarian vocal pedagogues such as Ivan Maksimov and Alice Bovaryan do consider vibrato to be a basic, inherent quality of the human voice, alongside strength, timbre, and range. However, they cite exclusions for choir singing and child singing (Koteterova-Dobreva, 2012), in favor of a more natural, polished, and homogeneous sound (Batkova, 2023). Yet, other sources caution against the use of Western classical singing vibrato in BFS, which traditionally employs vibrato differently, sparingly, and with specific purpose, favoring a more direct and steady entry into a tone. From the limited sources that discuss vibrato in BFS, when vibrato is used, it appears to be subtle and context-dependent, naturally emerging in certain regional styles or as an interpretative choice by individual singers.

In modern-day BFS, especially the kind disseminated across North America, there is no definitive vibrato practice; sources indicate many modern-day singers utilizing delayed onset of vibrato (at the ends of phrases), strong vibrato, vibrato-less (non-vibrato) production (Moneva, 2022), as well as vibrato conventions following the musical phrasing in some groups (Stankova, 2019).

5.3 Field Research: Acoustic, Ethnographic, and Pedagogical Exploratory Pilot Study

Exploratory field research including a pilot acoustic and ethnographic pedagogical study was undertaken, with the primary objective of investigating the presence, classification, characteristics, and artistic implementation of vibrato in BFS. Current vibrato research has been largely shaped by Western classical paradigms, which emphasize uniformity and consistency, often overlooking the natural variability present in non-Classical and non-Western singing traditions. This study seeks to address this gap by examining vibrato within the Bulgarian folk style, developing a more inclusive analytical framework that accounts for the unique spectral and temporal properties of vibrato beyond the Western art music tradition.

To achieve these goals, this research centered on fieldwork conducted at the Philip Kutev National School of Folk Arts in Kotel, Bulgaria, the foremost institution for the professional instruction and training of Bulgarian folk music. This study:

- Conducted on-site literature and archival searches, as well as interviews with leading folk song educators and scholars, to gather both historical and contemporary perspectives on vibrato in Bulgarian folk performance practice.
- Engaged with native vocal researchers, pedagogues, and singing professionals specialized in teaching BFS to learn more about vibrato pedagogy within this tradition.
- Collected and analyzed vibrato samples (using the same time-varying acoustic models and archetype framework defined and applied to Western singing in Chapter 3) from expert performers across various regional and stylistic variations of BFS.
- Identified and classified vibrato characteristics, integrating findings with expert discussions to gain a more culturally competent and responsive understanding of vibrato in this tradition.

By analyzing vibrato within the BFS tradition, this study aims to refine the proposed vibrato classification models introduced in Chapter 3 and offer an alternative analytical framework that more accurately represents the diverse vibrato behaviors in Bulgarian folk music.

5.3.1 Methods, Materials, Procedures, and Participants

This acoustic and ethnographic pedagogical study was approved by the McGill University Research Ethics Board (REB #22-06-071), and participation was voluntary. A total of 22 individuals, including 8 professional performers, 6 teachers, 6 students, 1 instrumentalist, and 1 ethnomusicologist were recorded and/or interviewed regarding vibrato in BFS. The equipment used to record these singers in the field included a Sound Devices 702 portable audio recorder connected to a DPA 4062 microphone with a pop filter/wind sock.

By conducting this research directly at the Philip Kutev National School of Folk Arts in Kotel, Bulgaria, local folk music pedagogues, researchers, and recording engineer specialists facilitated the collection of vocal vibrato sample recordings and information on vibrato ornamentation in this genre. This school is considered the premiere training center for Bulgarian folk music and houses experts from all major regions across Bulgaria. From the school's archive library and sound recording studio laboratories, several non-digitally accessible historical recordings were initially reviewed. These proved extremely valuable in gathering a comparison for historical to modern vibrato conventions in BFS. Subsequently, singing demonstrations and interviews provided by experienced practitioners and professional folk singers, including Galya Petrova-Kirkova, Iliyana Naydenova, and Polina Dimitrova Dimitrova, were gathered within this fully-immersive, on-site experience for field research.

During the field research, a major Bulgarian folk music festival provided a unique opportunity to document performances by professional folk singers and instrumentalists from diverse regions across Bulgaria. Notably, recordings were collected from esteemed musicians, including world-renowned Bulgarian *гайдар* (*gaidar* or bagpiper) Iliya Uchikov with Rhodope singer Mariana Angelova. Additionally, this fieldwork extended beyond the festival grounds to include visits to nearby villages, where local amateur singers, both soloists and ensemble groups, were recorded performing traditional songs specific to their respective regions. These recordings serve as valuable primary data for the study of regional folk music traditions in Bulgaria.

The pedagogical aspect of this study included participant observation of four voice lessons, each instructed by a distinguished BFS pedagogue, including Professors Zhenya Stancheva, Gonka Galukova, Antoaneta Lefterova, and Dimitrinka Dineva. A total of six student singers participated in these lessons, including four self-identifying female singers and two self-identifying male singers. These teaching observations provided valuable insights into the pedagogical approaches and instructional strategies employed for vocal production and technique within this distinct folk singing tradition, with particular attention to the characteristics and execution of vibrato.

Finally, ethnographic interviews with Stoyana Todorova Karaivanova and the school's director, Maria Gradeshlieva, was employed to examine social, cultural, and musical underpinnings of vibrato in BFS. This qualitative approach facilitated an in-depth exploration through firsthand narrative accounts and semi-structured interviews. This method allowed for flexibility in discussing individual experiences while ensuring consistency across key thematic areas, such as training techniques, artistic interpretation, and the role of vibrato in stylistic authenticity. These combined methods offered a comprehensive understanding of the ways vibrato is taught, learned, and performed within the tradition.

The semi-structured interviews were guided by a series of core questions designed to explore practitioners' conceptualization and pedagogical approaches to vibrato within the Bulgarian folk singing tradition. These questions served as flexible discussion prompts that allowed for in-depth exploration of participants' experiences while maintaining consistency across key thematic areas. The following questions formed the foundation of the interview protocol:

- What do you consider to be vibrato in Bulgarian folk singing? Are there different types? If so, which are they?
- How do you classify vibrato variability? As a technical element or singing or a musical ornament? Why & how so?
- How do you teach vibrato, if at all? Explicitly? Implicitly? How do you approach or address vibrato in the vocal studio?
- How do you train or practice vibrato in singing?
- How do you produce vibrato? What is the source of vocal vibrato?

5.3.2 Acoustic Analysis Results

For the acoustic analysis of this exploratory study's recorded vibrato samples, similar approaches to those detailed in Chapter 3 were undertaken. It was immediately evident that average rate and extent metrics applied in Western vibrato research studies would not be directly applicable to the vibrato in BFS. Therefore, the collected vibrato samples were tested using the methodological framework of Coefficient of Variation (CV) in conjunction with the vibrato half-extent time profiles (as defined in Chapter 3) to distinguish different shapes and patterns of the highly variable time-varying vibrato contours found in BFS.

From the acoustic evaluation, a small number of samples emerged that aligned with the simple (monophasic) and complex (multiphasic) vibrato half-extent time profiles proposed in Chapter 3. However, the majority of vibrato samples collected from this field research did not ascribe to either of these classifications, instead demonstrating new vibrato archetypes that may be added to the current proposed analytical framework. The section below demonstrates a series of examples of vibrato samples typical in BFS, gathered from professional Bulgarian folk singers recorded in this field research acoustic study. Each vibrato archetype is associated with a distinctive regional style (outlined in greater detail in Sec. 5.3.4); an essential variable in BFS vibrato that warrants future research. It is important to note that all vibrato samples were analyzed using both the Half-of-Extent (HoE) and Relative Half-Extent (rHE) metrics.

One example from the vibrato samples that did demonstrate a uniform simple (monophasic) half-extent time profile is illustrated in Fig. 5.3. This vibrato sample is recorded from a singer specializing in music from the Dobrudzha region in Bulgaria. As the singer described after performing this song, the relatively higher frequency, mixed vocal registration, and regional aesthetic conventions dictate the Western classical-aligning regularity of the vibrato here. The CV fits within the normative range (established in Chapter 3) of 20-25%, underlining the similarity to the Western classical tradition. When comparing the analysis conducted with the HoE metric to rHE, the CV increased (HoE = 16.6%; rHE = 25.9%). When analyzed with the rHE metric, the half-extent time profile calculated by the AIC/BIC criteria switched marginally from uniform (monophasic) to complex (multiphasic) due to the presence of one half-extent value in the beginning. However, the shape of the logistic function was uninterpretable by the proposed methodological framework. Therefore, this vibrato archetype switch from uniform to complex is declared to be artifactual and the rHE profile is ultimately classified as uniform simple (monophasic).

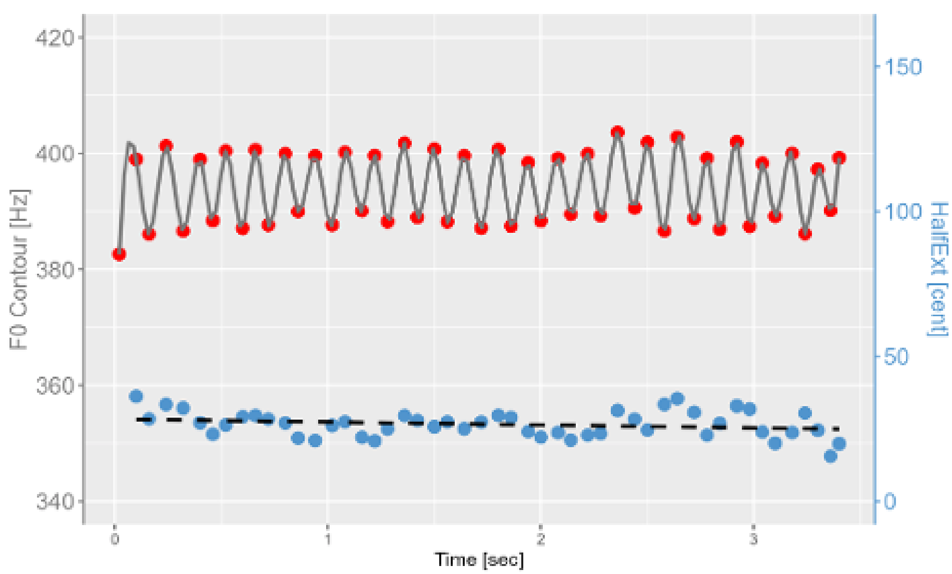


Figure 5.3: A typical uniform simple (monophasic) vibrato archetype / half-extent time profile analyzed with HoE and performed at approximately 392 Hz on the perceived pitch G4 (according to Western diatonic scales) by a singer of the Dobrudzha region.

The vibrato sample shown in Fig. 5.4, recorded from a singer originating from Bulgaria's Eastern Thracian / Strandzha region, demonstrates an atypical monophasic classification. This is a vibrato sample that the model declared to be uniform simple (monophasic), yet upon visual inspection, it is clearly diverging from the uniform vibrato samples classified in Chapter 3. The CV is much higher than any of the uniform type vibrato samples previously encountered; it increases from its HoE calculation (41.1%) to its rHE calculation (67.4%). While the classification methodology accommodates this highly irregular CV vibrato by classifying it as monophasic, there are time-varying changes in the symmetry of the vibrato peaks and troughs that are not captured by this model representation. The somewhat wider, alternating excursion of the vibrato is also characteristic of the region.

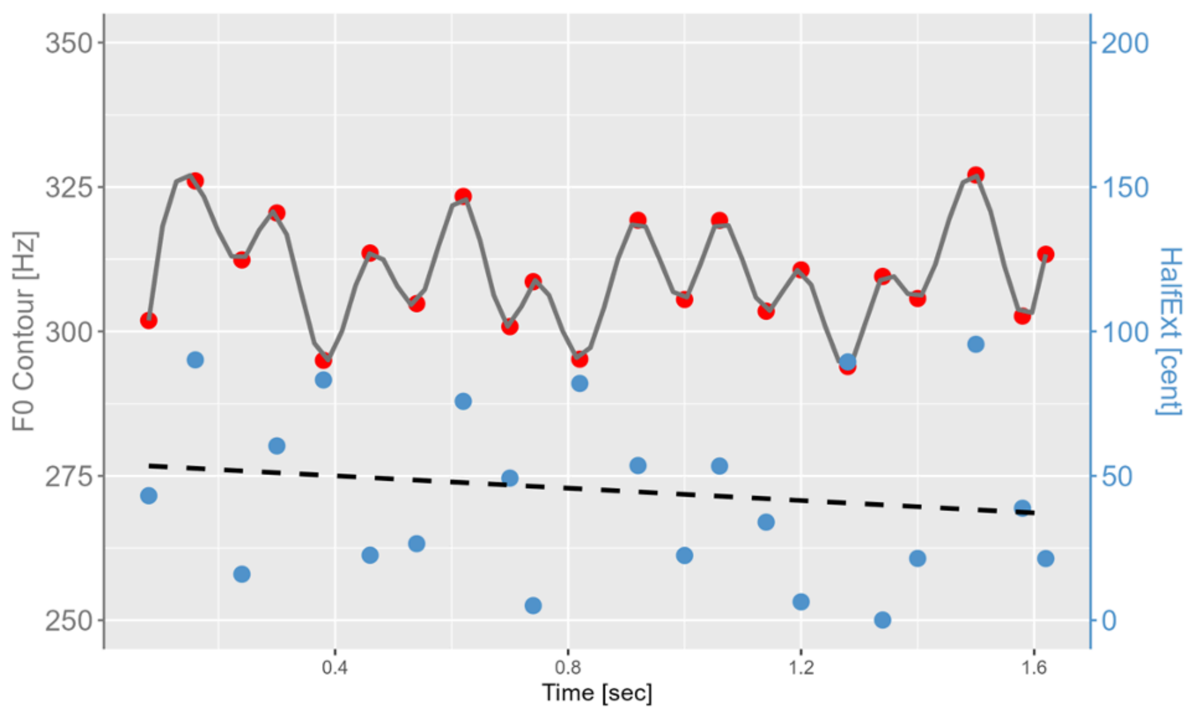


Figure 5.4: An atypical, irregular uniform simple (monophasic) vibrato archetype / half-extent time profile analyzed with rHE and performed between approximately 293.665-311.13 Hz alternating on the perceived pitches Eb4-D4 (according to Western diatonic scales) by a singer of the Strandzha region.

A third sample recorded from a singer from the Shopluk region demonstrated two highly typical vibrato-like ornamentations, one of which is quite distinctive. Тресене (henceforth referred to in its Romanized / Latinized version, *tresene*, roughly translated into English as "shaking"), is a rapid oscillation between pitches that blends elements of vibrato and trilling. Sounding akin to a yodel, this vocal ornament uses a non-legato, non-vibrato entry tone stroke transitioning into two main types (sometimes divided further into four subtypes) of vibrato-like *tresene*, *Граовско* / *Graovsko* and *Кюстендилско* / *Kyustendilsko* (Kerkelov, 2017).

This first type of *tresene* can be observed in Fig. 5.5. This *tresene* sample is one of the most interesting, yet highly irregular, vibrato half-extent time profile contour shapes encountered in this data set. The initial non-vibrato phase is followed by a vibrato phase. While the 4-PL regression using AIC/BIC correctly identifies this S-shape as a complex multiphasic type, the logistic fit is not as interpretable as it was for the Western music genres (Opera, Musical Theater, and Jazz styles) on which it was first applied. An S-shape here is difficult to interpret due to the highly asymmetrical yet quasi-periodic cycles at the onset of the *tresene*. Furthermore, the rate is too slow ($< 3Hz$) to be considered vibrato by Western music standards (normative Western vibrato rates lie between 5-7 Hz). Qualitatively, the shape of *tresene* departs from the Western vibrato paradigm, which categorizes vibrato as quasi-sinusoidal. In most *tresene* samples, including Figs. 5.5 and 5.6, there is a systematic trajectory of the pitch variation that is centered around a middle frequency, with a higher excursion above, and a lower excursion below.

The rHE calculation for the vibrato in Fig. 5.5 is much higher than its HoE counterpart (63.8 versus 42 cents), but the rHE CV is lower than the HoE CV (173% versus 221%). While these CV results are contrary to the general trend observed in Chapter 3, they are derived over the whole heterogeneous *tresene* sample, substantiating the claim that this pattern does not fully fit into the current methodological framework.

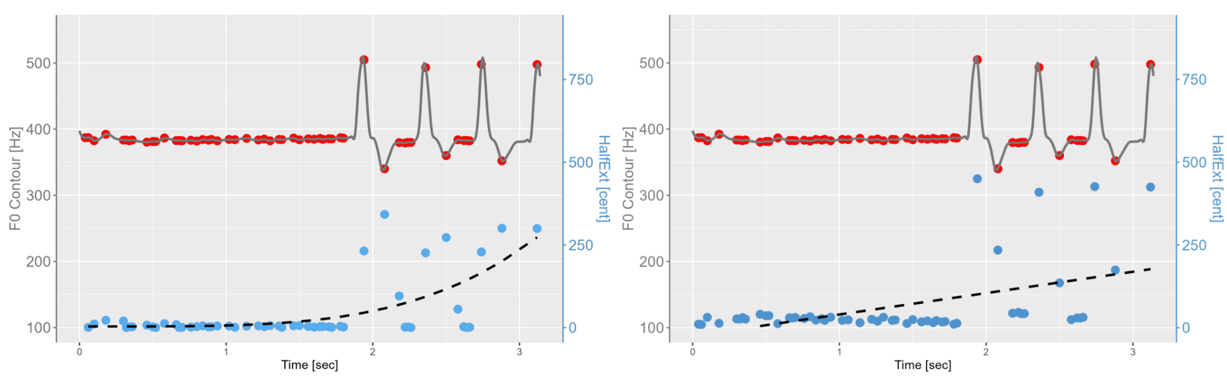


Figure 5.5: A side-by-side comparison of a *tresene* sample's complex multiphasic vibrato archetype / half-extent time profile performed at approximately 392 Hz on the perceived pitch G4 (according to Western diatonic scales) sung by a Shopska singer and analyzed with HoE (left) and rHE (right).

The second type of *tresene* appears in Fig. 5.6, a monophasic sample with the same vibrato-like ornament as the previous sample, classified correctly as simple (monophasic). In this case again the rHE is higher than the HoE (194 versus 176 cents), and the rHE CV is also somewhat higher (68% versus 64.7%). While this sample is also quite asymmetric, it retains quasi-periodic cycles, and the proposed classification framework somewhat accommodates this sample within the monophasic archetype. However, this sample significantly diverges from the uniform vibrato shape paradigm usually associated with monophasic contours in Western singing.

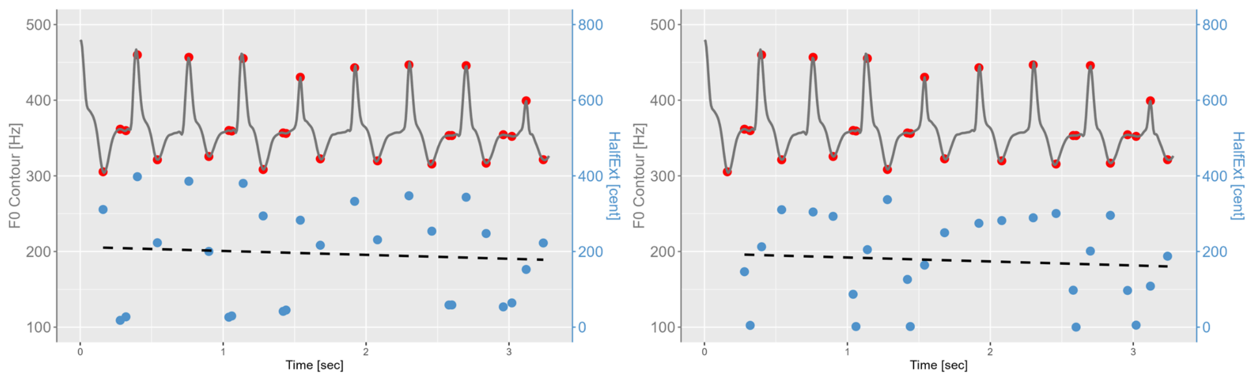


Figure 5.6: A side-by-side comparison of a *tresene* sample's complex multiphasic vibrato archetype / half-extent time profile performed at approximately 349.23 Hz on the perceived pitch F4 (according to Western diatonic scales) sung by a Shopska singer and analyzed with HoE (left) and rHE (right).

This distinctive ornamentation is achieved through subtle variations in laryngeal tension and breath support, creating a shimmering effect that adds emotional intensity to sustained notes. Some regions incorporate yodel-like breaks, in which singers rapidly shift between the chest voice and a falsetto-like quality, producing a percussive vocal leap reminiscent of Alpine yodeling. Another defining feature is the diaphonic singing style found in southwestern Bulgaria, northern Greece, and parts of North Macedonia. This practice involves one singer maintaining a steady drone while another performs a moving melody above it, creating a rich, pulsating harmonic texture. The precision required for diaphonic singing demands acute control over intonation and microtonal pitch shifts, as the harmonic interplay between the drone and melodic line is fundamental to the style's characteristic sound. For more on ornamentation in BFS, see the subsections within Sec. 5.3.4 below.

5.3.3 Acoustic Analysis Discussion

This research is the first known attempt to apply acoustic modeling of vibrato variability to BFS ornamentation. The analysis of BFS samples revealed distinctive and unique vibrato patterns and vibrato-like ornamentation, with certain types incorporating ornamentation that deviate significantly from the previously defined uniform monophasic and complex multiphasic vibrato. The proposed methodological framework demonstrated effectiveness in cases where the vibrato was monophasic, as indicated by correct classification and credible CV trends. However, when analyzing highly asymmetric, multiphasic vibrato patterns without accounting for their increased complexity, the methodology proposed in Chapter 3 proved insufficient. In such cases, an extension of the methodology is necessary to separate vibrato phases and derive a more consistent measure of variability for signals that do not conform to the uniform vibrato shape paradigm.

Conversely, a number of the BFS samples adhered to the uniform vibrato shape paradigm, for which the proposed methodological framework was fully applicable. Overall, BFS appears to employ a few time-varying vibrato contour patterns that overlap with those found in Western singing genres, but the majority of these time-varying vibrato contours feature region-specific ornamentation that does not align with the defined vibrato models. The presence of these distinct ornamentation patterns underscores the need for methodological refinement to accommodate the complexity of vibrato in BFS.

The use of vibrato half-extent in its rHE form implies that its CV reflects the variability of the vibrato extent while also accounting for the vibrato extent's asymmetry around a constant middle line (i.e., the intended f_o or heard pitch). Therefore, as seen in most of the BFS samples, it can be expected that rHE CV estimates for a vibrato sample are higher than the corresponding HoE-based CV estimates.

These findings demonstrate that in BFS, more complex types of the vibrato half-extent time profiles exist and this increased complexity is a consequence of inherent stylistic elements of vibrato-like ornamentation, reflected in the contour itself. One way to handle this situation requires more nuance and approach modification by decomposing each sample into several biphasic segments, and sequentially applying the 4-PL logistic regression to each of the segments (see Fig. 3.10 as an example). However, a more refined approach to fully characterize non-Western vibrato like the one found in BFS is needed, due to the asymmetric complexities integral to this specific BFS vibrato-like ornamentation. This is the subject of ongoing and future research work by the author of this thesis.

5.3.4 Qualitative Report of Ethnographic Pedagogical Study

Through this acoustic and ethnographic pedagogical study, it becomes evident that vibrato training in BFS prioritizes an approach centered on stylistic tendencies and ornamental practices dictated by regional and dialect differences. As with Western singing genres, many of the professionals and pedagogues interviewed and observed had stark, differing opinions on both vibrato production and usage. These were influenced by their empirical experiences and regional specializations. Some believe that vibrato is a natural aspect of ornaments, while others disagree.

The synthesized content of the following report was derived from a systematic thematic analysis of detailed field notes obtained during semi-structured interviews and narrative accounts, with themes prioritized based on their frequency and salience across participants' perspectives to reflect both common agreements and notable divergences.

Vowels in Regional Dialects: Foundations of BFS Pedagogy

It is typical for Bulgarian folk singers to specialize in the style (including vibrato practices) of the region from where they originate. Historically, singers were autodidacts or at most trained informally, through familial and community-based call-and-response type instruction that was passed down from generation to generation. This tradition continues in the present day, as students, even if they are trained in an institution located outside their region, tend to eventually specialize in the repertoire of their original region. While school-based study requires students to learn to sing with vocal exercise and repertoire assignments from all regions of Bulgaria, as students develop, they are encouraged to specialize, with less emphasis placed on versatility across regional styles. This is also clear in recommended pairing of students and teachers from similar regions in secondary and higher education.

Vowels and their dialectal differences based on regional specificities are regarded by some as the foundation of BFS pedagogy. From participant observations of voice lessons as well as ethnographic interviews, it became evident that this arises from the dynamism of the Bulgarian language itself, and its many geographic variations. In singing training, vocalises or warm-ups are conducted on vowels most typical for each region (i.e., Thracian regions vocalize more on /a/ than the Rhodope region, which involves more open /ɛ/ vowels, and Shopluk, which dialectally changes many final vowels). Vowel changes are explicitly taught based on repertoire, and ornaments are usually taught on single vowels, with attention given to which vowels are open and which are closed.

Bulgarian folk songs exhibit several specific linguistic characteristics, such as unconventional spellings that reflect phonology, distinctive morphological and syntactic structures, a unique vocabulary, and specialized poetic forms, oftentimes archaic and outdated. While Bulgarian folkloric texts themselves are often described as incorporating "dialectal" language, these songs contain relatively few linguistic elements that can be attributed to specific regional dialects— varieties of speech confined to particular geographic areas (Girvin, 2016). The singer is the one who imbues these texts with phonetic differences based on their region of origin or ideally the region of the song itself.

The Bulgarian language has two main dialects, as shown in Fig. 5.7. The Western dialects form the basis of Standard Bulgarian and are informally called *твърд говор* / *tvurd govor* or "hard speech", akin to Ukrainian, while Eastern dialects are informally called *мек говор* / *mek govor*

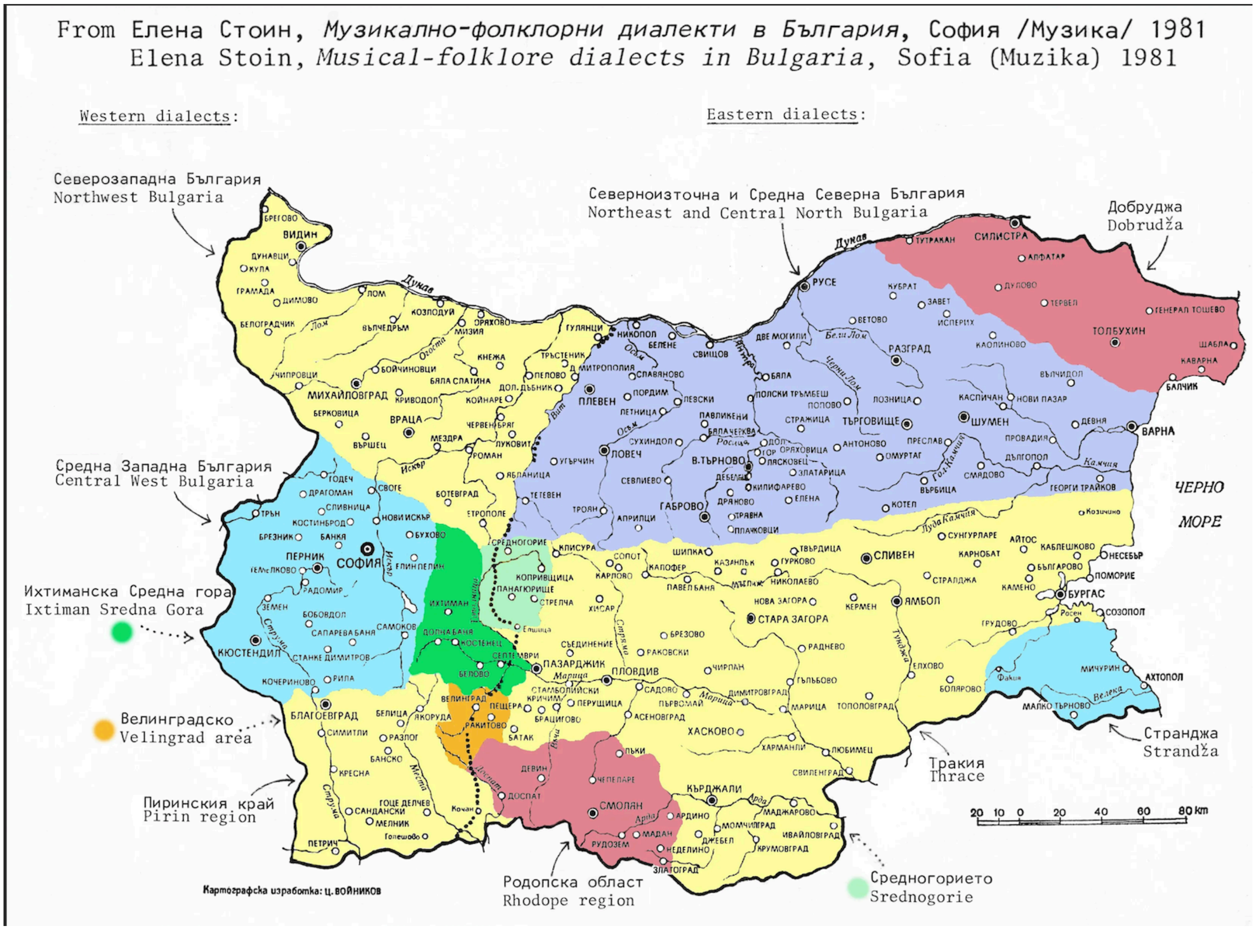


Figure 5.7: A map of dialect division in Bulgaria, originally by (Stoin, 1981) and modified with English translations by (Forsyth, 2016).

or "soft speech" and have a greater tendency to open vowels and palatalize consonants, akin to Russian (Nestorova et al., 2025a). The *ят* / *jat*, the historical Slavic vowel sound represented by the Old Church Slavonic letter, "Ѣ," whose pronunciation and spelling has undergone significant changes over time, is arguably one of the most important dialectal distinctions in Bulgaria. . This distinction is known as the "yat border", a linguistic divide (shown geographically in Fig. 5.8) between the Eastern and Western dialects of Bulgarian. This phenomenon appears quite prominently in the training of BFS, influencing not only vowel length and duration, but also vocal registration and ornamentation. Dynamic stress also differs by region, impacting vowel reduction and timbral quality in speech and singing alike.

Westernization of Institutionalized Training

In 1953, Filip Kutev established several foundational principles of Bulgarian folk music, which are still referenced and highly regarded as confirmed in the ethnographic pedagogical study. Kutev

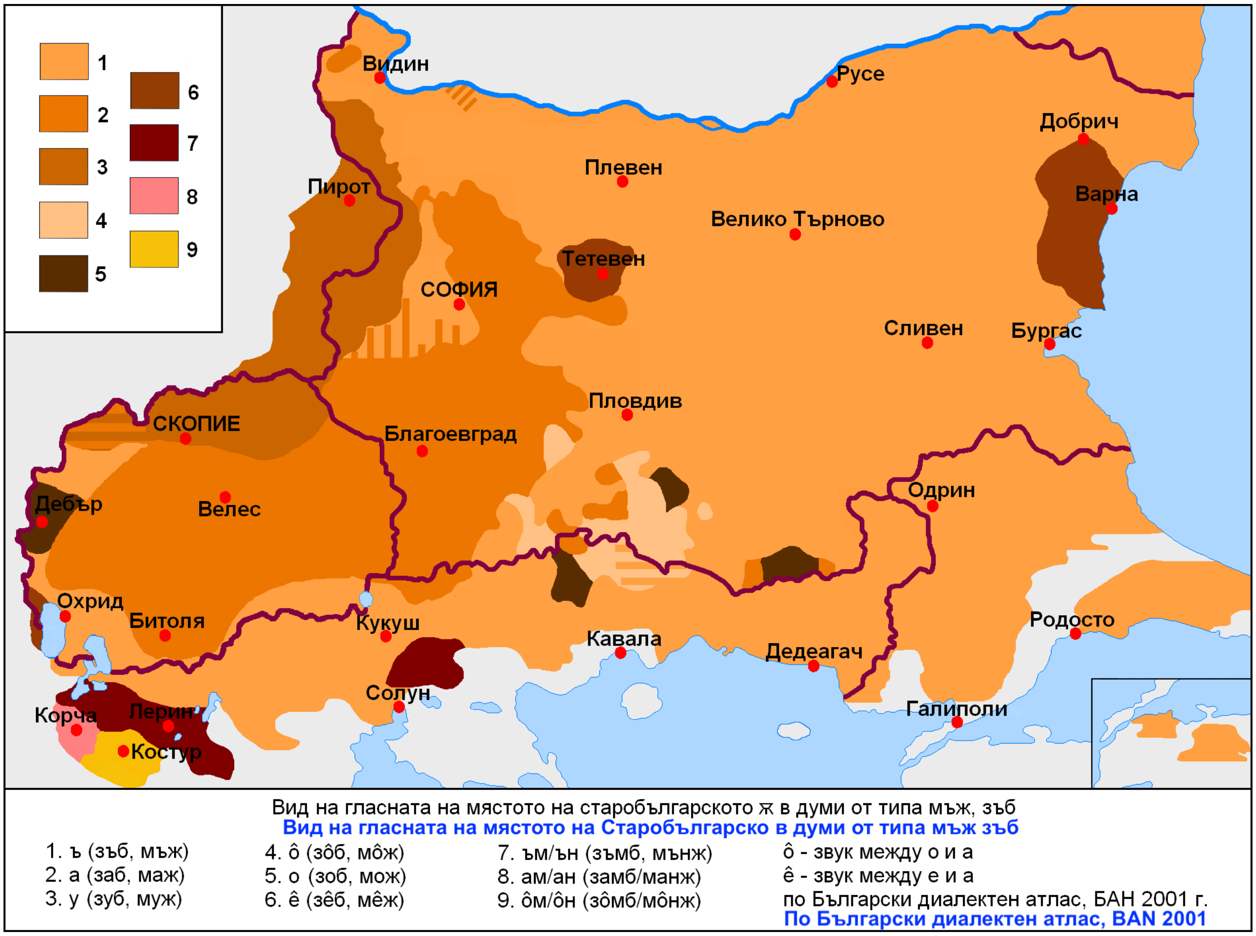


Figure 5.8: Map of the big *ям/jat/yat yus* isoglosses in Eastern South Slavic and eastern Torlakian according to the Bulgarian Academy of Sciences' atlas from 2001 (Kochev, 2016). Pronunciation of man and tooth, derived from Proto-Slavic words "мъж" and "зъб" on the map.

delineated the core tenets of the BFS "school" as follows: "open-throated singing, limited pitch ranges centered around B-flat for sopranos, even lower pitch ranges centered around D-flat and F-sharp for altos, a sonorous timbre, proper breathing, clear diction, a simple yet profound expression of the song, and most of all, beautiful, intricate ornamentation, etc..." (Borisova, 2013). This formalized approach to BFS led to the development of distinct stylistic features associated with specific ethnographic regions. As Bulgarian folk music became embedded within academic institutions, these features were increasingly systematized into standardized, "representative" markers of regional styles. Consequently, many of the idiosyncratic nuances of traditional folk singing may be overshadowed if they diverge from institutionalized standard norms. During the ethnographic interviews, several of the professionals indicated the Rhodopean breathing technique (characterized by specifically placed inhalations at the end of each physiological breath, which may appear mid-phrase in the song) as one of the instances of this, as it is deemed unsuitable for staged performances and is therefore excluded from formal training.

The majority of the pedagogue and performer professionals in the ethnographic interviews highlighted the transformation of BFS through the eras in its pedagogy, performance practice, and aesthetic ideals. Traditionally passed down orally within communities, folk singing was formalized in the mid-20th century through state-sponsored conservatory training programs. Influenced by Westernized methodologies, modern instruction of BFS now emphasizes standardized vocal technique, notation-based learning, and staged performance aesthetics. Several professional felt that these new practices diverge from the raw, regional authenticity of traditional singing. One notable shift observed in the voice lessons was the increasing use of piano accompaniment, replacing or supplementing traditional folk instruments like the *гъдулка* / *gadulka*[8], *кавал* / *kaval*[9], *тамбура* / *tambura*[10], and even the *каба гайда* / *kaba gaida* [11]. See Fig. 5.9 for an authentic photograph of this instrument. While this adaptation aligns with Western harmonic frameworks and facilitates structured training, it alters the modal, microtonal, non-equally tempered tuning nuances characteristic of more traditional performances of Bulgarian folk music.

However, the pedigree or master teacher lineage and preservation of oral and aural transmission are reflected in the modeling and imitation that still remain the primary didactic methods in voice lessons, especially for beginner students. For advanced students, the use of musically notated songs (as opposed to solely text) is encouraged. This impacts their vibrato production, as several of these notated songs include some symbolic notation or directives regarding vibrato-like ornamentation.

Much of the material used in the vocalises, warm-ups, and repertoire sections of the voice lessons were derived from published étude technique books, intended for the singer in training to develop vocal and technical skills in songs of the different regions. Several of these books (compiled by Diamandiev, Shopov, Grozdanova-Radeva, Karadzhova and Toncheva, and many others), while still predominantly hand-written rather than type-set, were adapted to Westernized solfège symbols and single-syllables. An example of this can be seen in Fig. 5.10, where the red box indicates a vibrato-based ornament, which is a hybrid between a trill and mordent in Western art music. Surrounding this, there are several notated versions of vibrato-like ornament vocalises and exercises with corresponding regions indicated.

A few of the exercises, harmonies, and principles of BFS have even been adapted into Western classical singing training in Bulgaria, with composers such as Zlatev-Cherkin publishing a

[8] A traditional Bulgarian bowed string instrument, held vertically and played with a bow; it has no fingerboard, and the strings are stopped by the sideways pressure of the player's fingers (Todorov, 1966).

[9] An end-blown flute common in the Balkans, traditionally associated with shepherds; it is fully open at both ends and played by blowing on the sharpened edge of one end (Todorov, 1966).

[10] A long-necked, metal-strung fretted lute used for rhythmic accompaniment and melodic solos in Bulgarian folk music; it resembles a mandolin and is played with a plectrum (Todorov, 1966).

[11] A traditional Bulgarian bagpipe made from animal skin, featuring a single drone and a chanter; it is a key component of Bulgarian folk music, often played solo or with a large drum (Todorov, 1966).

Figure 5.9: Photograph of a *гайдар* / *gaidar* / "Bulgarian bagpiper" with a Bulgarian folk singer, both in traditional garb from the *Шопската област* / "Shop region," after an in-the-field performance demonstration, acoustic recording, and interview.

vocalise book based on Folk Tonalities. Many of the pedagogues in the voice lessons incorporated cardinal vowel strings of /a-e-i-o-u/ inspired Western classical singing training, and, in a few cases, functional voice exercises, such as /m/ hums and other semi-occluded vocal tract positions.

While remaining true in origin, the pedagogical landscape for BFS is evolving, as admitted by the interviewed professionals, with an emphasis on interdisciplinary approaches that blend traditional methods with modern educational practices. In the increasingly digital 21st century, this includes the incorporation of technology and multimedia resources to enhance learning experiences and accessibility. Therefore, many of the voice lessons included reviewing recorded performances as inspiration or for critical analysis, exhibiting the integration of technology in teaching Bulgarian folk music (Petrova-Kirkova, 2023).

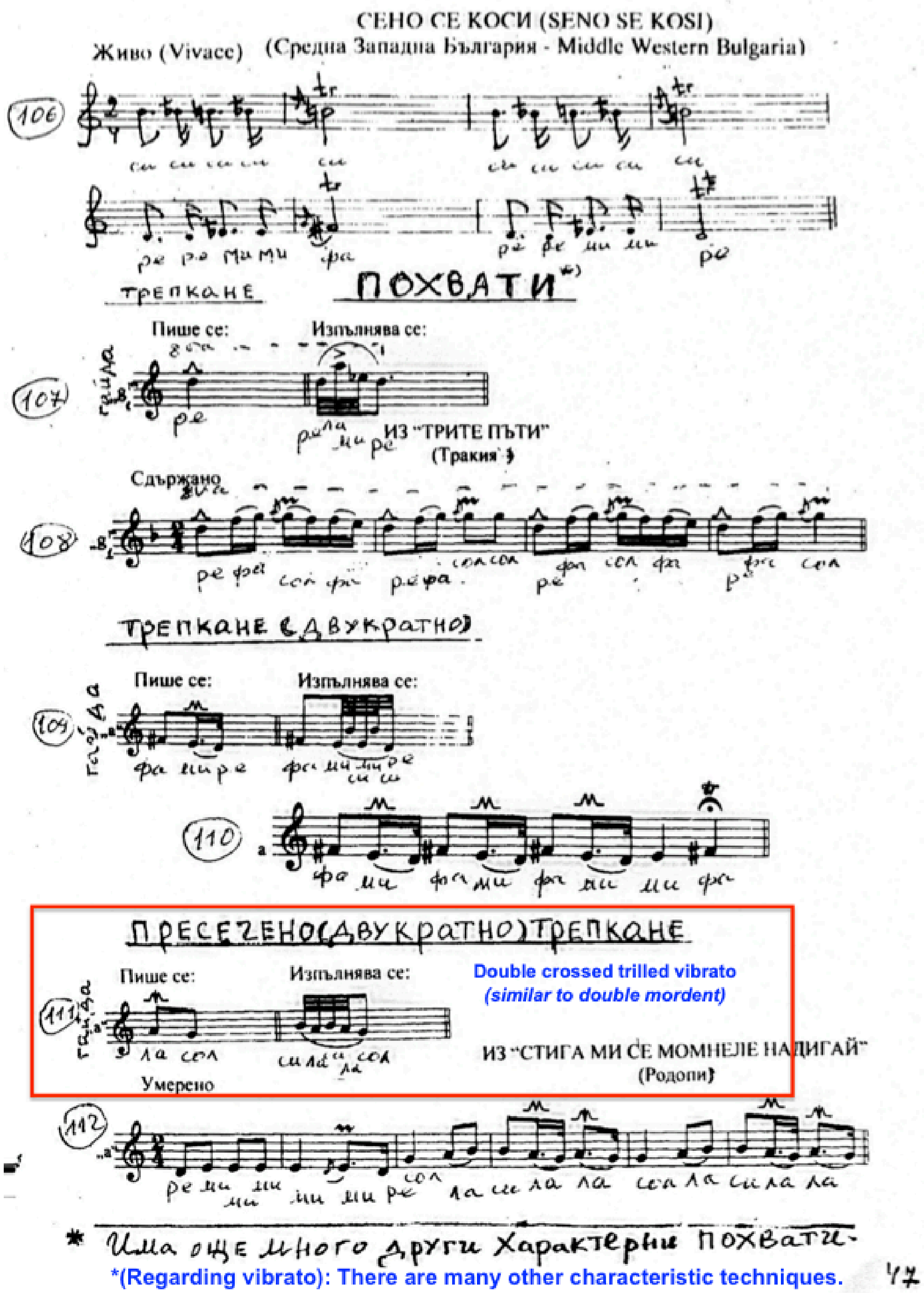


Figure 5.10: A source used in the participant observation voice lessons; an excerpt from a vocalise book compiled in 2016 by Konstantin Shopov.

Genre-Specific Technical Components

Divergent from the common postural standards of head, neck, and body alignment in Western singing, in Bulgarian Folk Signing, several divergent articulatory and postural specifications are necessary as a byproduct of production physiology. In many of the voice lessons, singers were instructed to retain an upward head angle and conical vocal tract shaping with the articulators (including narrower pharynx, higher larynx, horizontal mouth space, and spread lips) to facilitate vocal production within the sonic aesthetic.

The embodied nature of BFS pedagogy was another key finding from this study. Observations of pedagogical sessions revealed that many singers practiced step sequences and rhythmic body motions alongside vocal exercises, facilitating a deeper connection between bodily movement and vocalization. Ethnographic interviews further highlighted how this integration preserves regional stylistic nuances, enabling singers to internalize the asymmetrical rhythms innate to Bulgarian folk music.

A critical component of BFS is the use of a twang-like resonance (with a compacted spectral energy between 1–3 kHz, specific vowel articulation, and higher vertical laryngeal position). This strategy, akin to that typically ascribed to the Western Country Singing genre, tracks with prior spectral analyses and the voice lesson observations in this pedagogic study of BFS. One of the pedagogues interviewed attributed this twang-like sonority to the engagement of the aryepiglottic sphincter, enabling the pharyngeal production of a bright, metallic sound while ideally minimizing strain on the vocal folds. Ethnographic interviews further elucidated the cultural and musical significance of these techniques, highlighting the necessity of a powerful and penetrating vocal timbre for outdoor and communal performances.

One of the sonic identities of BFS was defined by many of the interviewed pedagogues and professionals as "white voice," described as an archaic vocal technique produced through speech-level phonation, high laryngeal position, forward resonance, and appropriate breath management strategies. A central tenet of this technique is the development and emphasis of the *passaggio*[12] as a register of its own, manipulated with various degrees of registral mix, alongside contributing factors such as nasality and timbral quality. This "white voice" resonance strategy is distinct from Western classical norms, which favor seamless perceptual resonance shifts and transitions facilitated through a blending of distinct yet unified registers.

An empirical conclusion regarding the vocal production was further drawn from this resonance strategy. Typical BFS vocal production likely utilizes laryngeal mechanism M1 (or M1.2, at most M1.5), physiologically relying on a high degree of cricothyroid (CT) and thyroarytenoid (TA) muscle engagement, with limited medialization of the vocal folds and minimal laryngeal tilt.

[12] A commonly used classical voice pedagogy term for the acoustic and/or physiologic transitions between registers. Sometimes referred to as a break when less smoothly executed (Hoch et al., 2022).

The thyroarytenoid (TA) muscle is predominantly active, resulting in high vocal fold contact, increased subglottic pressure, and a vibratory pattern involving the full vocal fold mass. Generally, one observes higher, thoracic breathing in BFS, as compared to diaphragmatic breathing in Western classical singing. Given the idiosyncratic nature of breathing strategies among individuals, a more precise investigation using respiratory inductance plethysmography (RIP) is warranted.

Since the turn of the 21st century, a modern trend of crossover genres and genre-mixing has heavily influenced the pedagogy of BFS and impacted vibrato, as well. With more Bulgarian folk singers performing folk songs with Contemporary Commercial pop/rock stylistic components (including microphones, electric instruments, etc.), the approach to vocal production is also changing.

Vibrato-Like Ornamentation

Ornamentation in BFS is a defining stylistic feature that varies significantly across regions, contributing to the rich diversity of the tradition. These embellishments serve not only as expressive devices but also as markers of regional identity (with a divide between the rich ornamentation of the northern compared to the southern regions of Bulgaria), each dialect of singing employing distinct melodic and rhythmic nuances. In the Shopluk region, for example, singers frequently use rapid, tight mordents and trills, creating a sharp, percussive vocal effect. By contrast, the Rhodopes are characterized by wide, slow, and undulating melodic ornaments that emphasize sustained, modal phrasing. The Thracian style incorporates intricate melismatic figures, often weaving between neighboring pitches in elaborate, cascading runs, while Dobrudzha singing features short, sharp glottal ornaments that enhance rhythmic precision. Findings from this acoustic and ethnographic pedagogical study suggest that vibrato is one of the most distinctive ornaments across all regions.

Vibrato in BFS was confirmed in this ethnographic pedagogical study to not necessarily be a default vocal quality, but rather an intentional ornament, often employed at the end of phrases or as a means of intensifying emotional expression. Some regional styles, such as those of the Pirin region, favor an open-throated, pulsating vibrato, while others maintain a predominantly non-vibrato, straight tone, reserving vibrato for moments of heightened musical tension. These varied approaches to ornamentation, deeply rooted in oral tradition, are integral to the authenticity of BFS and serve as a key differentiator from Western vocal techniques.

The following lists several common styles found in BFS, as gathered from the study, with particular attention given to vibrato-like ornamentation in each.

1. **Shopluk style:** Vibrato is not an ornamental characteristic, except as a facet of the most distinguished ornament of this style, *tresene*: rapid oscillation between pitches that blends

elements of vibrato and trilling (especially *Trunsko tresene* ("Trun-like shaking") . Vibrato may be preceded by the other Shop ornaments of *prihlutsvane* ("hiccuping") and *chukane* ("knocking").

2. **Rhodope style**: Vibrato is added occasionally as a part of the *grupetto* and *glissando* ornaments. Commonly sung in duets with two or more voices, the lower voice imitating the single-note drone of the Bulgarian folk instrument akin to the bagpipes, the *gaida*.

3. **Strandzha style**: Use of wide, swinging vibrato in combination with the most frequent ornament, the truncated mordent.

4. **Pirin style**: Non-vibrato entry tone, gradually transitioning to vibrant portion. A specific open sound production in combination with singing of *atzane* "yodeling" and sharp registral changes of falsetto vocalisation. Often sung in a duet-style with 2 voices.

5. **Thracian style**: Common *roulade* and *cascade* ornaments may include vibrato, but usually appear in a smaller pitch range and unmeasured songs.

6. **Dobrudzha, Northern, and Northwestern styles**: Vibrato is typically more refined, discreet, and omnipresent; closer to Western Art Music standards, with regularity and consistency.

Additionally, several professionals pointed to the transmission of accepted regional vibrato-like ornamentation practices through the legacy of notable Bulgarian folk singers. Many interviewed professionals referenced the uniquely vibrant timbre of Atanaska Todorova as well as the lighter, "head" voice (cricothyroid, CT-dominant) registration of Vulkana Stoyanova (Thracian region), the instrumentally influenced vibrato of Nadka Karadzhova and Radka Kushleva (Rhodope region), the narrow vibrato extent sung by Jovcho Karaivanov (Thracian region), the subtle, consistent, and persistent regular vibrato (quasi-Operatic) of Boris Mashalov (Northern regions), and the delayed onset vibrato (non-vibrato to vibrato) ornamentation performed by Kostadin Gugov (Shop and Macedonian regions).

Vibrato Impact on Voice Register, Classification, and Training

Voice classification in BFS is conducted according to the primary criterion— vocal pitch range— with two overarching groups, *tezhuk* "heavy" and *lek* "light" voices further divided into three main groups: high, medium, and low, which correspond to the traditional folk designations of first, second, and third voice, respectively. A unique factor in voice type classification based on vibrato emerged over the course of the voice lesson observations and ethnographic interviews.

Singers with ease and control over a versatile vibrato were encouraged to execute complex ornamental patterns, even at the extreme upper limits of their vocal range. In contrast, singers with a more naturally attenuated vibrato were inclined toward vocal effects in BFS such as "shouting," "hissing," and other extra-ornamental stylistic elements (Vekilova and Minkova, 2010).

In addition, observation of the voice lessons revealed that vibrato may be more accepted in songs sung by lighter voices with higher pitch ranges. This is corroborated mechanistically, as higher pitch ranges often necessitate the use of cricothyroid (CT) (alongside the opposing activation of the thyroarytenoid (TA)) production, which differs from the standard vocal registration approach in BFS. This balancing of agonist-antagonist laryngeal muscles are accepted as a physiological stimulus for the initiation of vibrato. On the other hand, the typical thyroarytenoid-predominant (without any activation of the opposing cricothyroid) production of the generally lower pitch ranges found in BFS may attenuate the initiation of vibrato.

Many pedagogues in the ethnographic pedagogical study defined a *самородно* / *samorodno*, or"purely native" vibrato that comes naturally to some students. While vibrato instruction in BFS appears more explicit (due to its link to ornamental practices) than in most Western training, this premise was accepted by many vocal pedagogues. For more on implicit versus explicit vibrato instruction, see Chapter 6.

The development of a controlled vibrato is a fundamental aspect of BFS vocal training. Targeted exercises, including sustained phonation on long tones without vibrato, were commonly observed in voice lessons. Given the individual nature of each student's vocal characteristics, strengths, and challenges, most instructors adopted a highly personalized approach to training, including addressing vibrato.

A recurring issue raised by pedagogues in the interviews was voice tremor, which is commonly associated with excessive tension, instability of the lower jaw, and involuntary trembling of the larynx. Pedagogues noted that this phenomenon often indicated over-extension of the vocal folds beyond their physiological limits. To address this, some implemented progressive exercises drawn from Meleki (2021): beginning with phonation on closed-mouth consonants /m/, /b/, /p/ transitioning to sustained vowel production (initially /a/ with a relaxed jaw), and gradually extending the duration of non-vibrato phonation. Further training incorporates all five cardinal vowels /a/, /ɛ/, /i/, /o/, /u/, and to enhance stability. Additionally, instructors reported that certain consonants /p/, /l/, /s/, /h/, /k/, /t͡ʃ/, /ʃt/, /ʃ/ present articulation challenges, which were mitigated through targeted phonetic and musical exercises embedded within the repertoire. These are exhibited in Fig. 5.11 as a part of a vocal pedagogy textbook referenced often by the teachers interviewed.

The mental, emotional, and physical state of the singer was also recognized as a critical factor in vocal stability (and vibrato) by many vocal pedagogues. Instructors reported that psycholog-

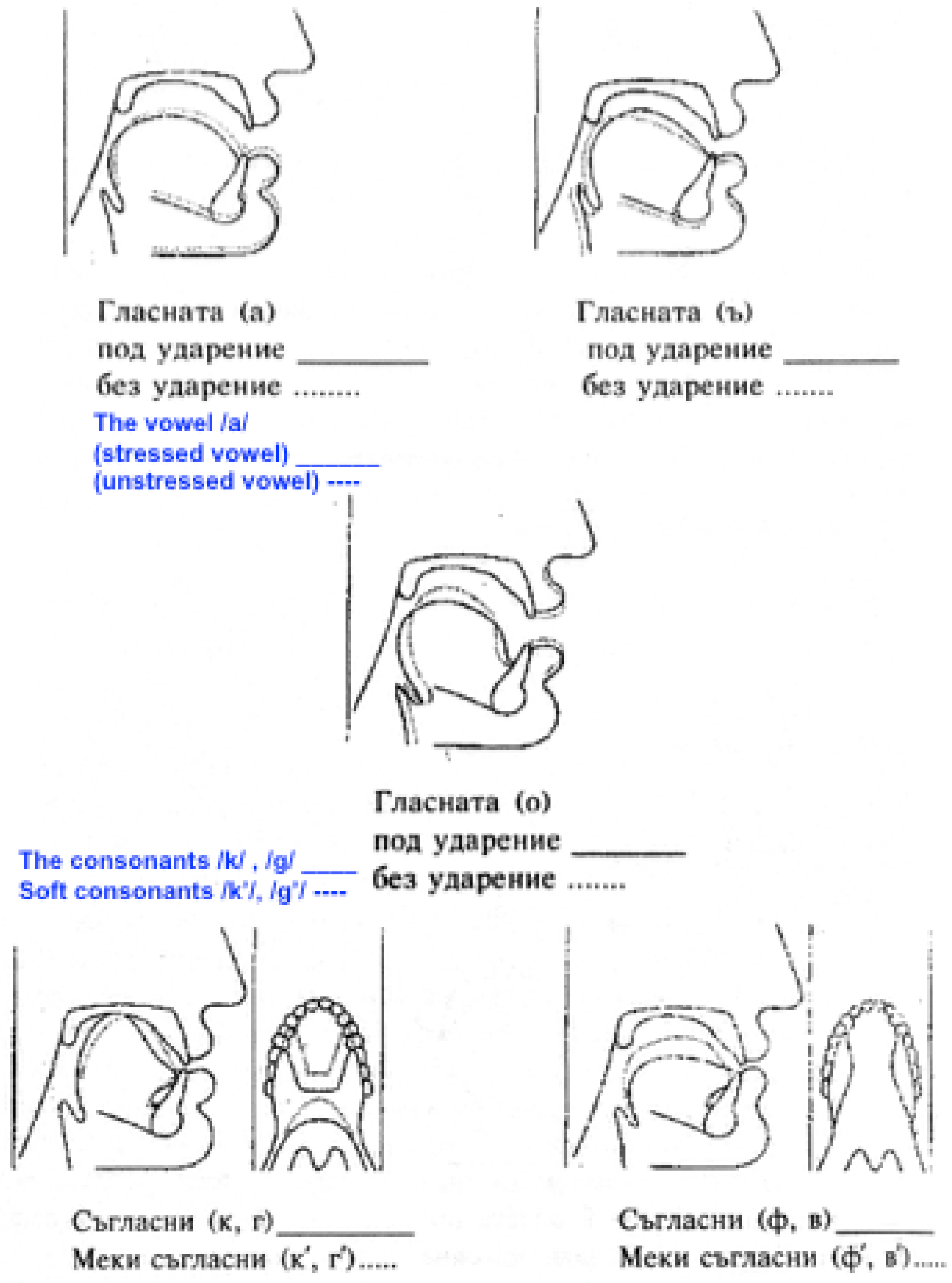


Figure 5.11: Sagittal view vocal tract and transverse view oral cavity diagrams of tongue contact and articulator approximation locations with movement trajectories of Bulgarian vowels (stressed and unstressed versions) and consonants (hard and soft versions) (Kushleva, 2000). See overlaid translations.

ical stress, along with hormonal fluctuations— notably during voice mutation in adolescence or menstruation in female singers— can significantly impact vocal function. Consequently, these instructors advised against excessive vocal strain during these physiological transitions.

Ensuring voice optimization, efficiency, and vocal health are key components of vibrato pedagogy that involve addressing stylistic authenticity. Given the intensity of BFS, vocal fatigue and hyperfunction[13] can be concerns, especially for singers transitioning from other vocal traditions. Care should be taken that flow instead of overly pressed phonation is practiced (especially when producing non-vibrato or straight tone singing), muscular maneuvers are often interspersed with opposing motions, and that any discomfort is quickly mitigated. A balanced pedagogical approach should therefore incorporate structured warm-ups, cool-downs, and targeted exercises to maintain vocal longevity while preserving the stylistic integrity of the genre.

5.3.5 Conclusions, Limitations, and Future Work

"Our folklore is vibration" was an apt quote from Prof. Maria Gradeshlieva, uttered at the opening ceremony for the new academic year at the Philip Kutev National School of Folk Arts in Kotel, Bulgaria. Vibrato appears as a core aesthetic feature in Bulgarian folk singing ornamentation, which is highly variable based on regional geographic, dialectal, and cultural factors. BFS itself represents a highly specialized vocal tradition that integrates unique physiological, acoustic, and stylistic elements, including vibrato.

This pilot acoustic and ethnographic pedagogy study built upon previous work presented in Chapter 3, which examined vibrato across multiple Western musical genres, including Opera, Musical Theater, and Jazz. Not only are commonly utilized average metrics unsuitable for analyzing vibrato or vibrato-like ornamentation in non-Western styles, but also the proposed methodological framework presented in Chapter 3 may not be able to accommodate the higher level of irregularity and variability inherent in non-Western styles. While the vibrato half-extent time profiling methodology captures the nuances found in BFS, the Westernized classification system does not fully and accurately represent the highly specialized use of vibrato in this genre. Since prevailing vibrato metrics assume stability and uniform periodicity, they do not adequately account for the stylistic, acoustic, and expressive variations observed in BFS.

A significant factor affecting vibrato in both the acoustic and pedagogical investigations was the diversity of the various regions, each of which has its own distinctive ornamental musical features, influencing the various vibrato types and divergent approaches to biomechanical production. Through observed voice lessons, it was noted that effective pedagogical approaches integrated traditional oral transmission methods with evidence-based voice science principles.

[13]Hyperfunctional phonation recruits the muscles of the larynx in a manner that is more taxing or straining than necessary to produce speech or singing (Hoch et al., 2022)

Aesthetics reinforcing the genre's unique sonic identity and deep historical continuity also pervade as key factors in vibrato development and instruction.

Advances in voice science and pedagogy, including the proposed vibrato or vibrato-like ornament modeling in this chapter, offer valuable insights into how this technique is produced and how it can be effectively taught. By combining traditional oral transmission with modern vocal research, pedagogues can ensure that this rich tradition is both preserved and adapted for contemporary singers, fostering both artistic expression and sustainable vocal practice.

Moreover, there is an increasing trend in the West, especially in North America, for the training and performance of BFS by non-native individuals. With over 200 ensembles in the United States (surveyed in 1998, with a presumably increasing number since then) either specializing in or including Bulgarian and/or Pan-Balkan Peninsula repertoire (Webster, 2004; Lausevic, 2015) and the world-renowned Koprivshtitsa Folk Festival inviting both amateur and professional Bulgarian folk music and dance ensembles from across the globe (MacMillan, 2015), there is a globally growing interest in Bulgaria's unique cultural art, yet very few sources address stylistic specificities, such as vibrato. This acoustic and ethnographic study sought to fill this gap.

As this was a pilot ethnographic pedagogical study, it was exploratory in nature, therefore, any definitive conclusions may not be generalizable across all BFS. However, the findings of this acoustic and ethnographic pedagogical study, derived from qualitative data gathered through participant observation of voice lessons and ethnographic interviews with professionals, provided valuable insights into the instructional methodologies and cultural underpinnings of vibrato in BFS. This emphasized the validity of this mixed methods acoustic and ethnographic pedagogy study with its quantitative parametric models and qualitative interview responses. These findings further contribute to comparative voice pedagogy and cross-cultural vocal research, expanding our understanding of vibrato's role in diverse global singing traditions.

Chapter 6

Pedagogy of Vocal Vibrato: Integrative Review and Teaching Recommendations

Author Contributions

This chapter includes restructured and reintegrated content from the following research articles: Nestorova, T. (2026, in press-a). To teach or not to teach vibrato?: Implicit versus explicit instruction in the vocal studio; Part I: Vibrato pedagogy in past to present pedagogical trends. *Journal of Singing*. [Manuscript accepted; expected publication in early 2026].
Nestorova, T. (2026, in press-b). To teach or not to teach vibrato?: Implicit versus explicit instruction in the vocal studio; Part II: Practical and applicable tools, tips, and tricks. *Journal of Singing*. [Manuscript accepted; expected publication in early 2026].

Author-specific contributions: *Theodora Nestorova:* conceptualization of the study, data acquisition, analysis of results, development of interpretation and conclusions, writing, review, and editing of the manuscript.

Chapter-specific contributions and differentiation from original journal publication: In this chapter, additional information that was not originally included in the article manuscript is included to connect the content and material originally found in the first and second article manuscripts. Finally, editing for grammar and flow were conducted to optimally incorporate the article contents into this thesis chapter.

6.1 Introduction

When it comes to vibrato in singing, “to teach or not to teach?” is indeed the question. Fiercely debated by singing teachers through the ages, the subject of vocal vibrato still raises unresolved pedagogical questions, including whether it should be implicitly or explicitly taught. It is commonly accepted that vibrato is a distinctive feature of singing, inextricably tied to both technical training and musical expression. As Hoch et al. (2016) state:

> One cannot be a singer without some understanding of vibrato. Absorbing some of [its] facts... is a first step toward unraveling some of the mysteries surrounding vibrato and unlocking its many "secrets."

It is a pedagogical imperative that voice teachers not only understand vibrato but also apply and evaluate its presence and function in their students' voices. It is, however, not commonly accepted how to teach vibrato, or whether to address vibrato in the voice studio at all. To this day, there exists no codified teaching methodology for vocal vibrato. Without a consensus on best practices, vibrato instruction remains a highly individualized process, varying widely across vocal studios and pedagogical philosophies. Vibrato's role in vocal pedagogy has been the subject of ongoing debate primarily due to the individual aesthetic preferences and collective traditional conventions staunchly upheld across a multitude of musical genres, stylistic contexts, and solo versus choral singing settings. In the vocal world, the threshold of vibrato "permissibility" may very well be the proverbial boundary that keeps genres siloed.

In examining past to present trends in vocal pedagogy, it becomes evident that the implicit versus explicit vibrato instruction debate is not one of absolutes but of pedagogical balance. Sections 6.1-6.4 of this chapter explore how historical and contemporary teaching philosophies across genres have shaped vibrato training, offering insights into effective practices and potential implications. In Sections 6.5-6.8, practical exercises and tools are recommended, as well as methods for integrating vibrato instruction as "FUNctional Vocal Pedagogy with a Student-Centered Approach" into the vocal studio, all while considering the factors of genre, style, and context.

As the fundamental tenet of this chapter, the purpose of the integrative literature review is to provide background on vibrato pedagogy in singing training from historical teachings to modern-day sources, across musical genres, and to synthesize pertinent vibrato research, offering targeted pedagogical tools to transform and modernize vibrato training as an integral and effective aspect of singing with versatility in any style, context, and setting.

Original audio and video recordings of the recommended vocal exercises, as well as teaching demonstrations are made available by the author of this thesis at `https://github.com/theodora-nestorova/vocalvibratovariability`

6.2 Historical Vocal Pedagogy Perspectives on Vibrato

As vibrato production physiology and acoustics are emergent research areas, it is evident that singing teachers of all styles need objective, tangible, and updated information and resources in order to appropriately address vibrato objectively and practically with their students. Production of vibrato requires a delicate "balance in the force" of airflow, sub-, supra-, and intra-glottic pressure, and simultaneously oscillating muscles and tissues (Ragan, 2023). There remain many factors and agents still unknown in vibrato causation, production, and control, yet singing professionals maintain stark aesthetic preferences. According to Sundberg (1987), within this muscular and airflow coordination, "it may be difficult to realize what is the hen and what is the egg... [as] it is also possible that the primary agent behind the vibrato is the entire pitch regulating system as a whole." This black box perpetuates the "mystery" of vocal vibrato.

Vibrato pedagogy has predominantly been framed through a Western classical lens, despite vibrato being a stylistic and temporally variable phenomenon across genres and traditions (Sundberg, 1994a; Nestorova, 2021; Nix et al., 2016). While recent voice science studies have expanded to include more genre diversity, most vibrato research still relies on Western classical norms for comparison. Western Operatic vibrato has received the most scholarly attention, yet little research explores how singers intentionally modify and alter their vibrato across genre contexts or how training influences vibrato characteristics. Furthermore, codified training approaches for vibrato within single genres is lacking, highlighting the need for more comparative studies and practical teaching recommendations across diverse musical and pedagogical traditions.

6.2.1 Vibrato Instruction in Historical Treatises & National Schools of Singing

Historical treatises discussing vibrato date back to the sixteenth century. Italian, French, and German treatises of the seventeenth and eighteenth centuries reveal a complex and varied understanding of vibrato, often referred to using terms such as *tremolo, trillo, groppo, tremoletto, tremolamento, fluttuazione, ondeggiamento, flattement, balancement, tremblement, cadence, ondulation, chevrotement, Bebung, Zittern der Stimme, Schwebung, and Beben* (and even the Spanish terms *temblor, trémulo*, and *ondulación*) (Zimmermann, 2019).

Italian writers like Giovanni Luca Conforti (1560–1608) and Lodovico Zacconi (1555–1627) described vibrato as part of other ornaments rather than a standalone feature (Zacconi, 1592; Conforti, 1989). Zacconi viewed it as essential to expressive singing, advocating its habitual use to facilitate vocal agility and emotion, suggesting that the continuous motion of the voice (through vibrato) facilitates smooth florid passagework. Zacconi emphasized making vibrato habitual—not just for its own sake, but as part of cultivating the vocal coordination needed for sustained legato. Luigi Zenobi (1547 or 1548–after 1602) and Giovanni Battista Doni (1593–1647) further distinguished between different types of vocal tremors, linking vibrato to both technical skill and aesthetic expression (Zenobi, 1856). Pier Francesco Tosi (ca. 1653–1732), however, cautioned against excessive or uncontrolled vibrato, suggesting a preference for stability in tone, leaving its role in Baroque singing somewhat ambiguous (Tosi, 1742).

French and German authors added further nuance to the vibrato discourse. French theorists, including Marin Mersenne (1588–1648) and Bertrand "Bénigne" de Bacilly (1621–1690), tied vibrato to tone quality and ornamentation, often using terms like *tremblement* and *cadence* (Mersenne, 1636; Bacilly, 1679). De Bacilly emphasized vibrato's aesthetic refinement while acknowledging its variability in natural talent and training. German treatises, such as those by Johann Andreas Herbst (1588–1666) and Christoph Bernhard (1628–1692), explored vibrato's physiological aspects, focusing on its deliberate use for emotional expression (Johann Andreas Herbst, 1653; Bernhard, 1650). Johann Friedrich Agricola (1720–1774) and others classified vibrato among optional ornaments, highlighting its contextual application (Agricola and Tosi, 1757). British scholars echoed these themes but leaned on practical guidance, such as John Playford (1623-1686)'s imitation exercises, or Roger North (1653-1734)'s hand-drawn vibrato contour transcriptions (such as in Fig. 6.1) over vocal exercises such as the *messa di voce* (Playford, 1654; North, 1728). Highly debated non-vibrato or straight-tone historical performance practices surrounding solo singing, also within choral and ensemble singing traditions, surfaced from modern interpretations of these treatises in these national schools. Examples of exercises from historical vocal pedagogues of various national schools can be seen in Fig. 6.2.

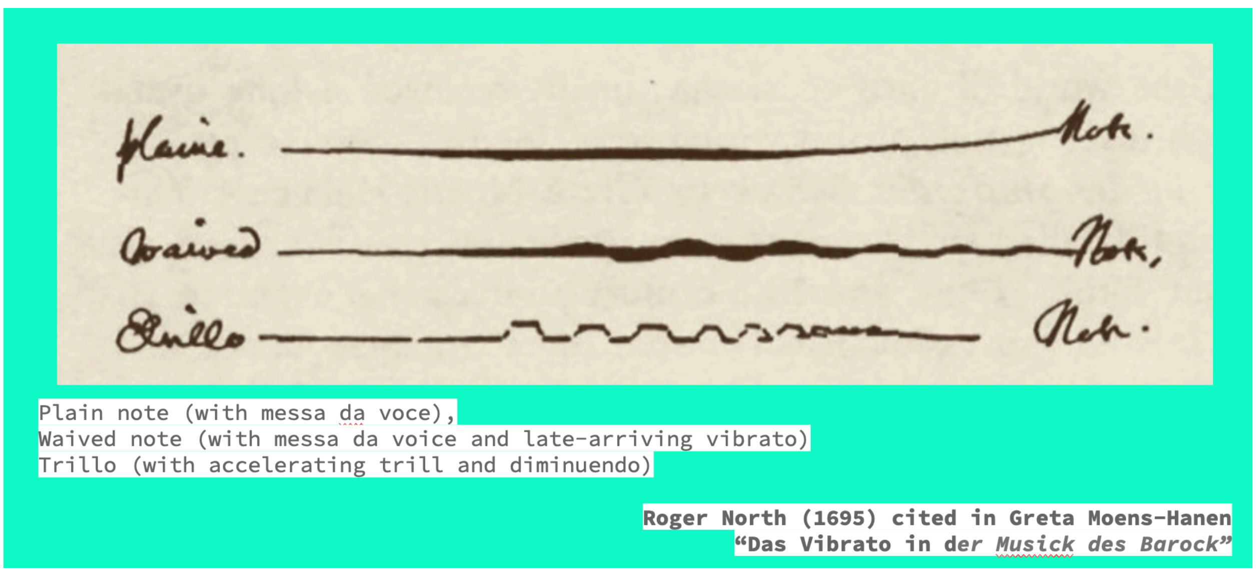


Figure 6.1: Visual representation from North (1728)'s treatise: Hand-drawn three different vibrato types (plain, waived, and trillo).

In *National Schools of Singing: English, French, German, and Italian Techniques of Singing*, Miller (1997) outlined the stylistic uses and physical aspects of the Italian, French, German, and English schools of singing. Many of the descriptions center around the vibrato demands in the Western classical canonic repertoire representative of each country. More recent attempts have explored these national schools in terms of contemporary repertoire; with the distinctive vibrato descriptions required by modern composers. Additionally, four kinds of vibrato (gentle glottal vibrato, hard glottal vibrato, goat vibrato, horse vibrato) were proposed to be produced primarily by the epiglottis, however there is no scientific evidence that the epiglottis moves in such a manner, nor are there any citations to support this claim (Isherwood, 2009). While globalism in the modern age as well as an increasing diversification of the repertoire canon may soon render these obsolete, current pervading vibrato conceptions likely emerged from generalizations of these national schools of singing.

Figure 6.2: Exercise examples targeting vibrato from various historical vocal pedagogues: Excerpted notated exercises from vocal method/étude books.

6.2.2 Vibrato Instruction in Historical Vocal Pedagogy Textbooks

From the dawn of scientific studies conducted on vibrato, authors were interested in pedagogical implications. Some of the earliest published manuscripts on vibrato include the inherent link "vibrato [as] a by-product of training, yet a by-product which is always present, and which contributes no small part to the effectiveness of the artistic singing voice" (Gray, 1926). In the seminal book *The Vibrato*, Seashore (1932) goes so far as to state qualities of "a good vibrato [as] a pulsation of pitch, usually accompanied with synchronous pulsations of loudness and timbre, of such extent and rate as to give a pleasing flexibility, tenderness and richness to the tone." Bozeman (2021) continues this notion with the statement that "ideal vibrancy is not a muscular sensation, rather an acoustic sensation." This points to vibrato's significant impact in the timbral dimension of perception and quality of singing, which has been and continues to be studied scientifically.

Venerated historical vocal pedagogues have long listened to vibrato as a way to troubleshoot biomechanical inefficiencies. "While empirical wisdom is not considered a benchmark for scientific validity," teachers have been linking students' underlying technical issues to vibrato for centuries (Rosenberg, 2023). In pedagogical textbooks from the mid-20th and early 21st century, vibrato is suggested as a potential indicator of strategies of resonance (Coffin, 1980; Doscher, 1994; Miller, 2004; McKinney, 2005) and breathing (Appelman, 1986; McKinney, 2005; Reid, 1965). Contemporary pedagogical texts have continued to outline vibrato archetypes such as *wobble, flutter,* and *bleat*, and suggested ways by which the voice teacher can correct vibrato issues for greater consistency and aesthetically pleasing sound output (Nix, 2014; Kirkpatrick, 2008).

Wobble is negatively associated with age, muscle fatigue, or overly weighty, dark vocal production, and used to describe the perception of wide fluctuations of pitch, or broad vibrato extent (in tandem with slow vibrato rate). Miller (2004) attributed this to a relatively unstable larynx. *Flutter*, on the other hand, is ascribed to a very fast rate (and probably co-occurring with a narrow extent) (Doscher, 1994). Vennard (1968) believed that whatever the type, exaggerations of the vibrato are potentially caused by muscle "overloading" in which muscular coordination is out of balance; the breath stream is not hooking up smoothly, and the greater the overload on that muscle, the greater the trembling (Doscher, 1994). Sympathetic vibrations of the tongue, jaw, and larynx during vibrato singing are considered with general agreement to be signs of excessive tension. *Bleat* refers to a rapid, uneven vibrato perceived as strained or overly pressed, often resulting from excess tension in the glottal or supraglottic structures. For example, Marchesi (1877) attributed an abnormally fast vibrato rate to the "undue pressure on the larynx," reflecting a misalignment between breath energy and glottal coordination.

However, there is caution against making students self-conscious or "badgering" them regarding their vibrato, with the thoughts that the problem may be likely as emotional and psychological as it is physical (Doscher, 1994; Vennard, 1968). "Fluid" and "free-flowing" are adjectives that are used to describe "appropriate vibrato" in many historical vocal pedagogy texts. These terms point to a Western classical lens prioritizing vibrato regularity and invariability, as the opposite may signal an underlying issue. Coffin (1980) claimed that irregular vibrato was a sign that "there is a fight going on between vocal cord vibration, vocal tract resonation, and the breath." Conceding the many still unknown forces in vibrato pinpoint a lacking area in singing voice science research; "thus far no one has studied the effect of resonance tract adjustments on vibrato, although the quantification of the acoustic properties of vibrato represents the greatest majority of studies done on vibrato" (Doscher, 1994). Reid (1965) proposed that vibrato indeed plays a role in "freeing the voice" and Doscher (1994) confirmed it as integral to the "functional unity of the singing voice" as the result of "intricate functional interdependence of the total singing instrument."

6.3 Influence of Genre on Vibrato Instruction

It is well documented that the overwhelming majority of historical vocal pedagogy literature was written with a Western classical bias. While recent years have seen a diversification of musical genres and styles being represented in published works on singing and at international singing conferences, the Contemporary Commercial Music (CCM) pedagogical output is still nascent when compared to Western classical (American Academy of Teachers of Singing et al., 2008). Yet, the historical focus on Western classical singing in vocal pedagogy overshadows vibrato training with a particularly strong impact.

Western classical singing is known for the convention of maintaining vibrato in between 80-95% of each song (Mason and Zemlin, 1971). It is generally accepted that within the umbrella of Western classical singing, only the Western Operatic aesthetic expects sung vibrato to be uniform, consistent, persistent, and omnipresent in the majority of tones. Moreover, the predominant aesthetic in this style also prefers vibrato to be relatively regular and simultaneously inconspicuous, as "the more dominant the vibrato, the smaller the appreciative audience" (Bozeman, 2021).

Since studies on vibrato have mainly been conducted with participants trained in this Western Operatic aesthetic, there has been little to no documentation on the volitional use of variable vibrato (or non-vibrato). Bjørklund (1961) claims that nearly all professional opera singers develop vibrato without conscious thought and without an active effort to acquire it, and Sundberg (1987) furthers this position by stating that "vibrato develops more or less by itself as voice training proceeds successfully," furthering the implicit or indirect teaching approach.

Bias toward this aesthetic is compounded by several perceptual studies demonstrating that ascribed vocal beauty in the Western classical singing voice favors singers who possess a consistent and even vibrato (Ekholm et al., 1998). This bias may be exclusionary, disregarding significant stylistic characteristics of many performed genres with stylistically appropriate vibrato features that may be deemed "non-normative" solely because they do not conform to the beauty standards outlined in Western classical singing.

Yet, over the course of examination of both historical treatises and contemporary vocal pedagogy texts that are steeped in the Western classical tradition, it becomes clear that "there has always been a plurality of styles of vocal tremulousness, and this allows for generosity toward and a tolerance of diverse interpretations" (Stark, 1999).

6.3.1 Contemporary Commercial Music (CCM) Perspectives on Vibrato Instruction

"Cultivating a vibrato for a specific style is by and large something that singers can (and must) train toward" as Hoch et al. (2016) put forth. This is truer now more than ever; as we progress through the twenty-first century, voice career sustainability and viability may be more and more predicated on cross-genre singing skills. Yet, "there is minimal research on the use and type of vibrato used in commercial music" while it is evident that vibrato aesthetics and production are dependent on musical style and are constantly evolving across genre contexts (Bartlett and Naismith, 2020; Nestorova, 2021).

Regardless of genre or style, it is still disputed whether vibrato can and should be voluntarily controlled. In pedagogical literature, it becomes clear that semantics drive this debate, as what is intended to mean balanced volitional control often gets mistaken for unbalanced manufactured or imitated control. In Musical Theater, "we don't always want vibrato in the sound… but the availability of vibrato is a good indicator of healthy vocal function" (Arneson and Brown, 2023). Similarly, in Jazz, though vibrato conventions have changed from the 1930s to modern-day, vibrato is used judiciously, subtly, and sparingly, always shaped and guided by the textual and emotional content of the music and often improvised or used in scatting (Shapiro, 2015).

Variable vibrato and non-vibrato singing are fully volitional, stylistic facets of Musical Theater, Jazz, and other CCM genres. Several scientific studies have demonstrated that singers possess some level of voluntary control over their vibrato, which can vary based on physiology and is certainly influenced by the singing style (Shipp et al., 1990; Carter et al., 2010). Western CCM and Eastern World Music singers may avoid or apply vibrato sparingly towards the offsets or onsets of tones, but the vast variety of genres and styles themselves may dictate these practices (American Academy of Teachers of Singing et al., 2008). These vibrato practices are intertwined with phonatory and acoustic strategies that tend to be closer to speech than those in Western classical music (Bartlett and Naismith, 2020; LoVetri and Weekly, 2003; Weekly and LoVetri, 2009).

Belting, one of those phonatory-acoustic strategies, is a hallmark component of the Musical Theater genre, and studies found that vibrato is less frequent and shorter in duration during belt singing than legit singing[1], a possible stylistic choice reflecting the influence of jazz and blues

[1] In Musical Theater, "legit" singing refers to a head-voice or chest-head-voice mix, Western classical-influenced style with a focus on clear diction and rounded vowels, while "belting" is a chest-voice technique producing a powerful and speech-like sound.

instrumentalists (Becker and Watson, 2022). The general convention, though currently up for debate, has been that 50% of the song (or the tone itself) may include non-vibrato (LeBorgne and Rosenberg, 2026). Elite Musical Theater belters may have a shorter delay of onset of vibrato than average belters, with greater intensity, magnitude, and intended emotional purpose. Furthermore, these findings indicate that the change over time (magnitude) of especially vibrato extent seem to be greater in belters than classical singers, but significantly different between elite versus average belters (LeBorgne, 2001). Thus, it is plausible that the physiological production of vibrato versus non-vibrato itself is likely also divergent (though this is subject to current and future research for the author of this thesis).

In Chapters 3 and 4 of this thesis, it was demonstrated that perceptually salient, time-varying aspects of vocal vibrato both acoustically and psychoacoustically distinguish various Western genres (Jazz, Musical Theater, and Opera) and Eastern World (Eastern European and Middle Eastern) musical styles. The extent of vibrato itself is the primary factor that is either reduced or augmented when vibrato is used ornamentally, as it is in many non-Western styles. As CCM singing is a very broad category distinct from Western classical singing, each style under the CCM "umbrella" warrants a dedicated pedagogical framework. Vibrato similarly merits its own pedagogical framework, as it arises as a critical genre-distinguishing component (Nestorova et al., 2024b).

6.4 To Vibrate or not to Vibrate?: The Straight Tone Debacle

While the purpose of this thesis is not to delve into an ontological discussion on non-vibrato (often referred to as "straight tone") singing, it must be acknowledged that the majority of styles beyond Western art music (even several subgenres) employ non-vibrato idiomatically (Howard et al., 2012; Bartlett and Naismith, 2020; LoVetri and Weekly, 2003).

6.4.1 Biomechanical Systems in Non-Vibrato Production

Titze emphasizes that non-vibrato is not necessarily a by-product of a healthy or unhealthy technique, but rather an indication of good muscle balance, explaining that "it may actually be a good exercise to learn how to disengage your vibrato... the extent is in the singer's control" (Olson, 2008).

Acoustically, there is no such thing as a completely straight tone; fluctuations to some degree are always present, dictated cortically. Perceptually, however, this fluctuation may be so minute and imperceptible, contributing to the lack of perceived vibrato. This phenomenon is colloquially termed "straight tone" in music. Titze et al. (2002) stated that when a singer sings non-vibrato or "straight tone," they do so not by reducing the vibrato rate, but by reducing the extent.

Aerodynamically, Iwata and Large (2005) found that airflow fluctuations were synchronized with amplitude (i.e., intensity) vibrato, and that airflow expenditure rates were approximately 10% higher for vibrato tones than for non-vibrato tones. Explained as a result of increased glottal resistance from the muscle activity needed to withhold the vibrato, it has also been posited that vibrato singing involves less glottal adduction than non-vibrato singing (Gauffin and Sundberg, 1989). This is corroborated by contemporary studies introducing the now parametrized airflow vibrato (Nandamudi and Scherer, 2019).

Perceptually, non-vibrato compared with vibrato in sung tones impacts the ratings of breathiness perceived by voice teachers. It has been suggested that fluctuations in vocal fundamental frequency may lead to vibrato-synchronous variations in sound pressure level by influencing the alignment of harmonics with vocal tract resonances. However, these variations could also stem from changes in respiratory effort, shifts in laryngeal voice quality, or articulatory adjustments affecting the positioning of vowel formants relative to harmonics. This highlights the complexity of vibrato production and the multiple physiological mechanisms that contribute to its acoustic properties (Oncley, 1971).

One key factor in vibrato (and arguably in non-vibrato) production is the oscillation of subglottal air pressure that drives vocal fold vibration. Additionally, variations in vocal fold abduction are more pronounced in different types of vibrato, with changes in the glottal airflow waveform significantly impacting perceived loudness. These findings suggest that the term "amplitude vibrato" requires further clarification to accurately reflect its physiological underpinnings.

A comparison of the aerodynamic properties of vibrato in Rothenberg et al. (1988) revealed that different muscle groups were engaged for vibrato at various levels of the vocal mechanism. If vibrato generally originate from neural processes, the substantial difference in vibrato rates implies that distinct sources of neural excitation are responsible for each muscle grouping. This distinction reinforces the idea that vibrato is not a singular phenomenon but rather a complex interplay of various biomechanical and neurological factors.

Research on airflow in vibrato suggests that vibrato productions are perceived as significantly less breathy than non-vibrato productions for the vowel /i/, with a less pronounced effect for the vowel /u/. In a related perceptual study, 48% of singing teachers reported that they "always" use vibrato as an indicator of vocal function (Nix et al., 2024a). Findings in the perceptual study in Chapter 4 of this thesis further suggest that the perception of vibrato as a measure of vocal function may vary by genre, with vibrato in non-Western classical styles tending to be interpreted more as an artistic feature rather than a physiological indicator. These differences in perception highlight the role of cultural and stylistic expectations in shaping how vibrato is evaluated and understood.

Physiologically, at the extremes of the singing voice, vibrato tends to diminish. At very high pitches, cricothyroid (CT) muscle activity is very high; simultaneously, the action of the respective antagonist muscle, the thyroarytenoid (TA), at that extreme point is very low, and vice versa for very low pitches. Titze et al. (2002) believes that this dominance of one vocal fold length changing muscle over another damps out the extent of the frequency swings, resulting in what is perceived as a more "straight" tone.

Dejonckere, Hirano, and Sundberg (1995) takes this a few steps further, citing the innate connection of both vibrato and non-vibrato production to phonation and registration. During vibrato, the full vocal mechanism oscillates, including: (1) laryngeal depressor muscles, such as the sternothyroid and sternohyoid muscles; (2) the lateral pharyngeal wall; (3) the velum or soft palate; (4) the base of the tongue; (5) and the epiglottis. These sympathetic oscillations may be useful for some singers in avoiding tightness or rigidity in adjustment of the vocal tract.

The Non-Vibrato vs. Vibrato Pedagogical Debate

The origins of contention surrounding non-vibrato are attributed less to styles beyond Western classical and more to a musicological factor within the Western classical genre— the fiercely debated "Vibrato Wars"— a late twentieth century conflict between pro- and anti-vibrato stakeholders in Historically Informed Performance (HIP), a practice which began with a period governed by the return to authentic, original instrument use (Neumann, 1991; Gable, 1992; Malafronte, 2015). Nevertheless, aesthetic preferences, national styles, and ideals have changed through different epochs and periods throughout history, alongside repertoire and context demands. One of the common pedagogical misconceptions regarding non-vibrato singing may have emerged amid this debate, during which there was a response from the voice teaching community denouncing "deliberate alteration" of natural vibrato as potentially "injurious to the vocal health and natural progress of young voice students" when used over extended periods (American Academy of Teachers of Singing, 1994).

Semantics may again be the culprit of this held belief; as "straight tone" is commonly used jargon in the singing community. The terminological use of "straight tone" as opposed to "non-vibrato" may bear the implication of rigidity in the vocal mechanism and ill-advised pressed phonation, suggesting undue strain that may fatigue the voice with prolonged use. Opposers claim that naturally-occurring vibrato produced through freedom in the vocal mechanism and flow phonation[2] can be beneficially inhibited as non-vibrato, if executed with skill: "if the use

[2]The term for a phonation in which transglottal breath pressure difference, airflow through the glottis, and glottal resistance are in an ideal balance for the particular pitch and intensity situation. Flow phonation is modeled as transglottal pressure/airflow = glottal resistance (ideally in a 1/1=1 balance). First coined by Johan Sundberg, the term flow phonation has been defined in the literature as a phonation type produced with the largest peak-to-peak flow amplitude, where the minimum still reaches zero (Hoch et al., 2022).

of straight tones summarily destroyed voices, thousands of singers would have lost their voices by now, since the use of non-vibrato is prevalent in some musical styles and cultures" (Mabry, 2002). For a more extensive discussion on flow phonation and production approaches to vibrato, see Sec. 6.7 below.

A shift in perspective is offered in this chapter: the action of perceived non-vibrato is inherently a functional practice in balancing coordination of controlling the degree (or extent) of vibrato's oscillation, rather than attenuating it fully. Non-vibrato singing still involves subtle oscillations as outlined above, since the neurologically dictated aspect of vibrato (cyclic rate) is ever present within the vocal instrument, and the extent of the frequency excursion is directly within voluntary control. Vibrato's underpinnings may be ascribed both intrinsically and extrinsically to the motor-vocal mechanism.

6.4.2 From HIP Vibrato Wars to Solo vs. Choir Singing Wars

Minimized-to-non-vibrato singing is a prominent feature in some ensemble or choral aesthetics. Two vocalists performing a duet or a few singers in a small ensemble may be able to synchronize their vibrato, larger ensembles tend to decrease vibrato extent, in particular, to preserve a better definition of pitch (Ternström and Sundberg, 1986). Many choral traditions favor a minimized vibrato compared to solo singing for the purposes of homogenized blending in ensemble singing. In a perceptual survey, Nix (2013) explored choral directors' instructions to student singers regarding vibrato adjustment. The majority of student singers responding indicated that they were asked to "sing without vibrato entirely" (81.7%) and/or asked to "sing without vibrato on selected notes, but not all the time" (60.4%). In another study conducted by Nix (2015), preferred vibrato extent in several synthesized voice samples was played to groups of choral directors, voice teachers, and student singers. It may come as no surprise that choral directors had the lowest preference threshold for vibrato extent, preferring a narrower vibrato extent fluctuation than both voice teachers and student singers.

A significant part of the Western classical voice teaching community cautions that habitual non-vibrato singing or a consistent attenuation of a soloistic vibrato rate and extent in the context of choir singing may be detrimental to still developing students' voice technique. There are concerns that in striving for a non-vibrato or dampened vibrato sound, singers may exert compensatory tension, forcing hyper- or hypo-adduction of the vocal folds, pressed phonation, and overactivity of the paralaryngeal musculature. This has contributed to the labels and poor reputation that non-vibrato singing has garnered in the modern vocal pedagogical climate (Katok, 2021). If physiological flexibility and acoustic context are considered when training vibrato, consilience may be brought into this dichotomous argument.

6.5 Approaches to Teaching Vibrato

6.5.1 Direct (Explicit) vs. Indirect (Implicit) Instruction

Direct (explicit) versus indirect (implicit) instruction is perhaps the most unresolved genre-agnostic pedagogical topic relating to vibrato. Some teachers staunchly believe that vibrato should never be deliberately taught, as it appears naturally with proper functional coordination and efficiency. These teachers often claim that uneven vibrato can be "cured" and opt to teach not tricks, but the technique by which vibrato is attained (Vennard, 1968; Doscher, 1994). Others oppose this, and address vibrato directly.

Nix (2013) conducted a survey of 350 both solo and choral singers, with results showing that 59.4% of singers received only indirect, non-explicit instruction in vibrato singing and primarily (70.4%) from a voice teacher in one-on-one instruction. On the other hand, if a singer received explicit instruction in non-vibrato singing (22.3% said yes), it was primarily from a choral director (53.2% indicated from a choral director, and 27.3% from both a choral director and a voice teacher). Within deliberate vibrato instruction given to singing students, the majority responded that the type of deliberate vibrato instruction they have received has been to both "adjust the vibrato rate" (51.6%) and "sing without vibrato on selected notes, but not all the time" (51.1%), with adjustment to vibrato extent, lessening vibrato at the onset of notes, or singing without vibrato entirely cited as less common instructions.

6.5.2 Literature on Impact of Voice Training on Vibrato

There is scientific evidence supporting the positive effect of training with vibrato development over time. In longitudinal studies lasting 2-3 years, a statistically significant correlation was noted between years of training in conservatory students and a stabilization in the periodicity of singers' vibrato rate. Specifically, students with an initial vibrato rate slower than the Western classical standard (below 5 Hz) exhibited an increase in vibrato rate, while those with an initial rate faster than the standard (above 7 Hz) showed a decrease. This suggests that extended training contributes to a more standardized vibrato rate among singers (Mürbe et al., 2007; Mitchell and Kenny, 2010; Mendes et al., 2003).

Studies on Western classical singing students have also displayed a linear interaction of vibrato and several aspects of technique relevant to training. Procuring breathing imagery was found to regulate singers' vibrato (both frequency and amplitude vibrato, as sound pressure level (SPL) measures were simultaneously collected with the acoustic recordings) in a manner consistent with that of a more proficient, warmed-up voice (Moorcroft et al., 2015). Vocal warm-up produced three notable changes in vibrato rate: (1) more regularity in vibrato rate within a note,

(2) more stability in mean vibrato rates from one sustained note to the next, and (3) a moderating of excessively fast and excessively slow mean vibrato rates (Moorcroft and Kenny, 2012). An acoustic and perceptual study asked singers to sing with varying degrees of the famed "open throat" or *gola aperta* technique and measured vibrato. Within the genre context, judges preferred "open throat" as the best tone quality and resulting vibrato. Sub-optimal "open throat" configurations reduced vibrato extent and delayed vibrato onset, though rate did not change (Mitchell and Kenny, 2004). Results from this study suggest that vocal tract shaping may have the greatest impact on vibrato extent and onset, two components of vibrato variability. Vocal tract shaping, vowel clarity, and precision form two essential aspects of vocal training and have been shown to be heavily influenced by vibrato and its variability (Sundberg, 1975).

Legato and vibrato have long been pedagogically intertwined; as Bozeman (2021) states "vibrancy is as normative in Western classical *bel canto* singing as legato line." Pedagogical tradition emphasized both elements, with early pedagogues discussing them in tandem. Singers have been shown to synchronize vibrato frequency modulation with pitch changes by adjusting note duration, vibrato rate (which tends to increase during the last three vibrato cycles of a tone) (Prame, 1994), or both to achieve smooth pitch transitions, with the end of the preceding note playing a significant role in pitch and smoothness perception (Titze, 1994a; Sundberg, 1994b). During melismatic singing, vibrato typically synchronizes with pitch changes in the fioritura or coloratura runs for smooth transitions, as Sundberg (1994b)) observed, noting that tempo influences vibrato rate, with singers either accelerating, eliminating vibrato, or adjusting note durations. Potentially connected to this pitch and smoothness perception is Bozeman (2021)'s "vibrant singleness of pitch percept" which he describes with words such as "steady" and "spinning" and "the ideal for Western classical timbre and evidence of excellent functional efficiency." Further research is needed to quantify and scientifically describe legato pitch changes, which is within the continued collaborative research conducted by the author of this thesis.

In a global inventory and similarity rating of singing voice assessment terms at English speaking academic institutions, vibrato was shown to be equally as important in training as in assessment of singing (Hausknecht et al., 2021). Yet, there have been only a few manuscripts predominantly focused on the methods or didactic practices surrounding the pedagogy of vibrato training from teachers of singing (Nix, 2014; Kirkpatrick, 2008; Katok, 2021). Most documented pedagogical texts regarding vibrato surround its presence or absence, and what is stylistically appropriate, rather than how and when to teach vibrato (or non-vibrato).

6.5.3 Vibrato as a Teaching Tool: A Generalist Approach

Finding existing literature that is definitive on the training of vocal vibrato proves challenging; there are few sources that reference any specific approaches to singing vibrato instruction. This may stem from a lack of precise knowledge for the exact cause of vibrato which "has not been indisputably determined, [so] the reason for its absence and the means of creating it are equally indefinite" (Reid, 1965).

Consequently, in vocal pedagogy texts, regardless of the targeted genre, style, or demographic, much of the literature points to a generalist approach to teaching vibrato. This approach is built upon the shared viewpoint of many singing teachers who consider vibrato to be one of the many indicators to verify the functionality of the student's vocal mechanism and efficiency of their singing technique. Vibrato has long been characterized as a by-product of a healthy, optimally functioning singing voice mechanism (Olson, 2008; Friedlander, 2012; Reid, 1965), though the question of voice health and optimal functionality may be prone to bias based on musical background and aesthetic preferences.

A more inclusive and comprehensive generalist approach, however, would center the indicative nature of vibrato more equally on both voice research and the student goals and perspectives in the Evidence-Based Vocal Pedagogy (EBVP) framework (Ragan, 2018), incorporating perception and cognition, as well as style-specific considerations. If, therefore, voice professionals can fine-tune how they listen to vibrato to identify what may be functionally occurring within the singer's vocal mechanism while also acknowledging the singer's notion of their own vibrato, this may lead to more effective and productive progress, while also helping the singer achieve specific artistic, aesthetic targets. In turn, this will transform the pedagogy of vibrato into a facilitation (midwife) approach (Jones, 2005), and make teaching vibrato more attainable and unambiguous in the vocal studio.

Examining not only vocal but also instrumental pedagogical sources becomes quite beneficial in crafting a vocal vibrato pedagogy. String and wind instrument pedagogies have many more codified vibrato training methodologies (with a plethora of resources from systematized schools like Suzuki and Baermann) than vocal pedagogy. Perhaps ironically, however, instrumental vibrato training is based on the voice, as timbrally, vibrato is meant to emulate the human voice's natural oscillation. However, it must be acknowledged that instrumental pedagogy texts on vibrato are likewise overwhelmingly steeped in the Western classical tradition. Studies involving vibrato on string instruments have explored non-vibrato (straight tone) and pitch production accuracy more deeply than studies involving the voice, another area of warranted research for vocal vibrato (Titze, 2014), with distinct significance to choral and ensemble singing stakeholders.

6.6 Translational and Practical Applications for Vibrato Training

"In the world of singing, there is still a myth that is perpetuated from time to time that vibrato is 'natural' and therefore is something that a singer cannot control. This is mostly nonsense" state Hoch et al. (2016). The components of naturalness and control co-exist in vibrato, and therefore are not mutually exclusive. Vocal vibrato is multi-factorial; it has elements which are non-volitional (involuntarily dictated) and volitional (voluntarily in the singer's control). This, combined with its highly variable nature across genres, styles, and contexts, makes vibrato instruction a disputed, relatively uncharted territory. If approached eclectically, from multidimensional and multi-modal perspectives, unity instead of controversy may be brought into the discussion of implicit versus explicit vibrato instruction. To that aim, Sections 6.5-6.8 of this chapter propose practical exercises, applied tools, and translational strategies for incorporating vibrato instruction into the vocal studio as "FUNctional Vocal Pedagogy with a Student-Centered Approach," while taking genre, style, and context into account.

6.6.1 The "Why"

Vibrato involves skillful coordination of airflow, subglottic pressure, vocal fold vibration, laryngeal muscular balance, and acoustic interactions, often stemming from a cumulative sense of vocal freedom and timbre awareness rather than isolated techniques (Erickson, 2021). This makes it challenging for voice teachers to micromanage their students' activity during vibrato production. Facilitation strategies may need to vary based on individual learning style and intended target acoustic product or genre. If it is accepted that vibrato is or ought to be the default procured through optimal muscular balance, explicit directives may not be as effective for vibrato as for non-vibrato. Yet, whether implicit or explicit in instruction, vibrato training is a fundamental tenet of vocal pedagogy across styles. Vibrato may even be regarded as inherently tied to overall vocal production itself, as the multi-factorial, phonatory centric nature of vibrato production is a perceptual phenomenon based on laryngeal and aeroacoustic balance.

Historically, instrumental pedagogy literature has offered specific methodologies for vibrato training, whereas in vocal pedagogy, such foundational frameworks have been largely absent. A large contingent of string pedagogues share D.C. Dounis's view that vibrato "is the basis of the left hand technique" and "cultivation of balance between the fingers" (Jensen and Chung, 2017). Exploration of vibrato is encouraged for students of all levels, with the unified goal of a free and balanced vibrato. Most critical of all in instrumental vibrato training, however, is the student's development of vibrato sensitivity, or a metacognition about vibrato. Viewed as an expressive tool, instrumentalists (especially string musicians) are encouraged to "use every degree, every shade of vibrato" rather than an "inexpressive one-gear vibrato" (Pleeth, 1992). Vibrato's integral role in technical and artistic development across string and wind instruments is made evident through the inclusion of abundant vibrato exercises in étude books and structured methodologies for vibrato training. In singing, the existing literature centers around correcting vibrato issues only once they arise, adopting a deficiency-centered perspective.

In addition, there is a need for a pedagogical approach that accounts for individual variability in vocal vibrato production. The majority of existing literature assumes vibrato develops naturally with technical refinement, yet this does not hold true for all singers. Factors such as physiological differences, stylistic demands, and even psychological barriers—such as fear of excessive vibrato or attachment to a specific tonal ideal—can all influence vibrato emergence and control. Furthermore, the differences in vibrato approaches across genres, from CCM to Western classical to World Music are broad and dichotomous, remaining disjunct. A comprehensive vibrato pedagogy should not only provide corrective strategies but also equip singers with the tools for vibrato versatility, adapted to different styles and expressive demands.

Several of the aforementioned authors in highly referenced pedagogical textbooks and systematic approaches outline their approaches (most often indirect) to teaching vocal vibrato. Some offer ideas or pedagogical strategies to remediate vibrato issues or resolve problems. However, in existing vocal pedagogy literature, there is a distinct lack of an organized framework including scaffolded exercises that specifically address vibrato (not only from a reactive, deficit-based perspective). Consequently, the purpose of this chapter is to fill this gap by recommending a framework of "FUNctional" vocal exercises for the purpose of vibrato versatility. By integrating both implicit and explicit instructional strategies, this proposed framework offers practical tools and exercises that are adaptable across various vocal styles and genres.

6.6.2 Vibrato Vocalises in the Literature

In existing vocal pedagogy literature, vocalise exercises targeting vibrato are scarce and rarely outlined systematically. Nix (2014) outlines three types of vibrato vocalise exercises targeted for vibrato training in singers; these exercises are intended to: (1) inspire interplay between the thyroarytenoid (TA) and cricothyroid (CT) muscles, (2) interface with adduction/abduction of the vocal folds (VFs), and (3) promote freedom within the vocal tract through lighter production and more efficient phonation. Nix also includes suggestions for facilitation, singing styles without vibrato, and problem-solving. Beyond this publication, there are only a few existing sources offering vibrato exercises (whether to elicit vibrancy (Kirkpatrick, 2008), or to encourage non-vibrato, as is the case in Katok (2021)). The majority of these vibrato exercises have been mapped onto common vocalises with smaller intervals and specific rhythms in mind. Melodic exercise patterns for vibrato practice are common, with chromatic intervals also recommended (Nix, 2014). Rather than providing vibrato vocalises or exercises for general application to vibrato across differing types, targets, and styles, sources from existing literature mostly focus on how to solve or correct common vibrato issues in Western classical singing, such as narrow vibrato, wobble, tremolo, and distinguishing vibrato from trill.

Narrow Vibrato: Potential Causes and Corrections

Existing literature suggests that narrow vibrato may result from excessive muscular tension, which can restrict natural laryngeal oscillation, potentially locking a specific laryngeal adjustment in place (Miller, 1986). Contributing factors are purported to include pressed phonation (high adductory activity with reduced airflow) (McCoy et al., 2019; Nix, 2014; Sundberg, 1987) or general imbalance between subglottic pressure and airflow (Kirkpatrick, 2008; McCoy et al., 2019; Vennard, 1968) and less expressive artistic delivery (Bozeman, 2021). Additionally, a vowel that is mismatched to the singer's fundamental frequency (f_o) and intensity (Nix, 2014) or an excessively high laryngeal position (Davids and LaTour, 2020; Kirkpatrick, 2008; Shipp et al., 1990) may contribute to this phenomenon. Stylistic preference has also been ascribed as the reason for narrow vibrato in singers (Davids and LaTour, 2020).

To address a narrow vibrato, pedagogical strategies in existing literature focus on promoting laryngeal flexibility and airflow balance. Exercises such as ascending glissandi with expressive intent (Miller, 2004) or imaginative vocal play, such as producing a "ghostly" or "scary" vibrato while gliding between pitches, are proposed to encourage freer vibratory motion. Engaging in dramatic speech patterns (Bozeman, 2021) or mimicking an operatic singer are offered to enhance vibrato fluidity. Motor imagery techniques, including visualizing a rapidly bouncing ball (Miller, 1986) or utilizing trill exercises with gradually increasing speed (Miller, 2004; Davids and LaTour, 2020), are recommended as yielding improvements. Sustained phonation until breath depletion are offered to prompt vibrato emergence (Kirkpatrick, 2008). Additionally, onset exercises— ranging from aspirated to balanced to glottal onsets— alongside semi-occluded vocal tract (SOVT) techniques (straw phonation, lip buzzes) are advocated to promote efficient vibratory function (Davids and LaTour, 2020; Nix, 2014).

Wobble and ***Tremolo***: Possible Causes and Solutions

Often highlighted as a vibrato issue in existing literature, a *wobble*[3] is often attributed to insufficient oscillatory energy (Dromey et al., 2003; Miller, 2004; Vennard, 1968), weak thyroarytenoid (TA) engagement (Titze et al., 2002), or excessive subglottic pressure due to hyper-adduction

[3] Excessive vibrato, usually in reference to a slower rate and wider excursion, in which the undulation of pitch is very apparent (Hoch et al., 2022).

(Kirkpatrick, 2008). Additional contributing factors are identified as muscular imbalances (Nix, 2014), overly darkened vocal production (Doscher, 1994; Vennard, 1968), or age-related changes (Nix, 2014; Titze et al., 2002). Tongue and pharyngeal inflexibility is also considered a potential factor (Davids and LaTour, 2020).

For a vibrato perceived as a *wobble*, exercises that restore vibratory coordination are recommended. Semi-occluded vocal tract (SOVT) postures (Davids and LaTour, 2020; Nix, 2014; Ragan, 2020) and onset variability (aspirated, balanced, glottal) drills (Davids and LaTour, 2020) are used to recalibrate phonatory effort. Narrow vibrato drills and alternating between vibrato extremes (Miller, 2004; Vennard, 1968) are suggested for refining control. Agility exercises— rapid embellishments concluding in sustained notes— are employed to enhance flexibility (Miller, 2004; Nix, 2014), while quick *crescendos* (*piano* to *forte* dynamics) are recommended to encourage appropriate subglottic pressure management (Kirkpatrick, 2008).

On the other hand, a *tremolo*[4] is problematically associated with increased muscular tension (Titze et al., 2002), pressed phonation, high subglottic pressure, and rigid articulators (Kirkpatrick, 2008; Miller, 2004). Registration imbalance and vowel distortion relative to pitch and intensity are also recognized as exacerbating factors (Nix, 2014).

For a vibrato perceived as a *tremolo*, strategies are advised to focus on soft, measured diminuendos and a gentle sighing quality in phonation (Davids and LaTour, 2020; Kirkpatrick, 2008; Vennard, 1968). Targeted articulation exercises and laryngeal lowering techniques are prescribed to alleviate excess tension (Kirkpatrick, 2008).

Distinguishing Vibrato from Trill

In the existing literature, vibrato and trill are frequently conflated, but many sources caution against this, as they exhibit distinct physiological characteristics:

- Vibrato is characterized by minimal laryngeal movement, maintaining a single perceived pitch. Descriptive cues include "stabilized vibrato" (Olson, 2008), "singleness of pitch" (Bozeman, 2021), and "collected vibrato" (Redman et al., 2023).
- Trill involves more pronounced laryngeal motion, with deliberate oscillation between a base pitch and an upper neighbor.

By recognizing vibrato's underlying biomechanical factors and implementing targeted solutions and exercises, voice instructors can be equipped to guide students toward a balanced, stylistically adaptable vibrato that enhances both technical function and expressive potential.

[4] A vibrato with a rapid rate and an extent/excursion large enough to detract from the intended pitch, causing a rapid pulsing or beating effect (Hoch et al., 2022).

6.7 Approaching Vibrato from a Voice Therapy Lens

When considering the complexity of vocal vibrato, addressing it multi-dimensionally is beneficial. Adapting evidence-based approaches that have proven efficacy with other multi-dimensional voice processes from the neighboring field of Speech-Language Pathology may be effective. In clinical voice habilitation and rehabilitation practice, a therapeutic plan combines direct symptom modification with indirect, holistic strategies to balance the physiologic subsystems of voice production, tailored to the patient's condition and goals (American Speech-Language-Hearing Association). Approaching vibrato training through a similar holistic method that combines both direct and indirect strategies has demonstrated value and it is the recommendation of the author of this thesis. As vibrato is a motor skill, using appropriate motor learning theory methods for optimal retention and transfer is also recommended (Madill et al., 2020).

Vibrato production itself involves both isometric and isotonic muscle movements. Isometric muscle movements are contractions occurring when a muscle generates tension without changing length (i.e., sustaining subglottic pressure from airflow on a single pitch when singing). There is no visible movement of joints or muscles, but force is being exerted to contract muscles. Isotonic muscle movements are contractions occurring when a muscle changes length while maintaining constant tension, leading to visible movement (i.e., vocal folds changing pitch, including both concentric (shortening) and eccentric (lengthening)). There is an abundance of vocal exercises in modern vocal pedagogy that are considered isotonic or calisthenic exercises, consisting of repetitive ascending scales and various forms of dynamic, flowing scale exercises. However, foundational techniques in historical vocal pedagogy placed emphasis on isometric exercises, such as the *messa di voce*, which involves a gradual *crescendo*, sustained *forte*, and controlled *diminuendo*. Unlike isotonic or calisthenic exercises that involve repeated muscle contractions, isometric training relies on sustained muscular engagement, though there is ongoing debate in vocal pedagogy about the role of extremely soft phonation as an exercise technique (Wrycza Sabol et al., 1995). Therefore, a balanced approach to vibrato training should include both isometric and isotonic exercises.

Applying scaffolded steps from healthcare and similar fields may be helpful to adapt and modify to the process of vibrato training. There are 5 essential steps in the vibrato training process; (1) identification, (2) assessment, (3) planning, (4) implementation, (5) evaluation. Whether an implicit or a explicit approach is taken, these 5 essential steps may help the singing teacher organize their strategy in addressing vibrato, which is multi-factorial and often so intrinsically tied to other concurrent processes of the vocal mechanism.

Finally, there is growing evidence that vibrato training itself may serve a habilitative purpose for those with voice disorders or inefficiencies (DeSilva, 2016; Lester et al., 2014). Scope of practice is critical in this discussion. As Ragan (2017, 2020) cautions: voice teachers are responsible for identifying technical inefficiencies but should not attempt to diagnose medical conditions in their students. If a student exhibits persistent vocal health concerns, they must be referred for proper clinical evaluation and treatment to a qualified otolaryngologist-voice specialist and, if necessary, a multidisciplinary voice team including a certified and licensed Speech-Language Pathologist and preferably a Singing Voice Specialist (American Academy of Teachers of Singing, 2022).

If, as in the realm of Communication Sciences and Disorders, it is accepted that vibrato is multifactorial rather than singular in origin, as are the various speech and hearing processes in the body, then the approach to training vibrato can likewise be multi-modal and eclectic.

6.7.1 Vibrato Problem-Solving Checklist

Voice instructors typically employ systematic "diagnostic" approaches when addressing vocal challenges (McKinney, 2005), and vibrato-related issues should be examined with the same level of analytical rigor. The following checklist provides a series of considerations for identifying and assessing vibrato irregularities, should this be a shared desire for the teacher and student. Holistically evaluating vibrato is of utmost importance, as "faults in vibrato can be traced to poor airflow, inefficient phonation habits, inadequate breath support (which is often coupled with poor postural alignment), and inadequate resonance strategies" (Chapman, 2017). Nix (2014) recommends dealing with the underlying causes of vibrato issues, not just the symptoms, as the quick fix is not necessarily the correct, long-lasting fix. For more detailed information on troubleshooting vibrato issues, see Nix, Nestorova, Glasner, and Redman (2024b).

1. **Respiratory Function:** Assess breath management to determine whether vibrato instability stems from an inconsistent or excessive airflow. Variations in subglottal pressure may contribute to fluctuations in vibrato rate and extent.

2. **Articulatory Tension:** Examine potential tension in the tongue or jaw, as excessive constriction in these structures can manifest as vibrato irregularities.

3. **Vowel Modification and Airflow Coordination:** Ensure that appropriate vowel shaping and breath flow are being utilized. In some cases, increased vibrato rate and reduced extent on front vowels may indicate excessive closure and elevated breath pressure, whereas a slower vibrato rate with greater extent on back vowels may suggest an overly open vocal tract configuration. Adjusting vowel modification strategies and optimizing airflow can contribute to more consistent vibrato production. Additionally, pairing vowels that facilitate stable vibrato with those that exhibit instability may serve as an effective intervention.

4. **Postural Alignment**: Evaluate the singer's overall body alignment, as vibrato irregularities may be symptomatic of broader postural imbalances.

5. **Extrinsic Muscular Engagement:** Observe the activity of the extrinsic laryngeal musculature. An overly elevated or depressed larynx may interfere with natural vibrato production.

Beyond examining the above five groups of factors in the student in the context of their voice lessons and performances, it is also crucial for the singing teacher to become acutely aware of what others may be demanding or preferring from their students' voice over the course of their training, especially when it comes to vibrato. This information, as well as the student's goals and perspectives, should be a central factor in the teacher's instructional approaches within the Evidence-Based Vocal Pedagogy (EBVP) framework (Ragan, 2018). Students often are tied to a particular tonal ideal which includes a certain type of vibrato. To foster change, voice instructors must carefully explain why vibrato flexibility is beneficial, using a nuanced and trust-building approach. Rather than enforcing prescriptive corrections, the goal should be to offer singers greater control and adaptability, much like the modulation of dynamics or pitch, allowing them to navigate multiple vocal styles. Moreover, singers are often deeply attached to their tonal concept, making any adjustments to vibrato a sensitive topic. As Vennard notes, it is appropriate to draw attention to the absence of vibrato and encourage awareness of its natural presence in well-produced voices, as a singer's ear may instinctively induce it (Vennard, 1968). However, excessive focus on vibrato irregularities, particularly tremolo, can be counterproductive. Pressuring a student in this regard may lead to emotional tension rather than technical improvement. It is important to constructively discuss the presence or absence of vibrato in a singer's voice, as simply bringing non-judgmental meta-awareness to vibrato often facilitates its natural development. In doing so, it may be most beneficial for teachers to emphasize the value of vibrato versatility as a part of both functional skill building and stylistic specificity.

6.8 Applied Exercises for Versatility in Vibrato

Planning and implementing general vibrato training and "interventions" toward vibrato targets based on style may be considered a daunting task for voice teachers, considering that vibrato is both volitional and non-volitionally controlled and is multi-factorial production process. Add to this the individual perceptual nature of vibrato for each student, and one may be tempted to avoid mentioning or working on vibrato at all. This further perpetuates the adopted approach of indirect vibrato instruction. Yet, the voice teacher should not be deterred from addressing vibrato, as should the student not be discouraged from working on vibrato.

The following framework offers what are termed "FUNctional" vocal exercises, referential to functional vocal exercises (Stemple et al., 1994; Bane et al., 2019), with the target being vibrato versatility. The acronym "FUN" should be kept in mind by both teacher and student; as explorative, play-based learning has proven efficacy for adult learners, as well as children. While this framework should be tailored to each individual singer and aligned with their vocal specificities, the overarching, ultimate goal of this teaching framework is for the singer to discover a versatile vibrato that has "F(low) U(nification) N(uance)" and is "F(lexible) U(nique) N(atural)".

6.8.1 Versatile Vibrato has "*F*(low), *U*(nification), *N*(uance)"*ctionality*

The "FUN" acronym is useful to remember the dimensions of versatile vibrato, consisting of flow, unification, and nuance, corresponding to vibrato's multidimensional aerodynamic, laryngeal, and acoustic properties. The author of this thesis has codified this related terminology and compiled these exercises, drawn from personal teaching and singing experience, as well as the inspiration of many pedagogical mentors (see the Acknowledgments). The exercises are derived from both traditional and non-traditional vocal pedagogy literature— from Western classical *bel canto* vocalise books to Eastern traditional "call and response" musical skill-building games. It is important that each of these exercises be adapted to the pitch range appropriate for the voice type of each individual singer; the notated exercises below are simply examples and begin on arbitrary pitches, but may be transposed accordingly. The recommended exercises as points of entry to vibrato versatility are organized under each category, as follows:

FLOW: ***S/MOVTEs, Stretch and Flow, and VFEs***

- Beginning in a comfortable range, pitch glide (as inspired by Vocal Function Exercises (VFEs) (Stemple et al., 1994; Bane et al., 2019; Wrycza Sabol et al., 1995)) on semi- or mostly-occluded vocal tract positions (which have repeatedly demonstrated benefit in re-balancing the vocal mechanism through more equalized pressure on both sides of the vocal folds (Hijleh and Pinto, 2021; Titze, 2013; Ragan and Kapsner-Smith, 2019; Guzman et al., 2018)), then sustain at random pitches that are within that glide. Keep interchanging sustained and glided tones and notice when vibrato might be inclined to activate or attenuate.

- On sung text from repertoire, cover the mouth externally (ideally with a strapped on medical anesthesia mask that has an opening to breathe and a strap around the head (Ragan, 2020)), but continue to move the articulators. Alternate this with opening the velopharyngeal port, singing with and without occlusion and with and without nasal resistance and notice the difference in vibrato. The aim is to "release the constrictor muscles of the lower pharynx" (Mürbe et al., 2007).

- Initiate flow phonation, through the most effective chosen modality for the student— whether straw submerged in water or in air (Bhatt et al., 2023), or simply consonants (both voiced and unvoiced) and semi-consonants/vowels (open, closed, and mixed), as there is often an over attention on vowels and straw phonation in mainstream vocal pedagogy, or simply starting the airflow before (as though fogging up glass; inspired by "Stretch and Flow" exercises, also called "Casper-Stone Flow Phonation" (Watts et al., 2015; Diaz et al., 2023)) onset pilot sound.

- Use inhalation / ingressive (reverse flow) phonation (Robb et al., 2001; Fornhammar et al., 2022; Gordon and Reed, 2019) for sound initiation, followed by regular, flow phonation, experimenting with the resulting vibrato.

UNIFICATION: *Registration Plays, Laryngeal Balance*

- Wobbling yodels: rapidly shift registers in progressive intervals. Suggest to begin on Major and minor 3rd intervals, then increase to 5th intervals and octaves, and then decrease to chromatic intervals with a wider excursion. Select the most optimal vowel or vowel combinations for the individual singer. Explore the resulting freedom. See Fig. 6.3 for a notated example.

Figure 6.3: "Wobbling yodels" vibrato exercise notated example.

- Initiate the sound in mechanism 0 registration, vocal fry (or "creaky" voice)[5] Glide on major or minor 3rd, 5th, and octave intervals in and out of this phonation mode, sustaining the

[5]Vocal fry is the lowest vocal register, produced through a loose glottal closure that permits air to bubble through slowly with a popping or rattling sound of a very low frequency. Also called M0 (mode zero) or pulse. Creaky voice is often used interchangeably with vocal fry, but more precisely, creaky voice is a phonation related to vocal fry which can occur at higher frequencies and is characterized by a perception of roughness and the presence of sub-harmonics in the glottal waveform. Creaky voice acquires its name because it gives the perception of a creaking door. It can be used to train competent glottal closure in M1 or M2 (Hoch et al., 2022).

tone and experimenting with the resulting vibrato.

- Alternate stationary, single-pitch, isometric "plank"-like exercises and mobile, multiple pitches, isotonic "push-up"-like exercises in quick succession. The goal is to explore vibrato on single and changing pitches while playing with and balancing registration, both laryngeally and acoustically. As previously established in Sec. 6.7 above, the singing voice mechanism consists of both isometric and isotonic muscular movements. Vibrato similarly involves both isometric and isotonic elements to be explored in discovering versatile vibrato. However, caution should be taken here, as the idea of "strength training" per se is not recommended for the voice, but rather a balance or coordination training is encouraged. See Fig. 6.4 for three notated examples of these, including:
  - Interweaving florid melismatic runs with sustained phonation
  - Shifting between arpeggiated sequences and prolonged tones
  - Juxtaposing agility exercises with rebounding articulations
  - Contrasting staccato agility figures with sustained pitches
  - Alternating florid, staccato, and held notes in sequence
  - Combining rebounding articulations with melismatic runs

Figure 6.4: Three notated examples of recommended vibrato exercises that alternate isometric with isotonic vocal muscular movements.

NUANCE: ***Resonance Exploration***

- Explore differing resonance strategies, particularly through cross-genre and cross-style practice. Sustain a pitch or series of pitches while alternating between convergent and divergent vocal tract shaping and postures to assess changes to vibrato response.

- Modify or reverse the pitch and/or rhythmic pattern of the existing exercise or phrase from existing repertoire. Ideas for altering existing note patterns include reversing note combinations or changing rhythms (such as over- or double-dotting Bach-like articulations). Change other variables (vowel, intensity) and observe the response in the vibrato. See Fig. 6.5 for a notated example of this.

Figure 6.5: A notated example of a vibrato exercise that explores resonance through modified and/or reversed pitch and/or rhythmic patterns.

6.8.2 Versatile Vibrato is "*F*(lexible) *U*(nique) *N*(atural)"*ctional*

FLEXIBLE: ***Non-Vibrato through Augmented Biofeedback (Low-to-High Tech. Approaches)***

- Audiation plays a critical role in vocal technique, particularly in controlling vibrato initiation versus non-vibrato. If this is challenging for the student to volitionally control, singing in differing acoustic spaces, some more reverberant than others, and comparing the acoustic feedback and according effect on vibrato may be helpful.

- Starting on slides and sirens, transition from a staccato arpeggio with vibrato onsets and offsets to sustained non-vibrato/straight tones on the same pitch and intensity level.

- Sustain "noodle tones" (*crescendo-decrescendo* pairs on single pitches), exploring dynamic (intensity) changes with vowels and consonants, allowing for vibrato to naturally ebb and flow, noticing at which point changes in vibrato occur.
  - To facilitate objective reflection, audio or video record these and listen back; journal about it—what observations can be made about the vibrato rate, extent, and stability? What is the target for both the student and the teacher?

- Visual biofeedback, such as spectrographic analysis, can significantly aid multi-modal learners with vibrato. Extensive evidence supports the efficacy of multisensory approaches in enhancing learning, particularly through the integration of visual input, which refines and coordinates motor responses (Brown, 1996). Feedback learning models suggest that successful biofeedback training relies on five critical factors: (1) perceptibility (the ability to perceive the biosignal), (2) autonomy (self-regulation of the output), (3) mastery (control over the biosignal), (4) motivation (satisfaction with the outcome), and (5) learnability (enhanced learning through biofeedback) (Gaume et al., 2016).
    - Often, however, vocal pedagogues have relied on a unidirectional relationship with the spectrographic display, in which the transmitted signal is simply projected for the interpretation of teacher and student. In vibrato training, it is critical to return the signal to the body by designing kinesthetic exercises for the singer to experience, based on the intended output in the visual display. One example of this is using a resistance band to emulate the singer's vibrato shown in the spectrogram. The tactile connection with their acoustic output allows for detection and recognition of vibrato changes, as well as an enhanced brain-body connection for intervention.

Unique: *Individual*

- Speech-level singing: On the words "I'm a scary ghost," the singer speaks and then sustains a sliding, *glissando*, ascending-descending pattern with the intention of releasing the vibrato. This is particularly effective for children, as a creative practice. An effective variation includes chant-like intonational prosody to bridge the gap into sustained-tone singing. See Figure 6.6 for a notated example of this.

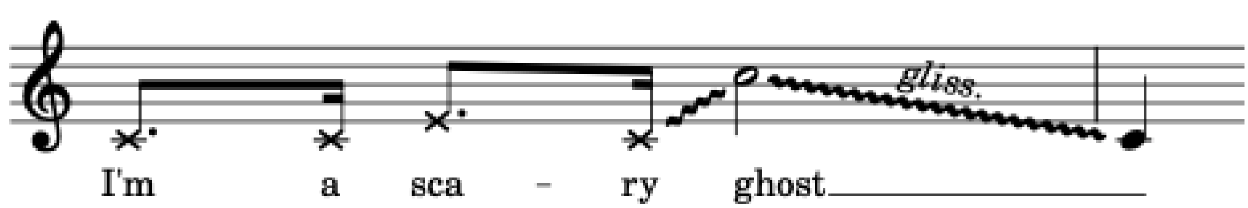


Figure 6.6: A notated example of a speech-to-singing, play-based learning vibrato exercise.

- Embodied cognition: Kinesthetic stimuli paired with full body movement to elicit vibrato, associated with the motion; activating more neuron pathways closer to motor cortices. Some examples include:

  - Opening and closing fingers to simulate the action of the vocal folds
  - Twinkling fingers or toes while moving arm or leg
  - For the metaphor-inclined, "shaking a door knob"
  - Bending, raising, and bouncing the knees according to vibrato type
  - Circling and/or rocking the pelvis to release the pelvic floor
  - Rapidly moving tongue while phonating or lightly palpating accessible extrinsic glossus muscles (geniohyoidus, genioglossus, hyoglossus, and styloglossus); it is important that this is done in a random and changing motion, so as not to encourage any static rigidity or compensatory functionality

NATURAL: *Musculoskeletal Release*

- Initiate spontaneous laughter or giggling (especially one that is pitched and increases in speed) and transition into a sung, sustained tone. See Fig. 6.7 for an approximated, notated example of this.

Figure 6.7: A notated example of a spontaneous laughter/giggling into sustained tone musculoskeletal release vibrato exercise.

- On a sustained tone, initiate abdominal isolation, with the epigastric region and hypogastric regions alternating motion, and notice the changes in the vibrato. If this is challenging to the singer, this indicates that they have habitual, excess tension in the abdominal region, likely affecting their air pressure regulation in vibrato production. If difficult to initiate without aids, the singer may work with their hands to massage the abdominal area and initiate this motion.

- An alternative option recommended for singers with extremely tense abdominal areas is the use of a massage gun on the epigastric and hypogastric regions, targeting muscular release, and noticing the according changes in vibrato.

6.9 Other Considerations in Vibrato Pedagogy

6.9.1 Functional Listening in Teaching

Functional listening is foundational to vibrato pedagogy. It is not merely an ancillary skill, but rather a primary driver in shaping how students understand, monitor, and refine their vibrato. Before any exercise can be effective– particularly those designed to develop coordination in parameters such as rate, extent, or variability of vibrato— singers and teachers must learn how to listen functionally. As Howell (2025) explains, functional listening is the process by which singers learn to perceive sound in a way that supports both artistic expression and efficient motor learning, distinguishing between what is heard externally and what is experienced internally.

This mode of listening provides the perceptual groundwork for every teaching strategy and exercise recommended above. If students cannot discern how their vibrato behaves— across different vowels, pitches, or dynamic levels— they may struggle more to make meaningful changes in coordination. Teachers should therefore begin not with correction, but with curiosity, using targeted perceptual prompts such as:

- What do you notice about the singer's vibrato on different vowels?
- How does the vibrato change across different pitch levels?
- What happens to the vibrato when the dynamic level shifts?

These questions invite singers to track three key acoustic-perceptual parameters of vibrato: rate (cycles per second), extent (pitch variation), and variability (consistency across time). By focusing on these variables across changing vocal conditions, students begin to understand that vibrato is not a fixed effect but a dynamic coordination influenced by vocal tract shaping, subglottal pressure, and register tuning.

In practice, functional listening might reveal that a singer's vibrato rate slows on high pitches, or that extent increases unintentionally on louder dynamics. When this awareness is cultivated, exercises— such as semi-occluded vocal tract (SOVT) work, vowel glides, or pitch pattern repetition— can be selected or modified in direct response to what the student is hearing and feeling. This increases the relevance and impact of the exercise, as it becomes a personalized tool for refining a specific perceptual-motor loop.

Functional listening also shifts the role of the teacher from master problem-solver to midwife perceptual guide (Gaunt, 2010). Instead of delivering assessments or corrections, instructors can use open-ended feedback prompts such as: "What changed in your vibrato when you opened the vowel?" or "Did the vibrato feel more stable on the descending pitch, and how so?" These questions help students form internal models of versatile vibrato production, grounded in their own sensations and acoustic feedback.

Additionally, fostering functional listening reinforces the broader goal of integrating vibrato technique into musical contexts. A singer who can monitor and adjust their vibrato across stylistic and expressive demands is more adaptable, expressive, and self-sufficient. Ultimately, functional listening is not a passive skill but an active, trainable, and attainable one.

Without functional listening, vibrato training risks becoming mechanistic and disconnected from repertoire singing. Through functional listening training, singers and voice teachers develop the perceptual acuity necessary to internalize coordination, refine artistic control, and transfer technique across repertoire. Moreover, functional listening allows for the parsing of complex, coexisting volitional versus non-volitional vibrato elements. As Nestorova (2021) underscores, cultivating a functional relationship with sound allows singers to sing the sound they intend to hear, transforming vibrato from an outcome to a choice.

6.9.2 From Exercise to Repertoire: Facilitating Effective Transfer in Vocal Pedagogy

A common misconception in vocal training is the notion of strength training, which implies that vocal progress stems from muscular hypertrophy. In reality, singing is a coordination-based activity rather than a strength-based one. While isolation exercises can serve a purpose in developing fundamental coordination, their primary function is not to build muscle but to refine neuromuscular efficiency (Johnson and Sandage, 2021). Exercises targeting vibrato are valuable, but only when taken out of individual iterations and contextualized within repertoire that is sung through. If an exercise, such as those suggested above, is learned in isolation without immediate application, the brain may recognize the skill as a separate task rather than as part of the singer's overall technique (Maxfield, 2011).

One of the key challenges in vocal pedagogy is adhering to motor learning principles by ensuring that technical exercises practiced in lessons and rehearsals translate seamlessly into repertoire for performance. Effective teaching goes beyond merely identifying and fixing issues; rather, it emphasizes finding and positively reinforcing what students are already doing well. This approach fosters confidence and builds upon existing coordination. By scaffolding with small, manageable steps— such as identifying and refining 60% of a skill before progressing— teachers can ensure steady development. Importantly, only one variable should be altered at a time; once a student demonstrates approximately 75% proficiency in a skill, adjustments in pitch range, vowel formation, or other parameters can be introduced to reinforce adaptability (Crocco et al., 2020; Maxfield, 2011; Anderson et al., 2019). This applies to vibrato training, as well.

A student-centered approach is most effective, at the appropriate stage of development, when directives are based on the singer's own feedback rather than solely on the teacher's perception. Encouraging students, particularly those who are intermediate and advanced, to assess their own performance helps develop self-awareness, autonomy, and agency (Goffi-Fynn, 2024). Simple prompts regarding specific skills such as vibrato are recommended, such as: "what was your experience of that vibrato, how do you know how to monitor it, and how can you practice it most effectively?" or "what part of a song would you like to apply that type of vibrato in?" encourage metacognition and active learning. However, care should be taken to separate knowledge of results from knowledge of performance (Steinhauer and Grayhack, 2000). By immediately applying an acquired skill within a musical context, students can solidify coordination and integrate new techniques into their repertoire with greater ease.

Thus, integrating the recommended vibrato exercises directly into repertoire in a student-centered, strengths-based scaffolded manner is essential for effective transfer. Without these holistic steps, students may struggle to bridge the gap between technical drills and practical performance. Especially when training vibrato, it is crucial that singing teachers not only determine technical challenges but also design targeted exercises that facilitate brain-body connection in their students. By ensuring that learned coordination ubiquitously transfers into singing across settings, voice teachers can cultivate singers who internalize vibrato as a natural and intuitive part of their artistry.

6.9.3 More Vibrato Teaching Resources

As discussed by Nix, Nestorova, Glasner, and Redman (2024b) in several recent National Association of Teachers of Singing (NATS) Chats and conference presentations on vibrato, a deeper theoretical understanding and practical application of assessment tools for vocal vibrato can be developed. The continuing education module "Ear Training for Voice Teachers: Do You Hear What I Hear?" enhances listeners' ability to discern and differentiate key aspects of vibrato in singing, while "Practical Voice Teaching Application: I hear this, what do I do now?" offered by the NATS Learning Lab provides guided listening questions to encourage critical thinking in translating what listeners' hear and observe regarding vibrato into actionable steps with singing students. Intervention options and recommendations, such as specific exercises eliciting versatility, habilitative exercises, and presenter facilitated participant discussions, are also offered and able to be viewed on demand on the NATS Learning Lab.

6.10 Conclusions

Elevating vibrato training in this manner as an essential aspect of voice teaching and learning will reveal deeper connections with excessive, compensatory tension in singing coordination. With the above proposed science-informed exercises as a part of the "FUNctional" framework for vibrato training, vibrato versatility becomes another tool in the voice teacher's toolbox to benefit vocal pedagogy. Versatile, variable vibrato training translates the scientific properties of vibrato while retaining the art, so both are honored and respected equitably and equally throughout.

Training vibrato is critical for singers cross-training different genres, even if that training constitutes solely bringing awareness to variability in vibrato across styles. Perhaps the answer to the question "to teach or not to teach" vibrato is best answered when considering the important genre agnostic skill of voice functionality; promoting vibrato versatility. If vibrato is to be "a bellwether of the stability and freedom of a singer's voice production" (Nix, 2023), then we ought to advocate for and promote the versatility and equal representation of vibrato and non-vibrato (and the variable transition in-between) training in the voice studio.

Non-vibrato may be a lucrative, necessary, practical, and valuable singing style that broadens employment avenues available to singers, allowing vibrato to be employed variably for technical and stylistic aims. Vibrato variability can be procured by alternating vibrato and non-vibrato singing, ensuring that the airstream and balance in the musculature remain consistent. A pedagogical approach of encouraging versatility in vibrato can artistically enrich voice warm-ups, cool-downs, and repertoire, just as a converging focus on balanced vibrato action bears potential clinical implications for diagnosis and use of vibrato in vocal habilitation and rehabilitation.

Perhaps, then, the most apt response to an age-old pedagogical philosophy question; "to teach or not to teach vibrato," is a wavy line best answered by: "why not both, to restore the 'balance in the force'?" Perhaps it is not a binary; the question of whether it should be deliberately taught or left to develop naturally are not mutually exclusive. Yes, vibrato often develops naturally, but all students develop at their own pace and with their own specificities. Deliberately teaching vibrato versatility to enable that development is not only wise, but encouraged, as a teacher's pedagogical expertise allows the student the support and guidance needed to progress and advance their vibrato development more effectively than by themselves.

By understanding the breadth and limits of scientific knowledge on vibrato production physiology, biomechanics, acoustics, and psychoacoustics, as well as volitionally variable vibrato based on genre, style, tradition, teachers of singing can better evaluate sources of excessive, compensatory tension or muscular imbalance in the vocal mechanism, evaluating vibrato more comprehensively to formulate an effective didactic intervention. This can be applied in a practical way, through several recommended exercises, eclectic and multi-modal.

Chapter 7

Diagnosability of Vibrato for Voice (Re)Habilitation

Author Contributions

This chapter includes restructured and reintegrated content from the following research article: Nestorova, T., Nestorov, I., and Chagnon, F. (2025). Diagnosability of vibrato for voice (re)habilitation: Clinical implications of vibrato variability. *Logopedics Phoniatrics Vocology*. [Manuscript submitted for publication].

Author-specific contributions: *Theodora Nestorova:* conceptualization, data acquisition and acoustic analysis, development of interpretation and conclusions, writing, review, and editing of the manuscript. *Ivan Nestorov:* coding of preliminary script, development and programming of original numerical and computational modeling for data and results processing, review and editing of the manuscript draft. *Françoise Chagnon:* supervision of the overall study project, results processing, and review and editing of the manuscript draft.

Chapter-specific contributions and differentiation from original journal publication: In this chapter, Secs. 7.2.3 and 7.3 contain additional information and were not originally included in the article, as they were out of the scope of the article itself. Additionally, Figs. 7.1 and 7.2 are reprinted in this chapter but only referenced in the article. Finally, editing for grammar and flow were conducted to optimally incorporate the article contents into this thesis chapter.

7.1 Clinical Implications: Vibrato as a Comprehensive Diagnostic Tool

The vocal mechanism is uniquely complex and dynamic, requiring both refined functional control and artistic sensitivity— especially in singing, where vibrato serves as a key technical and expressive feature across musical genres and cultures. Vocal vibrato, as a multidimensional phenomenon, is enigmatic and polemic among voice professionals, with many claiming that it "defies description except as a perceptual phenomenon" (Doscher, 1994). Despite vibrato's central role in vocal pedagogy, it remains under-examined in terms of its empirically-informed scientific analysis, with quantitative vibrato studies siloed from qualitative investigations. Furthermore, the lack of studies addressing temporal variance in vibrato deepen the isolated gap in between representative, ecologically valid experiences of vibrato in singing training and vibrato research.

This dissertation addresses that gap by developing acoustic and pedagogical tools to analyze naturally-occurring vibrato variability as a nuanced marker of genre-specific style. By integrating acoustic, perceptual, and ethnographic approaches, this research bridges scientific inquiry with pedagogical relevance, offering a more individualized, evidence-based framework for understanding and teaching vibrato. As a next step, it is crucial to connect the acoustic, perceptual, and ethnographic-pedagogical findings from this thesis to the biomechanical underpinnings of vibrato. In doing so, it contributes to a more inclusive and interdisciplinary model of vocal training— one that reflects the evolving needs of singers, pedagogues, and clinicians alike.

The variability of vibrato may be considered as a diagnostic biomarker of vocal function. When viewed through the lens of self-organization in vocal mechanics and physiology (Titze, 2021), vibrato may well serve as a signifier of this principle. The analytical methods and pedagogical approaches for time-varying vibrato proposed in this thesis offer a promising avenue when applied to clinical assessment and intervention to better understand and address voice disorders. While vibrato variability research is particularly relevant to occupational voice issues, such as functional conditions, there are significant implications of time-varying vibrato to neurogenic and psychogenic pathologies affecting the voice.

7.1.1 Functional Voice Disorders

Voice disorders can be classified into four broad categories: organic, functional, neurogenic, and psychogenic. See Fig. 7.1 below for a detailed voice disorder classification framework. Organic voice disorders result from physiological alterations in the respiratory, laryngeal, or vocal tract mechanisms. These can be further divided into structural and neurogenic disorders. Structural organic disorders are caused by physical changes to the vocal mechanism, such as vocal fold edema, nodules, or aging-related laryngeal changes. Neurogenic organic disorders arise from problems with central or peripheral nervous system innervation to the larynx, affecting vocal function; examples include vocal tremor, spasmodic dysphonia, and vocal fold paralysis. Functional voice disorders occur when the physical structure of the vocal mechanism is normal but inefficient use leads to symptoms like vocal fatigue, muscle tension dysphonia (MTD), diplophonia, or ventricular phonation (Sapienza and Hoffman, 2020; Sataloff, 2017). Psychogenic voice disorders, on the other hand, result from psychological stressors that lead to maladaptive voice production, such as psychogenic conversion aphonia or dysphonia (Rosen et al., 2021). These conditions may require interprofessional referrals to psychologists or psychiatrists for diagnosis and treatment, with collaboration from speech-language pathologists (SLPs) in subsequent behavioral treatment (Baker, 2003; Baker et al., 2013; Tezcaner et al., 2019; Rubin and Greenberg, 2002; Martins et al., 2014).

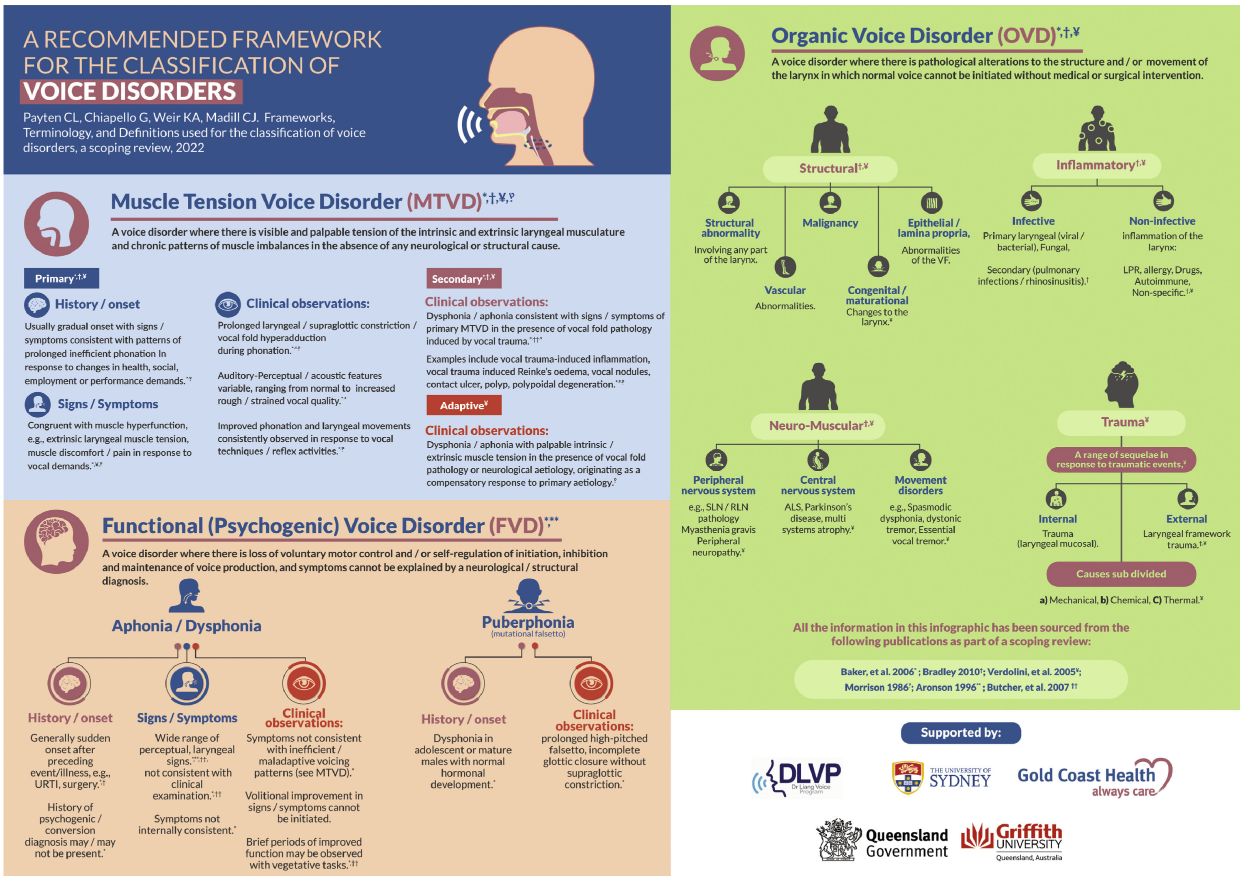


Figure 7.1: A new umbrella tree classification framework, describing the categories of voice disorders. Reprinted with permission from Payten et al. (2022).

The relationships between organic, functional, neurological, and psychogenic influences are often interwoven, and many voice disorders can involve contributions from multiple etiologies (Stemple et al., 2018; Verdolini et al., 2014). For instance, vocal fold nodules can result from habitual misuse of the voice (functional etiology), which can lead to structural damage (organic etiology) if left untreated. This complexity underscores the importance of a comprehensive, multifaceted approach to voice diagnosis and treatment. In singers, though the risk of developing these pathologies varies depending on many individual and environmental factors (including singing style and genre) (Kwok and Eslick, 2019; Jahn, 2013), the prevalence of functional voice disorders is notably high, with an estimated 46% of singers self-reporting dysphonia over the course of their careers (Pestana et al., 2017).

Dysphonia falls under the umbrella of vocal hyperfunction; in its phonotraumatic and non-phonotraumatic versions. Non-Phonotraumatic Vocal Hyperfunction (NPVH) is a multi-factorial condition presenting with habitual, chronic voice-related symptoms in daily life, such as increased vocal fatigue/effort, and anterior neck soreness/excess muscle activation, in the absence of any signs of phonotrauma or other phonation-disrupting structural or neurological impairments (Hillman et al., 2020). NPVH is also interchangeably referred to as primary Muscle Tension Dysphonia (pMTD or MTD-1), and is a diagnosis of exclusion (Kollbrunner and Seifert, 2017). pMTD is one of the most frequent diagnoses in patients seeking care from otorhinolaryngologists (ENTs), accounting for approximately 30% of clinical caseloads at voice centers, and is found primarily in women (Van Houtte et al., 2013).

Professional voice users, especially elite vocal performers such as singers, actors, broadcasters, have high occupational voice demands and vocal load and sensitivity to even minor voice dysfunction (Broaddus-Lawrence et al., 2000). Other suggested voice dysfunction risk factors for this demographic include personality traits common to this population (Toles et al., 2021; Rubin and Greenberg, 2002; Hoffman-Ruddy et al., 2001; Baker, 2003; Baker et al., 2013; Wodzak, 2022) due to the hyper-active, high-stress lifestyle and environment of occupational voice use (Kwok and Eslick, 2019). Nevertheless, many sources warn that psychological co-factors should not be considered as the primary cause for voice disorders in populations such as singers where exceptional vocal demands are present (Hunter et al., 2020a; Roy and Bless, 2000). As correlation does not imply causation, it remains unclear whether occupational singing itself serves as a risk factor for developing pMTD, or if singers are more sensitive to pre-existing, low-impact vocal health issues that, over time, lead to increased severity.

7.1.2 Paralaryngeal Disorders Related to the Voice

Another class of disorders related to the voice exists; these are similar to functional disorders but affect the paralaryngeal musculature rather than the larynx itself. One such condition is Temporomandibular Disorder (TMD), an umbrella term encompassing disorders that affect the masticatory muscles and/or the temporomandibular joints (TMJs). See Fig. 7.2 for a diagram of the TMJ and surrounding area. TMD is the second most common musculoskeletal pain disorder and may become chronic; it is estimated that approximately 33% of the U.S. population experiences at least one TMD symptom. The prevalence of self-reported painful TMD (pTMD) was 21.9% among vocalists compared to 12.0% in instrumentalists (van Selms et al., 2019). While factors such as female sex, oral behaviors, and the number of daily practice hours were identified as risk factors, it is non-conclusive whether singing itself is a direct risk factor for pTMD.

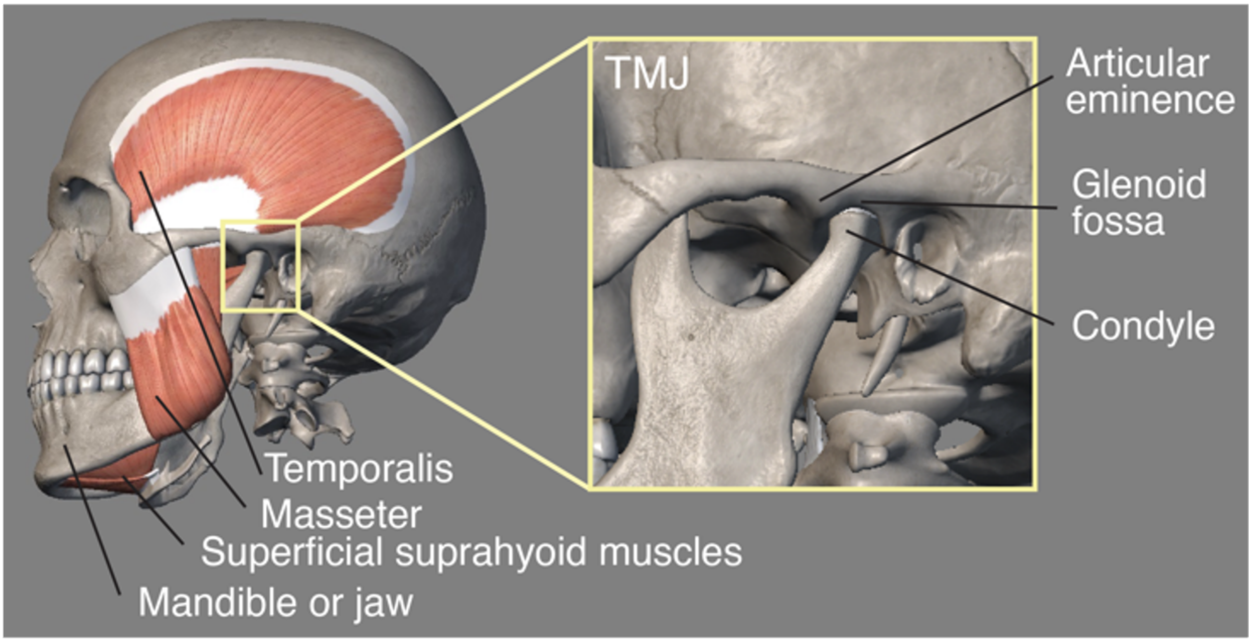


Figure 7.2: A diagram showing the location of the Temporomandibular Joint. Reprinted with permission from National Institute of Dental and Craniofacial Research (2018).

pTMD can significantly influence vocal function by altering mandibular mobility, increasing tension in the jaw and surrounding musculature, and affecting oral resonance and articulatory precision (Ohrbach et al., 2011; List and Axelsson, 2010). As occupational users of the head and neck, singers rely on coordinated action of the jaw, tongue, and associated musculature to shape the vocal tract and produce key acoustic features such as diction and vibrato. For singers, chronic pain or dysfunction in the TMJ can disrupt this coordination, potentially leading to maladaptive motor patterns and increased compensatory tension in the tongue, pharynx, or suprahyoid muscles, which can, in turn, impair vibratory freedom of the larynx (van Selms et al., 2019; Clukey, 2024).

Furthermore, in preliminary observations from two current studies conducted by the author of this thesis (Desjardins, Nestorova, Martel, and Chagnon (2024) and Zimmermann, Nestorova, Mac Master, Meloto, Komarova, Cossette, Hashemi, Moussa, and Manrique (2023)), there may be a connection between functional voice disorders like Muscle Tension Dysphonia (MTD) and the prolonged, maladaptive, miscoordinated articulatory patterns found in Temporomandibular Disorder (TMD).

7.1.3 Exploratory Findings and Interpretation: Functional and Paralaryngeal Disorders Apparent in Vibrato Variability

As exhibited in Chapters 3-6, time-varying analysis of vibrato is a valuable, intuitive diagnostic tool that impacts the perceptions and pedagogies of singers across Western and Eastern genre contexts. Anecdotally, many singing teachers and clinicians rely on vibrato as an indicator of vocal health and technique efficiency, using it to gauge the functionality of the vocal mechanism. While vibrato has traditionally been viewed as a byproduct of a healthy voice, it may be increasingly being considered a diagnostic marker of vocal dysfunction such as hyperfunction and stomatognathic issues.

A small number of the singers in the studies featured in this thesis self-reported and publicly disclosed voice or voice-related disorder diagnoses (including MTD and TMD) following each study's completion. Interestingly, this was corroborated both perceptually, as listeners remarked on potential functional issues, and acoustically, as the vibrato archetypes indicated high levels of complexity. These study findings accorded with these singers' self-disclosures of their post-study functional and/or paralaryngeal disorders (MTD and TMD). As revealed in the qualitative results in Chapter 4, expert listeners identified vocal excerpts with erratic vibrato variability as potentially indicative of compensatory tension conditions or voice-related disorders, using specific descriptors. In Chapter 3, the more complex / multiphasic vibrato half-extent time profile were noted with potential biomechanical implications for future study.

All four of the below vibrato samples were included in the recorded singer stimuli excerpts for the vibrato perception detailed in Chapter 4. These vibrato samples, within the context of their song, constituted the excerpts with the highest number of mentions from the listeners participating in the perceptual study. Most associated comments made note of the variability over time in these vibrato samples. Commonly used descriptive terms were related to functionality and production physiology, with phrases such as "overly tight," "forced by jaw," "muscle tension," and "imbalanced." Semantically, these descriptors align with common etiological and symptomatic characteristics of both MTD and TMD. Therefore, vibrato variability may indeed be both quantitatively and qualitatively used as a signifier for MTD and TMD in singers.

From the same data set presented in Chapter 3, when acoustically analyzing those specific singers' vibrato archetypes as case studies based on their vibrato half-extent time profile models, several interesting trends emerge.

As shown in Fig. 7.3, in two singers (Musical Theater and Opera) who self-reported their MTD diagnoses, there appears to be an internal, cycle-to-cycle instability and irregularity of the complex, biphasic, step-up and step-down vibrato half-extent time profiles.

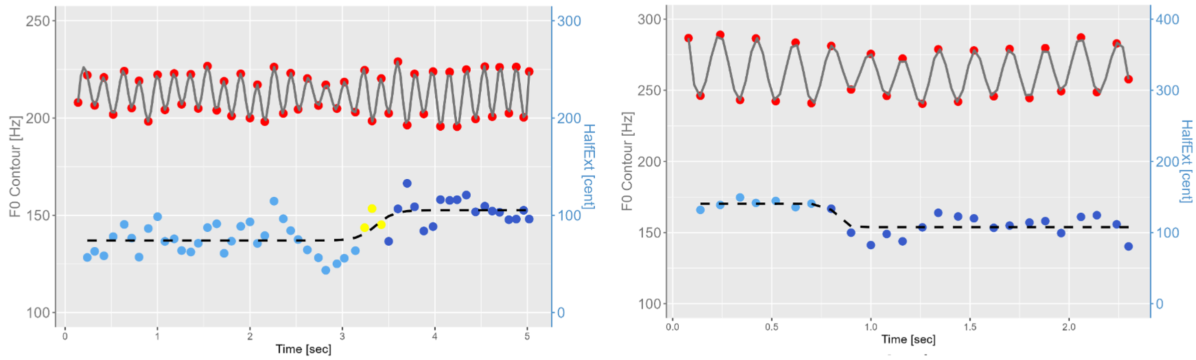


Figure 7.3: Two vibrato half-extent time profiles demonstrating internal, cycle-to-cycle instability and irregularity of complex biphasic vibrato of singers diagnosed with MTD post-factum to the acoustic study. The left panel shows the U-shaped deviation pattern in the step-up half-extent time profiles commonly found in the vibrato samples of this Musical Theater singer. The right panel shows the inverted S-curve pattern in the step-down half-extent time profiles that were quite typical in this Opera singer's vibrato samples.

Fig. 7.4 displays a vibrato archetype with an even more extreme internal, cycle-to-cycle instability and irregularity from a Jazz singer who self-reported their MTD diagnosis. In Chapter 3, the suggested classification of this sample was the quadriphasic (or highly complex) half-extent time profile.

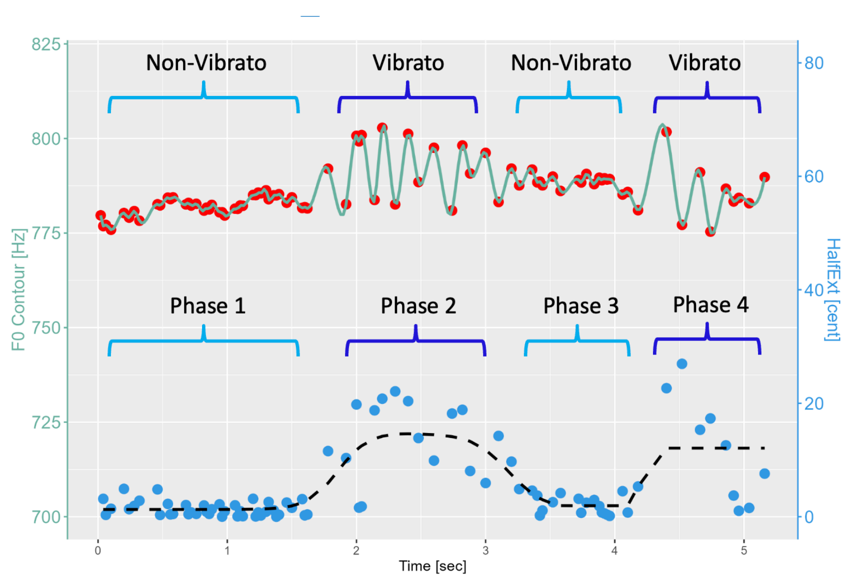


Figure 7.4: A vibrato half-extent time profile displaying internal, cycle-to-cycle instability and irregularity of a complex, quadriphasic vibrato of a Jazz singer diagnosed with MTD post-factum to the acoustic study.

As shown in Fig. 7.5, an Opera singer who self-reported their TMD diagnosis shortly after completion of the study, the classification according to the framework defined in Chapter 3 was the monophasic simple increasing half-extent time profile. This vibrato archetype was frequently found in this singer's vibrato samples, with speculations that their onset of vibrato was impaired or impeded due to biomechanical maladaptations in voice production caused by TMD.

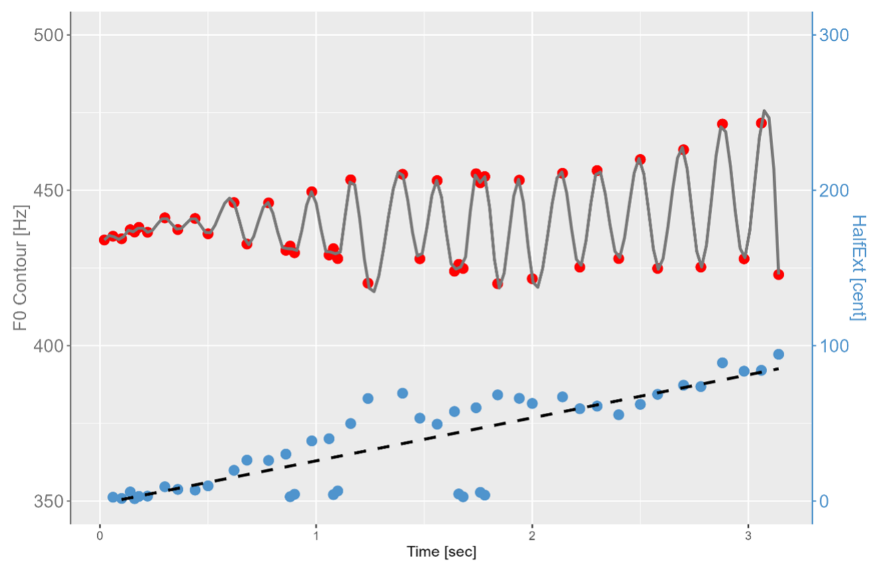


Figure 7.5: A vibrato half-extent time profile displaying a monophasic simple increasing vibrato of an Opera singer diagnosed with TMD post-factum to the acoustic study.

The hypothesis linking vocal vibrato variability to hyperfunctional and paralaryngeal disorders arose from the above case studies featured in this thesis research. While these preliminary observations highlight a potential tie between time-varying vibrato appearance and voice disorder onsets, it is important to approach this connection cautiously. As this represents the first documented connection between vibrato variability and functional paralaryngeal disorders in the literature, further quantitative and qualitative research is needed to elucidate the relationship between vocal vibrato and hyperfunctional and paralaryngeal disorders for generalizable and reliable conclusions.

Moreover, a previous study including thematic analysis of interviews with singers experiencing pMTD reveals a potential connection between vibrato and the functional disorder (Lepock, 2019). Several interview participants noted that their challenge with attempts to suppress their natural soloistic vibrato for choral singing may have contributed to the development of vocal dysfunction. However, Lepock (2019) further emphasizes the value of studies to investigate pMTD across all vocal tasks, and not just within the context of singing, or within a singular style of singing. Additionally, MTD and vibrato converge etiologically as idiosyncratic motor adaptations to physiological or psychological impulses, with the major distinguishing factor in MTD being the aversive stimuli disrupting instead of enabling muscle activation patterns via neural adaptations (Desjardins et al., 2022).

Based on the etiology of TMD, it is similarly plausible that TMD contributes to irregular vibrato characteristics such as amplitude instability or reduced modulation depth. Further research into TMD-related changes in jaw biomechanics, articulatory kinematics, and their acoustic correlates may reveal vibrato to be a sensitive, non-invasive indicator of paralaryngeal dysfunction in singers. TMD-associated pain or proprioceptive disruption may interfere with somatosensory feedback critical for pitch modulation and vibrato control, suggesting that early identification and management of such paralaryngeal dysfunctions could play a preventative role in vocal rehabilitation.

Vibrato, in conjunction with other vocal features, may serve as a diagnostic tool for preventatively identifying compensatory, maladaptive issues such as primary Muscle Tension Dysphonia (pMTD) and painful Temporomandibular Disorder (pTMD). This is grounded in the findings from this thesis that changes in vibrato over time reflect the coordination between multiple aspects of the vocal apparatus. By applying the vibrato variability models in this thesis to practical pedagogical and clinical contexts, voice professionals may refine their assessment of vibrato patterns to better monitor the internal health and coordination of the vocal system. Vibrato variability may even be applied to future auditory-perceptual scales that are the current gold standard in clinical voice assessments. This would allow teachers and clinicians to guide singers' vibrato more effectively in terms of technical development and vocal rehabilitation. Proper identification of maladaptive vibrato could help prevent and mitigate the onset of vocal disorders, offering a more practical and focused approach on compensatory, excessive tension in voice pedagogy. By integrating vibrato as a marker of vocal health, voice professionals can move beyond subjective aesthetic judgments, using vibrato as a tool for optimizing vocal performance and training while reducing the risk of vocal injury.

7.1.4 Vibrato Onset Time as (Re)Habilitative Therapy Marker

As seen in Groll et al. (2021), speakers with typical voices appeared to recruit distinct muscle groups for pitch elevation versus increased vocal effort, and did not exhibit the heightened Voice Onset Time variability characteristic of vocal hyperfunction, suggesting that such variability is more likely linked to disordered vocal motor control than to mere increases in vocal effort or strain. Similarly, the novel parameter, Vibrato Onset Time, proposed by the author of this thesis and highlighted in Chapter 3, may be considered as a marker for vocal effort. Though more research is warranted, Vibrato Onset Time trends across the specific singers post-diagnosed with MTD and TMD as compared to normophonic singers points to a connection between the Vibrato

Onset Time as a potential characteristic of MTD and TMD.

7.2 Other Considerations for Vibrato Variability

Vibrato is influenced by a confluence of physiological, psychological, and environmental factors—each of which can contribute to temporally-variant modulations in vibrato rate, extent, regularity, and shape. These moment-to-moment variations are not well captured by average-based vibrato metrics, underscoring the critical need for temporally sensitive vibrato analysis in both research and applied clinical or pedagogical settings. The sections below include synthesized information from a Critically Appraised Topic (CAT) literature review, alongside proposed interpretations and conclusions based on the available information and findings of this vibrato variability research.

7.2.1 Hormonal, Mental, Emotional, and Situational Influences

Hormonal changes— especially those associated with puberty, menstruation, pregnancy, and menopause— affect the structure and function of the vocal folds. Hormone receptors in the laryngeal tissue respond to fluctuations in estrogen and progesterone, which can influence tissue viscosity, fluid retention, and neuromuscular coordination (Abitbol et al., 1999). Vocal fold edema or reduced mucosal pliability during certain hormonal states may result in increased vibrato extent or reduced regularity, particularly evident in the time domain.

Psychological and emotional states, such as anxiety, stress, or emotional arousal, can also alter vibrato characteristics by modulating respiratory control, muscle tension, and autonomic nervous system activation (Scherer, 2003; Goldenberg et al., 2024; Sapienza and Hoffman, 2020). Emotional expressivity may enhance vibrato flexibility, while anxiety or over-effort can introduce unwanted tremor or irregularity. These changes often occur dynamically within musical phrases, requiring time-sensitive acoustic tools to meaningfully assess them. However, it is important to note that there is still a need in the voice science field for a more clarified, cross-disciplinary consensus on descriptions of vocal effort, vocal load, vocal loading, and vocal fatigue (Hunter et al., 2020b).

Heart rate and autonomic arousal have been linked to subtle modulations in vibrato behavior. In states of excitement or performance-induced stress, fluctuations in heart rate and respiration can interact with vibrato's cyclic modulations. Preliminary studies have observed possible synchronization or coupling between cardiovascular rhythms and vibratory oscillations of the voice (Müller and Lindenberger, 2011). Increased heart rate variability (HRV) may manifest as faster or less stable vibrato patterns—again supporting the importance of analyzing vibrato over time to capture these physiological echoes.

Hydration status and tissue viscosity are crucial but often underappreciated influences on vibrato stability. Proper systemic and superficial hydration reduces vocal fold viscosity, enabling more efficient mucosal wave movement and more regular, well-modulated vibrato (Nestorova et al., 2024a; Verdolini et al., 1994; Verdolini-Marston et al., 1990; Hartley and Thibeault, 2014; Miri et al., 2012). Dehydration or thickened mucus can increase biomechanical resistance, thereby reducing vibratory smoothness and potentially contributing to jitter in vibrato rate and shimmer in extent. Even transient dehydration— due to exertion, travel, or environment— may present perceptibly as a coarsened or unstable vibrato. Thus, vibrato may offer a non-invasive acoustic window into vocal fold viscosity and the efficiency of fluid transport within laryngeal tissues.

Situational and stylistic contexts also shape vibrato usage. Performers often adapt vibrato consciously or unconsciously to fit the acoustic setting, microphone presence, or aesthetic expectations of the genre, as shown in Bottalico et al. (2022); Redman et al. (2023). These adaptive strategies are time-bound and highly individualized, further reinforcing the inadequacy of averaged vibrato metrics.

While further research is needed, vibrato variability may serve as a sensitive marker of complex, multi-system physiological processes. Hormonal status, emotional condition, hydration, cardiovascular dynamics, and contextual demands all contribute to time-varying shifts in vibrato behavior. Time-resolved acoustic analysis captures these fine-grained changes, offering vital insight for vocal health monitoring, diagnostic voice science, and expressive performance training.

7.2.2 Vibrato across the Lifespan

Vocal vibrato also functions as a sensitive indicator of physiological development and neuromuscular integrity across the human lifespan. Its characteristics— presence, rate, extent, regularity, and variability— change considerably from childhood through older adulthood. These changes reflect a confluence of anatomical maturation, neuromuscular control mechanisms, and perceptual-motor coordination, all of which contribute to the production and perception of vibrato in singing. This section synthesizes acoustic, physiological, and pedagogical perspectives to chart the evolving morphology of vibrato over time.

Pediatric Vibrato through Childhood and Pubescence

In pediatric vocal development, vibrato is rarely intrinsic. According to Williams (2009), children— even those with balanced laryngeal activity and early training— typically do not exhibit naturally occurring vibrato prior to puberty. Instead, any vibrato-like oscillation in younger voices tends to be a consciously imposed “wobble” rather than an emergent product of neuromuscular coordination. As children approach puberty and gain more advanced technical skill, vibrato may begin to emerge organically, suggesting its dependence on both anatomical readiness and motor learning.

One proposed physiological explanation for the rarity of vibrato in prepubescent singers is the immaturity of the vocal ligament. Williams (2009) hypothesizes that the vocal ligament (the medial edge of the intermediate vocal fold layer) must reach a certain degree of development to provide sufficient body to the vocal fold, thereby enabling the vocalis muscle (medial portion of the thyroarytenoid (TA) muscle) to release adequately. This release is crucial for facilitating the agonistic-antagonistic interaction between the TA and cricothyroid (CT) muscles— a balance widely accepted as necessary for vibrato production. In the absence of a mature vocal ligament, this antagonistic relationship cannot be optimized, and vibrato remains underdeveloped or entirely absent. This hypothesis also suggests that some highly trained adult singers with minimal or imperceptible vibrato may have inherently slender vocal ligaments, limiting the extent to which vibrato can be produced.

Post-pubescent changes further alter vibrato production. As the vocal folds lengthen and the laryngeal cartilages begin to ossify, vibrato may emerge more reliably in technically proficient singers. However, its continued absence in adolescent or young adult singers may be diagnostically meaningful— often indicating excessive subglottal pressure, vocal fold stiffness, or other compensatory behaviors that inhibit vibratory freedom (Williams, 2009).

Adult Vibrato: Peak Neuromuscular Control and Expressive Modulation

In adult singers, vibrato rate and extent typically stabilize into a perceptually consistent and stylistically expressive oscillation. At this stage, vibrato becomes a highly trainable and controllable feature, modulated according to stylistic conventions and individual expressive intent. Even within normative parameters, vibrato exhibits considerable variability in relation to genre (as demonstrated by the work in this thesis), vocal technique, and somatic conditions such as fatigue or illness. Vibrato's volitional presence in healthy adult voices has led some researchers to regard the vibrato versatility in a singer as a sign of vocal efficiency and well-regulated phonatory control.

Aging Voices: Vibrato and Neurophysiological Decline

In later adulthood, vibrato undergoes perceptible acoustic and physiological changes. One of the most consistently reported phenomena is the gradual slowing of vibrato rate. Dejonckere et al. (1995) documented this trend in a cohort of 21 singers aged 20 to 65, observing a vibrato rate decrease from 5.4 Hz in the youngest to 4.7 Hz in the oldest participants. Similarly, Titze et al. (2002) references the longitudinal studies of H. Damsté et al. (1982) and Sundberg (1998), who found a comparable 1 Hz reduction in vibrato rate across the professional lifespans of trained singers performing the same repertoire. These studies collectively demonstrate that vibrato rate reduction may be considered a hallmark of aging in the singing voice.

The physiological mechanisms responsible for this decline are multifactorial. Titze et al. (2002) attributes the slowing to age-related reductions in nerve conduction velocity, increased muscle activation and contraction times, and diminished muscle tone or conditioning. These neuromotor degradations affect the timing and precision of laryngeal oscillations, which in turn manifest as slower, broader vibrato in the acoustic signal. Williams (2009) also highlights the impact of decreased neural innervation in the intrinsic laryngeal muscles, noting that such degeneration leads to delayed muscle response and increased vibrato extent.

Indeed, increased vibrato extent is another documented feature of the aging voice, particularly after age 40— a phenomenon often referred to as "senile vibrato." Dejonckere et al. (1995) report that this broader, slower vibrato is common in older singers, though interestingly, they found no correlation with voice type, suggesting that these changes occur consistently across *fächer* (voice types). While "senile vibrato" is not inherently pathological, it can deviate from stylistic norms and may require pedagogical adjustment or technical compensation, especially in performance contexts demanding more controlled vibrato characteristics. However, the inflammatory use of the term "senile vibrato" deserves re-examination, as do changes in vibrato extent variability across the aging process.

In addition to neuromuscular decline, age-related anatomical changes such as ossification of the laryngeal cartilages and alterations in bone conduction pathways may further impact vibrato production and perception (Makiyama and Hirano, 2017). Although the precise effects of ossification on auditory feedback mechanisms remain underexplored, these structural shifts likely contribute to the broader changes observed in vibrato acoustics among older adults.

Vibrato and Neurological Conditions

The variability of vibrato in later life also intersects with neurological health. Essential tremor (ET), a common movement disorder, can affect the larynx and introduce tremor-like modulations that may be mistaken for exaggerated or unstable vibrato. ET often co-occurs with neurodegenerative diseases such as Parkinson's disease, multiple sclerosis (MS), amyotrophic lateral sclerosis (ALS), and Alzheimer's disease, though it may also arise independently. In these cases, distinguishing between stylistic vibrato and pathological tremor becomes a clinical necessity.

The acoustic properties of vibrato—particularly its rate regularity and extent—can therefore serve as valuable indicators of neuromotor function and potential neurological decline. As such, vibrato holds promise as a non-invasive biomarker for tracking voice-related manifestations of neurodegenerative disorders. Continued research into the longitudinal patterns of vibrato in both healthy and pathological aging may offer new diagnostic tools and therapeutic strategies for clinicians, speech-language pathologists, and voice pedagogues alike.

7.2.3 Connection between Vibrato and Vocal Tremor

Vocal tremor, a neurological voice disorder that is a common symptomatic component of the aforementioned neurodegenerative diseases, often presents with involuntary oscillations in pitch and loudness that resemble vibrato but lack the consistency and control found in trained singers (Lester, 2014). Understanding the biomechanical and neural differences between vibrato and tremor is crucial for both differential diagnosis and treatment approaches.

Insights from vocal tremor perception research provide useful parallels for understanding vibrato characteristics in singing. Kreiman et al. (2003) found that listeners were most sensitive to changes in tremor rate when the oscillation was sinusoidal with small, stable amplitude—particularly when the rate was relatively fast, though tremor rate, regularity, and amplitude interact to impact perception. These results suggest that vibrato may also be perceived most clearly when it is steady, rapid, and exhibits minimal jitter or shimmer in both rate and extent. Such findings point to a psychoacoustic interaction between vibrato rate and extent (Glasner and Nix, 2023), where temporal regularity enhances perceptual clarity (Nestorova et al., 2025c).

Supporting this, studies in instrumental acoustics have shown that vibrato rate plays a greater role than extent in perceptual thresholds. For instance, research using synthesized string sounds found a just-noticeable difference (JND) for rate at around 6%, indicating fine perceptual resolution for temporal fluctuations. Broader pitch perception studies likewise demonstrate that frequency difference limens are influenced by duration, intensity, and fundamental frequency (f_o)—with longer and more stable tones facilitating greater sensitivity (Järveläinen, 2002; Zwicker et al., 1957; Moore, 1973, 2008; Friberg and Sundberg, 1994; Wier et al., 1977).

Speech science literature echoes these findings. Tremor becomes easier to detect in sustained phonation lasting over 500 milliseconds, as compared to connected speech (Lederle et al., 2012), reinforcing the importance of temporal length in identifying periodic modulations (Barkmeier-Kraemer and Story, 2010). Amplitude vibrato, or cyclic variations in loudness, is also relevant. While listeners typically do not perceive changes below a 15% threshold, they are more sensitive to loudness fluctuations than to spectral shifts. Amplitude modulations are perceived as approximately 1.3 times deeper than equivalent changes in spectral centroid, suggesting that loudness cues may contribute more strongly to vibrato perception than subtle shifts in timbre or harmonic structure (Almeida et al., 2021).

Together, these cross-disciplinary findings underscore the perceptual significance of the stability and regularity of vibrato rate and extent— factors that are critical not only for musical expression but also for differentiating healthy from pathological vibrato and tremor.

7.3 Conclusions and Discussion

In this chapter, the potential for vibrato and its time-varying properties to serve as a diagnostic tool was explored. The proposed vibrato variability analytical tools and instructional frameworks in earlier chapters of this thesis demonstrated practicality when applied to the clinical and medical case uses.

In particular, several of the vibrato samples used in the studies from earlier chapters of this thesis were acoustically re-examined, as the singers who recorded them self-reported diagnoses of Muscle Tension Dysphonia (MTD) and Temporomandibular Disorder (TMD) after the study completion. The trends explored in the temporally-variant archetypes and onset properties of their vibrato samples may serve as sensitive and intuitive biomarkers for diagnosing functional and paralaryngeal voice disorders such as MTD and TMD. However, these observations were drawn from individual, post-diagnosed cases and may not be broadly representative; future research must include controlled cohort studies to validate these exploratory findings. Expanding and refining these metrics will support more individualized, evidence-based voice training and rehabilitation.

A Critically Acclaimed Topic (CAT) synthesis revealed that time-varying vibrato is a dynamic phenomenon shaped by physiological, psychological, and environmental factors— including hormonal changes, hydration, emotional state, cardiovascular rhythms, and performance context— making it a sensitive marker of vocal health and function. Across the lifespan, vibrato reflects developmental, neuromuscular, and age-related changes, with implications for both expressive control and the early detection of voice and neurological disorders. The connection and differentiation of vibrato to vocal tremor highlights the need for more comprehensive, temporally-variant acoustic analysis baseline procedures in both clinical and pedagogical settings.

Through the integration of acoustic, perceptual, and clinical observations, the preliminary findings provide compelling support that disordered vibratory patterns— especially those exhibiting cycle-to-cycle instability— may reflect underlying neuromuscular and articulatory maladaptations. The potential application of vibrato variability models and metrics, including the half-extent time profiles and Vibrato Onset Time, in voice therapy and pedagogical monitoring offers a novel, non-invasive tool for early detection and personalized intervention. The exploratory observations derived from the studies conducted in this thesis position vibrato not merely as a stylistic or aesthetic element, but as a timbre-defining feature intricately linked to vocal function and health, both for singers and non-singers.

Chapter 8

Conclusion and Future Work

8.1 Summary of Findings and Contributions

The work presented in this thesis addresses existing analytical and pedagogical gaps by comprehensively exploring the diversity of vocal vibrato as it varies over time across musical genres. This acoustic, perceptual, ethnographic, and pedagogical research is distinguished by its multidimensional nature; vibrato variability is investigated scientifically and empirically while mixed-methods research studies are translated into practical pedagogical and clinical tools.

Currently accessible information and vibrato standards lead to Western-biased analytical and pedagogical conclusions regarding vocal technique. Only presenting average vibrato rate and extent does not fully characterize non-uniform vibrato that is pervasive in singing. Few studies have objectively examined vibrato's temporally-variant properties or involved singers intentionally varying their natural vibrato, an stylistic device typical of many genres, especially in Eastern folk music.

The cross-disciplinary research conducted in this thesis fills some of these gaps, with the aim to highlight a more equitable representation of musical demographics and cultural geographies alongside a deeper understanding of the production and functioning of the vocal mechanism in vibrato. With a more genre-inclusive perspective, this thesis contributes useful methodological, diagnostic, and instructional investigation and evaluation to transform and modernize vibrato analysis and training. With the current considerable lack of research in vibrato diversity, singing teachers of all styles need objective, tangible, and updated information about vibrato as an integral and feasible pedagogical device in order to effectively address it with their students. This work aims to enrich teaching capabilities and performance science for singers of all styles.

This dissertation has offered an interdisciplinary compilation of acoustic, perceptual, pedagogical, and ethnographic investigations into naturally-occurring, time-varying vocal vibrato across multiple musical genres. Vocal vibrato, a multifaceted element integral to vocal technique and artistry, is shaped by both voluntary and involuntary mechanisms. Traditional methods for analyzing vibrato rely heavily on average-based metrics, which fail to capture the dynamic, real-time fluctuations characteristic of natural singing. Moreover, these norms are largely informed by studies of Western Classical music, revealing a lack of stylistic and cultural diversity in current vibrato research. As a result, non-Western and non-classical traditions are often misrepresented or overlooked in analysis, interpretation, and pedagogy.

The first acoustic study applied acoustic analysis to 16 selected sung phrases from a previous dataset, employing sinusoidal modeling, f_o band-pass filtering, and long-term average spectra (LTAS). Relationships between vibrato characteristics and half-extent change over time were evaluated using linear, polynomial, and non-linear models. While the Coefficient of Variation (CV) effectively captured uniform vibrato, it overestimated variability for more complex, irregular patterns. To address this, a four-parameter logistic (4-PL) regression model is introduced as a more accurate tool for capturing the complexity of multiphasic vibrato. The study also outlined a new definition and classification system for monophasic to multiphasic vibrato as half-extent time profiles. Finally, the diverse analytical methods of Half-of-Extent (HoE) as compared to Relative Half-Extent (rHE) were elucidated and interpreted, as were novel metrics, such as Vibrato Onset Time (VOT), proposed.

The second perceptual study engaged 30 participants— 15 vocalists and 15 instrumentalists— who listened to 48 excerpts from a shared unaccompanied repertoire performed by 12 professional vocalists across Opera, Musical Theater, and Jazz. Participants provided open-ended descriptions, which were processed using linguistic lemmatization and thematic categorization. This secondary, extended analysis produced 65 unique descriptor categories grouped into 11 broader themes, with 240 total vibrato references. Notably, 94% of participants commented on vibrato, most often in relation to its variability, complexity, and evolving temporal patterns. Many listeners associated vibrato changes with underlying vocal production physiology, showing consistency across genres. Subtle differences also emerged between vocalists' and instrumentalists' descriptive vocabularies. The findings reinforce the perceptual importance of vibrato as a temporally dynamic phenomenon.

Averaging vibrato features inherently ignores such variability over time, which can provide significant information. The results of the acoustic methodology study demonstrated that vocal vibrato half-extent time profiles are an important analytical characteristic beyond using averages when evaluating and reporting vibrato. Variability in vibrato extent impacted the perception and accurate genre classification, revealing and reflecting acoustic trends and perceptual salience of Opera, Musical Theater, and Jazz genres, next extending to non-Western singing styles. In conclusion, these studies found that vibrato half-extent variability over time is an indicator of vocal genre and the proposed novel parametric models well-represent both simple and complex vibrato shapes inherent to genre.

A field research project was conducted with acoustic, ethnographic, and pedagogical elements, shifting focus to Bulgarian folk singing, an Eastern musical tradition known for its distinctive vocal techniques. Analysis revealed highly nuanced and complex vibrato-like ornamental forms that differ significantly from Western norms. The study also documented culturally embedded teaching methods, highlighting the pedagogical value of this tradition and the importance of acknowledging cultural context in vibrato training.

Given the absence of a unified framework or pedagogy for teaching vibrato, the pedagogical component of this research evaluated how scientific findings can be practically applied to "vibrato in the wild," (Nestorova et al., 2025d) as vibrato appears in singers in training. It contrasted direct (explicit) and indirect (implicit) instruction, advocating for the integration of vibrato variability for a style-specific training. Custom exercises under the proposed "FUNctional" Vibrato Pedagogy Framework are recommended to help singers develop versatile, flexible, and efficient vibrato across styles.

The final investigation explored naturally-occurring vibrato variability as both a stylistic indicator and a potential diagnostic marker of vocal health, with broader implications for the medical and clinical fields. Highlighting vibrato variability as a clinically and pedagogically meaningful marker of vocal function, the more complex vibrato patterns— particularly those showing cycle-to-cycle instability— showed potential to signal underlying hyperfunctional and paralaryngeal disorders such as Muscle Tension Dysphonia (MTD) and Temporomandibular Disorder (TMD). While these findings were purely exploratory and based on case studies, they suggest promising applications for vibrato metrics like half-extent profiles and Vibrato Onset Time in early detection and targeted intervention. A synthesis further illustrated vibrato's responsiveness to physiological, psychological, and environmental influences, reinforcing its diagnostic potential. The connections between vibrato and neuromotor and neuromuscular conditions such as vocal tremor, as well as the implications for vibrato variability analysis to neurodegenerative disorders, were examined. As more clinical insight, validation, and normative measures in cohort-based research are needed in these areas for both singers and non-singers, this is the current and future work of the author of this thesis.

Altogether, these interdisciplinary and mixed-methods studies demonstrated that vibrato variability is perceptually meaningful and acoustically quantifiable. This work proposes a more inclusive, genre-specific model for vibrato analysis— one that moves beyond static averages to embrace temporal complexity. The findings hold significant implications for vocal diagnostics, pedagogy, and therapeutic voice care, offering a foundation for more nuanced and effective approaches to vibrato in both teaching and clinical contexts.

8.2 Future Work and Directions

Beyond the aforementioned collaborative projects in progress including the author of this thesis (Zimmermann et al., 2023; Desjardins et al., 2024), future expansions of the work completed in this thesis include examination of the temporal variance of vibrato rate and extending the half-extent time profile models developed in this thesis on rate variability. This currently ongoing work by the author of this thesis may bring about valuable insights for periodicity analyses in the voice.

Furthermore, more robust perceptual testing on vibrato variability as it relates to vocal function can be undertaken, including use of standardized scales such as the Evaluation of the Ability to Sing Easily (EASE) (Phyland et al., 2014) and others. Extending the vibrato sample collection to include more genres from World Music is also highly recommended.

Ongoing work is currently underway by the author of this thesis center on the integration of biomechanical measures, such as airflow (Nandamudi and Scherer, 2019) and sub-, intra-, and supra-glottal pressure interactions with time-varying vibrato. An interesting area of application could be incorporating Ecological Momentary Assessment (EMA) within ambulatory voice monitoring (Cheema et al., 2024) to enhance ecological validity, capturing vibrato patterns and vocal behaviors in real-world, context-rich environments. The eventual goal of this work is the culmination of a structured framework of intervention options; the target exercise, desired outcome, and an estimation of timeline to address efficiently in a taxonomy similar to Van Stan et al. (2015).

Future directions include the continued development of the Embodied Music Lab (EML)'s Vibrato Tool by Howell and Nestorova (2025), designed as an open-source Praat plugin (Boersma and Weenink, 2025) for automatized acoustic vibrato analysis. This tool currently incorporates several vibrato metrics developed in this dissertation, with plans to expand its capabilities to support broader diagnostic and pedagogical applications. See Fig. 8.1 for a view of the graphical user interface (GUI) of the output of this plugin tool.

Building on the vibrato variability models developed in this thesis, a collaborative generative AI and machine learning project by Nestorova, Xie, Zhang, Nyanyo, Zhao, and Mongeau (2025d) is currently underway to classify and distinguish different variable vibrato types defined in this thesis, with particular attention to those associated with medically confirmed cases of Muscle Tension Dysphonia (MTD) and Temporomandibular Disorder (TMD). This pilot study applies machine learning models to a diverse dataset of vibrato samples, leveraging acoustic metrics, expert perceptual evaluations, and biomechanical correlates gathered. Preliminary results are highly promising: trained models have achieved high accuracy across training, validation, and test sets, successfully differentiating between functionally typical vibrato, TMD-related vibrato irregularities, and MTD-associated patterns. These early findings suggest that vibrato carries

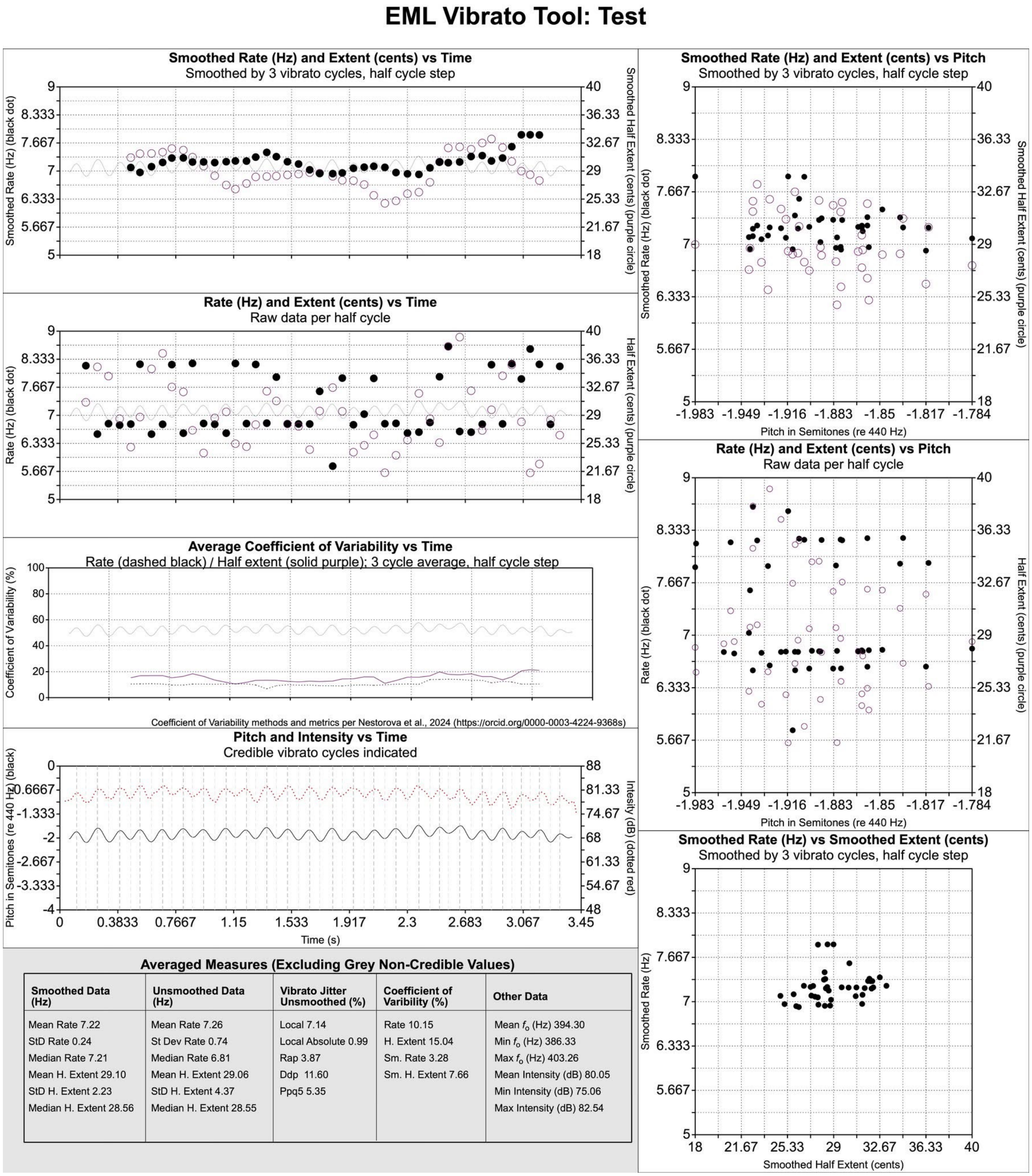


Figure 8.1: An output display of a vibrato sample analyzed with the EML Vibrato Tool, an open-source Praat plugin designed for automatized acoustic vibrato analysis (Howell and Nestorova, 2025).

embedded signatures of underlying neuromuscular control and paralaryngeal function, positioning it as a viable acoustic biomarker for early detection of voice disorders. The integration of these AI models into a diagnostic framework represents a novel translational application of this research work's vibrato analysis tools, with the potential to revolutionize clinical voice assessments by enabling non-invasive, data-driven detection of voice pathologies at earlier stages and with greater specificity.

Extending this further, a currently conducted pilot study by Nestorova, Maryn, Stark, McDowell, Pierce, Maxfield, and Barkmeier-Kraemer (2025b) investigates the biomechanical and neural mechanisms underlying vocal vibrato and vocal tremor, with a focus on automaticity and primary motor control. Using a multi-signal approach—combining acoustic, aerodynamic (pneumotachographic), Respiratory inductance plethysmography (RIP), electroglottographic, electromyographic, laryngoscopic and stroboscopic, and neural imaging data (EEG and/or fNIRS)— this research will examine how glottal airflow, subglottal pressure, production physiology, and cortical activity correlate with frequency and amplitude modulation in vibrato across trained and untrained singers, as well as individuals with vocal tremor. The study aims to identify where vibrato and tremor share common physiological pathways and where they diverge. Objectives of the study include examining the primary predictors of vibrato emergence and determining to what extent vibrato operates as an automatic or volitional process. In doing so, the study also seeks to clarify the physiological and neural distinctions between functional vibrato and pathological tremor. Insights gained may inform differential diagnosis and therapeutic strategies for neurogenic voice disorders and further elucidate the sensorimotor integration involved in expressive voice modulation.

Incorporating a structured approach to evaluating vibrato in both clinical and pedagogical settings could significantly enhance voice rehabilitation practices. By viewing vibrato not just as an aesthetic feature but as a functional marker of vocal health, professionals could improve their ability to diagnose and treat various voice disorders, particularly those involving tension and inefficiency. A deeper understanding of how vibrato interacts with conditions like MTD could help refine voice therapy, offering more targeted strategies for restoring vocal function. Ultimately, this would transform the study and teaching of vibrato from a passive observation of vocal artistry into an active diagnostic tool for maintaining vocal health and optimizing performance.

8.3 Concluding Remarks

This dissertation has offered a foundation for a more nuanced, inclusive, culturally-informed, and evidence-based understanding of vibrato as both an expressive feature and a diagnostic signal of vocal function. By integrating acoustic, perceptual, ethnographic, and pedagogical analyses across diverse genres and cultural contexts, it moves beyond traditional, Western-centric approaches and average-based metrics to reveal vibrato's dynamic, time-varying nature. The analytical tools and pedagogical frameworks developed here— spanning from psychoacoustic models to pedagogical and clinical applications— provide a new way to evaluate vibrato not only as a stylistic device, but also as a window into vocal health, biomechanical coordination, and artistry.

Through interdisciplinary collaboration and mixed-methods inquiry, this work bridges the gap between voice science, music education, and clinical practice. It highlights the voice's idiosyncratic complexity and advocates for a more equitable, data-informed, and culturally sensitive approach to training and diagnostics. As voice pedagogy begins to embrace more technological and scientific perspectives, the models and methodologies introduced in this dissertation can serve as a catalyst for future research and innovation.

By centering vibrato within the broader framework of vocal function and stylistic identity, this research repositions it as a valuable pedagogical and clinical tool. The contributions of this work are meant to empower researchers, educators, clinicians, and singers alike— to analyze more representatively, teach more precisely, listen more deeply, and understand more fully the rich variability and significance of the singing voice. Ultimately, this dissertation seeks to bridge the gap between the art, the science, and the health of the voice, fostering a more holistic and informed future for vocal training and care.

References

J. Abitbol, P. Abitbol, and B. Abitbol. Sex hormones and the female voice. *Journal of Voice*, 13(3): 424–446, 1999.

B. Ackermann and E. Altenmüller. The development and use of an anatomy-based retraining program (MusAARP) to assess and treat focal hand dystonia in musicians–A pilot study. *Journal of Hand Therapy*, 34(2):309–314, 2021. Publisher: Elsevier.

G. Acosta Martínez and H. Daffern. Complexity of Vocal Vibrato in Opera and Jazz Recordings: Insights From Entropy and Recurrence Analyses. *Journal of Voice*, 69(4), 2023a.

G. Acosta Martínez and H. Daffern. A Pilot Study of Vocal Vibrato Incorporating Nonlinear Time Series Analysis. In *Proceedings of the Stockholm Music Acoustics Conference*, pages 69–72, 2023b.

J. F. Agricola and P. F. Tosi. *Introduction to the Art of Singing by Johann Friedrich Agricola*. Cambridge University Press, 1757.

A. Almeida, E. Schubert, and J. Wolfe. Timbre Vibrato Perception and Description. *Music Perception*, 38(3):282–292, 2021.

American Academy of Teachers of Singing. Healthy Vocal Technique and the Performance of Early Music. *Journal of Singing*, 52(3):83–85, 1994.

American Academy of Teachers of Singing. The Singing Voice Specialist: An Overview of the History, Training Protocols, Career Opportunities, and General Considerations. *Journal of Singing*, 79(1):9–11, 2022. Publisher: National Association of Teachers of Singing.

American Academy of Teachers of Singing, A. Addison, L. Christiansen, P. Craig, J. E. Douglas, S. Emmons, J. W. Gregg, H. Harris, H. Hodam, B. Honn, M. Keenze, P. Kiesgen, A. Lavanne, J. LoVetri, E. Mannion, J. McCollum, S. McCoy, J. McLean, R. Miller, D. Moore, R. Oberlin, C. Owen, J. Powell, G. Shirley, R. Sjoerdsma, S. Verrett, and P. S. Wormhoudt. In support of Contemporary Commercial Music (nonclassical) voice pedagogy. *Journal of Singing*, 65(1): 7–11, 2008.

American Speech-Language-Hearing Association. Voice Disorders.

S. Anand, J. M. Wingate, B. Smith, and R. Shrivastav. Acoustic Parameters Critical for an Appropriate Vibrato. *Journal of Voice*, 26(6):820.e19–820.e25, 2012. Publisher: Elsevier.

D. I. Anderson, R. A. Magill, A. M. Mayo, and K. A. Steel. Enhancing motor skill acquisition with augmented feedback. In *Skill Acquisition in Sport*. Routledge, 3 edition, 2019. Num Pages: 17.

L. Antonova-Vasileva. Linguistic Geography's Role in Clarifying the Linguistic Situation in Specific Language Regions in Southeast Europe. In *Handbook of the Changing World Language Map*, pages 1–19. Springer, Cham, 2019.

D. R. Appelman. *The Science of Vocal Pedagogy: Theory and Application*. Indiana University Press, 1986.

C. Arneson and K. S. Brown. *Musical Theater Voice Pedagogy: The Art & Science*. Inside View Press, 2023.

A. E. Aronson and D. M. Bless. *Clinical Voice Disorders: Medicine & Health Science*. Thieme, 4th edition, 1992.

I. Arroabarren and A. Carlosena. Vibrato in Singing Voice: The Link between Source-Filter and Sinusoidal Models. *EURASIP Journal on Advances in Signal Processing*, 2004(7):1–14, 2004. Number: 7 Publisher: SpringerOpen.

B. d. Bacilly. *L'Art de bien chanter de M. de Bacilly . Augmenté d'un discours qui sert de réponse à la critique de ce traité, et d'une plus ample instruction pour ceux qui aspirent à la perfection de cet art...* Paris, 1679.

C. P. Baker, M. Brockmann-Bauser, S. C. Purdy, and T. O. Rakena. High and Wide: An *In Silico* Investigation of Frequency, Intensity, and Vibrato Effects on Widely Applied Acoustic Voice Perturbation and Noise Measures. *Journal of Voice*, 73(4), 2023.

J. Baker. Psychogenic voice disorders and traumatic stress experience: a discussion paper with two case reports. *Journal of Voice*, 17(3):308–318, 2003. Publisher: Elsevier.

J. Baker, D. Ben-Tovim, A. Butcher, A. Esterman, and K. McLaughlin. Psychosocial risk factors which may differentiate between women with Functional Voice Disorder, Organic Voice Disorder and a Control group. *International Journal of Speech-Language Pathology*, 15(6):547–563, 2013. Publisher: Taylor & Francis ;https://doi.org/10.3109/17549507.2012.721397.

E. Balaguer-Ballester, N. R. Clark, M. Coath, K. Krumbholz, and S. L. Denham. Understanding Pitch Perception as a Hierarchical Process with Top-Down Modulation. *PLOS Computational Biology*, 5(3):e1000301, 2009. Publisher: Public Library of Science.

M. Bane, V. Angadi, E. Dressler, R. Andreatta, and J. Stemple. Vocal function exercises for normal voice: The effects of varying dosage. *International journal of speech-language pathology*, 21(1): 37–45, 2019.

J. Barkmeier-Kraemer and B. Story. Conceptual and Clinical Updates on Vocal Tremor. *The ASHA Leader Archive*, 15(14):16–19, 2010. Publisher: American Speech-Language-Hearing Association.

I. Bartlett and M. L. Naismith. An Investigation of Contemporary Commercial Music (CCM) Voice Pedagogy: A Class of its Own? *Journal of Singing*, 76(3):273–282, 2020.

A. I. Batkova. *Pedagogical technologies for building singing skills in working with children aged 7 to 10 years in an out-of-school vocal group*. Ph.D., Academy of Music, Dance and Visual Arts “Prof. Asen Diamandiev”, Plovdiv, 2023.

A. S. Becker and P. J. Watson. The Use of Vibrato in Belt and Legit Styles of Singing in Professional Female Musical-Theater Performers. *Journal of Voice*, 62(4), 2022. Publisher: Elsevier.

C. Bernhard. *Von der Singe-Kunst oder Manier*. Frankfurt, 1650.

A. d. A. Bezerra, S. Cukier-Blaj, A. Duprat, Z. Camargo, and L. Granato. The Characterization of the Vibrato in Lyric and Sertanejo Singing Styles: Acoustic and Perceptual Auditory Aspects. *Journal of Voice*, 23(6):666–670, 2009. Publisher: Elsevier.

S. Bhatt, L. Scearce, D. Frank-Ito, T. Nixon, and K. A. Umstead. A Human-Centered Design Approach to SOVTE Straw Phonation Instruction. *Journal of Voice*, page S0892199723000188, 2023.

G. M. Bidelman, J. T. Gandour, and A. Krishnan. Musicians demonstrate experience-dependent brainstem enhancement of musical scale features within continuously gliding pitch. *Neuroscience letters*, 503(3):203–207, 2011.

A. Bjørklund. Analyses of Soprano Voices. *The Journal of the Acoustical Society of America*, 33(5): 575–582, 1961.

P. Boersma. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In *Proceedings 17*, pages 97–110, University of Amsterdam, 1993.

P. Boersma and D. Weenink. Praat: doing Phonetics by Computer, 2025.

A. Bonjyotsna and M. Bhuyan. Analytical Study of Vocal Vibrato and Mordent of Indian Popular Singers. *Journal of Voice*, 30(6):764.e11–764.e22, 2016. Publisher: Elsevier.

A. Borisova. The Ornaments of the Bulgarian Folk Song – Folkloristic and Pedagogical Aspects. In *Proceedings of Angel Kanchev University of Ruse*, volume LI of *6.3*, 2012.

A. Borisova. Folk singing – authentic and "academic". In *Collection of articles*, pages 676–687, Veliko Tarnovo, 2013. University of Veliko Tarnovo "St. Cyril and St. Methodius" Press.

M. K. Bosheva. Comparative Analysis of Voice Training in Classical, Folk and Pop Singing. *University of Plovdiv "Paisii Hylendarski" Science and Technologies: Education*, 3(8):86–89, 2013.

P. Bottalico, S. Graetzer, and E. J. Hunter. Effect of training and level of external auditory feedback on the singing voice: Pitch inaccuracy. *Journal of Voice*, 31(1):122.e9–122.e16, 2017.

P. Bottalico, N. Łastowiecka, J. D. Glasner, and Y. G. Redman. Singing in different performance spaces: The effect of room acoustics on vibrato and pitch inaccuracy. *The Journal of the Acoustical Society of America*, 151(6):4131, 2022.

B. Boyanov and S. Hadjitodorov. Acoustic analysis of pathological voices. A voice analysis system for the screening of laryngeal diseases. *IEEE Engineering in Medicine and Biology Magazine*, 16 (4):74–82, 1997. Conference Name: IEEE Engineering in Medicine and Biology Magazine.

K. Bozeman. *Kinesthetic Voice Pedagogy 2: Motivating Acoustic Efficiency*. Inside View Press, 2021.

P. L. Broaddus-Lawrence, K. Treole, R. B. McCabe, R. L. Allen, and L. Toppin. The effects of preventive vocal hygiene education on the vocal hygiene habits and perceptual vocal characteristics of training singers. *Journal of Voice*, 14(1):58–71, 2000. Publisher: Elsevier.

J. C. Brown. Fundamental Frequency Tracking and Applications to Musical Signal Analysis. In J. W. Beauchamp, editor, *Analysis, Synthesis, and Perception of Musical Sounds: The Sound of Music*, Modern Acoustics and Signal Processing, pages 90–121. Springer, New York, NY, 2007.

O. L. Brown. *Discover Your Voice: how to Develop Healthy Voice Habits*. Singular Publishing Group, 1996.

W. Brown, H. B. Rothman, and C. M. Sapienza. Perceptual and acoustic study of professionally trained versus untrained voices. *Journal of Voice*, 14(3):301–309, 2000.

E. M. Burns. Intervals, scales, and tuning. In *The psychology of music, 2nd ed*, Academic Press series in cognition and perception: A series of monographs and treatises, pages 215–264. Academic Press, San Diego, CA, US, 1999.

E. M. Burns and W. D. Ward. Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *The Journal of the Acoustical Society of America*, 63(2):456–468, 1978.

S. Capobianco, A. Nacci, O. Calcinoni, L. Bruschini, S. Berrettini, and P. Bottalico. Assessing acoustic parameters in early music and romantic operatic singing. *Journal of Voice*, 2023.

N. Carter, J. A. Hopkin, and C. Dromey. Volitional control of vibrato in trained singers. *Journal of Singing*, 67(1):9–18, 2010. Publisher: National Association of Teachers of Singing.

J. L. Chapman. *Singing and Teaching Singing: A Holistic Approach to Classical Voice.* Plural Publishing, Incorporated, 3rd edition, 2017.

A. J. Cheema, K. L. Marks, H. Ghasemzadeh, J. H. Van Stan, R. E. Hillman, and D. D. Mehta. Characterizing Vocal Hyperfunction Using Ecological Momentary Assessment of Relative Fundamental Frequency. *Journal of Voice*, 2024.

W. K. Chen. *The Electrical Engineering Handbook.* Elsevier, 2004. Google-Books-ID: qhHsSlaz-GrQC.

E. F. Clarke. Rhythm and Timing in Music. In *The Psychology of Music*, pages 473–500. Elsevier, 1999.

J. Clukey. Temporomandibular Disorders and the Singing Voice: A Summary of the Research. *Voice and Speech Review*, 0(0):1–24, 2024. Publisher: Routledge ;https://doi.org/10.1080/23268263.2024.2427547.

B. Coffin. *Coffin's Overtones of Bel Canto: Phonetic Basis of Artistic Singing : with 100 Chromatic Vowel-chart Exercises.* Scarecrow Press, 1980.

G. L. Conforti. *The Joy of Ornamentation: Being Conforto's Treatise on Ornamentation (Rome, 1593).* Pro-Am Music Resources, Incorporated, 1989.

I. Cossette, B. Fabre, V. Fréour, N. Montgermont, and P. Monaco. From Breath to Sound: Linking Respiratory Mechanics to Aeroacoustic Sound Production in Flutes. *Acta Acustica united with Acustica*, 96(4):654–667, 2010. Publisher: European Acoustics Association.

R. J. Crampton. *A Concise History of Bulgaria.* Cambridge University Press, 2005.

L. Crocco, P. McCabe, and C. Madill. Principles of Motor Learning in Classical Singing Teaching. *Journal of Voice*, 34(4):567–581, 2020.

H. Daffern, J. Brereton, and D. Howard. The impact of vibrato usage on the perception of pitch in early music compared to grand opera. In S. F. d'Acoustique, editor, *Proceedings of the Acoustics 2012 Nantes Conference*, Nantes, France, 2012.

C. d'Alessandro and M. Castellengo. Perception of vocal pitch vibrato in short tones. *The Journal of the Acoustical Society of America*, 89(4B_Supplement):1987–1987, 1991. Publisher: AIP Publishing.

J. Davids and S. LaTour. *Vocal Technique: A Guide to Classical and Contemporary Styles for Conductors, Teachers, and Singers, Second Edition.* Waveland Press, 2020.

A. de Cheveigné and H. Kawahara. YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917–1930, 2002.

P. H. Dejonckere, M. Hirano, and J. Sundberg. *Vibrato.* Singular Pub., 1995.

D. S. Delviniotis and S. Theodoridis. On Exploring Vocal Ornamentation in Byzantine Chant. *Journal of Voice*, 33(2):256.e17–256.e34, 2019. Publisher: Elsevier.

L. Demany and C. Semal. Learning to perceive pitch differences. *The Journal of the Acoustical Society of America*, 111(3):1377–1388, 2002.

B. E. DeSilva. *A Survey of the Current State of Contemporary Commercial Music (CCM) Vocal Pedagogy Training at the Graduate Level.* Doctor of Musical Arts, Temple University, 2016.

M. Desjardins, C. Apfelbach, M. Rubino, and A. K. Verdolini. Integrative Review and Framework of Suggested Mechanisms in Primary Muscle Tension Dysphonia. *Journal of Speech, Language, and Hearing Research*, 65(5):1867–1893, 2022. Publisher: American Speech-Language-Hearing Association.

M. Desjardins, T. Nestorova, A. Martel, and F. Chagnon. Chronicité dans les troubles musculosquelettiques de la voix : une étude préliminaire des profils cliniques et psychosociaux de patients avec symptômes aigus et chroniques. Manuscript in preparation for Logopedics Phoniatrics Vocology, 2024.

J. A. Deutsch and J. K. Clarkson. Nature of the Vibrato and the Control Loop in Singing. *Nature*, 183(4655):167–168, 1959. Number: 4655 Publisher: Nature Publishing Group.

J. A. Diaz, A. B. Rangarathnam, A. Medina, V. D. Khare, and H. R. Fisher. Preliminary Perceptual Findings of Singing After Flow Phonation Voice Exercises for Singers in Training. *Journal of Voice*, 72(4), 2023.

B. M. Doscher. *The Functional Unity of the Singing Voice.* Scarecrow Press, 2nd edition, 1994.

D. Doskov, T. Ivanov, and B. Boyanov. Comparative Analysis of Singer's High Formant in Different Type of Singing Voices. *Folia Phoniatrica et Logopaedica*, 47(5):291–295, 1995.

J. Driedger, S. Balke, and S. Ewert. Template-Based Vibrato Analysis of Musical Signals. *New York City*, 2016.

C. Dromey, N. Carter, and A. Hopkin. Vibrato Rate Adjustment. *Journal of Voice*, 17(2):168–178, 2003. Publisher: Elsevier.

S. Duvvuru. *The Effect of Timbre and Vibrato on Vocal Pitch Matching Accuracy.* Doctor of Philosophy, University of Tennessee Health Science Center, Memphis, TN, 2012.

T. J. Edmonds and D. M. Howard. An Investigation in the Measurable Differences between Pitch Perception in the Voice and Pitch Perception of External Sound Sources. *Journal of Voice*, 39 (3):656–663, 2025.

E. Ekholm, G. C. Papagiannis, and F. P. Chagnon. Relating objective measurements to expert evaluation ofvoice quality in western classical singing: Critical perceptual parameters. *Journal of Voice*, 12(2):182–196, 1998.

H. M. Erickson. Mobile Apps and Biofeedback in Voice Pedagogy. *Journal of Singing*, 77(4): 485–500, 2021.

M. L. Erickson. Can Inexperienced Listeners Hear Who Is Flat? The Role of Timbre and Vibrato. *Journal of Voice*, 30(5):638.e9–638.e20, 2016.

K. Evanini and C. Lai. The importance of optimal parameter setting for pitch extraction. *The Journal of the Acoustical Society of America*, 128(4):2291–2291.

I. Ferrante. Vibrato rate and extent in soprano voice: A survey on one century of singing. *The Journal of the Acoustical Society of America*, 130(3):1683–1688, 2011. Publisher: Acoustical Society of America.

L. Fornhammar, J. Sundberg, M. Fuchs, and L. Pieper. Measuring Voice Effects of Vibrato-Free and Ingressive Singing: A Study of Phonation Threshold Pressures. *Journal of Voice*, 36(4):479–486, 2022.

M. Forsyth. Folk song as an academic discipline?, 2016. Accessed via https://martha.forsyths.org/2016/06/22/folk-song-as-an-academic-discipline/.

A. Friberg and J. Sundberg. Just Noticable Difference in duration, pitch and sound level in a musical context. pages 339–340, 1994.

C. Friedlander. Vibrato Hell, 2012. Accessed via https://www.claudiafriedlander.com/the-liberated-voice/2012/05/vibrato-hell.html.

F. K. Gable. Some Observations Concerning Baroque and Modern Vibrato. *Performance Practice Review*, 5(1):90–102, 1992.

J. García, C. Taglialatela, P. Kudumakis, I. Barbancho, L. Tardon, and M. Sandler. Interactive Music Applications by MPEG-A Support in Sonic Visualizer. Audio Engineering Society, 2014.

J. Gauffin and J. Sundberg. Spectral correlates of glottal voice source waveform characteristics. *Journal of Speech and Hearing Research*, 32(3):556–565, 1989.

A. Gaume, A. Vialatte, A. Mora-Sánchez, C. Ramdani, and F. B. Vialatte. A psychoengineering paradigm for the neurocognitive mechanisms of biofeedback and neurofeedback. *Neuroscience and Biobehavioral Reviews*, 68:891–910, 2016.

H. Gaunt. One-to-one tuition in a conservatoire: the perceptions of instrumental and vocal students. *Psychology of Music*, 38(2):178–208, 2010. Publisher: SAGE Publications Ltd.

J. M. Geringer, R. B. MacLeod, and M. L. Allen. Perceived Pitch of Violin and Cello Vibrato Tones Among Music Majors. *Journal of Research in Music Education*, 57(4):351–363, 2010. Publisher: SAGE Publications Inc.

J. M. Geringer, R. B. MacLeod, C. K. Madsen, and J. Napoles. Perception of melodic intonation in performances with and without vibrato. *Psychology of Music*, 43(5):675–685, 2015. Publisher: SAGE Publications Ltd.

C. H. Girvin. *How to Build a "Folk" Song: Socialist Bulgarian Song Texts and Folkloric Language in the South Slavic Context*. PhD thesis, UC Berkeley, 2016.

J. D. Glasner and A. M. Johnson. Effects of Historical Recording Technology on Vibrato in Modern-Day Opera Singers. *Journal of Voice*, 36(4):464–478, 2020. Publisher: Elsevier.

J. D. Glasner and J. Nix. Perception of vibrato rate by professional singing voice teachers. *The Journal of the Acoustical Society of America*, 152(4):A54–A54, 2022. Publisher: Acoustical Society of America.

J. D. Glasner and J. Nix. Perception of Vibrato by Professional Voice Teachers: A Pilot Study. *Journal of Voice*, 64(3):S0892–1997(23)00352–1, 2023.

H. Gockel, R. P. Carlyon, and C. J. Plack. Across-frequency interference effects in fundamental frequency discrimination: Questioning evidence for two pitch mechanisms. *The Journal of the Acoustical Society of America*, 116(2):1092–1104, 2004.

J. Goffi-Fynn. Applying Student-Centered Strategies in the Applied Teaching Music Studio. *Visions of Research in Music Education*, 45(9):99–113, 2024.

R. Goldenberg, T. Nestorova, H. Bastidas, S. Mahalingam, and Z. Peddle. Respiratory Efficiency through Vocalization. In I. Titze and E. Johnson, editors, *Sing and Shout for Health: Scientific Insights and Future Directions for Vocology*. National Center for Voice and Speech, Clearfield, UT, 2024.

K. E. Gordon and O. Reed. The Use of Inhalation Phonation in the Voice Studio: A New Approach to Addressing MTD in Singers. *Journal of Singing*, 75(3):271–278, 2019.

S. Granqvist. The visual sort and rate method for perceptual evaluation in listening tests. *Logopedics Phoniatrics Vocology*, 28(3):109–116, 2003.

G. W. Gray. *An experimental study of the vibrato in speech.* Doctor of Philosophy, University of Iowa, Iowa City, IA, 1926. Publisher: NCA Website ;https://doi.org/10.1080/00335632609379644.

J. Green and N. Thorogood, editors. *Qualitative Methods for Health Research.* SAGE Publications Ltd, Los Angeles, second edition edition, 2009.

M. D. Groll, S. Hablani, and C. E. Stepp. The Relationship Between Voice Onset Time and Increase in Vocal Effort and Fundamental Frequency. *Journal of Speech, Language, and Hearing Research : JSLHR*, 64(4):1197–1209, 2021.

M. Guzman, G. Acuña, F. Pacheco, F. Peralta, C. Romero, C. Vergara, and C. Quezada. The Impact of Double Source of Vibration Semioccluded Voice Exercises on Objective and Subjective Outcomes in Subjects with Voice Complaints. *Journal of Voice*, 32(6):770.e1–770.e9, 2018.

M. A. Guzman, J. Dowdall, A. D. Rubin, A. Maki, S. Levin, R. Mayerhoff, and M. C. Jackson-Menaldi. Influence of Emotional Expression, Loudness, and Gender on the Acoustic Parameters of Vibrato in Classical Singers. *Journal of Voice*, 26(5):675.e5–675.e11, 2012. Publisher: Elsevier.

J. Gärtner. The vibrato, with particular consideration given to the situation of the flutist: historical development, new physiological discoveries, and presentation of an integrated method of instruction. 1981.

H. Damsté, A. Reinders, and S. Tempelaars. Why should voices quiver? *Vox humana, studies presented to Aatto Sonninnen*, Institute of Finnish Language and Communication, University Jyvaskyla, 1982.

J. Hakes, T. Shipp, and E. T. Doherty. Acoustic properties of straight tone, vibrato, trill, and trillo. *Journal of Voice*, 1(2):148–156, 1987. Publisher: Elsevier.

J. Hakes, T. Shipp, and E. T. Doherty. Acoustic characteristics of vocal oscillations: Vibrato, exaggerated vibrato, trill, and trillo. *Journal of Voice*, 1(4):326–331, 1988.

J. Hakes, E. T. Doherty, and T. Shipp. Trillo rates exhibited by professional early music singers. *Journal of Voice*, 4(4):305–308, 1990. Publisher: Elsevier.

Q. Han and R. Zhang. Acoustic Analyses of the Singing Vibrato in Traditional Peking Opera. *Journal of Voice*, 31(4):511.e1–511.e9, 2017. Publisher: Elsevier.

N. A. Hartley and S. L. Thibeault. Systemic Hydration: Relating Science to Clinical Practice in Vocal Health. *Journal of voice*, 28(5):652.e1–652.e20, 2014.

J. B. Hausknecht, K. M. Murdaugh, E. Nagl, and C. T. Herbst. Global Inventory and Similarity Rating of Singing Voice Assessment Terms Used at English Speaking Academic Institutions. *Journal of Voice*, 0(0), 2021. Publisher: Elsevier.

K. Heidemann. A System for Describing Vocal Timbre in Popular Song. *Music Theory Online*, 22 (1), 2016.

N. Henrich, M. Kiek, J. Smith, and J. Wolfe. Resonance strategies used in Bulgarian women's singing style: A pilot study. *Logopedics Phoniatrics Vocology*, 32(4):171–177, 2007. Publisher: Taylor & Francis ;https://doi.org/10.1080/14015430600891504.

M. Henry. *A Perceptual Study of Amplitude Modulation Vibrato.* Master of Arts in Music Technology, McGill University, Montréal, Québec, Canada, 2021. Publisher: McGill University.

C. T. Herbst. Registers—The Snake Pit of Voice Pedagogy. *Journal of Singing*, 77(2):175–190, 2020.

C. T. Herbst, S. Hertegard, D. Zangger-Borch, and P.-A. Lindestad. Freddie Mercury—acoustic analysis of speaking fundamental frequency, vibrato, and subharmonics. *Logopedics Phoniatrics Vocology*, 42(1):29–38, 2017. Publisher: Taylor & Francis ;https://doi.org/10.3109/14015439.2016.1156737.

D. J. Hermes. Measurement of pitch by subharmonic summation. *The Journal of the Acoustical Society of America*, 83(1):257–264, 1988.

P. Herrera and J. Bonada. Vibrato Extraction and Parameterization in the Spectral Modeling Synthesis Framework. 1999.

K. Hijleh and C. Pinto. Realizing the Benefits of SOVTEs: A Reflection on the Research. *Journal of Singing*, 77(3):333–344, 2021.

J. Hillenbrand. Acoustic Analysis of Voice: A Tutorial. *Perspectives on Speech Science and Orofacial Disorders*, 21:31, 2011.

R. E. Hillman, C. E. Stepp, J. H. Van Stan, M. Zañartu, and D. D. Mehta. An Updated Theoretical Framework for Vocal Hyperfunction. *American Journal of Speech-Language Pathology*, 29(4): 2254–2260, 2020. Publisher: American Speech-Language-Hearing Association.

M. Hoch, L. Lister, and N. Cabell. *Voice Secrets: 100 Performance Strategies for the Advanced Singer*. Rowman & Littlefield Publishers, Incorporated, Blue Ridge Summit, UNITED STATES, 2016.

M. Hoch, K. Bozeman, K. Leigh-Post, M. Kennedy-Dygas, D. M. Howell, K. Harris, R. Goldenberg, and M. Randall. Science-Informed Terminology and Definitions for Voice Pedagogy. In *NATS July 2022 Pedagogy Workshop Working Group Three*, pages 1–53, Chicago, IL, 2022.

B. Hoffman-Ruddy, J. Lehman, C. Crandell, D. Ingram, and C. Sapienza. Laryngostroboscopic, Acoustic, and Environmental Characteristics of High-Risk Vocal Performers. *Journal of Voice*, 15(4):543–552, 2001. Publisher: Elsevier.

S. O. Holmes. *The Effects of Emotion on Acoustic Characteristics of Vocal Vibrato in Trained Singers*. Master of Science, Brigham Young Unversity, Provo, Utah, 2013.

Y. Horii. Acoustic analysis of vocal vibrato: A theoretical interpretation of data. *Journal of Voice*, 3(1):36–43, 1989. Publisher: Elsevier.

D. M. Howard, H. Daffern, and J. Brereton. Quantitative voice quality analyses of a soprano singing early music in three different performance styles. *Biomedical Signal Processing and Control*, 7(1):58–64, 2012.

I. Howell. *Hearing Singing: A Guide to Functional Listening and Voice Perception*. Rowman & Littlefield, 2025.

I. Howell and T. Nestorova. EML Vibrato Tools, 2025. Accessed via https://www.embodiedmusiclab.com/eml-tools-sign-up.

P. Howes, J. Callaghan, P. Davis, D. Kenny, and W. Thorpe. The relationship between measured vibrato characteristics and perception in Western operatic singing. *Journal of Voice*, 18(2):216–230, 2004. Publisher: Elsevier.

T.-Y. Hsiao, N. P. Solomon, E. S. Luschei, and I. R. Titze. Modulation of fundamental frequency by laryngeal muscles during vibrato. *Journal of Voice*, 8(3):224–229, 1994. Publisher: Elsevier.

E. J. Hunter, C. L. C. Cantor, L. E. van, M. M. van, C. D. Nanjundeswaran, P. Bottalico, M. J. Sandage, and S. Whitling. Toward a Consensus Description of Vocal Effort, Vocal Load, Vocal Loading, and Vocal Fatigue. *Journal of Speech, Language, and Hearing Research*, 63(2):509–532, 2020a. Publisher: American Speech-Language-Hearing Association.

E. J. Hunter, L. C. Cantor-Cutiva, E. van Leer, M. van Mersbergen, C. D. Nanjundeswaran, P. Bottalico, M. J. Sandage, and S. Whitling. Toward a consensus description of vocal effort, vocal load, vocal loading, and vocal fatigue. *Journal of Speech, Language, and Hearing Research*, 63 (2):509–532, 2020b.

N. Hérold. Timbre et analyse musicale : les possibilités d'intégration du timbre dans l'analyse formelle des øeuvres pour piano du dix-neuvième siècle. In M.-N. Masson, editor, *L'interprétation musicale*, Pensée Musicale, pages 79–103. Delatour France, 2012.

N. Isherwood. Vocal vibrato: new directions. *Journal of Singing*, 65(3):271–284, 2009. Publisher: National Association of Teachers of Singing.

T. Ivanova, V. Ganeva-Raicheva, Y. Bosseva, and D. Dimitrova. Singing the nature - ethnobotanical knowledge in Bulgarian folk songs. *Botanical Sciences*, 99(2):321–334, 2021.

S. Iwata and J. Large. Aerodynamic Study of Vibrato and Voluntary "Straight Tone" Pairs in Singing. *The Journal of the Acoustical Society of America*, 49(1A_Supplement):137, 2005.

A. F. Jahn. *The Singer's Guide to Complete Health*. OUP USA, 2013.

H. J. Jensen and M. R. Chung. *CelloMind: Intonation and Technique*. Ovation Press, Ltd., 2017.

Johann Andreas Herbst. *Arte prattica & poëtica*. Frankfurt, 1653.

A. M. Johnson and M. J. Sandage. Exercise Science and the Vocalist. *Journal of Voice*, 35(4): 668–677, 2021.

R. Jonck, T. Nestorova, and A. Del Simone. Legato Then and Now, Vibrato Edition: A close-up of what happens between the pitches in the classical bel canto tradition, 2024.

G. Jones. Gatekeepers, Midwives, and Fellow Travellers: The Craft and Artistry of the Adult Educator. 2005.

J. A. Jones and D. Keough. Auditory-motor Mapping for Pitch Control in Singers and Nonsingers. *Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale*, 190(3): 279–287, 2008.

H. Järveläinen. Perception-based Control of Vibrato Parameters in String Instrument Synthesis. In *Proceedings of the International Computer Music Conference*, pages 287–294, Gothenburg, Sweden, 2002.

S. Kaludova-Stanilova. Bulgarian Folklore Singing– Modernity and Trends. In *Collection of reports*, pages 14–21, Plovdiv, 2022. Academy of Music, Dance and Visual Arts "Prof. Asen Diamandiev".

D. Katok. Healthy Minimization of Vibrato: An Exploration of "Straight Tone". *Choral Journal*, 2021.

D. Kehaiov. *Мъжко народно пеене от България / Male folk singing from Bulgaria*. FastPrintBooks, 2018.

P. Kerkelov. Ornamentation in Bulgarian vocal folklore music according to various researchers. Classification perspectives. *Bulgarian Musicology*, (1):103–116, 2017. Publisher: Институт за изследване на изкуствата, Българска академия на науките.

J. W. Kim, J. Salamon, P. Li, and J. P. Bello. CREPE: A Convolutional Representation for Pitch Estimation, 2018. arXiv:1802.06182 [cs, eess, stat].

J. B. King and Y. Horii. Vocal matching of frequency modulation in synthesized vowels. *Journal of Voice*, 7(2):151–159, 1993.

D. K. S. Kirilov. *Bulgarian Harmony: In Village, Wedding, and Choral Music of the Last Century.* Ashgate Publishing, Ltd., 2015.

B. Kirkov and T. P. Zieliński. Formant Analysis of Traditional Bulgarian Singing from Rhodope Region. In *2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pages 148–152, 2019. ISSN: 2326-0319.

A. Kirkpatrick. Teaching methods for correcting problematic vibratos: using sustained dynamic exercises to discover and foster healthy vibrato. *Journal of Singing*, 64(5):551–557, 2008. Publisher: National Association of Teachers of Singing.

D. H. Klatt. Voice Onset Time, Frication, and Aspiration in Word-Initial Consonant Clusters. *Journal of Speech and Hearing Research*, 18(4):686–706, 1975. Publisher: American Speech-Language-Hearing Association.

S. Z. Knezevic, J. C. Streibig, and C. Ritz. Utilizing R Software Package for Dose-Response Studies: The Concept and Data Analysis. *Weed Technology*, 21(3):840–848, 2007.

I. Kochev. *Bulgarian Dialect Atlas.* Bulgarian Academy of Sciences Press, Sofia, 2016.

J. Kollbrunner and E. Seifert. Encouragement to Increase the Use of Psychosocial Skills in the Diagnosis and Therapy of Patients With Functional Dysphonia. *Journal of Voice*, 31(1):132.e1–132.e7, 2017. Publisher: Elsevier.

B. I. Koteterova-Dobreva. *Model for teaching song folklore to children from 4 to 14 years old.* Ph.D., Southwest University ”Neofit Rilski”, Blagoevgrad, 2012.

R. Kozakand. Folk Costumes & Embroidery: Costume of Sofia area, Shope region, Bulgaria, 2012.

J. Kreiman, B. Gabelman, and B. R. Gerratt. Perception of vocal tremor. *Journal of speech, language, and hearing research: JSLHR*, 46(1):203–214, 2003.

B. Kremenliev. Some Social Aspects of Bulgarian Folksongs. *The Journal of American Folklore*, 69 (273):310–319, 1956a. Publisher: University of Illinois Press.

B. A. Kremenliev. Types of Bulgarian Folk Songs. *The Slavonic and East European Review*, 34(83): 355–376, 1956b. Publisher: Modern Humanities Research Association.

C. L. Krumhansl. Music Psychology: Tonal Structures in Perception and Memory. *Annual Review of Psychology*, 42(1):277–303, 1991. ;https://doi.org/10.1146/annurev.ps.42.020191.001425.

F. Kuhlewind. *Vibrato and Open Quotient in the Singing Voices of Female Jazz and Classical Singers.* Master of Arts in General Linguistics, University of Amsterdam, Amsterdam, 2014.

K. Kukushev. *Rhythm and Melody. Towards the Theory and History of Rhythmic and melodic phenomenon.* Varna, 2008.

A. M. Kushleva. *Production problems in the development of the folk singer's voice from childhood to high professionalism: Tradition and contemporary trends.* University Press Paisii Hilendarski, Plovdiv, 2000.

M. Kwok and G. D. Eslick. The Impact of Vocal and Laryngeal Pathologies Among Professional Singers: A Meta-analysis. *Journal of Voice*, 33(1):58–65, 2019. Publisher: Elsevier.

A.-M. Laukkanen, E. Vilkman, and U. K. Laine. Aspect of the physiological sources of vocal vibrato: A study of fundamental-period-synchronous changes in electroglottographic signals obtained from one singer and two excised human larynges. *Scandinavian Journal of Logopedics and Phoniatrics*, 17(2):87–93, 1992. Publisher: Taylor & Francis.

M. Lausevic. *Balkan Fascination: Creating an Alternative Music Culture in America.* American Musicspheres. Oxford University Press, Oxford, New York, new to this edition:, new to this edition: edition, 2015.

W. D. LeBorgne. *Definint the Belt Voice: Perceptual Judgements and Objective Measures.* PhD thesis, University of Cincinnati, 2001.

W. D. LeBorgne and M. Rosenberg. *The Vocal Athlete.* Plural Publishing, Incorporated, third edition edition, 2026.

A. Lederle, J. Barkmeier-Kraemer, and E. Finnegan. Perception of Vocal Tremor During Sustained Phonation Compared With Sentence Context. *Journal of Voice*, 26(5):668.e1–668.e9, 2012. Publisher: Elsevier.

S. W. Lee and M. Dong. Singing Voice Synthesis: Singer-Dependent Vibrato Modeling and Coherent Processing of Spectral Envelope. pages 2001–2004, 2011.

S. W. Lee, M. Dong, and P. Y. Chan. Analysis for Vibrato with Arbitrary Shape and its Applications to Music. In *APSIPA ASC*, 2011.

E. Lepock. *No Space to Sing: A Narrative Inquiry into the Experiences of Classical Singers with Primary Muscle Tension Dysphonia.* Doctor of Musical Arts, The University of Western Ontario, 2019.

R. A. Lester. *Modulation of Voice Related to Tremor and Vibrato.* Ph.D., University of Arizona, 2014.

R. A. Lester, B. H. Story, and A. J. Lotto. Acoustical bases for the perception of simulated laryngeal vocal tremor. *The Journal of the Acoustical Society of America*, 136(4_Supplement):2293, 2014.

R. A. Lester-Smith, J. H. Kim, A. Hilger, C.-L. Chan, and C. R. Larson. Auditory-Motor Control of Fundamental Frequency in Vocal Vibrato. *Journal of Voice*, 39(1), 2021. Publisher: Elsevier.

R. A. Lester-Smith, A. Hilger, K. E. Dunne-Platero, J. H. Kim, C. L. Chan, and C. R. Larson. The Effects of Masked and Delayed Auditory Feedback on Fundamental Frequency Modulation in Vocal Vibrato. *Journal of Voice*, 43(4), 2022. Publisher: Elsevier.

C. Levy. Folk in Opposition? Wedding Bands and the New Developments in Bulgarian Popular Music. *Music and Politics*, III(1), 2009.

C. Leydon, J. J. Bauer, and C. R. Larson. The role of auditory feedback in sustaining vocal vibrato. *The Journal of the Acoustical Society of America*, 114(3):1575–1581, 2003. Publisher: Acoustical Society of America.

M.-Y. Liao and J. W. Davidson. The effects of gesture and movement training on the intonation of children's singing in vocal warm-up sessions. *International Journal of Music Education*, 34 (1):4–18, 2016. Publisher: SAGE Publications Ltd.

L. Lisker and A. S. Abramson. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD*, 20(3):384–422, 1964.

T. List and S. Axelsson. Management of TMD: evidence from systematic reviews and meta-analyses. *Journal of Oral Rehabilitation*, 37(6):430–451, 2010. ;https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2842.2010.02089.x.

D. Y. Loni and S. Subbaraman. Timbre-Vibrato Model for Singer Identification. In S. C. Satapathy and A. Joshi, editors, *Information and Communication Technology for Intelligent Systems*, Smart Innovation, Systems and Technologies, pages 279–292, Singapore, 2019. Springer.

J. L. LoVetri and E. M. Weekly. Contemporary Commercial Music (CCM) Survey: Who's Teaching What in Nonclassical Music. *Journal of Voice*, 17(2):207–215, 2003. Publisher: Elsevier.

Lumivero. NVivo, 2025. Accessed via https://lumivero.com/products/nvivo/.

S. Mabry. *Exploring Twentieth-Century Vocal Music: A Practical Guide to Innovations in Performance and Repertoire*. Oxford University Press, Incorporated, Oxford, UNITED STATES, 2002.

I. MacMillan. Fascination, Musical Tourism, and the Loss of the Balkan Village (Notes on Bulgaria's Koprivshtitsa Festival). *Journal of the Society for Ethnomusicology*, 59(2):227–261, 2015. Publisher: [University of Illinois Press, Society for Ethnomusicology].

C. Madill, A. McIlwaine, R. Russell, N. J. Hodges, and P. McCabe. Classifying and Identifying Motor Learning Behaviors in Voice-Therapy Clinician-Client Interactions: A Proposed Motor Learning Classification Framework. *Journal of Voice*, 34(5):806.e19–806.e31, 2020.

D. E. Mager, E. Wyska, and W. J. Jusko. Diversity of mechanism-based pharmacodynamic models. *Drug Metabolism and Disposition: The Biological Fate of Chemicals*, 31(5):510–518, 2003.

K. Makiyama and S. Hirano. *Aging Voice*. Springer Nature Singapore, 2017.

S. Makov, A. Minaev, I. Grinev, D. Chernyshov, B. Kudruavcev, and V. Mladenovic. A spectral-based pitch detection method. volume 2188, page 050005, 2019.

J. Malafronte. Vibrato Wars. *The Magazine of Early Music America (EMAg)*, 21(2), 2015.

C. Manfredi, D. Barbagallo, G. Baracca, S. Orlandi, A. Bandini, and P. H. Dejonckere. Automatic Assessment of Acoustic Parameters of the Singing Voice: Application to Professional Western Operatic and Jazz Singers. *Journal of Voice*, 29(4):517.e1–517.e9, 2015. Publisher: Elsevier.

F. Marandola. *Les polyphonies vocales des pygmées Bedzan du Cameroun: une approche expérimentale du système scalaire*. Ph.D., Université Paris-Sorbonne, 1970-2017, France, 2004a. Type: Reproduction de University: Université Paris-Sorbonne.

F. Marandola. The Study of Musical Scales in Central Africa: The Use of Interactive Experimental Methods. In *Computer Music Modeling and Retrieval*, pages 34–41. Springer, Berlin, Heidelberg, 2004b. ISSN: 1611-3349.

M. Marchesi. *L'art du chant; méthode pratique en quatre parties (Op. 21)*. Legare Street Press, 1877.

R. H. G. Martins, E. L. M. Tavares, P. F. Ranalli, A. Branco, and A. B. B. Pessin. Psychogenic dysphonia: diversity of clinical and vocal manifestations in a case series. *Brazilian Journal of Otorhinolaryngology*, 80(6):497, 2014.

R. M. Mason and W. R. Zemlin. Physiological Components of Vocal Vibrato. *The Journal of the Acoustical Society of America*, 49(1A):136–136, 1971. Publisher: Acoustical Society of America.

M. Mauch and S. Dixon. PYIN: A fundamental frequency estimator using probabilistic threshold distributions. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 659–663, Florence, Italy, 2014. IEEE.

M. Mauch, C. Cannam, R. Bittner, G. Fazekas, J. Salamon, J. Dai, J. Bello, and S. Dixon. Computer-aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency. 2015.

L. M. Maxfield. *Application of principles from motor-learning theory to the studio voice lesson: effects of feedback frequency on retention of classical singing technique - University of Iowa*. Doctor of Philosophy, University of Iowa, 2011.

S. McAdams. Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, 4(1):181–198, 1989. Publisher: Routledge ;https://doi.org/10.1080/07494468900640281.

S. McAdams and X. Rodet. The Role of FM-induced AM in Dynamic Spectral Profile Analysis. In A. Press, editor, *Basic Issues in Hearing*, pages 359–369. Academic Press, 1988.

S. McCoy, M. L. Halstead, and L. Helding. *Your Voice: An Inside View.* Inside View Press, 3rd edition edition, 2019.

J. C. McKinney. *The Diagnosis and Correction of Vocal Faults: A Manual for Teachers of Singing and for Choir Directors.* Waveland Press, 2005.

V. A. Meleki. *Genre Specifics and performance issues of the actor's vocal training.* Ph.D., National Academy for Theatre and Film Arts "Krastyo Sarafov", Sofia, 2021.

A. P. Mendes, H. B. Rothman, C. Sapienza, and W. S. Brown. Effects of vocal training on the acoustic parameters of the singing voice. *Journal of Voice*, 17(4):529–543, 2003. Publisher: Elsevier.

A. P. Mendes, A. F. Rodrigues, and D. M. Guerreiro. Acoustic and Phonatory Characterization of the Fado Voice. *Journal of Voice*, 27(5):655.e9–655.e15, 2013. Publisher: Elsevier.

M. Mersenne. *Harmonie universelle.* Paris, 1636.

G. F. Messner. *Die Schwebungsdiaphonie in Bistrica: Untersuchungen der mehrstimmigen Liedformen eines mittelwestbulgarischen Dorfes.* Schneider, 1980.

P. Miksza and K. Elpus. *Design and Analysis for Quantitative Research in Music Education.* Oxford University Press, 2018.

R. Miller. *The Structure of Singing: System and Art of Vocal Technique.* Schirmer, Boston, 1st edition edition, 1986.

R. Miller. *National Schools of Singing: English, French, German, and Italian Techniques of Singing Revisited.* Scarecrow Press, 2nd edition edition, 1997.

R. Miller. *Solutions for Singers: Tools for Performers and Teachers.* Oxford University Press, USA, 2004.

A. K. Miri, F. Barthelat, and L. Mongeau. Effects of Dehydration on the Viscoelastic Properties of Vocal Folds in Large Deformations. *Journal of Voice*, 26(6):688–697, 2012.

H. F. Mitchell and D. T. Kenny. The impact of 'open throat' technique on vibrato rate, extent and onset in classical singing. *Logopedics Phoniatrics Vocology*, 29(4):171–182, 2004. Publisher: Taylor & Francis ;https://doi.org/10.1080/14015430410001033.

H. F. Mitchell and D. T. Kenny. Change in Vibrato Rate and Extent During Tertiary Training in Classical Singing Students. *Journal of Voice*, 24(4):427–434, 2010. Publisher: Elsevier.

E. A. Mohammed, C. Naugler, and B. H. Far. Chapter 32 - Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. In Q. N. Tran and H. Arabnia, editors, *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology*, Emerging Trends in Computer Science and Applied Computing, pages 577–602. Morgan Kaufmann, Boston, 2015.

V. S. Moneva. *Contemporary Interpretations of songs for women folklore choir.* Ph.D., Academy of Music, Dance and Visual Arts “Prof. Asen Diamandiev”, Plovdiv, 2022.

L. Moorcroft and D. T. Kenny. Vocal Warm-Up Produces Acoustic Change in Singers’ Vibrato Rate. *Journal of Voice*, 26(5):667.e13–667.e18, 2012. Publisher: Elsevier.

L. Moorcroft, D. T. Kenny, and J. Oates. Vibrato changes following imagery. *Journal of Voice*, 29 (2):182–190, 2015.

B. Moore. *An Introduction to the Psychology of Hearing: Sixth Edition.* BRILL, Leiden, Boston, 2013.

B. C. J. Moore. Frequency difference limens for short-duration tones. *The Journal of the Acoustical Society of America*, 54(3):610–619, 1973.

B. C. J. Moore. Pitch perception and frequency discrimination in normally hearing and hearing-impaired people. In B. C. J. Moore, editor, *Perceptual Consequences of Cochlear Damage*, page 0. Oxford University Press, 1995.

B. C. J. Moore. The Role of Temporal Fine Structure Processing in Pitch Perception, Masking, and Speech Perception for Normal-Hearing and Hearing-Impaired People. *Journal of the Association for Research in Otolaryngology*, 9(4):399–406, 2008. Company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 4 Publisher: Springer-Verlag.

M. S. Morelli, S. Orlandi, and C. Manfredi. BioVoice: A multipurpose tool for voice analysis. *Biomedical Signal Processing and Control*, 64:102302, 2021.

J. M. Morse. Determining Sample Size. *Qualitative Health Research*, 10(1):3–5, 2000. Publisher: SAGE Publications Inc.

D. Myers and J. Michel. Vibrato and pitch transitions. *Journal of Voice*, 1(2):157–161, 1987.

M. Müller, T. Schulz, T. Ermakova, and P. P. Caffier. Lyric or Dramatic - Vibrato Analysis for Voice Type Classification in Professional Opera Singers. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29:943–955, 2021. Conference Name: IEEE/ACM Transactions on Audio, Speech, and Language Processing.

V. Müller and U. Lindenberger. Cardiac and Respiratory Patterns Synchronize between Persons during Choir Singing. *PLoS ONE*, 6(9):e24893, 2011.

D. Mürbe, T. Zahnert, E. Kuhlisch, and J. Sundberg. Effects of Professional Singing Education on Vocal Vibrato—A Longitudinal Study. *Journal of Voice*, 21(6):683–688, 2007. Publisher: Elsevier.

S. Nandamudi and R. C. Scherer. Airflow Vibrato: Dependence on Pitch and Loudness. *Journal of Voice*, 33(6):815–830, 2019. Publisher: Elsevier.

National Institute of Dental and Craniofacial Research. Prevalence of TMJD and its Signs and Symptoms | Data & Statistics | National Institute of Dental and Craniofacial Research, 2018.

T. Nawka, W. Seidner, and M. Cebulla. Experiments on intentionally influencing the singer's vibrato. In A. Friberg, J. Iwarsson, E. Jansson, and J. Sundberg, editors, *Proceedings of the Stockholm Music Acoustics Conference 1993*, volume 79, pages 221–227, Stockholm, Sweden, 1993. Royal Swedish Academy of Music.

T. Nestorova. *Does Vibrato Define Genre or Vice Versa?: A Novel Parametric Approach to Complex Vibrato Patterns.* Master of Music in Vocal Pedagogy and Music-in-Education, New England Conservatory of Music, Boston, 2021.

T. Nestorova. Vibrato, 2022. Accessed via https://timbreandorchestration.org/writings/timbre-lingo/2022/6/27/vibrato.

T. Nestorova and J. D. Glasner. What About Extent?: Examining Current Vibrato Metrics. *Pan American Vocology Association InFormant Newsletter*, 2(3), 2021. Accessed via https://cdb7d7c9-7da9-491b-ba11-e8058101017a.filesusr.com/ugd/0af83e_41b1eec5ed0848f2a5d80e604f5b25cd.pdf?index=true.

T. Nestorova, M. Brandner, B. Gingras, and C. T. Herbst. Vocal Vibrato Characteristics in Historical and Contemporary Opera, Operetta, and Schlager. *Journal of Voice*, 58(2), 2023. Publisher: Elsevier.

T. Nestorova, M. DeFrancesco, H. Bastidas, and K. James. Vocalization Influences Fluid Transport with Reduced Viscosity. In I. Titze and E. Johnson, editors, *Sing and Shout for Health: Scientific Insights and Future Directions for Vocology*. National Center for Voice and Speech, Clearfield, UT, 2024a.

T. Nestorova, I. Nestorov, J. B. Gilbert, and I. Howell. Does Vibrato Define Genre Or Vice Versa?: A Novel Approach To Vocal Vibrato Analysis. *Journal of Voice*, In press, 2024b. https://doi.org/10.1016/j.jvoice.2025.02.046.

T. Nestorova, Y. Huang, and J. Dahman. Singing in Cyrillic: A Guide to the Accessibility of Bulgarian Lyric Diction. *Journal of Singing*, Manuscript under review, 2025a.

T. Nestorova, Y. Maryn, A. Stark, S. McDowell, J. Pierce, L. Maxfield, and J. Barkmeier-Kraemer. Vocal Vibrato and Tremor: Multi-Signal Pilot Study on Biomechanical and Neural Mechanisms of Automaticity and Primary Motor Control. Manuscript in preparation for Journal of Speech, Language, and Hearing Research, 2025b.

T. Nestorova, L. Reymore, J. Marchand Knight, and K. Soden. Speaking of Vocal Vibrato: Timbre Perceptions and Semantic Descriptions with Pedagogical and Rehabilitative Implications. *Voice and Speech Review*, Manuscript accepted, 2025c. DOI forthcoming.

T. Nestorova, J. Xie, C. Zhang, F. Nyanyo, Y. F. Zhao, and L. Mongeau. Vibrato Machine Learning Models as Diagnostic Biomarkers: AI for Vocal Health. *NCVS Insights*, 3(4), 2025d. DOI forthcoming.

F. Neumann. The Vibrato Controversy. *Performance Practice Review*, 4(1):14–27, 1991.

J. Nix. New Voices in Research: Vibrato and Non-Vibrato Singing: Who teaches it? How do they teach it? Does it make a difference? *The Choral Journal*, 53(9):57–66, 2013. Publisher: JSTOR.

J. Nix. Shaken, Not Stirred: Practical Ideas for Addressing Vibrato and Nonvibrato Singing in the Studio and the Choral Rehearsal. *Journal of Singing*, 7(4):8, 2014.

J. Nix. On Shaky Ground Part II: Navigating Challenges in Vocal Vibrato. *The Voice Foundation's The Voice Newsletter*, 28(1), 2023.

J. Nix, N. Perna, K. James, and S. Allen. Vibrato Rate and Extent in College Music Majors: A Multicenter Study. *Journal of Voice*, 30(6):756.e31–756.e41, 2016. Publisher: Elsevier.

J. Nix, R. A. Lester-Smith, N. Cummings, E. Eidson, and E. Derrick. Singing Teacher Perception of Breathiness in Vibrato and Non-Vibrato Tones. *Journal of Singing*, 81(3):315–325, 2024a.

J. Nix, T. Nestorova, J. D. Glasner, and Y. G. Redman. A Balance in the Force: Perceiving and Assessing Vibrato, 2024b. Accessed via https://nats.sclivelearningcenter.com/MVSite/MVStore.aspx?ItemID=75638.

J. P. Nix. Listener preferences for vibrato rate and extent in synthesized vocal samples. page 035002, Pittsburgh, Pennsylvania, 2015.

R. North. *Roger North's The Musicall Grammarian 1728*. Cambridge University Press, 1728.

J. Oates and A. Winkworth. Current knowledge, controversies and future directions in hyperfunctional voice disorders. *International Journal of Speech-Language Pathology*, 10(4):267–277, 2008. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17549500802140153.

R. Ohrbach, R. B. Fillingim, F. Mulkey, Y. Gonzalez, S. Gordon, H. Gremillion, P.-F. Lim, M. Ribeiro-Dasilva, J. D. Greenspan, C. Knott, W. Maixner, and G. Slade. Clinical findings and pain symptoms as potential risk factors for chronic TMD: descriptive data and empirically identified domains from the OPPERA case-control study. *The Journal of Pain*, 12(11 Suppl):T27–45, 2011.

M. Olson. Vibrato vs. Nonvibrato: the solo singer in the collegiate choral ensemble. *Journal of Singing*, 64(5):561–565, 2008. Publisher: National Association of Teachers of Singing.

P. B. Oncley. Frequency, Amplitude, and Waveform Modulation in the Vocal Vibrato. *The Journal of the Acoustical Society of America*, 49(1A):136–136, 1971. Publisher: Acoustical Society of America.

H.-S. Pang, J.-s. Lim, and S. Lee. Discrete Fourier transform-based method for analysis of a vibrato tone. *Journal of New Music Research*, 49(4):307–319, 2020. Publisher: Routledge ;https://doi.org/10.1080/09298215.2020.1784959.

P. M. Patinka and J. Nix. Changes in Vibrato Rate, Vibrato Extent, and Vibrato Jitter in Soprano Voices in Response to Changes in Mouth Opening: A Pilot Study. *Journal of Singing*, 77(2): 213–219, 2020. Publisher: National Association of Teachers of Singing.

C. L. Payten, G. Chiapello, K. A. Weir, and C. J. Madill. Frameworks, Terminology and Definitions Used for the Classification of Voice Disorders: A Scoping Review. *Journal of Voice*, 39(2), 2022. Publisher: Elsevier.

G. Pecoraro, D. F. Curcio, and M. Behlau. Vibrato rate variability in three professional singing styles: Opera, Rock and Brazilian country. In *Proceedings of Meetings on Acoustics*, volume 19, pages 1–6, Montréal, Québec, Canada, 2013. Acoustical Society of America.

P. M. Pestana, S. Vaz-Freitas, and M. C. Manso. Prevalence of Voice Disorders in Singers: Systematic Review and Meta-Analysis. *Journal of Voice*, 31(6):722–727, 2017.

G. Petrova-Kirkova. The Repertoire In Folk Singing Training As An Opportunity For Creating Creativity In Students' Minds – The Future Teachers. *Knowledge: International Journal*, 30(2): 453–458, 2019.

G. Petrova-Kirkova. Opportunities for optimization of the teaching process in the discipline of Folklore singing through the application of digital resources and digital transformation of the musical content of the study program. In *Collection of abstracts*, page 52, 2023.

D. J. Phyland, J. F. Pallant, S. L. Thibeault, M. S. Benninger, N. Vallance, and J. A. Smith. Measuring vocal function in professional music theater singers: construct validation of the Evaluation of the Ability to Sing Easily (EASE). *Folia phoniatrica et logopaedica: official organ of the International Association of Logopedics and Phoniatrics (IALP)*, 66(3):100–108, 2014.

J. Playford. *An Introduction to the Skill of Musick*. London, London, 1654.

W. Pleeth. *Cello*. Kahn & Averill, 1992.

T. Porcello. Speaking of Sound: Language and the Professionalization of Sound-Recording Engineers. *Social Studies of Science*, 34(5):733–758, 2004. Publisher: SAGE Publications Ltd.

E. Prame. Measurements of the vibrato rate of ten singers. *The Journal of the Acoustical Society of America*, 96(4):1979–1984, 1994. Publisher: Acoustical Society of America.

E. Prame. Vibrato extent and intonation in professional Western lyric singing. *The Journal of the Acoustical Society of America*, 102(1):616–621, 1997. Publisher: Acoustical Society of America.

R Core Team. R: The R Project for Statistical Computing, 2021. Accessed via https://www.fon.hum.uva.nl/praat/.

N. Radhakrishnan, R. C. Scherer, and S. Bandyopadhyay. Laryngeal Dynamics of Pedagogical Taan Gestures in Indian Classical Singing. *Journal of Voice*, 25(3):e139–e147, 2011. Publisher: Elsevier.

K. Ragan. Understanding Voice Doctors: Whom to Call and When to Call Them. *Journal of Singing*, 74(1):57–62, 2017.

K. Ragan. Defining Evidence-Based Voice Pedagogy: A New Framework. *Journal of Singing*, 75 (2):157–160, 2018.

K. Ragan. *A Systematic Approach to Voice: The Art of Studio Application*. Plural Publishing, 2020.

K. Ragan. A Balance in the Force: Producing, Perceiving, Measuring and Assessing Vibrato – NATS Chat; March 19, 2023, 2023.

K. Ragan and M. Kapsner-Smith. Vehicular Vocalizing to Start Your Day. *Journal of Singing*, 76 (2):165–170, 2019.

D. A. Ramsdell. The Psychophysics of Frequency Modulation. *The Journal of the Acoustical Society of America*, 7(3_Supplement):236, 1936.

R. A. Rasch. Jitter in the Singing Voice. In *Proceedings of the Tenth International Congress of Phonetic Sciences*, pages 288–292. De Gruyter Mouton, 1984.

A. A. Reddy and U. Subramanian. Do Rate And Extent Of Vibrato Vary With Gender In Untrained Singers? *International Journal of Interdisciplinary Research and Innovations*, 2(4):5, 2014.

A. A. Reddy and U. Subramanian. Singers' and Nonsingers' Perception of Vocal Vibrato. *Journal of Voice*, 29(5):603–610, 2015. Publisher: Elsevier.

Y. G. Redman, J. D. Glasner, D. D'Orazio, and P. Bottalico. Singing in different performance spaces: The effect of room acoustics on singers' perception. *The Journal of the Acoustical Society of America*, 154(4):2256–2264, 2023.

C. L. Reid. *The Free Voice: A Guide to Natural Singing*. 2018 | 50th Anniversary Edition. Oxford University Press, 1965.

L. Reymore and D. Huron. Using auditory imagery tasks to map the cognitive linguistic dimensions of musical instrument timbre qualia. *Psychomusicology: Music, Mind, and Brain*, 30(3): 124–144, 2020. Place: US Publisher: Educational Publishing Foundation.

L. Reymore, J. Noble, C. Saitis, C. Traube, and Z. Wallmark. Timbre Semantic Associations Vary Both Between and Within Instruments: An Empirical Study Incorporating Register and Pitch Height. *Music Perception*, 40(3):253–274, 2023.

L. Reymore, T. Nestorova, J. Marchand Knight, and K. Soden. Timbre Semantics of the Singing Voice. *Music Perception*, Manuscript under review, 2025.

T. Rice. *May It Fill Your Soul: Experiencing Bulgarian Music*. University of Chicago Press, 1994.

C. Ritz, F. Baty, J. C. Streibig, and D. Gerhard. Dose-Response Analysis Using R. *PLOS ONE*, 10 (12):e0146021, 2015. Publisher: Public Library of Science.

M. P. Robb, Y. Chen, H. R. Gilbert, and J. W. Lerman. Acoustic Comparison of Vowel Articulation in Normal and Reverse Phonation. *Journal of Speech, Language, and Hearing Research*, 44(1): 118–127, 2001. Publisher: American Speech-Language-Hearing Association.

D. W. Robinson and R. S. Dadson. A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics*, 7:166–181, 1956. ADS Bibcode: 1956BJAP....7..166R.

D. C. Rosen, J. B. Sataloff, and R. T. Sataloff. *Psychology of Voice Disorders, Second Edition.* Plural Publishing, Inc., San Diego, 2021.

M. Rosenberg. On Shaky Ground Part II: Navigating Challenges in Vocal Vibrato. *The Voice Foundation's The Voice Newsletter*, 28(1), 2023.

M. Rothenberg, D. Miller, and R. Molitor. Aerodynamic Investigation of Sources of Vibrato. *Folia Phoniatrica et Logopaedica*, 40(5):244–260, 1988. Publisher: Karger Publishers.

H. B. Rothman, J. A. Diaz, and K. E. Vincent. Comparing historical and contemporary opera singers with historical and contemporary Jewish cantors. *Journal of Voice*, 14(2):205–214, 2000. Publisher: Elsevier.

N. Roy and D. M. Bless. Personality and voice disorders: A multitrait-multidisorder analysis. *Journal of Voice*, 14(4):521–548, 2000. Publisher: Elsevier.

RStudio Team. RStudio: Integrated Development Environment for R, 2023. Accessed via https://posit.co/.

J. S. Rubin and M. Greenberg. Psychogenic Voice Disorders in Performers: A Psychodynamic Model. *Journal of Voice*, 16(4):544–548, 2002. Publisher: Elsevier.

C. Saitis and S. Weinzierl. The Semantics of Timbre. In K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, and R. R. Fay, editors, *Timbre: Acoustics, Perception, and Cognition*, volume 69, pages 119–149. Springer International Publishing, Cham, 2019. Series Title: Springer Handbook of Auditory Research.

C. Saitis, B. L. Giordano, C. Fritz, and G. P. Scavone. Perceptual evaluation of violins: a quantitative analysis of preference judgments by experienced players. *The Journal of the Acoustical Society of America*, 132(6):4002–4012, 2012.

C. Saitis, C. Fritz, and G. P. Scavone. Categorization and lexicon in verbal descriptions of violin quality by performers. *The Journal of the Acoustical Society of America*, 135(4_Supplement): 2214, 2014.

S. Samad, A. Hussain, and L. Fah. Pitch detection of speech signals using the cross-correlation technique. volume 1, pages 283–286 vol.1, 2000.

C. Sapienza and B. Hoffman. *Voice Disorders, Fourth Edition.* Plural Publishing, 2020.

R. T. Sataloff. *Treatment of Voice Disorders, Second Edition.* Plural Publishing, 2017.

R. T. Sataloff, F. Chowdhury, and M. J. Hawkshaw. *Surgical Techniques in Otolaryngology-Head and Neck Surgery: Laryngeal Surgery: Head & Neck Surgery.* Jaypee Brothers Medical Publishers, New Delhi, 2014.

G. P. Scavone, A. Lefebvre, and A. R. da Silva. Measurement of vocal-tract influence during saxophone performance. *The Journal of the Acoustical Society of America*, 123(4):2391–2400, 2008.

K. Scherer. Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40(1-2):227–256, 2003.

H. J. Schultz-Coulon. The neuromuscular phonatory control system and vocal function. *Acta Oto-Laryngologica*, 86(1-2):142–153, 1978.

C. E. Seashore. *The Vibrato.* Iowa City, Ia. : The University, 1932.

C. E. Seashore. *Psychology of the Vibrato in Voice and Instrument.* The University Press, 1936.

J. L. Sebaugh and P. D. McCray. Defining the linear portion of a sigmoid-shaped curve: bend points. *Pharmaceutical Statistics*, 2(3):167–174, 2003. ;https://onlinelibrary.wiley.com/doi/pdf/10.1002/pst.62.

J. Shapiro. *So You Want to Sing Jazz: A Guide for Professionals.* Rowman & Littlefield Publishers, Incorporated, Blue Ridge Summit, UNITED STATES, 2015.

T. Shipp, R. Leanderson, and J. Sundberg. Vocal vibrato/rate and extent of vibrato as a function of vowel, effort and frequency. In *Transcripts of the eighth symposium on care of the professional voice*, pages 50–56, 1979.

T. Shipp, R. Leanderson, and J. Sundberg. Some Acoustic Characteristics of Vocal Vibrato. *Journal of Research in Singing*, 4:18–25, 1980.

T. Shipp, E. T. Doherty, and S. Haglund. Physiologic factors in vocal vibrato production. *Journal of Voice*, 4(4):300–304, 1990. Publisher: Elsevier.

J. I. Shonle and K. E. Horan. The pitch of vibrato tones. *The Journal of the Acoustical Society of America*, 67(1):246–252, 1980.

C. Silverman. *Romani Routes: Cultural Politics and Balkan Music in Diaspora.* Oxford University Press, 2011.

M. V. Stankova. *Globalization, Nationalism and Cultural Translation in Bulgarian Folk Singing.* Ph.D., New York University, United States – New York, 2019. ISBN: 9781392004395.

J. Stark. *Bel Canto: A History of Vocal Pedagogy.* University of Toronto Press, 1999.

K. Steinhauer and J. P. Grayhack. The role of knowledge of results in performance and learning of a voice motor task. *Journal of Voice*, 14(2):137–145, 2000. Publisher: Elsevier.

J. C. Stemple, L. Lee, B. D'Amico, and B. Pickup. Efficacy of vocal function exercises as a method of improving voice production. *Journal of Voice*, 8(3):271–278, 1994.

J. C. Stemple, N. Roy, and B. K. Klaben. *Clinical Voice Pathology: Theory and Management, Sixth Edition.* Plural Publishing, 2018.

C. E. Stepp, G. R. Merchant, J. T. Heaton, and R. E. Hillman. Effects of Voice Therapy on Relative Fundamental Frequency during Voicing Offset and Onset in Patients with Vocal Hyperfunction. *Journal of Speech, Language, and Hearing Research*, 54(5):1260–1266, 2011.

S. Stoikov. *Bulgarian Dialectology.* State Publishing House "Science and Art", 1962.

E. Stoin. *Musical and Folklore Dialects in Bulgaria.* Muzika, 1981.

J. Sundberg. Vibrato and vowel identification. *Dept. for Speech, Music, and Hearing Quarterly Progress and Status Report*, 16(2-3):49–60, 1975.

J. Sundberg. *The Science of the Singing Voice.* Northern Illinois University Press, DeKalb, Ill, 1987.

J. Sundberg. Acoustic and psychoacoustic aspects of vocal vibrato. *Dept. for Speech, Music, and Hearing Quarterly Progress and Status Report*, 35(2-3):45–68, 1994a.

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

J. Sundberg. Expressivity in singing. A review of some recent investigations. *Logopedics Phoniatrics Vocology*, 23(3):121–127, 1998. Publisher: Taylor & Francis ;https://doi.org/10.1080/140154398434130.

J. Sundberg. The Perception of Singing. In *The Psychology of Music.* Gulf Professional Publishing, 1999.

J. Sundberg, E. Prame, and J. Iwarsson. Replicability and accuracy of pitch patterns in professional singers. *Dept. for Speech, Music, and Hearing Quarterly Progress and Status Report*, 36(2-3):052–062, 1996.

J. Sundberg, L. Gu, Q. Huang, and P. Huang. Acoustical study of classical Peking opera singing. *The Journal of the Acoustical Society of America*, 131(4):3376–3376, 2012. Publisher: Acoustical Society of America.

S. Ternström and J. Sundberg. Acoustics of Choir Singing. In *Acoustics for choir and orchestra: papers given at a seminar organized by the Music Acoustics Committee of the Royal Swedish Academy of Music [on January 28th, 1985, in Stockholm]*, Stockholm, 1986. The Academy [Musikaliska akad.].

Z. C. Tezcaner, M. F. Gökmen, S. Yıldırım, and G. Dursun. Clinical Features of Psychogenic Voice Disorder and the Efficiency of Voice Therapy and Psychological Evaluation. *Journal of Voice*, 33(2):250–254, 2019. Publisher: Elsevier.

I. R. Titze. *Principles of Voice Production.* National Center for Voice and Speech, Iowa City, IA, 1994a.

I. R. Titze. Workshop on Acoustic Voice Analysis, 1994b.

I. R. Titze. One More Small Step in Solving the Mystery of the Benefits of Semioccluded Vocal Tract Exercises. *Journal of Singing*, 69(3):305–306, 2013.

I. R. Titze. Pitch Accuracy With and Without Vibrato. *Journal of Singing*, 71(2):207–208, 2014.

I. R. Titze. Self-Organization in Vocal Mechanics and Physiology. *Journal of Singing*, 78(2):233–234, 2021. Publisher: National Association of Teachers of Singing.

I. R. Titze, N. P. Solomon, E. S. Luschei, and M. Hirano. Interference between normal vibrato and artificial stimulation of laryngeal muscles at near-vibrato rates. *Journal of Voice*, 8(3):215–223, 1994. Publisher: Elsevier.

I. R. Titze, B. Story, M. Smith, and R. Long. A reflex resonance model of vocal vibrato. *The Journal of the Acoustical Society of America*, 111(5):2272–2282, 2002. Publisher: Acoustical Society of America.

M. Todorov. *On Bulgarian Folk Music.* Sofia University Press, 1966.

L. E. Toles, N. Roy, S. Sogg, K. L. Marks, A. J. Ortiz, A. B. Fox, D. D. Mehta, and R. E. Hillman. Relationships Among Personality, Daily Speaking Voice Use, and Phonotrauma in Adult Female Singers. *Journal of speech, language, and hearing research: JSLHR*, 64(12):4580–4598, 2021.

J. C. Toscano, B. McMurray, J. Dennhardt, and S. J. Luck. Continuous Perception and Graded Categorization: Electrophysiological Evidence for a Linear Relationship Between the Acoustic Signal and Perceptual Encoding of Speech. *Psychological Science*, 21(10):1532–1540, 2010. Publisher: SAGE Publications Inc.

P. F. Tosi. *Observations on the florid song; or, sentiments on the ancient and modern singers.* 1742.

C. Traube. *An interdisciplinary study of the timbre of the classical guitar.* Ph.D., McGill University, Montréal, Québec, Canada, 2004.

E. Van Houtte, S. Claeys, E. D'haeseleer, F. Wuyts, and K. Van Lierde. An examination of surface EMG for the assessment of muscle tension dysphonia. *Journal of Voice*, 27(2):177–186, 2013.

M. K. A. van Selms, J. W. Wiegers, F. Lobbezoo, and C. M. Visscher. Are vocalists prone to temporomandibular disorders? *Journal of Oral Rehabilitation*, 46(12):1127–1132, 2019.

J. H. Van Stan, N. Roy, S. Awan, J. Stemple, and R. E. Hillman. A Taxonomy of Voice Therapy. *American Journal of Speech-Language Pathology*, 24(2):101–125, 2015.

M. Vatti, S. Santurette, N. H. Pontoppidan, and T. Dau. Perception of a sung vowel as a function of frequency-modulation rate and excursion in listeners with normal hearing and hearing impairment. *Journal of speech, language, and hearing research: JSLHR*, 57(5):1961–1971, 2014.

P. Vekilova and S. Minkova. Specificity of sound formation and ornamentation in the Bulgarian Folk Song. *University of Ruse "Angel Kanchev" Research Works*, 49:198–202, 2010.

W. Vennard. *Singing: The Mechanism and the Technic.* Carl Fischer, L.L.C., 1968.

K. Verdolini, I. R. Titze, and A. Fennell. Dependence of phonatory effort on hydration level. *Journal of Speech and Hearing Research*, 37(5):1001–1007, 1994.

K. Verdolini, C. A. Rosen, and R. C. Branski, editors. *Classification Manual for Voice Disorders-I.* Psychology Press, New York, 2014.

K. Verdolini-Marston, I. R. Titze, and D. G. Druker. Changes in phonation threshold pressure with induced conditions of hydration. *Journal of Voice*, 4(2):142–151, 1990.

V. Verfaille and C. Guastavino. Perceptual Evaluation of Vibrato Models. In *Proceedings of the Conference on Interdisciplinary Musicology (CIM05) / Actes du Colloque interdisciplinaire de musicologie (CIM05)*, Montréal, Québec, Canada, 2005.

V. Vutov. *Phonetics and Phonology of the Bulgarian Language.* University Publishing House "St. Cyril and Methodius", Veliko Tarnovo, 2002.

M. Wadie, S. I. Adam, and C. T. Sasaki. Development, Anatomy, and Physiology of the Larynx. In *Principles of Deglutition*, pages 175–197. Springer, New York, NY, 2013.

Z. Wallmark. Body and Emotion in the Sonic Act. In Z. Wallmark, editor, *Nothing but Noise: Timbre and Musical Meaning at the Edge*, page 0. Oxford University Press, 2022.

C. R. Watts, S. S. Diviney, A. Hamilton, L. Toles, L. Childs, and T. Mau. The Effect of Stretch-and-Flow Voice Therapy on Measures of Vocal Function and Handicap. *Journal of Voice*, 29(2): 191–199, 2015.

J. L. Webster. The Mysterious Voice! American Women Singing Bulgarian Songs. *Anthropology of East Europe Review*, 22(1):155–169, 2004. Number: 1.

E. M. Weekly and J. L. LoVetri. Follow-Up Contemporary Commercial Music (CCM) Survey: Who's Teaching What in Nonclassical Music. *Journal of Voice*, 23(3):367–375, 2009. Publisher: Elsevier.

C. C. Wier, W. Jesteadt, and D. M. Green. Frequency discrimination as a function of frequency and sensation level. *The Journal of the Acoustical Society of America*, 61(1):178–184, 1977.

J. Williams. Vibrato in singers, 2009.

E. Willis and D. D. Mehta. Personal Conversation following presentation of "Adding vibrato-based ambulatory singing measures in the clinical assessment of vocal hyperfunction" at MGH Vocal Hyperfunction Clinincal Research Center Monthly Research Seminar, 2024.

C. R. Willson. *The Mystery of The Mystery of Bulgarian Voices*. Bachelor of Arts with Departmental Honors, Wesleyan University, Middletown, CT, 2017.

F. Winckel. Physikalische Kriterien für objektive Stimmbeurteilung. *Folia Phoniatrica et Logopaedica*, 5(4):232–252, 1953. Publisher: Karger Publishers.

J. Wodzak. Past and Current Issues in Functional Voice Disorder Classification. *Voice and Speech Review*, 0(0):1–18, 2022. Publisher: Routledge ;https://doi.org/10.1080/23268263.2022.2140913.

R. Wooding and J. Nix. Perception of Non-Vibrato Sung Tones: A Pilot Study. *Journal of Voice*, 30(6):762.e15–762.e21, 2016. Publisher: Elsevier.

J. Wrycza Sabol, L. Lee, and J. C. Stemple. The value of vocal function exercises in the practice regimen of singers. *Journal of Voice*, 9(1):27–36, 1995.

L. Yang, K. Z. Rajab, and E. Chew. AVA: A Graphical User Interface for Automatic Vibrato and Portamento Detection and Analysis. Utrecht, The Netherlands, 2016.

C. Yarbrough, J. Bowers, and W. Benson. The Effect of Vibrato on the Pitch-Matching Accuracy of Certain and Uncertain Singers. *Journal of Research in Music Education*, 40(1):30–38, 1992. Publisher: SAGE Publications Inc.

L. Zacconi. *Prattica di musica.* Bibliotheca musica Bononiensis. Forni, Venice, 1st edition edition, 1592.

A. Zacharakis, K. Pastiadis, and J. D. Reiss. An Interlanguage Unification of Musical Timbre. *Music Perception*, 32(4):394–412, 2015. Publisher: University of California Press.

S. A. Zahorian and H. Hu. A spectral/temporal method for robust fundamental frequency tracking. *The Journal of the Acoustical Society of America*, 123(6):4559–4571, 2008.

J. M. Zarate, C. R. Ritson, and D. Poeppel. Pitch-interval discrimination and musical expertise: Is the semitone a perceptual boundary? *The Journal of the Acoustical Society of America*, 132(2): 984–993, 2012.

L. Zenobi. *The Perfect Musician.* Musica Iagellonica, Katedra Historii i Teorii Muzyki Uniwersytetu Jagiellońskiego, 1856.

E. Zimmermann, T. Nestorova, J. Mac Master, C. Meloto, S. Komarova, I. Cossette, S. Hashemi, M. Moussa, and L. Manrique. Singing as TMD Therapy & TMD in Trained Singers: Analysis of Articulatory Kinematics, Muscle Tension, and Vibrato Quality in Singers & Non-Singers, 2023. Manuscript in preparation.

M. Zimmermann. *The Ornamentation of Baroque Music: A Guide for Independent Embellishing.* BoD - Books on Demand, Ettlingen, 2019.

E. Zwicker, G. Flottorp, and S. S. Stevens. Critical Band Width in Loudness Summation. *The Journal of the Acoustical Society of America*, 29(5):548–557, 1957.

Appendix A

Acoustic Study Protocol Task List

The following page includes the Voice Study Task List Protocol from Nestorova (2021) and Nestorova et al. (2024b).

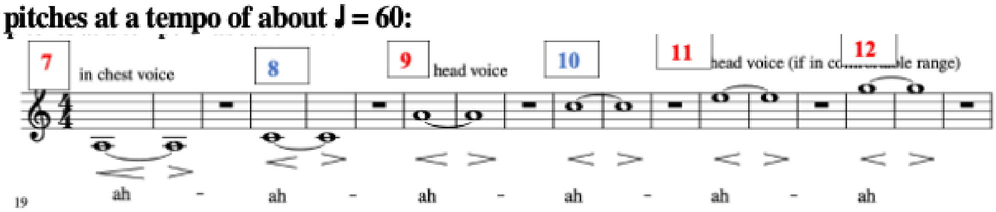
Voice Study Task List Protocol

Singing Voice Study Task List:

For the following tasks, please sing each phrase in your "best, most natural operatic / musical theater / jazz voice." No need to use a metronome for suggested tempo markings; sing at moderate tempo for all.

1. **Expressively sing the following excerpt from "Summertime" at a tempo of about ♩ = 70:**

2. **Sing a sustained *messa di voce* tone (crescendo to decrescendo on one note) on the following pitches at a tempo of about ♩ = 60:**



Repeat this task one more time with approximately two seconds between each repetition.

3. **Expressively sing "Amazing Grace":**

Appendix B

Reported Singing Vibrato Characteristics across Genres

B.1 Reported Singing Vibrato Characteristics Across Western to Eastern Musical Genres

The summary table on the next page details specific reported vibrato acoustic metrics (commonly derived vibrato parameters) of human subject studies conducted with singers across different genres / styles, for direct parametric comparison.

MUSICAL GENRE / STYLE	REFERENCE	SAMPLE SIZE / REPORTED DEMOGRAPHIC OF SINGERS	STUDY DESIGN / METHODS	RESULTS & CONCLUSIONS
Western Classical (Opera/Art Song)	Prame (1994) & Prame (1997)	10 singers (unknown m. vs. f.)	Vibrato rate measured on sonograms from commercial recordings	Vibrato rate increased at end of each tone. Vibrato extent had negative correlation w./tone duration, positive correlation w./intonation.
	Bezerra et al. (2009)	20 males (10 in each group)	Vibrato rate and extent & corresponding perceptual measures from live recordings	Vibrato rate lower on average in Sertanejo than in Western classical singers. No statistical difference for mean vibrato extent. More irregularity in frequency oscillation for Sertanejo.
	Ferrante (2011)	75 female singers	Vibrato rate, extent, tone length (duration), intonation & correlations from commercial recordings	Vibrato rate and extent correlated. Vibrato rate decreased over century (validated from previous studies).
	Nix et al. (2016)	75 students (29 m., 46 f.)	Mean vibrato rate, peak-to-peak extent, jitter, & mean SPL extent across 5 cardinal vowels and 3 production conditions in varied live recordings	High vs. low pitch and female vs. male mean vibrato rate support previous literature. Both vibrato rate and extent reduced in non-vibrato. Little difference in vibrato across vowels.
	Glasner and Johnson (2020)	20 (5 of each voice type)	Vibrato rate, extent, jitter (ddp), shimmer (dda), & f_o across each live recording condition (wax cylinder phonograph or microphone)	No significant effect of recording condition on vibrato rate or f_o, but overall higher vibrato extent, jitter, and shimmer (more cycle-to-cycle variability) in acoustic signals wax cylinder versus modern recording equipment.
	Arroabarren and Carlosena (2004)	4 (unknown m. vs. f.)	Vibrato rate, extent, and intonation across both frequency &amplitudemodulation realms	No significant differences in vibrato rate across singers, excellent correlation between rate, extent, and intonation and their corresponding perceptions.
Western Classical Early Music	Hakes et al. (1987)	4 (2 m., 2 f.)	Presence & amount of f_o oscillation, frequency location, vocal ornament compared to target tone from live recordings	Wide range of variability/individuality in vibrato rate and extent of f_o oscillations during ornament performances.
	Hakes et al. (1988)	10 (6 m., 4 f.)	Oscillation rate, frequency extent, & jitter of vibrato, exaggerated vibrato, whole-tone trill, & trillo measured from live recordings	Trillo rate much faster than other ornaments with vibrato. Regular vibrato least variable of all. Extent increased progressively from regular vibrato, exaggerated vibrato, whole-tone trill, to trillo.
	Hakes et al. (1990)	9 (5 m., 4 f.)	Rate of repeated trillo across both slow and fast from live recordings	Exaggerated vibrato had least jitter, trillo involved the most. No difference in trillo rates by pitch level according to range, voice type, experience, or age.
	Capobianco et al. (2023)	10 (5 m., 5 f.)	Use of BioVoice software (Morelli et al., 2021) to measure vibrato rate (Vrate), vibrato extent (Vext), vibrato jitter (Jvib), vibrato shimmer (Svib) and quality ratio (QR)	Vibrato in EM singing was characterized by a higher rate, a smaller extent, and less regular cycle-cycle period duration (higher Jvib) compared to Romantic Opera singing. As in previous studies, Romantic Opera rather than Early Music singing presented a more prominent singer's formant, as indicated by a smaller QR.
Musical Theatre	Becker and Watson (2022)	5 well-known MT treble performers' recordings of one legit and one belt performance	Downloaded recordings analyzed for average pitch, duration, the proportion of a note sung with vibrato, for vibrato rate (Hz), vibrato extent (semitones), and cycle-to-cycle perturbation (jitter-local and shimmer-local).	Average vibrato rate somewhat slower in belt versus legit production. No significant difference between legit and belt production for vibrato extent or cycle-to-cycle perturbation. Belt production had lesser proportion and shorter duration of notes sung with vibrato than in legit production.
Jazz	Manfredi et al. (2015)	48 total; 29 Western classical & 19 jazz– 29 f. & 19 m.	Vibrato rate, extent, duration, and jitter & shimmer regularity from live recordings	Vibrato rate did not differ across Western classical and jazz singing. Extent and duration significantly larger (especially females) and were longer in Western classical than Jazz singers. Jitter and shimmer were lower in Western classical than Jazz singers.
	Kuhlewind (2014)	6 female Conservatory mezzo-soprano voice students	Note duration, vibrato time (onset and duration of the vibrato portion), vibrato rate, number of vibratory cycles, average vibrato extent, maximum vibrato extent, minimum vibrato extent, position of minimum vibrato extent, position of maximum vibrato extent, difference between minimum and maximum vibrato extent, mean pitch, desired pitch, and deviation from desired pitch	Lower average vibrato rate and extent and higher variability for jazz versus Western classical singers. No significant differences between jazz and Western classical regarding vibrato onset and offset time, correlation between vibrato rate and note duration, and as in previous studies, same positive correlation between vibrato rate and vibrato extent.
Rock	Herbst et al. (2017)	Case study on 1 male	Acoustical analysis of Freddie Mercury's isolated *a cappella* vocals from commercial recordings	High frequency vibrato rate in vocal tremor range and highly irregular vibrato in Freddie Mercury's voice.
Brazilian Sertanejo	Pecoraro et al. (2013)	15 males (5 per group)	Vibrato rate variability from commercial recordings of professional Opera, Rock, & Brazilian Sertanejo singers	Mean rate values for opera and Brazilian Sertanejo singers higher overall than for rock singers.
Portuguese Fado	Mendes et al. (2013)	15 amateurs (10 m., 5 f.)	Acoustic and phonatory profile, jitter, shimmer, HNR of live recordings	Vibrato rate and extent within range described in previous Western classical singing studies.
Chinese Peking Opera	Sundberg et al. (2012)	7 males	Sound pressure level, f_o, & spectrum characteristics on live recordings	Vibrato rate of Peking opera considerably slower than average Western classical vocal vibrato rate.

	Han and Zhang (2017)	7 roles (3 m., 4 f.)	f_o, electroglottographic, and inductive plethysmography signals in live recording	Mean vibrato rate of Peking opera much slower than that of Western classical singing, but mean vibrato extent demonstrated no obvious differences in comparison.
Indian Classical	Radhakrishnan et al. (2011)	Case study on 1 male	f_o, alternating current (AC) glottal flow, and electroglottographic width measures on live recording	Voluntarily controlled *taan* rates significantly lower than frequency oscillation rates of involuntary vibrato production and trill, but on lower end for trillo.
Indian Popular	Bonjyotsna and Bhuyan (2016)	10 professionals (unknown m. vs. f.)	Vibrato parameters and mordent parameters are computed from the fundamental frequency contour of the samples using the Praat software (Boersma and Weenink, 2025)	Average Indian popular singing vibrato rates were comparative to Western classical norms, but dependent on song tempo, usually four times the bpm for a range of 80–100 bpm.
Byzantine Chant	Delviniotis and Theodoridis (2019)	4 male chanters	Slope [SS], relative speaker's formant [SPF] level, formant frequencies [Fi] and bandwidths [Bi], and noise-to-harmonics ratio [NHR] difference values between the vocal ornament and its neighbor steady note, and the rate and extent, were compared with those of vibrato	Mean frequency oscillation extent values of Byzantine ornament were almost double corresponding vibrato values. Byzantine ornament frequency oscillation rate was slightly higher than vibrato.
Jewish Cantorial	Rothman et al. (2000)	16 majority males (4 per group)	Frequency & amplitude vibrato rate & extent from commercial & live historical (early 20th C.) & contemporary (late 20th C.) recordings	Mean extent values of ornaments almost double corresponding vibrato values. Ornament rate slightly higher than vibrato.

Appendix C

f_o Contour Derivation and Pre-Processing Algorithm

C.1 f_o Contour Derivation and Pre-Processing Algorithm

Task list for the singers: The 16 items task list including the first verse of two genre-fluid songs ("Summertime" and "Amazing Grace") and several *messa di voce* (crescendo-decrescendo) vocal exercises is given in Table C.1.

Sample Index	Sample Text	Sample Type	Sample Pitch
1	time	song	C5
2	easy	song	C4
3	jumpin	song	Bb4
4	high	song	G4
5	rich	song	C5
6	lookin	song	C4
7	[*a*]	exercise	A3
8	[*a*]	exercise	C4
9	[*a*]	exercise	A4
10	[*a*]	exercise	C5
11	[*a*]	exercise	E5
12	[*a*]	exercise	G5
13	grace	song	F4
14	me	song	A4
15	found	song	A3
16	see	song	D4

Table C.1: Vibrato Sample Data with Text, Type, and Pitch

Pre-processing of data: The fundamental frequency (f_o) contour for each sample was isolated and extracted using the "Filter bandpass" and "Get pitch" commands in the Praat software platform. Each sample was then subjected to sinusoidal extraction of the mean pitch frequency (Hz) at 0.02 seconds from a standardized start time = 0 via a custom script. Using a custom code written in RStudio, the peaks and troughs (or tops and bottoms of the sinusoid) of each sample's f_o contour were derived.

Both onsets and offsets in the f_o contours were subsequently manually and visually incorporated into the vibrato tokens for analysis. Inclusion of the entire acoustic signal, including beginning, middle, and end, ensured fully representative vibrato tokens for analysis of each sample.

The f_o contour data was subsequently converted into cents via the transformation variable equation first reported in Herbst et al. (2017), and modified here so that the reference point is a rounded and even 100, permitting non-even temperament scales (instead of the original approximation of $C_4 \approx 261.63Hz$).

$$c[i] = 1200\frac{\log\left(f_o[i]/100\right)}{\log(2)} \tag{C.1}$$

Vibrato extent was then calculated as the average absolute deviation c from the mean musical pitch by using the following equation from Herbst et al. (2017), modified by Glasner and Johnson (2020):

$$\overline{\Delta c} = \frac{1}{n}\sum_{i=0}^{n-1}\frac{|\overline{c} - c_i[f_{o\_min}]| - (\overline{c} - c_i[f_{o\_max}])}{2} \tag{C.2}$$

where $c_i[f_{o\_min}]$ is the indexed minimum f_o of the deviation of a vibrato cycle expressed in cents, and $c_i[f_{o\_max}]$ is the opposite peak of the same vibrato sample. Half instead of full-extent was calculated by first converting each local f_o maximum and minimum to cents in relation to middle C (C_4, 261.63 Hz).

Appendix D

Statistical Analysis of the Half-Extent Time Profiles

D.1 Statistical Analysis of the Half-Extent Time Profiles

The time profiles of the half-extent are fit with a linear regression (function lm()) or a 4-parameter logistic regression line (Mager et al., 2003) (functions *L.4* or *LL.4* respectively from the R package *drc*) (Knezevic et al., 2007). The mathematical expressions for the two logistic regression curves are as follows:

$$f(x) = c + \frac{d - c}{1 + \exp\{b(x - e)\}} \tag{D.1}$$

$$f(x) = c + \frac{d - c}{1 + \exp\{b(\log x - \log e)\}} \tag{D.2}$$

where:

- *b* – slope
- *c* – lower asymptote
- *d* – upper asymptote
- *e* – the middle time value (time at which the value of f(x) is half of (d-c)).

The specific interpretation of the logistic function parameters for the analysis of the half-extent time profiles is given in the body of the manuscript. For the detection of the transition phase between the lower and upper asymptote, the inverse functions of the logistic expressions (3) and S(4) are used as functions *L.4* or *LL.4* respectively, as follows:

$$x = e + \frac{1}{b} \log \left[\frac{d - c}{y - c} - 1 \right] \tag{D.3}$$

$$x = \exp\left\{\log e + \frac{1}{b}\log\left[\frac{d-c}{y-c} - 1\right]\right\} \quad \text{(D.4)}$$

where:

- t_1 – the time of the start of the transition phase when *y = c + 0.2*(d-c)*
- t_2 – the time of the end of the transition phase when *y = d - 0.2*(d-c)*,

The transition phase itself is defined as the profile between 20% and 80% of the distance between the lower and upper asymptote, which coincides approximately with the linear ascend of the logistic curve.

The model discrimination between the linear and logistic approximations is done via the Akaike Information Criterion (AIC) and the Bayes Information Criterion (BIC) (Mohammed et al., 2015; Sebaugh and McCray, 2003), as detailed below:

$$AIC = n \cdot \ln\frac{SSE}{n} + 2k \quad \text{(D.5)}$$

$$BIC = -n \cdot \ln L + k \cdot \ln(n) \quad \text{(D.6)}$$

where:

- *n* – sample size
- *k* – the number of parameters of the approximation (n = 2 for linear model, k = 4 for the logistic model)
- *SSE* – the Sum of the Squared Errors from the approximation
- *L* – the likelihood of the approximation

Both the AIC and the BIC are automatically calculated from the approximation functions in R. The model selection favors the approximation with the lower AIC and BIC.

As noted in the pertinent sections of the thesis above, the detailed data analysis procedures as well as the annotated scripts of the original source code material used to develop these models and parameters are made available by the author of this thesis at `https://github.com/theodora-nestorova/vocalvibratovariability`

Appendix E

REB Consent Forms

The following pages consist of the Ethics Approved Consent Forms from the Research Ethics Boards at McGill University for several of the studies included in this thesis.

McGill

INFORMED CONSENT FORM | MUSIC PERCEPTION AND COGNITION LABORATORY | SCHULICH SCHOOL OF MUSIC, McGILL UNIVERSITY

REB 201-1114 | *Timbre of the Singing Voice*

WHY ARE WE DOING THIS RESEARCH? Our aim is to establish scientific knowledge about how people hear and interpret the sound characteristics of the singing voice. We also seek to understand how musicians from various domains understand the notion of musical timbre.

PRIVACY. We know that you value your privacy. All data will also be held securely in a password-protected computer system (for electronic data) or a locked office (for paper-based data). Only the principal investigator of the study and participating students will have access to identifiable information. You will not be identified as an individual in any scientific report of this research, and your name will not be linked to your responses in this study unless we have explicitly asked you to provide written consent to be named and quoted. You may discontinue your participation in this study at any time either during the interview process or in the future.

WHAT WILL HAPPEN DURING THE EXPERIMENT? The interviewer will ask you a number of questions to which you may respond freely. Your responses to the questions will be recorded. The recordings will be transcribed for analysis. The experiment will take approximately **1** hour. You are free to withdraw from the experiment at any time without prejudice. If you withdraw, you can ask to have the data collected be destroyed. However, once data have been aggregated or published, they can't be withdrawn. They can only be removed from use in further analyses. Identifiable materials will be kept for 7 years following publication. Once the data has been de-identified, they can't be withdrawn. You have the right to refuse to do anything that you find disturbing or uncomfortable in this study. You will be free to discontinue your participation at any time. Feel free to ask any questions you may have of the interviewer.

This research is funded by the Social Sciences and Humanities Research Council.

The Research Ethics Board II of McGill University has reviewed this study for compliance with ethical standards. If you have any ethical concerns or complaints about your participation in this study, and want to speak with someone not on the research team, please contact the Associate Director, Research Ethics at 514-398-6831 or lynda.mcneil@mcgill.ca citing REB file number 201-1114.

PARTICIPANT'S STATEMENT:
"I have read the preceding details and voluntarily agree to participate and to have my responses recorded. I do ____ do not ____ [check one] agree that my name be cited and any direct quotes from the recording of the interview with me may be reproduced. I understand that my approval will be obtained before a direct quote is sent for publication. I understand that by consenting, I do not waive any legal rights."

______________________ ______________________ ______

Signature Printed Name Date

I would like to participate in other experiments in the Music Perception and Cognition Lab ☐

I would like to receive a summary sheet of the experimental findings ☐

E-mail Address: ____

Lindsey Reymore, Postdoctoral Researcher, Schulich School of Music, McGill University
lindsey.reymore@mail.mcgill.ca
Kit Soden, Postdoctoral Researcher, Schulich School of Music, McGill University
kit.soden@gmail.com
Theodora Nestorova, PhD Candidate, Schulich School of Music, McGill University
theodora.nestorova@mail.mcgill.ca
Juanita Marchand Knight, PhD Student, Concordia University
juanita.marchand@mail.concordia.ca
Prof. Stephen McAdams, Schulich School of Music, McGill University
stephen.mcadams@mcgill.ca

McGill Schulich School of Music École de musique Schulich

Research Consent Form

Title of Project: *Bulgarian Folk Singing Vibrato Analysis*

Principal Researcher: *Theodora Nestorova, Interdisciplinary Ph.D. Student, Schulich School of Music, McGill University, +1-425-943-1556, theodora.nestorova@mail.mcgill.ca*

Supervisor: *Gary Scavone, Chair, Department of Music Research, Professor of Music Technology, Schulich School of Music, McGill University, 514-398-4400 ext. 089834, gary.scavone@mcgill.ca*

Purpose of the Study: Vocal vibrato in non-Western musical genres remains an under-researched topic. Published research on the Bulgarian Folk Singing Style in particular is almost non-existent. The objective of this study is to begin this research; evaluating the presence, classification, characteristics and artistic implementation of the vibrato in the Bulgarian women's folk singing style. As a part of this inquiry, an essential aspect is meeting with native vocal research experts and pedagogues with experience in teaching and performing Bulgarian women's folk singing style and discussing the level of understanding and knowledge about the vibrato through guided interviews as well as collecting audio samples of various types Bulgarian women's folk singing vibrato to derive the characteristics of Bulgarian folk singing vibrato.

Study Procedures: The interview and/or audio recording session will take place at The National School of Folk Arts Philip Kutev in Kotel, Bulgaria or elsewhere at an appropriate location in Bulgaria. You will be asked specific questions pertaining to vibrato or to perform specific music from your repertoire that involves vibrato. A portable, high-quality microphone and digital audio recorder will be recording this session and you will be compensated for your time.

Your participation is entirely voluntary and you can choose to withdraw at any point from the study's interviews and/or audio recordings, for no given reason, and with no consequences. If you choose to withdraw from the study, any information and data related to you will be destroyed.

Voluntary Participation*:* Participation in this study is fully voluntary. If you decide to withdraw at any time (consent may be revoked at any time), your data will be destroyed unless you give explicit permission otherwise. Should you wish, you may be given additional information about the study after your participation in both lessons is complete, in a debriefing discussion. If you decide to withdraw after your participation has ended, once the data has been combined for publication, it may not be possible to withdraw it in its entirety.

Confidentiality & Data-Use: Your personal identifiable data will be encoded and turned into numerical data for confidentiality in the processing and analysis. *Due to the nature of the audio recordings, anonymity cannot be guaranteed.* However, you may withdraw from the study at any time, for no given reason, with no consequences, and your data as well as any identifying information will be destroyed. All audio recordings from your participation will be kept for up to 7 years and only shared for publication purposes if you consent to this. Any paper forms you submit to the researchers will be scanned, and the paper copy will be destroyed within 24 hours of your participation. All data will be digitized and kept in a password-protected McGill OneDrive folder (confirmed to be on the list of McGill IT's approved cloud services: https://www.mcgill.ca/it/cloud-services/approved-cloud-services) and will *only be accessible by the PI and her faculty co-investigator and supervisors and advisors.*

Potential Risks: There are no reasonably foreseeable harms or discomfort that you or others might be subject to during or as a result of this research.

Compensation: You will receive compensation at a rate of 15 BGN per hour after the session(s).

Potential Benefits: Your participation in this study may contribute towards a more in depth understanding of Bulgarian women's folk singing style and knowledge using vibrato for the advancement of singing voice science, pedagogy, and performance.

Dissemination of Data: The results from these interviews and audio recordings will be gathered for analysis that will be solely for the use of the PI's Ph.D. research work, supervised by Dr. Gary Scavone. This will only be published publicly for educational purposes of the PI and her supervisors and advisors. Data collected from the interviews and audio recordings might serve as insights into the PI's doctoral dissertation.

Questions: Contact the PI, Theodora Nestorova, at theodora.nestorova@mail.mcgill.ca or the faculty project supervisor, Gary Scavone, at gary.scavone@mcgill.ca

If you have any ethical concerns or complaints about your participation in this study, and want to speak with someone not on the research team, please contact the Associate Director, Research Ethics at 514-398-6831 or lynda.mcneil@mcgill.ca citing REB file number 22-06-071.

Thank you for your interest and time in participating in this project!

WRITTEN CONSENT:

Please sign below if you have read the above information and consent to participate in this study. Agreeing to participate in this study does not waive any of your rights or release the researchers from their responsibilities. To ensure the study is being conducted properly, authorized individuals, such as a member of the Research Ethics Board, may have access to your information. A copy of this consent form will be given to you and the researcher will keep a copy.

YES: ___ NO: ___ I consent to be audio-recorded while singing various samples.
YES: ___ NO: ___ I consent to be audio-recorded during the interview.
YES: ___ NO: ___ I consent to have my audio-recording(s) shared for publication purposes.

Participant's Name: (please print) ____

Participant's Signature: ____________________________________ **Date:** _______________

Principal Researcher's Signature: ____________________________ **Date:** _______________
