

COMPLEXITY AND SYNCHRONICITY OF VIBRATO IN THE
SINGING VOICE AND DUET SINGING

GERARDO ACOSTA MARTÍNEZ

PHD

UNIVERSITY OF YORK
PHYSICS, ENGINEERING AND TECHNOLOGY
SEPTEMBER 2025

Abstract

Vibrato is a musical ornament that is used extensively in the singing voice. It is defined as modulation of fundamental frequency from the mean perceived pitch. It has been currently studied using parameters that explain the amplitude and frequency of modulation, as well as mean variation metrics to calculate regularity. Vibrato is an important and controversial topic in the context of group singing with the extent to which it should and should not be utilized or can even be consciously controlled in multi-voice singing often debated. It is speculated that vibrato affects choral blend and that singers can modify or adapt in an effort to improve choral blend.

This thesis explores the potential of nonlinear metrics to understand vibrato complexity. Additionally, the application of recurrence metrics measured from coupled vibrato tones during duet blended singing is studied. Results confirmed that singers can adapt and modify their vibratos while singing in a blended duet situation. Instances of high phase synchronisation were found between choral singer participants singing in a blended context. The results give insights into how music produces events of synchronisation that future research can further explore, and the potential well-being benefits that arise from synchronisation in music and the perception of choral blend.

Contents

List of Tables	vii
List of Figures	ix
Acknowledgments	xii
Declaration	xiv
1 Introduction	1
1.1 Vibrato in the Singing Voice	2
1.2 Vibrato Complexity	2
1.3 Synchronisation	3
1.4 Statement of Hypothesis	4
1.5 Contributions	4
1.6 Thesis Structure	6
1.7 Statement of Ethics	7
2 A Literature Review of Vibrato in Group Singing	8
2.1 Introduction	8
2.1.1 Vibrato in the The Singing Voice	8
2.1.2 Vibrato in Group Singing and Choral Singing	10
2.2 Methodological Approaches to Vibrato in Group Singing	11
2.3 Thematic Categories of Vibrato in the Singing Voice	14

2.3.1	Adjustment: Intrapersonal Adaptations and Vibrato Matching	15
2.3.2	Identification: Auditory Isolation of Voices	15
2.3.3	Interaction: Interpersonal Relationship of Voices	16
2.3.4	Intonation: Fundamental Frequency Trajectories and Accuracy	17
2.3.5	Style: Solo Singing, Choral Singing, and Choral Sound	18
2.4	Studies Focused on Group Singing and Vibrato	20
2.4.1	Production and Perception	22
2.4.2	Review of Common Methods	23
2.4.3	Vibrato and Choral Singing	25
2.4.4	Analysis Techniques: Complex Vibrato	28
2.4.5	Virtual Choirs	30
2.4.6	Implications of Choral Vibrato Research	30
2.5	Conclusion	31
3	Nonlinear Time Series Analysis of Vibrato	32
3.1	Introduction	34
3.2	Methodology	35
3.3	Results	37
3.4	Discussion	47
3.5	Conclusions	51
4	Vibrato Complexity From Jazz and Opera Singers	52
4.1	Introduction	52
4.2	Materials and Methods	54
4.2.1	Music Recordings	55
4.2.2	Voice Isolation	56
4.2.3	Vibrato Notes	56

4.2.4	Time Series Analysis	57
4.2.5	Data Analysis	61
4.3	Results	63
4.3.1	Vibrato Metrics	63
4.3.2	Dimensionality Reduction	64
4.3.3	First Principal Component	67
4.3.4	Recurrence Plots	69
4.3.5	Deterministic Controls	71
4.3.6	Low Determinism	74
4.4	Discussion	74
4.4.1	Regularity Metrics	76
4.4.2	Vibrato Usage	78
4.4.3	Limitations	79
4.4.4	Future Work	80
4.5	Conclusion	81
5	Vibrato Synchronicity in Blended Duet Singing	83
5.1	Introduction	84
5.2	Background	85
5.3	Methods	86
5.3.1	Musical Material	87
5.3.2	Pre-Recorded Singer	88
5.3.3	Vibrato Modification	89
5.3.4	Participants	90
5.3.5	Data Analysis	95
5.4	Results	98
5.4.1	Pre-Recorded Singer Vibrato	98
5.4.2	Participants' Vibrato	100
5.4.3	Vibrato Synchronisation	102

5.5	Discussion	111
5.5.1	Limitations	118
5.5.2	Future Directions	120
5.6	Conclusion	121
6	Vibrato in Blended Live Singing Duets	122
6.1	Introduction	122
6.2	Methods	123
6.2.1	Duets	123
6.2.2	Musical Material	124
6.2.3	Room and Equipment	126
6.2.4	Procedure	126
6.2.5	Data Analysis	127
6.2.6	Statistical Tests	128
6.3	Results	129
6.3.1	Vibrato Metrics	129
6.3.2	Synchronisation	133
6.4	Discussion	140
6.5	Conclusion	142
7	Conclusion	143
7.1	Further Work	144
7.2	Final Remarks	145
A	Systematic Review Data	146
B	Vibrato Dynamics	149
B.1	Introduction	149
B.2	Methods	150
B.2.1	Rössler System	150

B.3 Results	151
C Experimental Supplementary Material	158
Abbreviations	164
Bibliography	165

List of Tables

3.1	Descriptive characteristics of the notes analysed, vibrato parameters, and regularity metrics.	44
3.1	(cont.)	45
4.1	Singers and music material selected for collection of vibrato notes.	55
4.2	Descriptive characteristics of vibrato notes from opera and jazz singers.	63
4.3	Comparison of vibrato notes from opera singing and jazz singing using the first three principal components.	67
4.4	Comparison of two notes with high score on Component 1 from Opera (note 291), Jazz (note 431), and other highly deterministic signals adjusted to oscillate within standard vibrato parameters.	71
5.1	Musical notes and vowels used as singing material.	88
5.2	Descriptive characteristics of the prerecorded singer and the participants in the study.	91
5.3	Timeline and sequence of the experimental session.	95
5.4	Vibrato parameters of the prerecorded singer notes.	98
5.5	Vibrato parameters from the participants in the Solo and Blending conditions.	100

5.6	Rate post-hoc comparison of the participant vibrato notes from the Blending and Solo singing conditions.	100
5.7	Extent post-hoc comparison of the participant vibrato notes from the Blending and Solo singing conditions.	101
5.8	Jitter post-hoc comparison of the participant vibrato notes from the Blending and Solo Singing conditions.	101
5.9	Shimmer post-hoc comparison of the participant vibrato notes from the Blending and Solo singing conditions.	101
5.10	Mean phase synchronisation coefficients calculated from the time series of participants blending to the prerecorded singer. . .	103
5.11	Mean sound level calculated from the DPA close microphone notes across the different singing conditions (Solo and Blending) in the study.	110
6.1	Descriptive characteristics of Singer A and the participants that conformed the duets in the study.	125
6.2	Assignment of duet number based on the singer labels from the experiment in Chapter 5, Table 5.10. (Singer A took part in all duets.)	125
6.3	Rate live adaptations.	129
6.4	Extent live adaptations.	130
6.5	Jitter live adaptations.	131
6.6	Shimmer live adaptations.	131
6.7	Vibrato coefficient live adaptations.	132
6.8	Phase synchronisation in the live context and the unedited stimuli values for comparison.	133
A.1	Methodology approaches of studies included in the review. . .	146
A.2	Objective category of each study.	147
A.3	List of assessed records and exclusion criteria.	147
A.3	(cont.)	148

List of Figures

2.1	Proportions of the methodology approaches of studies addressing vibrato in group singing context.	12
2.2	Thematic categories of studies addressing vibrato in choral singing contexts.	14
3.1	Flowchart of data extraction and analysis.	36
3.2	Phase space of note 46.	38
3.3	Recurrence plot of note 46.	39
3.4	Phase space of note 67.	40
3.5	Recurrence plot of note 67.	41
3.6	Phase space of note 62.	42
3.7	Recurrence plot of note 62.	43
3.8	Phase space of note 17.	45
3.9	Recurrence plot of note 17.	46
3.10	Phase space of note 43.	47
3.11	Recurrence plot of note 43.	48
4.1	Overview of the methods utilised for vibrato analysis in this study.	54
4.2	Flowchart of time series analysis and data visualisation performed in MATLAB.	58
4.3	Two vibrato notes of the same length as analysed by the implemented vibrato code.	62

4.4	Scree plot of the PCA model performed with the normalised vibrato data.	64
4.5	Loadings of the first three principal components.	65
4.6	Score plot of the first three principal components with the 780 vibrato time series used in this study.	66
4.7	Differences across singers within Opera and Jazz on the first principal component.	68
4.8	Opera vibrato with high score on Component 1 from Carreras (note 291).	70
4.9	Jazz vibrato with high score on Component 1 from Fitzgerald (note 431).	72
4.10	Dynamics from one of the states from the Rössler system with $a = 0.1$, $b = 0.1$, and $c = 12$	73
4.11	Opera vibrato with low score on Component 1 from Callas (note 86).	75
5.1	Overview of the experimental processes in this study.	87
5.2	Five modulation modifications edited with Flex Pitch in Logic Pro X used as different vibrato conditions.	90
5.3	Experimental setup and participant during recording and blending to the prerecorded singer (reproduced from the Genelec speaker).	92
5.4	Cross Recurrence Plot between the Lx and DPA microphone signals recorded from a participant to assess their similarity. The upper graph represents the underlying time series from the Lx (black) and the DPA (red). The diagonal Line of Identity indicates both signals contain the same frequency information.	99
5.5	Example of paired vibrato notes with higher probability of phase synchronisation between the prerecorded singer and the participant time series.	105

5.6	Example of paired vibrato notes with lower probability of phase synchronisation between the prerecorded singer and the participant time series.	106
5.7	Example of paired vibrato notes with negative probability of phase synchronisation between the prerecorded singer and the participant time series.	107
5.8	Correlation analysis of jitter (Solo condition), shimmer (Solo condition), and mean CPR (Blending conditions) data from the 23 participants, 24 notes per condition, in the study. . . .	108
6.1	High sync with pitch drifting.	134
6.2	A singer leading vibrato production.	135
6.3	High sync blending.	136
6.4	Both singers singing with a straighter tone.	137
6.5	Both singers adding late vibrato.	138
6.6	Participant singing with a straight tone whilst Singer A added vibrato.	139
6.7	Attractor from a straighter tone engulfed by an attractor with higher extent, depicted on the same phase space.	140
B.1	In-phase synchronisation.	152
B.2	Anti-phase synchronisation.	153
B.3	Ninety-degree synchronisation.	154
B.4	Chaos synchronisation.	155
B.5	Time series of four synthesised vibrato waves.	156
B.6	Four-dimensional vibrato dynamics: The Tesseract vibrato. .	157

Acknowledgments

Firstly I would like to thank my supervisor, Prof Helena Daffern, and express my deep gratitude for her guidance throughout my PhD. All of the help with her comments and revisions on my writing was invaluable for me and the completion of this PhD.

I would also like to thank the members of my Progression Panel, Prof Gavin Kearney and Prof Damian Murphy, who provided me with really useful comments and advice on my research.

Many thanks to the National Council of Humanities, Sciences, and Technologies and the Foundation Institute of Fine Arts for the support with a PhD scholarship; and to the Society for Education and Music Psychology Research for the support with the Arnold Bentley New Initiatives Award.

Heartfelt thanks to the singers who took part in my study. Doing the singing experiments was my favourite part of this PhD, and I'm grateful for the experience of listening to these amazing singers and learning from the comments they shared regarding vibrato and choral singing - deep gratitude for their time and interest in the study. Also, special thanks to the anonymous Singer A! Singer A kindly agreed to participate in dozens of my experiments, and I am indebted to her for all the support she provided with her singing expertise and deep interest in my research.

I would like to thank the AudioLab and everyone there, especially the tech staff for the help with the setup and gear for the experiments. Also, thanks to the MINERVA and SAFFIRE teams - being able to be part of those projects was a highlight of my degree. And thanks to everyone at the lab for providing a really friendly and fun work environment - lunch time

there is hilarious.

I should thank all the wildlife and animals that live around campus - they make the experience of studying here and living on campus a hundred times better.

Y lo más importante! Qué habría sido de este doctorado sin el apoyo y confianza de mi mamá y papá, Sonia y Miguel. Mi mamá me dejó convertir sus recursos en esta tesis, aparte de siempre ayudarme con mis nuevos treinta problemas cada que le llamaba. Gracias a ella es que este libro existe. Y mi papá, él es el culpable de mi curiosidad por la interpretación y tecnología musical, compartiendo buena música desde mi niñez. La investigación de este doctorado me ayudó a entender porque es que nos gusta tanto escuchar aquellos artistas y cantantes. Yo pensaría que entre los dos me transmitieron interés por el aprendizaje y la música - yo nada más medio que quise combinar esos dos en uno. (Por extensión, también agradezco a mi abuela y abuelo, María y Juan!) Y ahora sí que como diría ya saben quién: Cómo les puedo pagar todo lo que hacen por mí, todo lo feliz que soy, todo este grande aprendizaje. Solamente con mi tesis? Tengan mi tesis. Se las doy.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references.

I also declare that parts of this research have been presented as part of the following conference abstract proceedings:

- Gerardo Acosta Martínez and Helena Daffern, “Vibrato dynamics in group singing: Understanding vibrato behaviour through dynamical systems theory,” in Abstracts of the 14th Pan-European Voice Conference, Tallinn, Estonia, 2022, p. 46.
- Gerardo Acosta Martínez and Helena Daffern, “Analysis of vibrato variability and adaptability in a duet singing experiment: A pilot study,” in Abstracts Book of the 15th Pan-European Voice Conference, Santander, Spain, 2024, pp. 87–88.

Additionally, parts of this thesis have been presented as peer-reviewed full papers in conference proceedings as follows:

- Gerardo Acosta Martínez and Helena Daffern, “A pilot study of vocal vibrato incorporating nonlinear time series analysis,” in *Proceedings of the Stockholm Music Acoustic Conference*, S. D’Amario, S. Ternström, and A. Friberg, Eds., Stockholm, Sweden, 2023, pp. 182–189. DOI: 10.30746/978-91-8040-865-3.

Furthermore, parts of this thesis have been presented in scientific journal papers as follows:

- Gerardo Acosta Martínez and Helena Daffern, “The role of vibrato in group singing: A systematic review,” *Journal of Voice*, vol. 39, no. 3, 847.e21–847.e39, 2025. DOI: 10.1016/j.jvoice.2022.11.004.
- Gerardo Acosta Martínez and Helena Daffern, “Complexity of vocal vibrato in opera and jazz recordings: Insights from entropy and recurrence analyses,” *Journal of Voice*, 2023. DOI: 10.1016/j.jvoice.2023.11.020.
- Gerardo Acosta Martínez and Helena Daffern, “Vibrato synchronicity in sustained notes from female singers blending to a prerecorded singer,” *Journal of Voice*, 2025. DOI: 10.1016/j.jvoice.2025.09.006

To Apollo

Chapter 1

Introduction

*Since the early days, I've always used vibrato on chords
as a way of adding real depth and heaviness to the sound.*

—Tony Iommi

The best musicians in the world can often be distinguished by their creative and extensive use of artistic vibrato. This phenomenon has been extensively studied in *solo instruments* and instruments that can produce single notes with one fundamental frequency—like the singing voice, violins, or guitar. However, the presence of vibrato in multiple, simultaneous notes or chords produced by a single instrument is more a complex, and in some cases controversial, phenomenon.

One of the first definitions of vibrato comes from the work of Seashore in 1931, *The Natural History of the Vibrato*. He provided a definition of this ornament as:

An artistic vibrato consists of a periodic oscillation in pitch in which the extent of oscillation for the best singers averages approximately a half-tone and for string instruments approximately a quarter-tone, at an average rate of approximately six or seven cycles per second, and is usually accompanied by synchronous intensity and timbre oscillations which play a secondary role. [1]

This definition provides key quantitative parameters that are paramount in our understanding of vibrato to this date: The extent of vibrato, and the rate of vibrato, which will be explored thoroughly throughout this thesis. But given that vibrato, and music in general, is of a more perceptual nature, a qualitative definition was also provided as: “A good vibrato in music is a periodic pulsation, generally involving pitch, intensity, and timbre, which produces a pleasing flexibility, mellowness and richness of tone.” This definition integrates the frequency modulation, amplitude modulation, and spectral characteristics of vibrato in the singing voice.

1.1 Vibrato in the Singing Voice

In the singing voice, further research has explored the artistic, stylistic, and emotional implications of vocal vibrato. One of the more comprehensive reports is the *Acoustic and psychoacoustic aspects of vocal vibrato* [2] by Johan Sundberg, 1994. In this research, Sundberg goes further into describing the acoustic parameters used to define vibrato, such as rate and extent, but also incorporates “regularity”, and “waveform”. This work reports the generally observed parameters of rate of 5 to 7 undulations per second (Hz), and extent of ± 1 semitones. In this paper, it is described that regularity “shows how similar the frequency excursions are to one another. It is considered a sign of the singer’s vocal skill: the more skilled the singer, the more regular the undulations” and that the waveform is similar to a sine wave.

1.2 Vibrato Complexity

The acoustic approach used to study vibrato, in the form of rate and extent, provided exceptional insights into the nature of vibrato oscillations. These metrics study specifically the amplitude (or depth) of the fundamental frequency variation, as well as the frequency (or reciprocal of wavelength period) of those variations within a single second. However, for the regularity and waveform aspects of vibrato, a different perspective could be beneficial

for both the qualitative and quantitative understanding of it. From the quantitative side, the greatest advancement in understanding vibrato regularity was made by Horii [3] with the introduction of jitter and modulation shimmer to study regularity of vibrato. In this context, jitter reflects the mean variation of vibrato rate, and modulation shimmer reflects the mean variation of vibrato extent. Although these metrics are useful, they are derived from analysis used in the fields of acoustics. The use and application of metrics derived from dynamical systems and complex systems to bio-signals has been extensively studied in the past but not in the study of vibrato regularity. The study of vibrato complexity through the lens of nonlinear dynamics opens up possibilities of further understanding vibrato as a biological signal.

1.3 Synchronisation

The case of multiple and simultaneously produced vibratos in the context of group singing and choral singing poses an interesting questions regarding entrainment and human interactions. The concept of adjustment of rhythms and *frequency entrainment* describes that:

If two nonidentical oscillators having their own frequencies f_1 and f_2 are coupled together, they may start to oscillate with a common frequency. Whether they synchronize or not depends on the following two factors. (1) Coupling strength: This defines how weak (or how strong) the interaction is. [...] (2) Frequency detuning or mismatch $\Delta f = f_1 - f_2$ quantifies how different the uncoupled oscillators are [4].

Another benefit of the application of complexity and dynamical metrics in vocal vibrato, is that synchronisation has been widely researched and studied using these metrics. The field of synchronisation points out that there are many benefits that come from entrainment. The author Buzsáki describes of the term synchronisation,

Syn (meaning same) and chronos (meaning time) together make

sure that everyone is up to the job and no one is left behind, the way the conductor creates temporal order among the large number of instruments in an orchestra [5].

A sort of schism between the field of synchronisation and entropy has also been suggested. Where the laws of thermodynamics and entropy would dictate that things tend towards entropy and disorder. But synchronisation has given evidence of how in living, and non-living things, there seems to be a powerful tendency for synchronisation to occur at many levels. Although this thesis will go further into how previous research has addressed the problem of multiple vibratos, this has not been extensively studied. This constitutes the main area and focus of this thesis.

The main motivation for this thesis was to understand if synchronisation would be present (and possible to analyse) in the context of people singing together. Although the main interest of this thesis is the study of vocal vibrato as a way to understand singers interaction and synchronisation, it tells us more about human beings fundamentally and more about musicians and singers specifically.

1.4 Statement of Hypothesis

The main hypotheses that is considered in this thesis is as follows:

Singers adapt their vibrato when aiming to achieve blend with another singer, often resulting in synchronisation as observed through the application of nonlinear dynamics and recurrence analysis.

1.5 Contributions

The purpose of this thesis is to highlight and establish the importance of complexity metrics in the study of vibrato regularity, and to use these metrics

to further assess vibrato synchronicity in blended singing. This contributes the research field of voice science and music signal analysis with original knowledge. This thesis tested new metrics to analyse vibrato signals and has shown the value of understanding vibrato through dynamical systems theory lenses.

Below is a list of specific contributions to the field.

1. **A review of vibrato in group singing:** This provides a detailed review of research that has addressed simultaneous vibrato tones, and the main approaches and themes that researchers have used and addressed.
2. **Application of dynamical systems approaches to understand the complexity of vibrato signals:** This thesis presents analysis that shows that vibrato can be understood in a new way through the lens of nonlinear dynamics. It presents the first body of work applying dynamical methods to vibrato signals, including those derived from musical instruments such as string and wind instruments.
3. **Entropy and recurrence metrics in the study of vibrato signals:** The results presented in this thesis unveil how these metrics, in conjunction with the traditional vibrato metrics can distinguish differences and characteristics of vibrato from two different stylistic backgrounds.
4. **Phase synchronisation metrics from recurrence applied to vibrato:** This thesis presents a new method for assessing and quantifying phase synchronisation from vibrato tones produced simultaneously, where it is possible to understand the relationship and interaction between the coupled pairs.
5. **Experimental dataset:** A new dataset was created that includes recordings from choral singers singing in solo and duet environments, whilst manipulating vibrato stimuli as different conditions.
6. **Vibrato sync:** Results of synchronisation and application of the new analysis protocol to the created datasets to show the complexities and

adaptations of vibrato when singers aim to blend together. This was studied using pre-recorded stimuli and live sync blending methodologies.

7. **Vibrato coefficient:** This thesis presents a new vibrato metric defined as the coefficient of vibrato, where it was shown that it is possible to quantify the amount of vibrato that a single singing note contains, independent from the measured rate or extent. This metric, in theory, can be applied to vibrato signals derived from any musical instrument.

1.6 Thesis Structure

This PhD thesis is split into seven chapters, starting by providing a relevant review of the literature, and continuing by presenting studies of vibrato under different scenarios and analyses. The chapters are summarised as follows:

- Chapter 2 provides a literature review of the connection between group singing and vibrato. It is based on a systematic review, performed under the Preferred Reporting Items for Systematic reviews and Meta-Analyses guidelines. It presents various methodological approaches and areas in which the field has been studied, and others that can be used as future work and potential research areas.
- Chapter 3 presents a pilot study of vocal vibrato which incorporates tools and mathematical methods from Nonlinear Time Series Analysis. This analysis provides evidence of the potential of these tools for further analysis, unveiling information regarding complexity and regularity of vibrato. The chapter ends by detailing how future work can incorporate these metrics.
- Chapter 4 expands on the work presented in the previous chapter. The aforementioned nonlinear metrics used in the application of vibrato in the singing voice are used to characterise and understand differences between world-class singers from Jazz and Opera backgrounds. The

chapter brings insights into how vibrato plays a crucial role in the stylistic differences between singers and genres, as well as the diversity of complexity that individual notes can contain.

- Chapter 5 presents the main experiment in this thesis. The experiments address how choral singers adapt their vibrato parameters under different singing conditions of Solo singing and Blend singing—singing along to a pre-recorded stimulus. It provides a mathematical way to analyse vibrato synchronisation in duet blended singing, and it provides further questions and potential research implications for future work.
- Chapter 6 expands on the work presented in the previous chapter. It explores, using a similar methodology, vibrato adaptations and synchronisation but using a two-way, live singing, blending protocol. It provides results and individual cases from scenarios that can arise during blended duet singing.
- Chapter 7 presents the conclusions derived from this thesis. It addresses and discusses the hypothesis and how it is supported by the findings in this research. It additionally brings further insights into potential future work and the implications for group singing practices.

1.7 Statement of Ethics

The research for Chapters 3 and 4 were approved by the Physical Science Ethics Committee of the University of York, with reference number Acosta20221114. And the experiments undertaken for the research in Chapters 5 & 6 and the corresponding data management were approved by the Physical Sciences Ethics Committee & Data Protection of the University of York, with reference number Martinez20240402. Appendix C shows the Information Sheet and Consent Form materials that were provided to the singers prior to participation in the study.

Chapter 2

A Literature Review of Vibrato in Group Singing

In order to better understand the concepts and research of this thesis, this chapter presents a relevant literature review of vocal vibrato in choral singing. The review highlights the main methodological approaches and research themes that emerge from the study of simultaneously produced vibrato tones. In this chapter potential ideas and possible research gaps are concluded that could further advance the field in this area. The following research was published in Acosta Martinez and Daffern (2025) [6].

2.1 Introduction

2.1.1 Vibrato in the The Singing Voice

The singing voice is a complex instrument that can create a wide variety of sounds encompassing many musical styles and conveying complex emotional expressions [7]–[10]. Among these, arguably one of the most controversial features is the phenomenon of vibrato—an oscillation of the fundamental frequency in the singing voice that has been studied from perspectives of physiological aspects, acoustic factors, and psychoacoustic attributes. In 1931, Seashore sought to define vibrato from a psychoacoustic perspective,

identifying “a periodic pulsation, generally involving pitch, intensity, and timbre” adding “which produces a pleasing flexibility, mellowness and richness of tone” [1], and also “an interesting, warming vocal quality, full of life and sparkle, giving a deviation from a perfectly steady tone which has emotional and artistic value” [11]. In this reference, it was also suggested that the frequency modulation is “usually accompanied by synchronous intensity and timbre oscillations which play a secondary role” [1].

A review by Sundberg in 1994 [2] provided a comprehensive study of the four acoustical attributes of vibrato—*extent*, *rate*, *regularity*, and *waveform*—and defined it as “a fundamental frequency undulation at a rate of 5 to 7 undulations per second and an extent of about ± 1 semitone.” Additional studies have reported parameter values within this range, although there is also evidence of slower rates of 4 Hz (starting to be perceived more as a *wobble*) and faster rates of 8 Hz (as found in opera singers [12]). Regarding the physiological aspects, the exact mechanisms of vibrato production are still debated. Research has shown that there exists laryngeally mediated vibrato, characterised by the dynamics of the cricothyrioid muscle; and abdominally mediated vibrato, affected by the amount of subglottal pressure [13]. A more extensive description by Hirano et al [14] provides four major aspects that play roles in the physiology of vibrato:

1. Neural mechanisms in the Central Nervous System;
2. Muscular activities of the voice organ;
3. Movements of the voice organ;
4. Aerodynamics.

Further understanding from the mechanisms underlying vibrato production may shed light on this category of adjustment. A study by Titze in 1984 [15] investigated the effects of vibrato stimulation at a physiological level based on a neural oscillator. This was done by superimposing an artificial, mechanical oscillation to the laryngeal muscles of a trained tenor, the hypothesis being that if the stimulus was inert, it would result simply

in an aggregate of the oscillations, but if the stimulus affected the neural oscillator, “some form of phase locking or entrainment should be expected.” A model of vocal vibrato, with the main components of the system being a neural oscillator and a mechanical oscillator (in this case, the stimulus) was presented and results showed that vibrato could be altered by the stimulus. This finding provide relevant, fundamental insights regarding the neural adjusting mechanisms of vibrato.

Although not directly related to the main hypothesis of this thesis, it can be noted that vibrato can be present in other instruments, such as string and wind instruments. Since the early research on vibrato in 1931, it was described: “At the present time it is used on sustained tones in nearly all string instruments. [...] Occasionally it is also used with wind instruments” [1].

2.1.2 Vibrato in Group Singing and Choral Singing

Currently voice research has widely studied the production and perception of vibrato through measures of acoustic parameters, loudness, mathematical models, synthesis, and listening tests [16]–[35]. On the other hand, numerous pedagogical articles, mostly written by experienced choral singers or directors, discuss the importance of vibrato, sharing anecdotal insights about its control (often advocating reduction) in choral blend and vocal health [36]–[41]. For instance, in the article *Good Vibrations: Vibrato, Science, and the Choral Singer* [38], the author described that:

Experts offer a wide range of view points regarding vibrato in choral singing, from those who believe in a free, soloistic vibrato to those who advocate a minimal vibrato or even straight tone. [...] Given the polarity of these opinions, with healthy vocal technique and sound quality in the balance, it behooves choral conductors to learn as much as possible about the phenomenon of vibrato.

The challenges associated with researching vibrato are multiplied singer-

fold in the context of choral singing. This is caused by the fact that studies need to observe not one but multiple vibrato oscillations at once, and as Neubauer put it: “Coupling several oscillators, even just two, can generate highly complex temporal patterns of motion for the overall system, even more complex than the individual oscillators” [9]. Furthermore, empirical research on choir acoustics faces numerous practical challenges, including for instance, how to isolate the individual vibrato contours of participants from the combined sound of the choir for analysis. New technological developments have shown the ample potential of tools and analyses needed to address some of these problems [42]–[47].

Although certainly not abundant, existing literature related to vibrato in choral singing unveils scientific findings [48]–[52]—which are elaborated throughout this chapter—that could support singing and choral pedagogy and performance practices. Examining empirical research to identify findings and approaches to inform our understanding of vibrato and choral singing could highlight areas for further enquiry to advance, and hopefully encourage, best practice in choral performance, pedagogy, and research.

This literature review aims to identify the main questions posed in the study of the relationship of vibrato and choral singing, and how they have been addressed to date through empirical investigations. The objective was to assess the role of the production and perception of vibrato in choral singing in the published scientific literature. The research questions were: Are there common findings associated with vibrato in the current choral literature? Is existing empirical evidence sufficient to support common recommendations regarding vibrato as contributing to blend and music style in choral practice?

2.2 Methodological Approaches to Vibrato in Group Singing

In the study of vocal vibrato and its role in the singing voice, three main categories can be identified: Studies of vibrato production, studies of vibrato

perception, and studies combining the two.

Researchers have utilised a number of common tools and methods in order to study the effects and behaviour of vibrato in choral singing. Four main approaches have been identified to the methodologies: Synthesis, stereo / binaural recordings, multi-track recordings, and contact-microphones / electro Laryngograph (Lx) or electroglottograph recordings. Figure 2.1 illustrates the proportion of the methodology approaches from a set of studies included in the Appendix A, Table A.1.

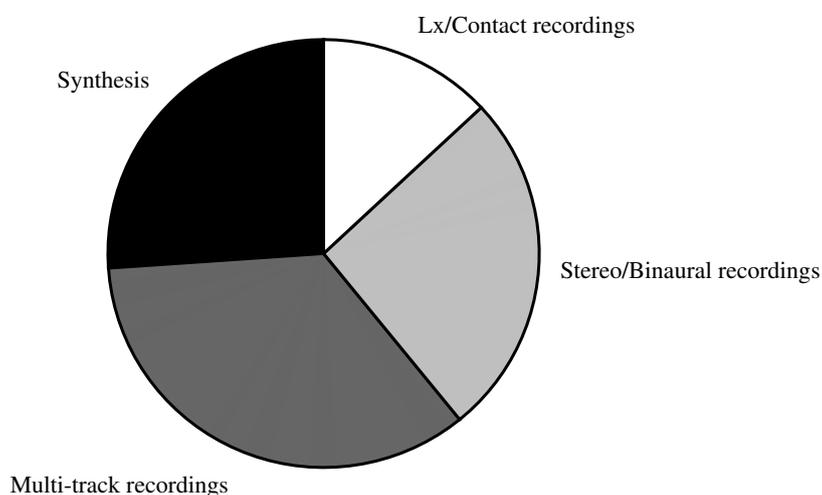


Figure 2.1: Proportions of the methodology approaches of studies addressing vibrato in group singing context.

Vocal synthesis provides a controlled and isolated environment from which to study singing voice characteristics. It is used both directly studying synthesised signals with vibrato, and as a stimulus in singing voice experiments. Stereo and binaural recordings have also been used and can be useful in studies which either deal with few subjects or are interested in analysis where isolated signals are unessential (for example perceptual testing). Multi-track recording was the most used method in the mentioned studies (Figure 2.1). Here the definition of multi-track recordings is used from studio techniques,

which means to record individual signals either *independently* (using pre-recorded stimulus as a “backing track”) or *simultaneously* (as in live performance of multiple singers with some method to capture isolated voices with minimal bleed from other singers). This method is useful for acquiring individual signals (to some extent in the case of simultaneity) from ensembles with large numbers of participants, and also for pre-recorded signals that are later used as stimuli.

Another approach to capturing isolated signals from individuals as they sing together is to use contact microphones and Lx recordings. Both rely not on air vibrations but direct-contact measurements. Lx a noninvasive method that measures the vibratory cycles of the vocal folds during phonation via electrodes placed in the front of the neck. It measures the variation of impedance from the oscillations of the vocal folds. Up to now, Lx was utilised in only one of the studies on vibrato and choral singing [53]. Research on choir acoustics shows, however, this tool offers a practical solution to isolate and study directly the individual contributions of each voice within ensembles of high number of participants [46], [54].

The above methods are not employed exclusively, and often studies used a combination of them in a single experiment. The study by Ternström and Sundberg [55], for instance, combines the four methods in a set of experiments that aimed to analyse the difficulty of intonation using different vowels and different harmonic intervals, with the hypothesis that vibrato facilitates intonation due to reduction of beats and roughness. For this they first used (1) a 10-second long synthetic stimulus of sung vowels under different conditions (e.g. vibrato), then (2) sung stimulus recorded with Sennheiser binaural microphones, (3) a dummy head in a Revox stereo tape recorder, and (4) contact microphones “fastened to the throat below the larynx” for the singing participants.

The study by Sacerdote [56] did not explicitly state which method was used to acquire the choral signals that are described in his article’s last experiment: A female choir singing in unison and a monodic choir of “around 15 men”. However, given the graphs and descriptions presented (as well as

the date of the study) it's inferred the author recorded the signals from the choir singers using a single microphone or a stereo setting.

2.3 Thematic Categories of Vibrato in the Singing Voice

The main aims/objectives each study sought to accomplish (Figure 2.2) were categorised to consider similarities across the research. In total, five thematic categories were found as described below. This categorisation provides a perspective of the landscape of the research overall, but naturally some studies are extensive and their objectives may overlap between categories. (Full data of studies in Appendix A, Table A.2.)

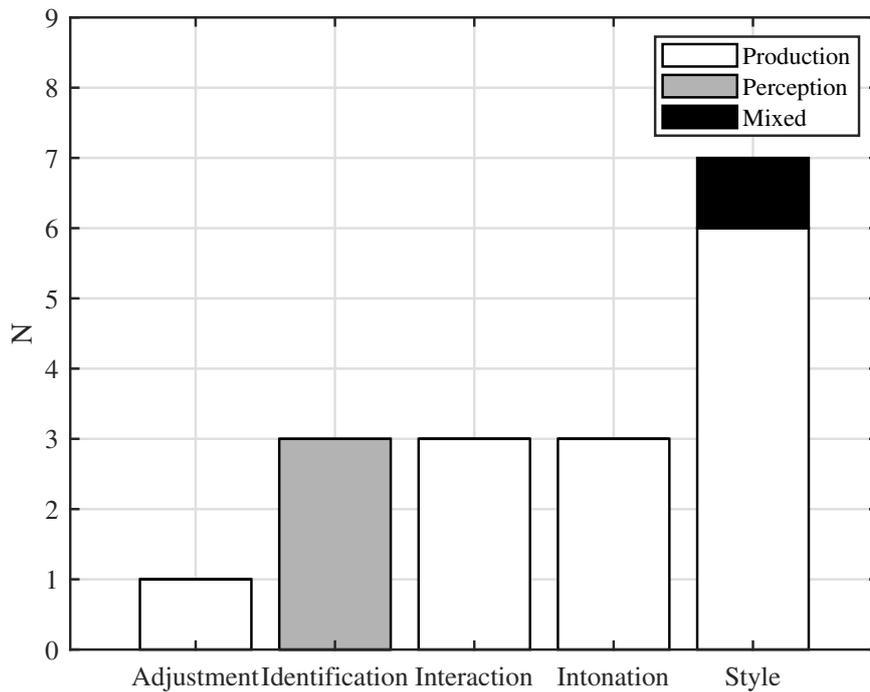


Figure 2.2: Thematic categories of studies addressing vibrato in choral singing contexts.

2.3.1 Adjustment: Intrapersonal Adaptations and Vibrato Matching

One of the included records studied adjustment and adaptation of vibrato [57]. This category deals not with the complex interactions of multiple voices with each other, but with the level at which individual voices can intrinsically alter and match their properties to another or other voices. This was studied using an external, pre-recorded stimulus that a singer attempts to match or adapt to. Dromey *et al.* studied the acoustic characteristics of singers who attempted to match vibrato stimuli with different rates (slow and fast). Results indicated that trained singers were able to adjust their vibrato to the external stimuli, and the study raised the question of whether this is also possible in the long term and without external stimuli.

2.3.2 Identification: Auditory Isolation of Voices

The ability to distinguish and identify individual voices within a sound (Auditory Segregation) was the main objective of the three studies in this category. Erickson [58] performed listening tests to find the effect of vibrato on the perception of two simultaneous singers, the hypothesis being that voice pairs with vibrato would be easier to segregate, particularly at higher pitches. This prediction was derived from the previous studies by McAdams and Culling [59], [60] showing that frequency modulation [61] increases *prominence* or source identification within a complex sound, regardless of coherence and “differences in the pattern of that modulation”.

Participants of Erickson’s study were both listeners with experience in music and singing, and listeners with no experience in any activity related to music. The task consisted of hearing two sets of synthesised singing stimuli with different characteristics, one with a pair of voices singing simultaneously with vibratos 180° out of phase, another without vibrato. The study provided no comprehensive conclusion as authors reported that vibrato “significantly improved listener’s perceived ability to hear two simultaneous voices” at some combination of pitches and formants, but no improvement at others. The

discrepancy between this and the findings from McAdams and Culling could be related to additional timing onset conditions in Erickson's study.

2.3.3 Interaction: Interpersonal Relationship of Voices

Three studies have analysed the inter-participant interaction of voices in a group singing situation [53], [62], [63]. In this category, the aim transcends the assessment of individual properties of the voice. Rather, these studies focus on the degree to which singers can adapt or modify their voices based on their interactions with each other as they sing together. The three articles introduce interesting concepts in group singing such as vibrato synchronisation, generally referring to a form of:

- *coherence*, “the measure of the state in which two signals maintain a fixed phase relationship with each other”
- *phase-locking*, “the mutual interaction among oscillators by which the phase difference between any two oscillators is fixed” [5];

which form the basis for studying vibrato behaviour in multiple voices. The study by Sacerdote, in 1957 [56], hypothesised about the possibility of vibrato synchronisation in choirs, but experimental evidence and measurements would only come in future studies.

In the study by Duncan *et al.* [62], the authors used CD stereo recordings from internationally renowned singers singing in pairs. The chosen piece of the recordings allowed the authors to extract the vibrato oscillations from musical notes sung under different conditions (solo singing and simultaneous singing) and also different pitches and durations. Results indicated that it is possible that singers are able to adapt their vibrato in order to achieve synchronisation. This gives rise to patterns of interaction such as in-phase synchronisation, a positive correlation; or anti-phase synchronisation, a negative correlation. This last case means the extent of vibrato would have inverted amplitude or be 180° out of phase (as in the listening tests of Erickson and Gaskill [58]).

Daffern in 2017 [53] reported similar results in an ecological study of a Soprano, Alto, Tenor, Bass (SATB) quartet ensemble through 10 weeks of training. With an ensemble of four singers, this study utilised multi-track and Lx recordings to isolate individual voices. The results provided evidence of both the employment of vibrato as a collective feature potentially contributing to choral blend and its adaptation across multiple sessions as the quartet worked towards a recital performance.

2.3.4 Intonation: Fundamental Frequency Trajectories and Accuracy

Intonation has long been one of the main features of interest in choral acoustics [48], [64], [65]. This doesn't come as a surprise given that in choral singing and *a capella* singing pitch accuracy is imperative and complicated as “vocal tuning is completely variable to a fine degree; singers are non fretted instruments” [51]. The studies included in this category focused on understanding the role of vibrato on: The intonation precision of choral singers [55], the intonation of a synthesised choral system controlled by hand movement [66], and a report of the absence of vibrato in analyses of intonation trajectories from an unaccompanied SATB quartet experiment [67].

The study by Ternström [55] mentioned above which used voice synthesis and real voice recordings, analysed the intonation precision of choral singers to a stimulus. The study found that lack of vibrato facilitated the mentioned f_o accuracy.

The conference paper by Le Beux *et al.* [66] presents a synthesis system for SATB choral singing, where the vocal intonation of each singer could be controlled by four game joysticks. The system allowed users to perform a musical piece simultaneously by varying the micro-intonation (5 to 10 % of a semitone) of their vocal part with the joystick. In this way, a “waving motion” wrist would create oscillatory patterns that were associated with vibrato. The study reports recordings from the authors' performances in the system and a set of graphics illustrating the f_o waveforms. They described

the “bass player had on all trials a vibrato in the same order of magnitude as real singers, around half a semi-tone.” And the “other players [had] a vibrato of approximately a quarter of semi-tone.”

Finally, in the study of intonation trajectories performed by Dai & Dixon [67], the authors actively reported a lack of vibrato in a singing model of unaccompanied soprano, alto, tenor, bass quartet singing. The experiments consisted of participants singing their parts in different conditions and recording/modelling the intonation trajectories, in which there was no vibrato—at least not even and prominent enough to be modelled. The authors suggest that choral singers avoid vibrato perhaps in their attempt to contribute to blend. The study discusses that “although vibrato is a feature of many singing pitch trajectories, we did not explicitly model it in this work. The use of vibrato is less marked in unaccompanied ensemble singing where the voice does not need to be projected over instrumental parts, and the stylistic goal is for the voices to blend rather than stand out.”

2.3.5 Style: Solo Singing, Choral Singing, and Choral Sound

Seven studies [56], [68]–[73] focused on stylistic differences of the voice between music genres (e.g. classical, opera, liturgical) and the characteristics that create the *choral sound* or *choral effect*. Sacerdote, in 1957 [56], laid the groundwork in choral singing science exposing the interest in the question of what defines the choral sound. Researchers have been interested in understanding the acoustic properties of *solo singing* versus *choral singing*, which some of the studies here define as *singing modes*, stating that experienced singers are able to use a different type of voice when singing in a solo or a choral context.

One of the earlier studies to measure the phenomenon of singing modes was Goodwin in an experiment [69] that analysed the spectral characteristics of sustained vowels of sopranos first singing in their solo voice, and next attempting to blend to a pre-recorded unison ensemble sound of soprano

voices, to yield an alternative choral mode. The results showed the singers employed vibrato reduction in the choral condition alongside other blend techniques such as vowel modification and dynamic level adjustment. In a later study, Rossing *et al.* [70] extended this research by analysing Sound Pressure Level (SPL), Long-Term Average Spectrum, and vibrato characteristics in four sopranos and one mezzo-soprano singing in solo and choir modes. Regarding vibrato production differences, the study concluded that “the extent of vibrato appears to be somewhat greater in solo singing.”

On the other hand, the study by Ternström [68] addressed the subject of choral sound using vocal synthesis of voices signals with different types of frequency modulation in f_o . His objective was to replicate the subtle nuances that multiple singers have by using random-noise generators and band-pass filters. The cut-off frequencies of the filters were different—in the ballpark of the commonly accepted vibrato band [2]—so that each one added different components to the sound to emulate the choral effect once summed together. The rationale was that signals that are identical to each other create no “sensation of ensemble” as they would be perceived simply as a single sound with doubled amplitude. This is a key element in choral sound according to the author, which described, “The slight tremor in F [f_o] generally varies in character from singer to singer, especially with regard to the vibrato content.”

In this chapter, the main questions that arise in the study of vibrato in choral singing are investigated and highlighted, and how they have been addressed through empirical investigations. It’s been shown that (a) relatively few investigations have studied vibrato in vocal ensembles; (b) the majority, from the defined set of experiments described, focused on vibrato production rather than perception; (c) methodological approaches include Synthesis, Multi-track recordings, Stereo/Binaural recordings, and Lx / Contact recordings; and (d) the studies focused on main themes such as Adjustment, Identification, Interaction, Intonation, and Style. Furthermore, vibrato has an important impact in choral singing from both production and perception perspectives, and that there is evidence of singers adapting/synchronizing vibrato in choral conditions.

2.4 Studies Focused on Group Singing and Vibrato

Seventeen articles have been identified as investigating vibrato and singing voice in the choral context. But perhaps more important than the volume was the heterogeneity of the findings. From a meta-analysis perspective, findings from 17 articles could yield acceptable data to conduct informative statistics, but given our qualitative synthesis results, it was found that the overall metrics and method variations prevent this; particularly considering the perceptual studies, in which no agreement currently exists for studying the degree of choral blending and vibrato effects.

The number of participants in the mentioned studies had high variability, across both production and perception experiments. In the case of repeated measures, studies such as Coleman 1995 have assessed a pair of singers as satisfying the choral context. Additionally, an adequate amount of vibrato samples can be obtained by either collecting multiple takes of vibrato singing within participants, or also by having multiple singers, as the case of Mann 2014. On the other hand, the study by Duncan *et al.* also studied duo singing with combinations of four participants. This allowed for the data collection of different duo combinations to study whether vibrato behaviour could favour a specific pair of singers; an analysis that could be extended to studies with more participants. Likewise, the longitudinal study by Daffern included only four participants, where one of the aims was to analyse possible changes in the level of singers interaction over time and rehearsals. For perception studies, the study by Culling & Summerfield had three participants for an experiment, and four for a subsequent within the same article; numbers that could be considered low although the listeners were highly experienced in the area.

Potential differences in treatment of vibrato parameters across vocal ranges during choral singing has been touched upon but not thoroughly explored. In the included study by Rossing *et al.* 1987, he found that vibrato extent tended to be higher in solo singing; and in his early paper in 1986 [74], he

studied the same conditions of choral singing and solo singing in male singers, and provided no findings regarding vibrato production. Research has indicated potential differences in vibrato characteristics across voice classes such as soprano, alto, tenor, or bass. The study by Daffern [53], for instance, reported differences in vibrato production and synchronisation between voice categories as the soprano tended to trigger or “lead” the onset of vibrato in other voices. Also, descriptive differences in vibrato were observed, for instance, between the soprano (more vibrato) and the bass (less vibrato) of the quartet. This could suggest that voices can have different roles of vibrato behaviour depending on the frequency range. Seashore has stated that “No great violinist of today plays without a vibrato” [1], while no such statement was made regarding double bass playing. Anecdotally, in contemporary music, producers mostly advocate a stable or “tight” low-end sound from electric basses to provide a solid ground for the higher-pitched instruments to ornament. In short, the role of frequency range and vibrato production in choral singing is something that could prove beneficial to study particularly from a perceptual perspective.

Only five production studies report the age of participants for production studies. The age of choral singers could have research implications as additional studies have reported differences in vibrato behaviour of singers over time and choral training. A study by Howard [75] performed a longitudinal study of three female choristers during puberty concluding there was an “improvement” in vibrato, among other vocal parameters of the voice. The variation of vibrato parameters, however, was not reported. On the other hand, Berghs *et al.* [76] studied the effects of age on voice parameters from choral singers. Because of the wide range of ages the study wanted to address, rather than a longitudinal approach, the authors chose a comparison of a population of singers from a choir in Netherlands that ranged from 20 to 60 years old. They recorded and analysed the vocal parameters and divided them into age categories for statistical analyses between male and female subjects. The study infers that changes in the acoustic characteristics are associated with age as the chosen choral singers shared similarities in their training and accustomed singing practices. The standard deviation of vi-

brato amplitude across participants was “moderately” negatively correlated with age in females but not in males. However, participants, in subjective questionnaires, commented that “in their experience vibrato control is one of the main consequences of more technical skills over the years, gained by studying and singing regularly in vocal ensembles”. These studies illustrate the relevance of reporting age, training, and experience characteristics for future studies.

2.4.1 Production and Perception

Through systematic inspection, there is common agreement between the production studies to report findings using the well-known vibrato parameters rate and extent. From the 17 studies, six of them used at least one of these parameters [53], [57], [62], [72], [73], [77]. In future work, reporting vibrato extent and rate along with sample size, mean, standard deviation, mid-range, or mid-quartile range would prove useful for cross-study analyses and statistics. Other metrics were sometimes reported with similar phenomena to vibrato such as shimmer and jitter. Another agreement between production and perception studies was to report the vibrato waveform [14]. Although not explicitly reporting rate or extent values, some studies showed the trajectories over time of f_o to describe the oscillatory characteristics.

Less than 25% of the studies included addressed vibrato perception in choral singing. Of the four perception studies, three of them [58]–[60] analysed identification of characteristics from synthesised voices and vowels. In this case, the methodologies used vibrato in simultaneous singing, but the central research question was focused on segregation of tones in sound perception more generally. (Here, segregation is a term that is used to refer to the perceptual identification of multiple sources from a single sound.) The argument for the inclusion of voice identification and segregation within this review comes back to the Style category of understanding the choral sound. The three studies in this category suggest that vibrato and the coherence of the signals can help to discern that a complex sound is the sum of various voices. This could be one of the reasons why choral directors might advocate

for vibrato minimization. The concept of blend in choirs would dictate that the “auditory segregation” of the voices is in fact something undesirable in choral singing, where the real objective is to come up with a fused, cohesive sound in which individual voices cannot be identified.

The remaining study [73] that contained perceptual measures corresponds to a mixed study, where the main experiment was focused on vibrato production, and the perceptual aspect came from questionnaires filled in by the singers immediately afterwards. It is useful to emphasize the difference of these measures from those obtained from listening tests. Though both provide useful insights, results show the possibility of more listening tests to complement and consolidate the work done on production—auditory perception being of essential consideration in studying the acoustics of the voice. As a listening study of vibrato puts it, “singing is predominantly a perceptual phenomenon [...] tailored for human perception” [24]. However, perception is a more complex field that can be easier to address once the objective analysis from audio and acoustics have been applied to establish a knowledge basis.

2.4.2 Review of Common Methods

There were common methodologies used to approach vibrato in choral scenarios across studies, with frequent use of voice synthesis particularly in perceptual studies [58]–[60]. A rationale could be that synthesis facilitates fine control over the frequency modulation coherence of coupled vibratos. This allows for purely sinusoidal modulation with perfect periodicity. Although possible to some degree, this would be harder to find in real-voice simultaneous recordings and the phase-locking event might not be independent from changes in other variables, thus the preference to control and isolate coherence through synthesis. However, given that the three perception studies of coherence relied strictly on voice synthesis [58]–[60], the possible expansion to the case of real-voice choral signals can be highlighted. Additionally, interaction effects that can occur during real voice, simultaneous singing is not present in individual voice synthesis.

The two multi-tracking methods of pre-recording or live recordings have

both advantages and limitations of their own, depending on the research questions. Pre-recordings were favoured in studies as in this way it is very easy to isolate the stimulus signal from the experimental recordings of participants. One disadvantage is that it can add variables (such as auditory feedback) that can heavily influence the results. For instance, the self-to-other ratio has been shown to influence intonation and vibrato behaviour [49]. In other words, the monitoring levels of the singers influence the vocal response to the stimulus. This can also depend on whether the stimulus is presented through speakers or headphones (and their quality and frequency response); and in the case of speakers, special attention would have to be given to acoustics as well.

For live multi-tracked recording, studies successfully reported the use of close omnidirectional microphones [53], [63], although a justification for the specific choice of omnidirectional—instead of directional—microphones has not been addressed. One possible reason could be the proximity effect introduced by directional microphones, which in choral singing applications may result in increased low-end spectrum recordings, depending on each particular microphone. (Some manufacturers can often compensate the frequency response for the close proximity applications that the microphones are designed for.)

As additional studies have emphasised the importance of isolated signals [46], [78], and data analyses and mathematical methods strongly rely on signal quality and noise reduction [74], [79], the trade off between (1) a flat response but increased bleed from omnidirectional microphones and (2) a non-ideal frequency response but slightly-increased isolation from directional microphones must be considered. This can depend on the post production treatment of data. For the case of f_o analyses, it would seem plausible that the proximity effect could actually be beneficial as it increases low-frequency response of close sources (usually where the f_o of interest is found), while attenuating the low frequency response of far sources (where the f_o of others may be found) and limiting bleed to higher frequencies, if any. However, for other types of analyses, like looking at the ratio of f_o to other frequencies, the

proximity effect could be undesirable. Another possible explanation for the use of omnidirectional microphones may be plosive or distortion artefacts.

In simultaneous singing it is important to obtain isolated signals to study effectively the individual vibratos from multiple singers. However, Duncan [62] successfully extracted frequency information of individual voices from stereo recordings given that the singers were performing a piece in polyphony, with different frequency components that didn't have obstructive harmonics. This is of course restrictive in terms of potential musical material and would not be appropriate for unison voices. The use of multi-tracking techniques and also technologies such as Lx or contact microphones as demonstrated by Jers and Daffern [53], [63] specifically aim to support multi-voice signal analysis and remove this limitation, successfully recording polyphonic and unison singing in real time.

2.4.3 Vibrato and Choral Singing

Five categories of aims and objectives were identified. It is interesting to note that some studies reported objectives while others reported aims or research questions. The category addressed by most studies was the theme called Style, which aimed to study the choral sound of individual singers compared to their solo singing and other genres such as opera or early music. Only one of these included perception analyses [73], which is surprising considering it could be argued that the essence of this category is a perceptual standpoint: What individual factors contribute to the choral sound?

The category of Style illustrates how different types of choral singing can have different aesthetic ideals—perhaps tied to the music genre or regional context. In the study by Coleman [71], the author writes: “A Broadway chorus has a different sound quality from an excellent church choir; a barbershop quartet produces sound differently than does a quartet of opera singers.” And another study [67] from the same category describes that “choral style favours minimal vibrato, and barbershop style generally forbids vibrato.” This latter comparison is entangled with the Intonation category, where in barbershop singing “great care is devoted to intonation” [80] and harmonization to avoid

beats and create the *lock and ring* effect. Likewise, the effects of vibrato in the consonance and/or salience of harmonies, especially close harmonies, is a topic that could bring further insights to choral research.

Studies in the Adjustment category demonstrate that singers have the capability to change their vibrato behaviour. The degree to which this is a conscious or unconscious behaviour is still not clear. However, as mentioned above, studies have shown that age, vocal training, singers' experience singing together, or music genre (style) can influence vibrato behaviour. This is another point that emphasises the continuum between the theme categories presented here.

Studies that focus on the adaptations of single voices, whilst not conducting an ecological study of choral singing per se, provide foundational evidence around the possibilities of vibrato synchronisation in choral singing interaction: If it's possible for one singer to adjust their vibrato in real time to a stimulus, a pair of singers simultaneously adjusting could behave as a coupled oscillator sensitive to changes in both vibrations. It could be questioned, however, if research could transpose the effects obtained from pre-recorded material to those from live group singing setups due to the higher complexity needed to adapt in two-way or multiple-way interactions—"Complexity is not just complicated stuff but a unique quality that emerges from the relationship or interaction among elements" [5]. Further research could seek to answer this question by comparing the level of adaptation with a static stimulus versus that with another performer singing in real time.

This is the main essence of the interaction category: The interpersonal relationship of singing voices. This category is specially relevant given that it directly addresses our research question of how voices relate to each other in group singing in terms of vibrato. In this regard, three studies provided compelling evidence of forms of oscillatory synchrony [53], [62], [63]. The common interest of vibrato synchronisation between these studies is not surprising given that it could be seen as a proxy for the *sense of ensemble* between performers. Singers being in high synchrony probably suggest they are aware and receptive to each other, which may in turn have positive im-

plications on the output sound. This is, however, still an open question. The studies by Goodwin and Daffern [53], [69] hypothesised that vibrato synchrony could improve blend, a perceptual concept and key objective in choral sound. Both production studies provided evidence of vibrato synchronisation within singers to that end, although the perceptual significance of this hypothesis is yet to be confirmed through perceptual tests of choral blend.

Regarding choral singing, a major body of work that considers the role of vibrato is a series of pedagogical articles by *The Choral Journal* [37]–[40]. The early article *Factors Related to Choral Blend* describes blend in the choral sound as being an essential achievement that in turn is affected by vocal tone (a psychoacoustic term) and vibrato. In a survey to choral directors about the factors associated with blend, 40% graded vibrato “very important”, 38% “important”, and 20% “a factor but not of primary importance” [37]. This agrees with findings in this review that deal with the choral sound and choral singing modes. Studies analysing vibrato production showed that singers have different characteristics in choral singing modes and suggest that research could trace effects of vibrato behaviour as a factor also associated with blend.

The more recent study by Hinkley [77] studied the vibrato response to stimuli recorded from a pair of professional choral singers who served as *vocal models*: a baritone with experience in audio engineering and a soprano with experience as choral director. The two vocal models recorded the stimuli with “legato articulation at approximately 60 bpm on the neutral syllable ‘tah’” with two vibrato conditions—vibrato and minimal vibrato—and also a set of sung melodies to consider a more choral-voice scenario. The stimuli were presented to 76 high school and undergraduate choral singers in a matching-based experiment, female students responding to the female model, and male students to the male model (a method based on previous studies referenced in Hinkley [77]). This experiment expands on the findings by Dromey [57] in the Adjustment category, which questioned if vibrato adjustments could stand after simulation ceased: In Hinkley’s experiment, participants heard

the voice model example, then sung themselves right afterwards. Results showed that choral models did have an effect in the vocal vibrato of the students, finding differences in rate and extent.

In the introductory chapter of *Fascinating and Intriguing Vibrato* by Philippe Dejonckere from the book *Vibrato* [14], it is stated: “It has been claimed that vocalists singing a duet can synchronise their vibrato rates, but there is no experimental evidence for this” [81].

Vibrato oscillations seem a prominent research avenue in group singing. Some of the potential strengthening aspects are the need of more data and extensive/robust analyses; and regarding the experimental methods, further exploration of how to best acquire clean signals in choral context and how to best isolate or control the many variables that take place in group singing to understand the individual relationships of each one. As such, in the next subsections it is presented a new perspective for data analysis, and an additional experimental method in choral singing research.

2.4.4 Analysis Techniques: Complex Vibrato

In his vibrato review [2], Sundberg acknowledges that research on the extent and rate of vibrato has been extensive while not much attention has been given to waveform and regularity. This reliance on extent and rate, which in the case of a single note, involving averaging using either the arithmetic or geometrical mean, has proven insightful in terms of understanding the parts that conform vibrato, but also limited in terms of the variability and relationship understanding.

In 1987, Rothman and Arroyo [82] described that “Our ultimate goal in studying vibrato is to understand more fully the three parameters that comprise it [here referring to extent, rate, and amplitude modulation], their interrelationship, and their relationship to perception”, and concluded that, “we are immensely engrossed in scientific exploration of large amounts of fascinating, highly variable data.” In this vein, studies have successfully utilised the standard deviation. For instance, in their investigations of *good* versus

poor quality vibrato, Diaz & Rothman in 2003 [17] measured the extent and rate to calculate the time variability and standard deviation to estimate the vibrato regularity.

As future work, perhaps further insights could be obtained by thinking of vibrato oscillations not from an acoustical but more of a biological perspective, from where it's really derived. Research on biological rhythms has long been baffled by the complexity of “change, flow, and rhythm, mostly in things that are alive” [83]. The book *Sync: The Emerging Science of Spontaneous Order* [84] describes how scientists have used modelling of nonlinear differential equations to study rich behaviour and synchronisation, where the nonlinear properties of a system denote that the whole is different than the sum of the parts. “In science, we do reductionism”, mentioned the author Strogatz about Sync [84], “and that has been phenomenally successful for every branch of science. But the great frontier in science today is what happens when you try to go back, to put the parts together to understand the whole. That's the field of complex systems.” This turns particularly interesting in the context of choral singing.

Studies in the category of interaction between voices report the vibrato contours in the time domain—given that they are correlated in time. An alternative way to observe the evolution of coupled f_o could be found in dynamical systems by observing the qualitative interaction of vibrato in a phase space, illustrating the overall pattern of the f_o oscillations. Dynamics and nonlinearity are by no means novel concepts in voice science, where studies have investigated perturbations *within* the voice [85]–[91], but are not yet explored in the context of vibrato and/or group singing.

The coherence or incoherence of coupled vibratos can be studied using nonlinear dynamics to gain insight about potential adaptative interactions. Differences in phases such as the 180° vibrato stimuli in Erickson [58] or the anti-phase synchronisation in Duncan [62] would resemble a periodic limit cycle, recalling that these signals were purely sinusoidal. Far much harder to characterise—if possible—would be the complex interactions that may arise in real voice recordings.

2.4.5 Virtual Choirs

Other avenues of future research into vibrato in group singing could be found from emerging methods for evaluating other methods of choral singing, such as those explored in Daffern *et al.* 2019 which include the implementation of virtual reality technologies to improve the ecology of the experiments [92].

2.4.6 Implications of Choral Vibrato Research

This last mentioned study is a continuum to our final perspective which brings us to the fundamental aspects of choral singing and vibrato—the experiences of choral singers and the well-being benefits that can accompany choral singing, and how we can further understand them. According to Sundberg, choral singing is “probably the most widespread type of singing” [93], and studies have shown that during performances singers can experience synchronisation of cardiac and respiratory patterns [47]. (Also brain rhythms synchronisation in the case of duet music improvisation [94].)

Synchronisation of bio-measures might support qualitative evidence of well-being increases that have been reported in choral singing [95], which would agree with a biomathematical perspective that considers the human body as a collection of oscillations—an orchestra of oscillations—that in persistent cases and events tends to intra- or inter-synchronise, be it in phase or out of phase. “For reasons we don’t yet understand, the tendency to synchronise is one of the most pervasive drives in the universe”, describes Strogatz [84], “and, for some reason, it often gives us pleasure. We like to dance together, sing in a choir, play in a band.”

The limitless scenarios that can arise from vibrato oscillations in the voice might be yet another view from where to study synchronisation and well being in choral singing, a view that may in turn spotlight distinct sounds to auditory perception—*the sound of cycles in sync* [84].

2.5 Conclusion

In this chapter, it has been shown that there are relatively few empirical studies of vibrato production and perception that specifically address vibrato behaviour during group singing tasks. However, this body of work sits within a vast amount of scholarly and performance-based literature around vibrato and choral singing and is highly relevant in complementing and contributing to the broader scope of the work on this topic.

The common approaches that were identified highlight the challenges of data capture and analysis of multi-voice audio signals and highlight the value of the specific approach of simultaneous multi-tracking and/or Lx/EKG capture of relative vocal fold contact area. High variability in methodologies, research questions and limited data-sets reduce the value of any meta-analysis or synergy of findings without future research.

In summary, with the current body of literature, main subareas and methods have shown to be fruitful and in which research can continue advancing have been recognised. The relevance of this work to singing practice is well evidenced in light of its prominence as a vocal feature in choirs and further questions that will collectively shed light on the complex relationships of this oscillatory pattern called vibrato may prove valuable beyond singing practice to understand joint action behaviours and broader concepts of choral blend and singing together.

Chapter 3

Nonlinear Time Series Analysis of Vibrato

This thesis aims to study the different kinds of vibrato dynamics that can be produced during group singing and choral singing. To this end, the first step was to explore which metrics are well-suited for the analysis of vibrato wave signals. In an exploratory study (Appendix B), it was found that Dynamical System Theory was a promising field that could provide new metrics and insights into vibrato analysis.

Dynamical systems study the evolution and change in variable states in mathematical models, usually described through sets of differential equations. Dynamics are more elegantly described as:

[...] The subject that deals with change, with systems that evolve in time. Whether the system in question settles down to equilibrium, keeps repeating in cycles, or does something more complicated, it is dynamics that we use to analyze the behavior [96].

The rationale behind the work presented in Appendix B was that if vibrato waves could be derived or synthesised directly from mathematical nonlinear systems, then vibrato signals from real-voice recordings could be analysed using nonlinear dynamics qualitative and quantitative methods (with the caveat of noise effects). One of the benefits from this approach is that a

number of well-known mathematical models with low dimensional dynamics has already been extensively studied previously. As such, the characteristics of the signals derived from such systems are known and can be compared to other vibrato signals. This chapter extracts a number of different metrics—from classically utilised metrics in vibrato analysis to dynamical systems metrics—from a small dataset of notes.

To test the methods used throughout this thesis, a pilot study of entropy and dynamics in vibrato analysis was conducted. Part of this chapter and research was published in Acosta Martinez and Daffern (2023) [97]. The contribution of this work is relevant given that it is one of the first reports testing the application of Nonlinear Time Series Analysis to vibrato signals. For this, a small pool of vibrato tones was used, and the chapter also provides a graphical, qualitative representation of vibrato waveform, phase spaces, and Recurrence Quantification Analysis. The concept of recurrence applied to singing voice signals is compelling as it illustrates the process of reaching a state; deviating from that state—to exploit the joy of motion and add depth to the trajectory; and recurring to the state—but altered by the trajectory and under slightly different conditions, depending on complexity and dimensionality.

*Turn your back on everything
Banging heads with prophecy
Lavish in uncertainty
[...]*

*What's the confliction between heart and mind
How to bridge the void, fuse the two in kinship
To understand your prospects and vistas
And yet to see the love here too
[...]*

*Lavish in uncertainty
Turn your back on everything
—Tesseract: Altered State*

3.1 Introduction

Research on vocal vibrato suggests that regularity and periodicity can affect its auditory perception and pleasantness. In one of the earliest empirical studies of music acoustics, Seashore defined vibrato as “a periodic pulsation, generally involving pitch, intensity, and timbre” [1], the word periodic referring to the fact that its waveform can resemble that of a perfectly sinusoidal wave. In terms of vibrato production, parameters are associated with singing aesthetics—a rate of between 5 Hz to 8 Hz and extent of up to 200 cents peak to peak [98]—and research suggests that “samples with better quality are the most periodic ones, [...] more constant and regular over time” [17].

More recently, vibrato regularity has been addressed by jitter and modulation shimmer analyses [99], which are a percentage measure of variability in rate and extent, respectively [3]—the concept of shimmer is applied to vibrato modulation throughout this thesis and it does not refer to intensity. In other words, jitter considers all the measured rate values from a single note, and compares how different they are between them. Likewise for shimmer, the measured values from extent are averaged and compared to obtain a percentage of how different those values are. However, these features may not thoroughly capture and illustrate the intricacies of the time varying nature of vibrato (or the evolution in respect to time). Nonlinear time series analyses associated with complexity and periodicity are tools that have been used in voice science, and music rhythm and synchrony [9], [90], [100]–[107] that could bring further insights on vibrato stability and its production and perception.

Utilising a qualitative approach, phase spaces and recurrence plots are tools that illustrate the dynamics of time series. Phase space reconstruction of time series can be achieved using a time delayed signal, whereby the observation of the geometric pattern is considered the first step in studying the dynamical behaviour [96]. In the phase space, the cycles of vibrato would create orbits bounded between the extent of the undulations, with the density of the orbits affected by its regularity.

Derived from phase spaces, recurrence plots show how the reconstructed trajectory repeats itself. As such, recurrence plots can reveal features of time series and contain structures such as homogeneity, disruptions (white bands), periodic/quasi-periodic patterns, single isolated points, or diagonal lines parallel to the line of identity [108]. These different structures provide information about the dynamics of a signal. For instance, disrupted structures such as *white areas* or *bands* can often be associated with transitions or nonstationarities in a time series.

Recurrence quantification analysis, as the name suggests, quantifies the information provided by recurrence plots using measures like determinism or laminarity. These and other recurrence metrics can collectively (along with other methods such as surrogate techniques) suggest periodicity/regularity of time series [109]. Another metric from the family of nonlinear time series analysis, sample entropy, has also been found to be a robust method to quantify complexity in even short and non-stationary time series [110] and shows “promise as a metric of perceptually relevant voice instabilities” [111].

This preliminary, descriptive study investigates vibrato tones from published musical pieces to question whether (a) phase spaces and recurrence plots are valuable for observing patterns of vibrato behaviour in solo singing, and (b) visual differences or patterns would be reflected in nonlinear features such as sample entropy and recurrence quantification analysis and / or other measures of vibrato regularity.

3.2 Methodology

Ethical approval for the research was obtained from the Physical Sciences Ethics Committee of the University of York (reference N^o Acosta20221114). Sixty-eight notes were selected from three published pieces—La Serenata, 'O Sole Mio, Nessun Dorma—recorded by the world-famous operatic tenor Luciano Pavarotti for vibrato analysis. Voice isolation was performed from wav files of the recordings using iZotope Rx 9, and vibrato notes were manually selected and exported as pitch listings using Praat software [112]. Figure 3.1

shows a flow diagram of the process.

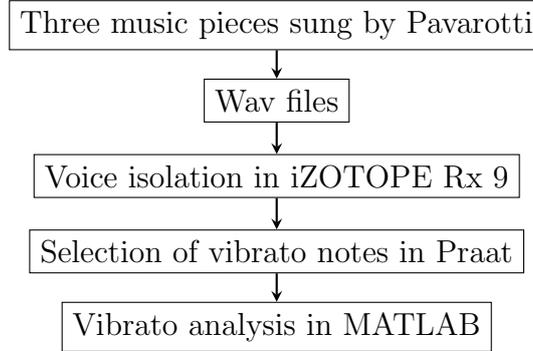


Figure 3.1: Flowchart of data extraction and analysis.

For this study, a note containing vibrato was defined as phonation on a single vowel and single note as defined by the score, with no less than five visible oscillations to allow sufficient length for the analyses. The f_o time series obtained from Praat had a timestep of 0.01 s ($f_s = 100$ Hz), and each note was imported to MATLAB and smoothed using a low-pass filter with a cutoff frequency of 15 Hz.

A MATLAB script was created to calculate from each note the mean fundamental frequency (Hz), length (s), number of cycles, rate (Hz), extent (cents), jitter, shimmer, sample entropy, and determinism. The values of jitter and shimmer of vibrato were calculated as described in Horii (1989) [3]. These metrics were calculated as:

$$\text{Jitter} = \frac{\sum |\text{period}_{i+1} - \text{period}_i| / (N - 2)}{\text{average of period}} \times 100 \quad (3.1)$$

$$\text{Modulation Shimmer} = \frac{\sum |\text{extent}_{i+1} - \text{extent}_i| / (N - 1)}{\text{average of extent}} \times 100 \quad (3.2)$$

where N is the number of modulation cycles (≥ 2) in the time series, and $i = 1 : N - 1$. It is important to note that this equation defines the concept and calculation of shimmer as used throughout this thesis. This clarification is needed given that in voice science, shimmer is not the extent variation

of frequency vibrato but amplitude perturbation of acoustic intensity. From Horii (1989) [3]: “It should be noted that although the terms jitter and shimmer are often used to describe cycle-to-cycle variations of fundamental periods and amplitudes, respectively, these terms refer to cycle-to-cycle variations of modulation periods and modulation extent in this paper. Conceptual parallels here should be apparent”.

Sample entropy (SampEn), defined as “the negative natural logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, where self-matches are not included in calculating the probability” was calculated using the SampEn toolbox [113]. For sample entropy calculation, the tolerance was set to $r = 0.1$. As entropy calculation results are heavily dependent on parameter settings, this was mitigated by defining a constant value for the parameters across the notes, based on recommended values [110], [113].

The CRP toolbox [108] was used for the recurrence analysis. For the phase space reconstruction of vibrato notes, the embedding dimension was $m = 3$ and embedding delay $\tau = 4$. And for recurrence analysis, the adaptive threshold ϵ was set to ensure that the recurrence rate $RR \lesssim 0.05$ [108].

3.3 Results

The mean fundamental frequency across notes from the three pieces varied from 175 Hz to 501 Hz, which roughly correspond to the pitches F3 to B4. The numerical analyses of rate, extent, jitter, shimmer, sample entropy, and determinism for all the notes are presented in Table 3.1. The time series $\text{Vibrato}[n]$ were used for the phase space reconstruction, which indicates the extent of the vibrato, i.e. the deviation from the mean f_o of the studied note in cents. Visual inspection of the phase spaces and recurrence plots revealed observable trends of vibrato behaviour. Figures 3.2 to 3.11 present some exemplary notes of phases spaces and recurrence plots (along with time series graphs) that illustrate these trends in dynamics and regularity.

In the phase space plots, the axes from the 3D space correspond to the

extent of vibrato and the two delayed signals for the reconstruction, usually contained within -100 and 100 cents. For the recurrence plots, the upper graph shows the f_o time series of the notes, with the horizontal axis corresponding to time (seconds \times 100) and the vertical axis to a normalised version of the extent (z -score normalisation). The bottom graphs show the obtained recurrence plot of the signals, which are symmetrically mirrored around the line of identity.

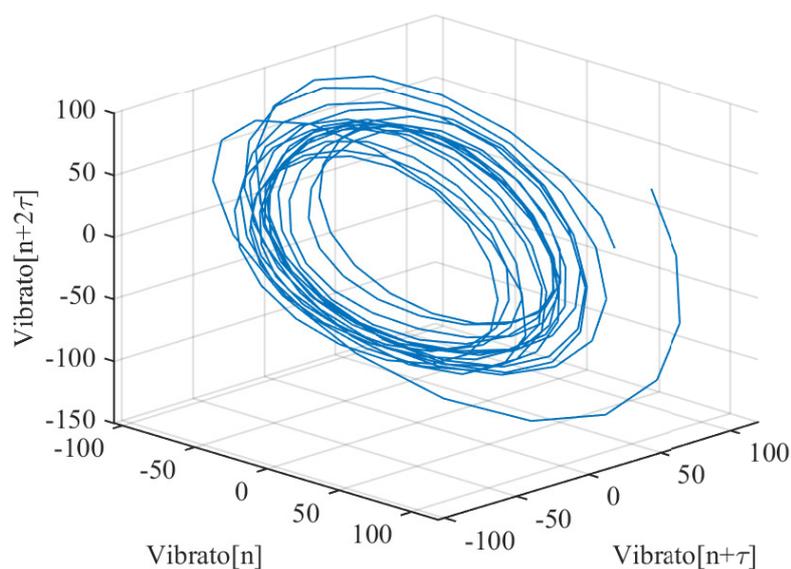


Figure 3.2: Phase space of note 46.

Figure 3.2 shows the reconstruction of note 46, which is one of the longest extracted notes, having 19 cycles and a length of 3.3 seconds. The recurrence plot for this note in Figure 3.3 shows that, with the parameters used in this study, it is possible to detect some of the common structures described by Marwan [108] corresponding to diagonal lines parallel to the line of identity. The same patterns associated with periodicity can be seen in the recurrence plot of note 67, which is also 19 cycles in length (Figure 3.5), with the main difference that here another structure is present in the form of white bands associated with transitions. This makes it possible to detect three transitory windows in Figure 3.5 separated by the vibrato cycles that occurred around

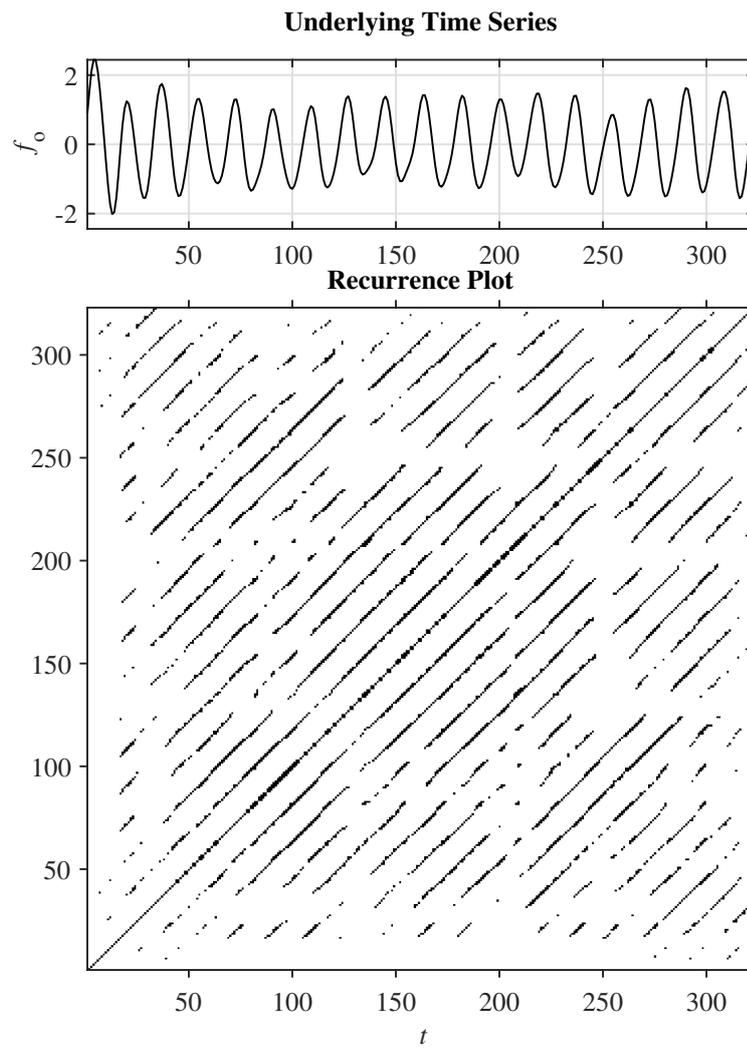


Figure 3.3: Recurrence plot of note 46.

140 and 210 (t between 1.4 s and 2.1 s).

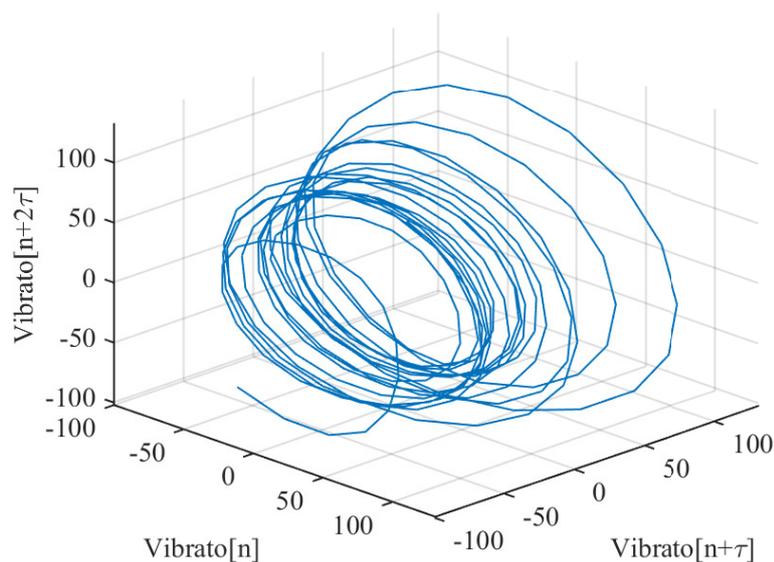


Figure 3.4: Phase space of note 67.

The results from note 62, with only 5 cycles of vibrato, in contrast to those from the two previous long notes, show that recurrence plots are also suitable for short notes. This can be observed by the point scatters in Figure 3.7 that reflect the periodicity from the very circular orbit in Figure 3.6. This phase space reconstruction of the note, along with its low value of sample entropy ($\text{SampEn} = 0.273$) show that these two analyses are also appropriate for short time series.

Figure 3.8 shows another example of a short vibrato tone. In this case, the note has a high sample entropy value of 1.08 and high determinism of 0.919. As compared to the phase space in Figure 3.6, the orbits created by this attractor look denser, meaning that the distance between them has higher separation from each other. In other words, the extent varied more over each vibrato cycle. However, it is also observable that the orbits are very smooth, which may be related to its high value of determinism. The recurrence plot of this note (Figure 3.9) also reflects high complexity, as contrary to note 62, there are fewer diagonal/continuous lines.

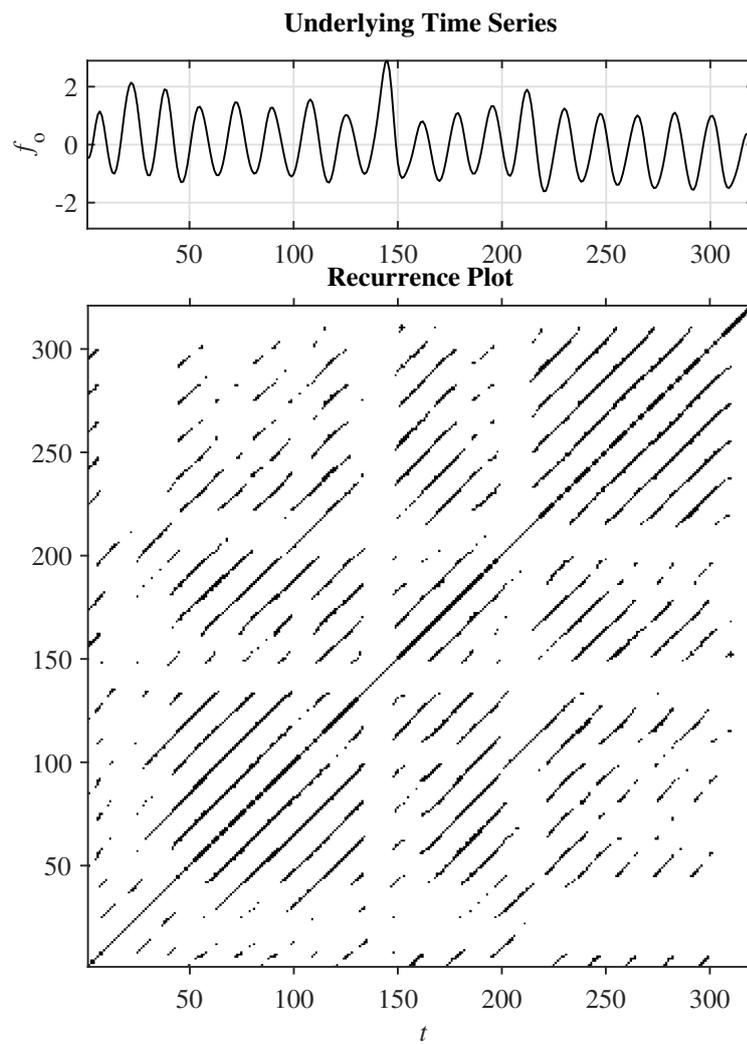


Figure 3.5: Recurrence plot of note 67.

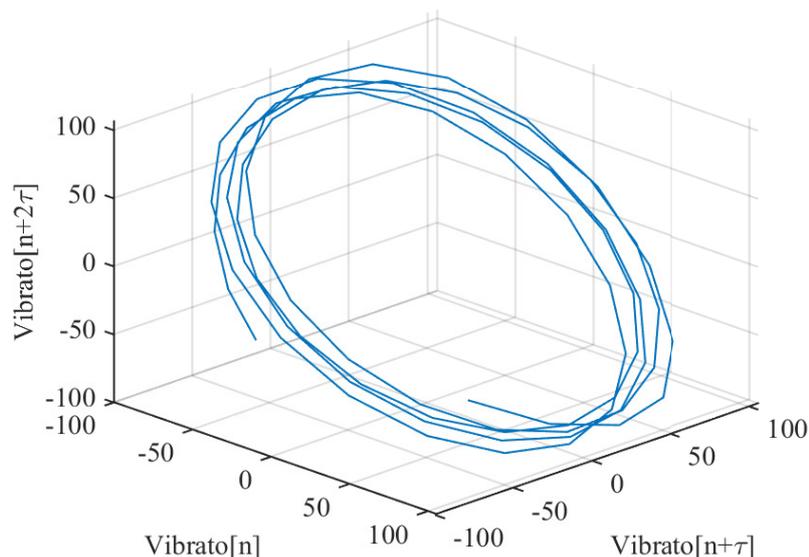


Figure 3.6: Phase space of note 62.

Sample entropy values from notes 62 and 17—both approximately one second long—indicate a possible difference in vibrato regularity between the two notes ($\text{SampEn} = 0.273$ for note 62, and $\text{SampEn} = 1.08$ for note 17). However, the measure of determinism did not reflect such difference between the notes ($\text{DET} = 0.967$ for note 62, and $\text{DET} = 0.919$). This comparison is relevant given that the number of cycles and length of those two notes are very similar, however, the higher sample entropy value of note 17 as well as the visually apparent separation of orbits suggest that note 17 is more irregular than note 62.

A number of vibrato tones were excluded from the above analysis because they were not restricted to a single musical note, that is, the singer performed a *portamento* between two adjacent musical notes whilst continuing to phonate on a single vowel and producing vibrato. Examples of such instances are shown in Figure 3.10 and 3.11 and illustrate the property of recurrence plots to reveal the windows between transitions or non-stationary events and their potential to allow analysis of transitory behaviour of vibrato across musical notes.

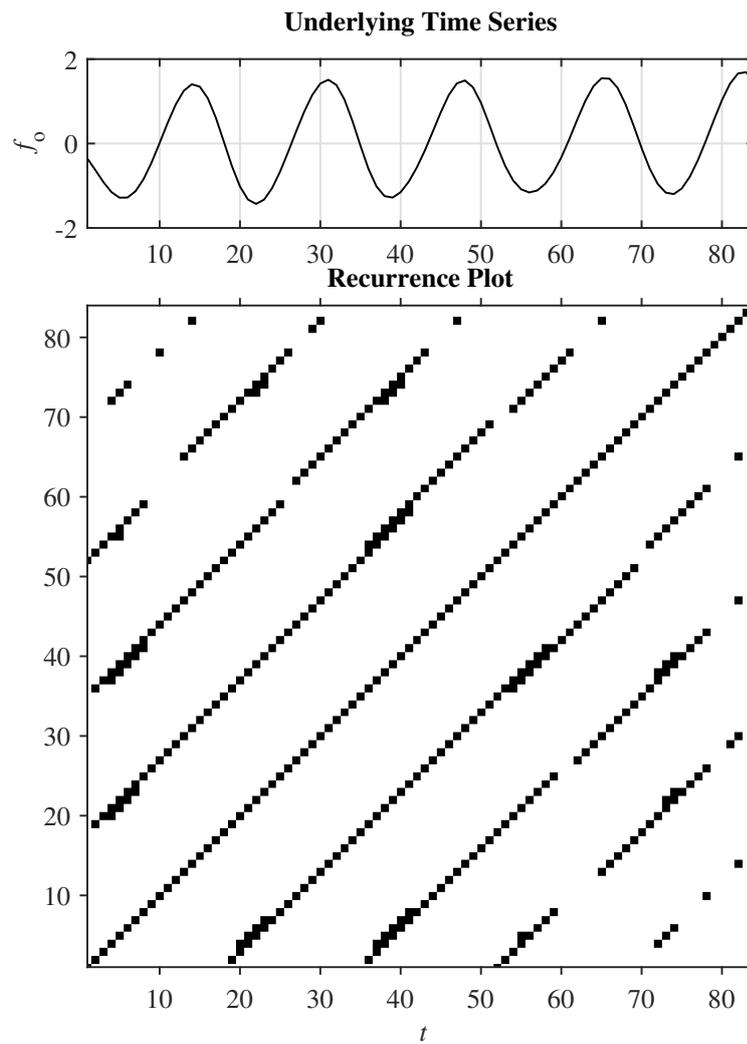


Figure 3.7: Recurrence plot of note 62.

Table 3.1: Descriptive characteristics of the notes analysed, vibrato parameters, and regularity metrics.

N	(Hz) f_o	(s) Length	Cycles	(Hz) Rate	(cents) Extent	(%) Jitter	(%) Shimmer	SampEn	DET
<i>La Serenata</i>									
1	298	2.83	17	6	117	6.59	22.54	0.558	0.93
2	193	1.6	9	5.6	121	4.63	36.52	0.872	0.908
3	297	2.79	16	5.7	107	6.47	20.97	0.432	0.928
4	297	1.46	9	6.2	121	5.93	35.29	0.542	0.927
5	216	1.72	11	6.4	93	14.49	36.08	0.836	0.869
6	245	2.09	12	5.7	117	3.46	30.16	0.639	0.878
7	394	1.03	6	5.8	128	11.49	27.08	0.371	0.936
8	336	1.03	6	5.8	143	4.57	10.05	0.5	0.859
9	354	1.54	10	6.5	105	7.92	40.69	0.539	0.95
10	332	2.26	13	5.8	91	4.4	32.82	0.701	0.911
11	293	1.28	8	6.3	102	6.98	33.56	0.318	0.912
12	268	1.81	11	6.1	112	4.5	42.21	0.557	0.906
13	195	1.34	7	5.2	149	7.92	22.76	0.768	0.897
14	302	3.26	18	5.5	128	5.12	23.83	0.464	0.92
15	297	2.28	12	5.3	119	7.15	25.97	0.706	0.908
16	300	0.99	6	6.1	117	7.81	23.44	0.478	0.922
17	246	1.1	6	5.5	114	5.49	44.31	1.08	0.919
18	403	1.13	7	6.2	127	3.53	40.69	0.542	0.885
19	371	1.23	7	5.7	125	6.86	21.9	0.816	0.92
20	357	1.44	8	5.6	128	8.27	24.73	0.54	0.953
21	333	1.59	9	5.7	112	9.73	23.45	0.591	0.906
22	291	1.12	7	6.3	89	3.46	46.75	0.467	0.915
<i>'O Sole Mio</i>									
23	211	0.87	6	6.9	153	4.87	29.32	0.359	0.912
24	175	0.82	6	7.3	113	6.58	45.7	0.737	0.881
25	238	1.32	9	6.8	103	15.06	40.25	0.39	0.915
26	210	1.33	8	6	108	4.7	38.71	0.704	0.921
27	432	1.77	10	5.6	131	2.15	14.74	0.408	0.907
28	321	0.95	6	6.3	128	7.18	31.83	0.448	0.946
29	286	1.6	10	6.2	151	5.92	23.9	0.611	0.896
30	287	0.91	6	6.6	123	7.44	43.38	0.591	0.965
31	318	2.69	14	5.2	129	6.22	16.83	0.392	0.939
32	322	1.63	10	6.1	108	5.96	28.67	0.418	0.964
33	314	0.84	5	6	115	1.93	18.15	0.438	0.877
34	317	0.94	6	6.4	141	7.27	35.12	0.468	0.939
35	319	1.83	12	6.6	119	12.01	21.57	0.491	0.913
36	237	1.24	8	6.5	167	10.42	15.72	0.519	0.928
37	235	0.81	6	7.4	117	15.15	30.38	0.521	0.859
38	235	1.22	8	6.6	103	8.41	28.25	0.964	0.892
39	428	0.84	5	6	136	8.74	58.44	1.362	0.881
40	321	0.91	6	6.6	151	12.05	39.26	0.474	0.899
41	284	0.81	5	6.2	135	11.76	36.38	0.562	0.892
42	316	2.65	15	5.7	144	6.95	20.5	0.449	0.916
43	320	1.31	8	6.1	120	9.05	31.34	0.238	0.94
44	315	1.47	9	6.1	148	15.85	24.44	0.451	0.870
45	474	1.51	9	6	136	4.08	30.83	0.464	0.922
46	432	3.3	19	5.8	130	2.97	15.65	0.415	0.935

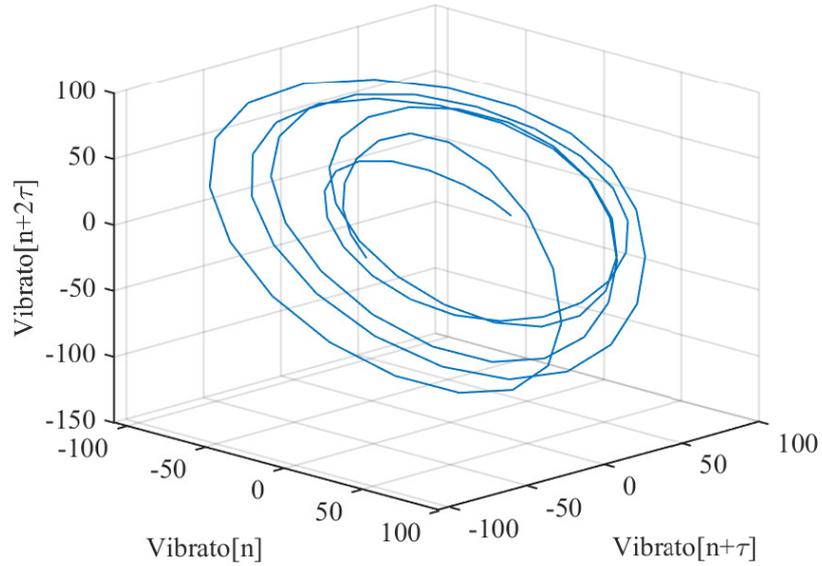


Figure 3.8: Phase space of note 17.

Table 3.1: (cont.)

N	(Hz) f_0	(s) Length	Cycles	(Hz) Rate	(cents) Extent	(%) Jitter	(%) Shimmer	SampEn	DET
<i>Nessun Dorma</i>									
47	255	0.84	6	7.1	121	10.14	43.73	0.764	0.879
48	253	1.06	7	6.6	97	4.75	52.12	0.674	0.859
49	250	1.33	8	6	106	6.14	20.47	0.411	0.897
50	197	1.29	7	5.4	192	3.6	17	0.487	0.886
51	197	1.16	7	6	149	10.59	34.48	0.479	0.895
52	252	1.3	8	6.2	192	5.93	15.5	0.657	0.913
53	449	1.68	10	6	104	5.05	18.81	0.45	0.897
54	451	0.97	6	6.2	103	1.4	35	0.336	0.948
55	437	0.94	6	6.4	81	1.51	53	0.632	0.913
56	380	2.24	13	5.8	143	3.64	23.98	0.39	0.923
57	219	0.93	6	6.5	158	4.63	38.97	0.383	0.839
58	335	3.21	19	5.9	113	5.43	25.54	0.455	0.927
59	301	0.92	6	6.5	129	5.95	48.82	0.501	0.889
60	299	1.16	7	6	116	5	20.76	0.611	0.912
61	449	0.85	6	7.1	97	8.01	50.81	0.477	0.882
62	381	0.91	5	5.5	179	1.93	4.93	0.273	0.967
63	373	0.73	5	6.8	100	6.25	35.56	0.665	0.895
64	330	1.88	11	5.9	134	4.58	22.55	0.414	0.897
65	304	1.55	9	5.8	158	4.17	10.7	0.568	0.92
66	409	2.27	14	6.2	151	4.99	16.76	0.443	0.945
67	501	3.28	19	5.8	120	5.45	24.64	0.537	0.933
68	455	3.26	19	5.8	118	2.66	24.95	0.426	0.942
Mean±SD				6.1±0.5	125±23				

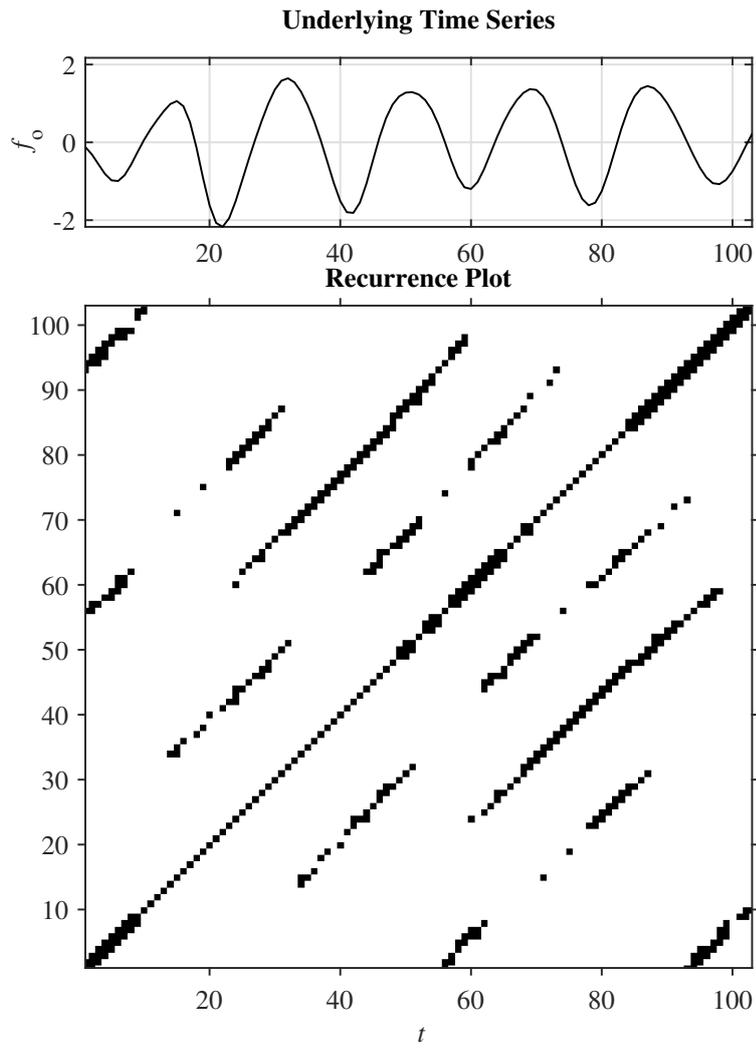


Figure 3.9: Recurrence plot of note 17.

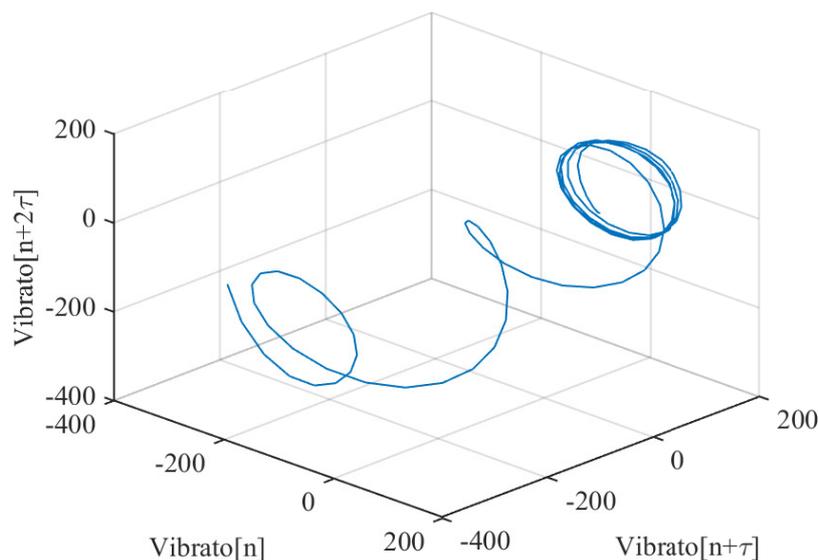


Figure 3.10: Phase space of note 43.

3.4 Discussion

This study investigates the feasibility to utilise nonlinear time series analyses to better understand vocal vibrato. Sixty-eight notes were obtained from the three music pieces as described in the previous section. The length and number of cycles per note was diverse; minimum five cycles and no maximum, resulting in examples ranging from 5 to 19 cycles. The results for the rate and extent show, as expected, that whilst some variation is present, all the notes analysed fall inside the commonly reported values for rate of 5 to 8 Hz and within a peak to peak extent of 200 cents [2]. The mean and standard deviation of notes (last row of Table 3.1) provide information about the variability across notes but no insight of small changes in regularity throughout a note.

Phase spaces, recurrence analysis, and sample entropy provided a different perspective on the treatment of vibrato across notes. Both approaches were sensitive to the selection parameters (embedding dimension and delay, tolerance, and threshold) which were set constant across notes based on

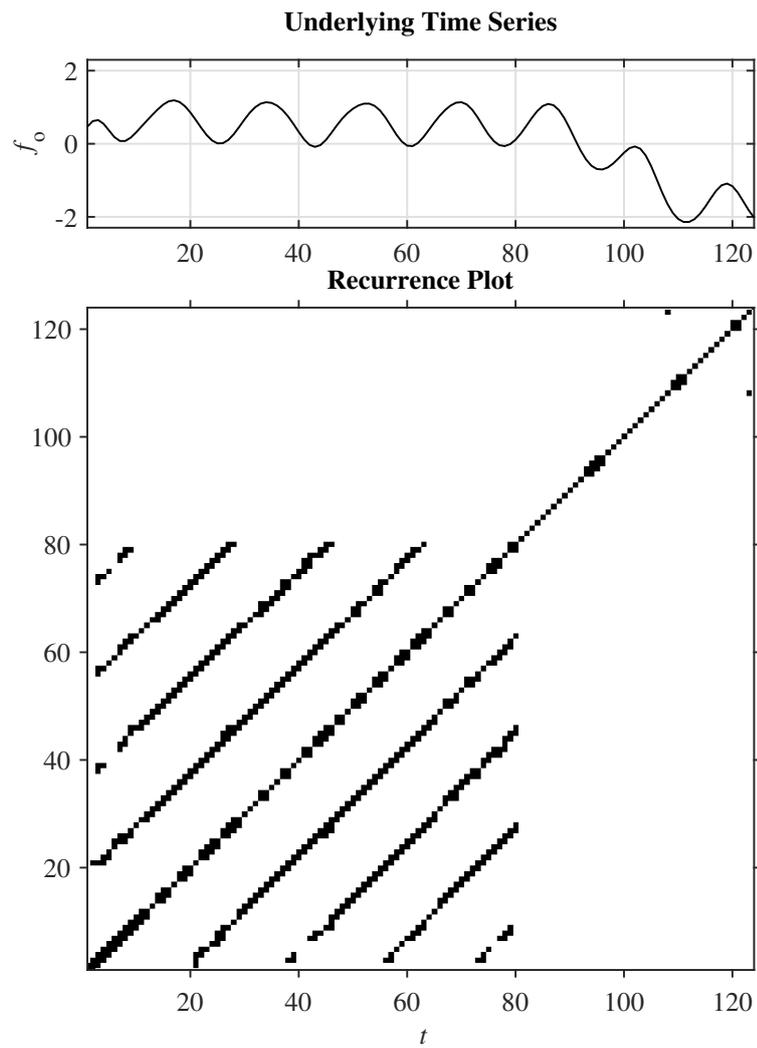


Figure 3.11: Recurrence plot of note 43.

standard recommendations [108], [113]. As suggested by these results, phase spaces and recurrence plots could be a visual aid in assessing dynamic vibrato behaviour and a useful resource in vocal pedagogy. A possible application, for instance, would be the detection of vibrato tones with more than one sinusoid [85].

The presented phase space and recurrence plots showed trends that indicate different types of vibrato dynamics being employed in Pavarotti's singing. Based on the sample entropy and determinism values, Pavarotti produced vibrato notes with different levels of complexity. However, the relationships between jitter and shimmer, entropy, and determinism are not straightforward with no clear trends observed in the current data set. This is because sample entropy / determinism can be sensitive to changes in both rate and extent as well as the shape of the cycles, and they therefore present a different picture of the vibrato behaviour of a note.

A clear advantage of jitter and shimmer over nonlinear metrics can be that these metrics offer information specifically about the rate / extent of vibrato: Jitter reflects only variation of rate, without being affected by variation of extent. On the other hand, shimmer explains only variation of extent. An advantage of the presented integrated regularity metrics is that spotlighting vibrato under nonlinear science can allow for more robust and better established methods for quantification of complexity. And also better understanding of which methods are best suited for short and non-stationary time series from experimental data. For instance, regardless of parameter selection, entropy and recurrence analysis have been favoured over other measures like Correlation Dimension or Lyapunov Exponent Estimation [114].

As it has been noted before [109], the value of DET can be affected by smoothing of time series, high values reflecting the smooth shape of the waveform. As such, determinism alone cannot be used to infer nonlinearity in time series. But collectively with jitter, shimmer, and sample entropy, it suggests that vibrato is not perfectly periodic as it has sometimes been referred [85]. And these measures from complex systems may allow the more nuanced characteristic of vibrato to be understood.

One limitation of the study is a lack of statistical analyses within any of the numerical analyses performed. Another is that the patterns and trends presented in this study do not necessarily represent a comprehensive study of all different vibrato scenarios that can occur on different genres, or even different singers. Further studies need to include a more extensive data set of vibrato notes, as well as a closer examination of the relationship between all the metrics studied in this pilot study. Additionally, since vibrato is defined as frequency modulation, accurate algorithms for fundamental frequency computation are imperative for future studies. And a rigorous approach for note selection is needed, as this selection crucially influences the regularity vs. perturbation characteristics of the analysed data. Nonlinear analysis is particularly affected by non-stationarity. Lastly, further studies are needed to provide insights into data requirements regarding minimum sample size—for which recommendations are a minimum of 200 data points [115] and this threshold can be confirmed in the specific application of vibrato.

Our understanding of these results, for instance, could benefit from the application of Principal Component Analysis (PCA) to study which metrics are more relevant in terms of variability and how they interact with each other [116] (as explored in the next chapter). A group of the variables that come into play in vibrato production can be added, such as musical intention / expression, prescribed dynamics (intensity), vowel, or musical style. Using dimensionality reduction techniques, groups of notes with similar values on principal components could be clustered or classified to study their perceptual significance. Investigating the perceptual correlates of these nonlinear metrics could lead to applications in pedagogical settings.

These features could also be used to classify individual singers or music styles using larger data sets. Recurrence plots could be used to contrast singers with regular vibrato with other singers with a more irregular pattern—or patients with indications of vocal tremor [117], [118]—possibly striving to show statistical differences of *DET* between the two cases.

Finally, more metrics and surrogate techniques could be used to provide a more robust assessment of nonlinearity and periodicity in the time series. As

future work, more comprehensive research can question if the level of vibrato chaos/irregularity could relate to the underlying physiological mechanisms of the voice. As neural mechanisms have been associated to physiological aspects of vibrato [14], [15], it would be of interest to understand what these metrics mean physiologically. For instance, if they potentially point at coupled oscillators (e.g. via different neural feedback loops) within the voice production system.

3.5 Conclusions

In this chapter, it was sought to analyse steady-state vibrato tones sung by Luciano Pavarotti in published music material. Classic nonlinear time series analyses methods have been shown here to be promising tools to better understand characteristics of vibrato dynamics, allowing consideration of the time-varying characteristics. The rationale behind this study is that the biological systems and mechanisms responsible for vibrato production might present nonlinear interactions, thus using nonlinear signal analysis seems appropriate. Application of these metrics to vocal vibrato could be valuable to pedagogy and understanding stylistic traits of different genres.

Once it was established that these metrics could be applied to vibrato tones, the next step was to apply them for further characterisation and understanding of voice signals. The following chapter expands on both the description and the application of these metrics. It uses a larger set of vibrato notes that allowed for a more thorough description and characterisation of the data, which was explored using classifier-like methods from PCA.

Chapter 4

Vibrato Complexity From Jazz and Opera Singers

This chapter applies the recurrence and entropy measures introduced in Chapter 3 to the analysis of the complexity of vocal vibrato in different styles of singing. The goal of this was to measure the degree of periodicity of vocal vibrato, which has been shown to be linked with perceived voice quality. It presents an extension of the previous methodology tested in the pilot study presented in Chapter 3. Additional sung notes from more singers were extracted from published recordings to allow statistical testing. The study presented in this chapter was published in Acosta Martinez and Daffern (2023) [119].

4.1 Introduction

Vibrato has been defined as an oscillation in frequency, intensity, and timbre of the singing voice [1], [120]. Previous studies have found a “relationship between the periodicity of the vibrato wave and its perceived quality” [17] from which periodicity, or regularity, can vary given the biological mechanism of phonation [20]. Depending on the conscious or unconscious control of the voice by a singer, there are a range of vibrato possibilities that can

convey their expressive intention and those of the musical piece. As such, variations in its usage and characteristics have been employed by singers from different styles and musical genres [98], [99], [121]. For instance, studies have observed inclusion of vibrato with the onset of the note in classically trained singers, while keeping vibrato throughout the whole note. On the other hand, different singing styles can introduce vibrato progressively at a given point of the note and decrease it in the same fashion [122]. The flexibility and diversity of vibrato enriches the music and singer's performances, but create challenges for quantifying and capturing the dynamic characteristics that contribute to achieving these expressive goals.

This complexity across vibrato tones has been addressed using mainly four parameters: Rate, extent, regularity, and waveform [14]. For rate and extent, "a vibrato range of 5 - 8 Hz and peak to peak extent of up to a tone are [...] indicative of good vocal quality" [123]. As Sundberg wrote, "awaiting that regularity and waveform will be studied in future investigations" [2], research has vastly focused on rate and extent, or measures derived from them. Although one of the first formal studies of vibrato defines it as "a periodic pulsation" [11], rate and extent, and their variability averages might not be sufficient to capture the intricacies of vibrato. As Seashore stated about vibrato shapes, "The vibrato pulsation may show a symmetrical trough and crest, it may show a larger trough than crest, or vice versa, it may show variations in length of connecting glides. The generalised artistic form presents a fairly symmetrical appearance" [11]. Such descriptions of the variability, shape, and form of vibrato can be found in literature, but no comprehensive quantifiable analysis exists to date. Jitter and shimmer, measures originated from the analysis of glottal signals, have successfully been used to study regularity. Research on these metrics is important because they have given insight into the large variability of vibrato [19], [57], [82], [124]. However, metrics utilised within different disciplines that specialise in periodicity and complexity could provide further insights of vibrato characteristics and have yet to be tested.

In previous preliminary research, it is described and tested the applica-

tion of geometric representation and nonlinear time series analysis to study vibrato dynamics [97]. The present study assesses whether nonlinear metrics are relevant features in characterizing and illustrating differences in vibrato behaviour between note, singer, or genre in Opera and Jazz singing, as well as considering the relationship of nonlinear metrics to other vibrato parameters. To this end, a larger dataset of vibrato tones from published music material of world-class singers from Opera and Jazz was analysed with metrics from entropy, recurrence quantification analysis, and the established parameters of rate, extent, jitter, and shimmer. Dimensionality reduction was employed to consider the relationship and significance of each of the metrics in collectively characterizing vibrato.

4.2 Materials and Methods

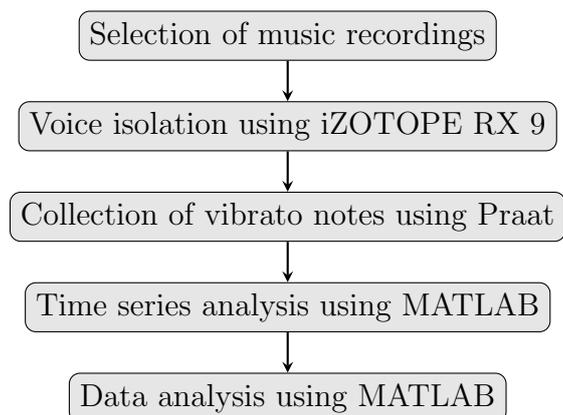


Figure 4.1: Overview of the methods utilised for vibrato analysis in this study.

This study was approved by the Physical Sciences Ethics Committee of the University of York (reference number Acosta20221114). An overview of the materials and methods utilised in this investigation is shown in Figure 4.1, which is described in depth in the next subsections.

4.2.1 Music Recordings

Music recordings from published material from Opera and Jazz were selected for the study. These two genres were selected given the contrast between classical operatic singing and its relatively well-studied vibrato parameters and contemporary commercial music [125], from which Jazz singing style is an established subgenre. A total of ten renowned singers, five per category, were selected from each genre. No exclusion criteria were defined in terms of nationality, age, or any characteristic of the singers. However, the singers were selected to have a representative number of biological females and males.

Table 4.1: Singers and music material selected for collection of vibrato notes.

N	Singer	Recordings
Opera		
1	Maria Callas	<i>La Gioconda: Suicidio!</i> [126] — <i>I Vespri Siciliani: Merci, Jeunes Amies</i> [127] — <i>Gianni Schicchi: O Mio Babbino Caro</i> [127]
2	José Carreras	<i>Turandot: Nessun Dorma</i> [128] — <i>Pagliacci: Vesti La Giubba</i> [128] — <i>Carmen: La Fleur Que Tu M'avais Jetée</i> [128] — <i>Roméo et Juliette: L'amour! L'amour!</i> [128]
3	Renée Fleming	<i>Gianni Schicchi: O Mio Babbino Caro</i> [129] — <i>I Vespri Siciliani: Merci, Jeunes Amies</i> [129] — <i>Norma: Casta Diva</i> [129]
4	Luciano Pavarotti	<i>La Serenata</i> [130] — <i>O' Sole Mio</i> [130] — <i>Turandot: Nessun Dorma</i> [130] — <i>Torna a Surriento</i> [130]
5	Joan Sutherland	<i>Turandot: In Questa Reggia</i> [131] — <i>La Traviata: È Strano! Ah Fors'è Lui</i> [132]
Jazz		
6	Nat King Cole	<i>Unforgettable</i> [133] — <i>For Sentimental Reasons</i> [133] — <i>Mona Lisa</i> [133] — <i>Nature Boy</i> [134]
7	Ella Fitzgerald	<i>Summertime</i> [135] — <i>Cry Me a River</i> [136] — <i>Bewitched, Bothered, and Bewildered</i> [137]
8	Billie Holiday	<i>Blue Moon</i> [138] — <i>Solitude</i> [138] — <i>You Go to My Head</i> [138] — <i>Moonglow</i> [138] — <i>I'll Be Seeing You</i> [139] — <i>Strange Fruit</i> [139] — <i>On the Sunny Side of the Street</i> [139] — <i>Fine and Mellow</i> [139]
9	Frank Sinatra	<i>In the Wee Small Hours of the Morning</i> [140] — <i>Strangers in the Night</i> [141] — <i>Mood Indigo</i> [140] — <i>Summer Wind</i> [141]
10	Sarah Vaughan	<i>Misty</i> [142] — <i>That's All</i> [143] — <i>Black Coffee</i> [143] — <i>Broken Hearted Melody</i> [144]

Table 4.1 shows the songs and the singers chosen for the study, with two to eight songs per singer. The number of songs was adjusted to obtain at least 76 vibrato tones per singer. Where available, the same songs were

sought in case possible comparisons between songs and singers could be made. (Although this was not included as essential criteria as in this study we did not perform song-to-song comparison). The audio files in .wav format for all the songs were obtained from the published discography of the singers.

4.2.2 Voice Isolation

To isolate the voice signal from the audio files we used the audio editing software iZotope RX and its Music Rebalance function. This allows the user to deconstruct a music signal into four components (Vocal, Bass, Percussion, and Others) and to increase or decrease the gain of each of them individually. Although currently there are various methods for source separation in music signals, iZotope RX was used as it has been successfully tested and validated in voice science applications [145].

The parameters used in iZotope RX (version 9) for the voice separation were: Voice kept at 0 dB while the level of Percussion, Bass, and Others was set to minus infinity; separation and quality for rendering were set at 50 and best, respectively. The songs were exported as wav files containing the vocal signal.

4.2.3 Vibrato Notes

The isolated voice signals were imported into Praat [112] to extract the fundamental frequency of vibrato tones. For the purpose of this study, the criteria for vibrato note selection were:

- Continuous vowel phonation on a single musical note as defined by the score
- Presence of at least five modulation cycles in the note
- Whenever present, exclusion of straight pitch segments at either the beginning or end of a musical note.

These criteria were evaluated subjectively through visual/auditory inspection in Praat. As will be explained below, however, the adequate application of the criteria was also confirmed and optimised numerically in MATLAB. It can be pointed out that Praat inadequately calls fundamental frequency “pitch”, which was computed using the auto-correlation method.

The Pitch Listing function in Praat was used to obtain the vibrato time series, which contained two vectors: a time stamp from the audio file, and fundamental frequency (f_o) from the voice signal. The time step of the vectors was $t_s = 0.01$ seconds, which corresponded to a sampling frequency of $f_s = 100$ Hz and was theoretically high enough for sampling the natural frequencies of vibrato ($\lesssim 8$ Hz).

A total of 760 vibrato notes were extracted and included in the study, having an equal number of 380 notes per genre. The scored pitch of the selected tones ranged from around G2 to E6, and had on average around 10 cycles per note. (As noted above, five cycles being the minimum for a note to be included in the collection.)

4.2.4 Time Series Analysis

The vibrato notes were imported into MATLAB for post-processing and analysis. Figure 4.2 summarises the MATLAB processing and analysis of vibrato notes, detailed in the next subsections.

Preprocessing

The vibrato notes were smoothed with a low pass filter at frequency of $f_{\text{cutoff}} = 20$ Hz. After preliminary tests, it was deemed appropriate to increase the upper limit beyond the commonly reported maximum frequencies of vibrato to retain high frequency perturbations and nonlinearities of the notes.

Another aspect that was considered was if there could be a potential issue with low-pass filtering the fundamental frequency contours with a cutoff-frequency of 20 Hz: While this may be suitable to get solid initial estimate

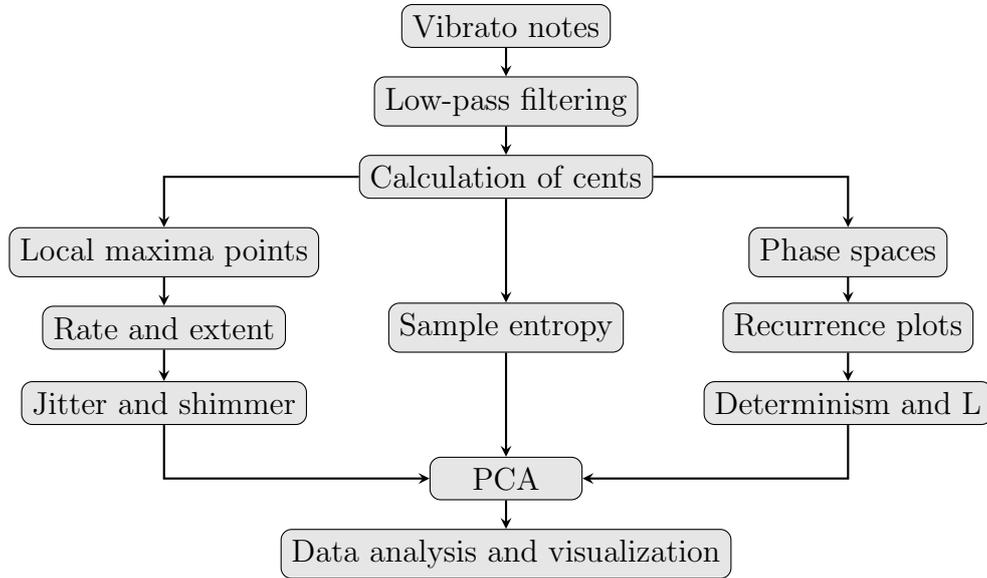


Figure 4.2: Flowchart of time series analysis and data visualisation performed in MATLAB.

of the average modulation rate, this could remove crucial high-frequency information from the analysed signals and might affect the outcome of any nonlinear analyses. It could be argued, however, that since vibrato is a phenomenon well defined between 4 Hz to 8 Hz, this high frequency information would reflect another spectral characteristic of the fundamental frequency and not vibrato per se.

The mean fundamental frequency ($\overline{f_o}$) from each note was calculated using the *mean* MATLAB function. This allowed for the calculation of vibrato expressed in terms of cents as:

$$\text{vibrato}[t] = 1200 \times \log_2(f_o[t]/\overline{f_o}) \quad (4.1)$$

where $\text{vibrato}[t]$ represents a new vector indicating the number of cents measuring the interval from f_o on a given sample t to $\overline{f_o}$ of the note.

Rate and Extent

The calculation of rate and extent was made based on the MATLAB function *findpeaks*, which provides the peak location, prominence, and width (proxy for period / 2) from the note cycles. With this function it was also possible to define minimum values to ensure:

- Rate ≥ 2 Hz
- Extent ≥ 12 cents.

This prevented false cycle detections from small fluctuations with less than 12 cents of prominence. The selection of 12 cents for the lower extent threshold was made according to results from Wooding and Nix 2016 [146]. (Additionally, this threshold helped to confirm the adequate visual selection of the vibrato note in Praat.)

Jitter and Shimmer

The measures of jitter and shimmer are derived from cycle-to-cycle variations of rate and extent, respectively, and are represented as a percentage [3]. Using the notation described above, these metrics were calculated in MATLAB as described in Chapter 3.

Entropy

Sample entropy, defined as “the negative natural logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, where self-matches are not included in calculating the probability” [113], was obtained using the MATLAB function *sampen* by Víctor Martínez-Cagigal, which computes the algorithm according to Richman and Moorman [113]. The parameters used in the function were defined and set constant across vibrato notes: $m = 3$, $r = 0.1$ (tolerance value), and Chebyshev distant type (a measure of distance between two points as the maximum difference over any of their axis values).

Recurrence

Recurrence plots are derived from phase spaces. In a phase space illustrating a vibrato note, it is possible to see the trajectory of the vibrato wave in a three-dimensional space, where the evolution of the wave creates a geometrical figure in which the cycles/orbits reflect how similar they are compared to the previous ones. In recurrence plots, “periodic motion is reflected by long and non-interrupted diagonals. The vertical distance between these lines corresponds to the period of the oscillation” [108]. Therefore, the longer the solid diagonal lines appear in the recurrence plot, the more periodic the vibrato is.

Determinism, or predictability, has been defined as “the ratio of recurrence points that form diagonal structures (of at least length l_{min}) to all recurrence points”; and Line Length (L) as “the average time that two segments of the trajectory are close to each other, and can be interpreted as the mean prediction time” [108]. For formal definitions and interpretations of recurrence plots and recurrence quantification analysis, a thorough and intriguing description can be found in Marwan et al. [108].

Recurrence plots and the measures based on diagonal lines used in this study (Determinism & Line Length) were obtained using the MATLAB CRP Toolbox by Norbert Marwan, with parameters: $m = 3$, $\tau = 4$, and adaptive threshold ϵ to ensure recurrence rate $RR \leq 0.05$. Phase spaces were plotted using the same time delay τ , which specifies the number of points that the original time series is delayed by to create the reconstruction.

Control Signals

To test the algorithm and also provide baseline values from dynamical systems, two highly deterministic signals were included in our time series analysis. The first was a sine wave created in MATLAB with the *sin* function. The second was one of the states from the Rössler system, a nonlinear system whose recurrence characteristics have been previously described and illustrated in Marwan et al. [108]. Both the sine wave and the Rössler signal

were adjusted to oscillate with similar modulation parameters to vibrato. This is possible given that the dynamics of these two signals are independent of the amplitude of the carrier signal, and depth/rate of modulation.

4.2.5 Data Analysis

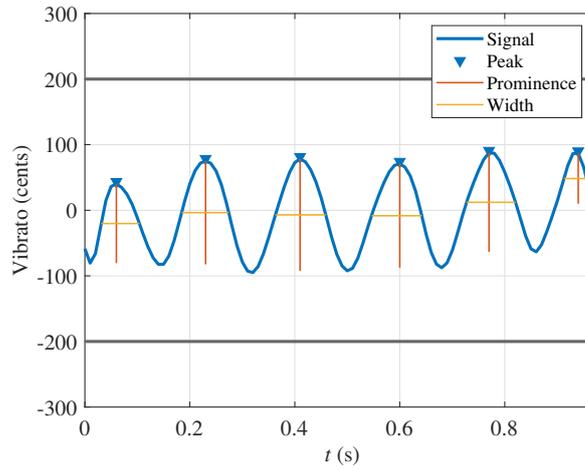
Data for Rate, Extent, Jitter, Shimmer, Sample Entropy, and Determinism are reported using the median and interquartile range as they were not normally distributed based on Anderson-Darling tests and inspection of histograms and normality plots [147]. An exception to this was the variable L, which presented normality across all singers. Therefore this value is summarised using the mean and standard deviation. Access to the dataset can be found in Mendeley Data [148].

Principal Component Analysis

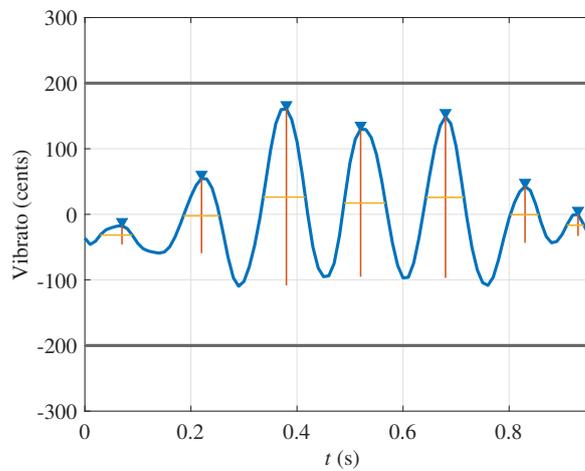
After the calculation of these seven metrics (Rate, Extent, Jitter, Shimmer, Sample Entropy, Determinism, L), the data were normalised using the *zscore* function to autoscale the variables. Dimensionality reduction was performed on the data using the *pca* MATLAB function of PCA.

Statistical Tests

After transformation, the scores from the PCA model did not present normality according to simple visualization of the data and histograms. Therefore, Wilcoxon rank sum tests with $\alpha = 0.05$ and box plots were used to analyse and describe the scores, respectively.



(a) Opera (note 7)



(b) Jazz (note 546)

Figure 4.3: Two vibrato notes of the same length as analysed by the implemented vibrato code.

4.3 Results

4.3.1 Vibrato Metrics

Figure 4.3 shows an example of two vibrato notes with the same length as analysed by the implemented MATLAB algorithm. This figure illustrates a note collected from Pavarotti and another from Holiday. The vertical axes (extent) of the graphs were standardised in the code to be between -300 cents to 300 cents, which it was found aided visual comparisons across notes. Additionally, it was found that some outlier notes could oscillate with extent as high as 400 cents (which was not found to be an f_o measurement error).

Table 4.2: Descriptive characteristics of vibrato notes from opera and jazz singers.

Singer	(Hz) Rate	(cents) Extent	(%) Jitter	(%) Shimmer	Sample Entropy	Determinism	L
Opera							
Callas	6.5 (6.0 to 7.0)	127 (107 to 146)	8 (6 to 11)	25 (21 to 32)	0.77 (0.63 to 1.02)	0.92 (0.90 to 0.95)	5.60 (1.2)
Carreras	5.7 (5.5 to 6.0)	142 (130 to 153)	7 (5 to 9)	23 (18 to 28)	0.58 (0.50 to 0.68)	0.95 (0.93 to 0.96)	7.16 (1.2)
Fleming	6.9 (6.6 to 7.1)	115 (98 to 136)	6 (5 to 8)	21 (17 to 26)	0.71 (0.57 to 0.85)	0.92 (0.89 to 0.94)	5.90 (1.4)
Pavarotti	6.0 (5.8 to 6.3)	130 (120 to 144)	6 (4 to 8)	24 (19 to 30)	0.64 (0.53 to 0.77)	0.93 (0.91 to 0.95)	6.35 (1.2)
Sutherland	5.8 (5.5 to 6.2)	118 (100 to 134)	9 (7 to 14)	25 (18 to 34)	0.86 (0.72 to 1.10)	0.89 (0.84 to 0.91)	4.86 (1.1)
Jazz							
Cole	5.9 (5.6 to 6.0)	127 (101 to 164)	8 (6 to 11)	30 (26 to 38)	0.79 (0.65 to 0.90)	0.93 (0.90 to 0.95)	5.68 (1.0)
Fitzgerald	6.6 (6.4 to 6.9)	131 (100 to 172)	6 (5 to 9)	26 (23 to 32)	0.68 (0.60 to 0.83)	0.93 (0.90 to 0.94)	5.96 (1.0)
Holiday	6.4 (6.2 to 6.7)	175 (107 to 255)	9 (6 to 14)	37 (31 to 49)	0.71 (0.54 to 0.89)	0.92 (0.90 to 0.94)	5.66 (0.9)
Sinatra	6.5 (6.2 to 6.8)	129 (103 to 157)	11 (7 to 15)	34 (27 to 47)	0.80 (0.66 to 1.14)	0.90 (0.85 to 0.93)	5.35 (1.2)
Vaughan	6.8 (6.5 to 7.1)	187 (160 to 234)	9 (7 to 15)	34 (24 to 44)	0.74 (0.59 to 0.93)	0.87 (0.84 to 0.91)	5.10 (1.0)

Note: Data reported as median (interquartile range)—except for L, which is reported as mean (standard deviation) as it presented normality across all singers.

Table 4.2 shows the quantitative results from the time series of the 780 vibrato notes. Results show that extent and rate of both Opera and Jazz singers agreed with the standard values reported in the literature. For regularity, different metrics yielded slightly different results. For instance, Fleming had low values of both jitter and shimmer compared to the other singers, but also higher values of sample entropy and lower determinism. Sinatra had the highest values of jitter and shimmer, but did not score highest on sample entropy. 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.

4.3.2 Dimensionality Reduction

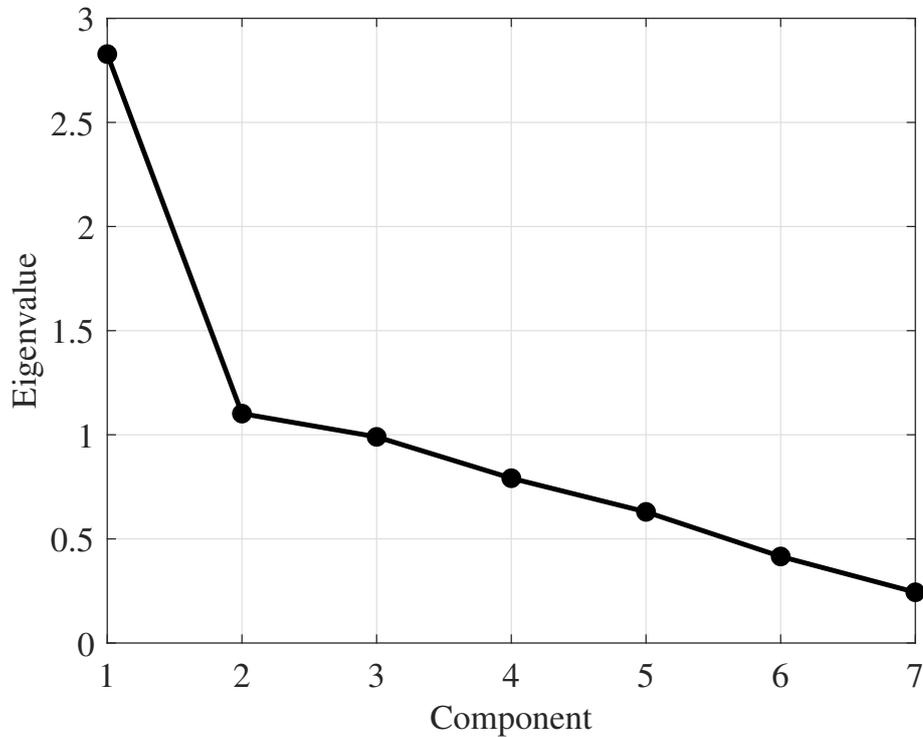


Figure 4.4: Scree plot of the PCA model performed with the normalised vibrato data.

The PCA revealed that a high percentage of variance across the data could be summarised using three components (Figure 4.4). The derived scree plot showed a pronounced elbow at the second component, and the components above the third had eigenvalues lower than 1 (eigenvalue of the third factor = 0.99). The percentage of the variance explained by Component 1 was 40%, Component 2 was 16%, and Component 3 was 14%. Together these three components explained 70% of the variance.

Figure 4.5 shows the loadings of these three components. They indicate what each component (a new variable made out of a linear combination of the original metrics) represents. Positive bars on the loadings represent values above the mean of that metric, while negative bars represent values below the mean.

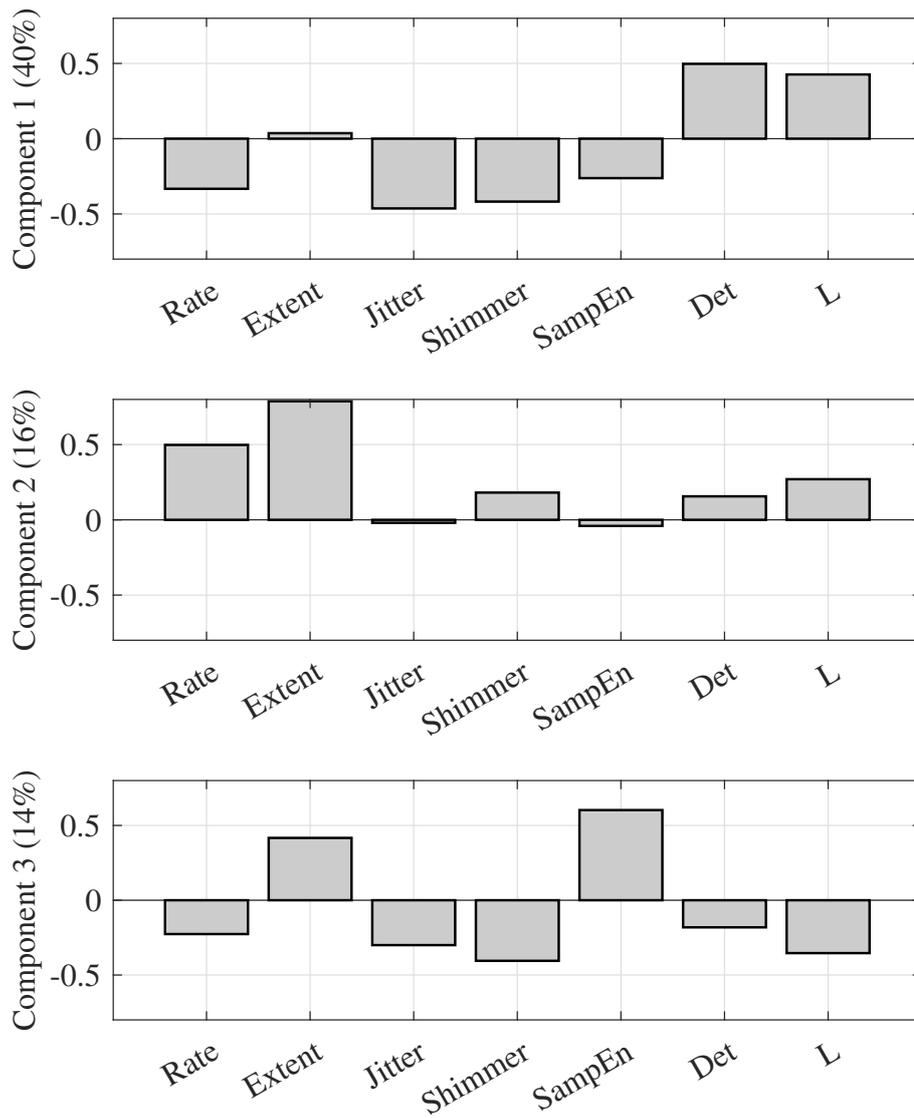


Figure 4.5: Loadings of the first three principal components.

Component 1 shows the pattern of regularity metrics that would be expected based on their definition—a relationship of determinism and L that is inverse to the relationship of jitter, shimmer, and sample entropy. We also see that rate had negative weight on this pattern, and extent had almost no weight, being neutral about regularity variance. Although this was the first principal component, this trend represents 40% of the variance. Component 2 (16%) was more heavily influenced by both extent and rate. This made sense given that these two modulation parameters define vibrato and it is expected that they might differ highly across notes or singers. Component 3 (14%) had a high sample entropy and extent weight, whilst having highly negative jitter, shimmer and L.

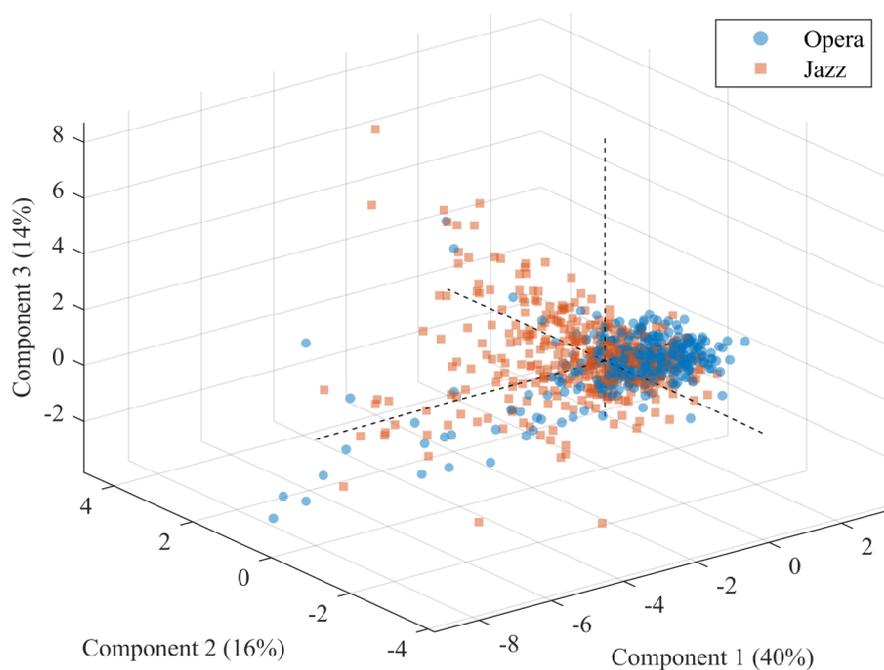


Figure 4.6: Score plot of the first three principal components with the 780 vibrato time series used in this study.

Figure 4.6 shows the scores on these three components from all the vibrato notes in this study. Positive scores on a particular component represent a

high value of what the loading represents. Highly negative scores imply that the note has a large amount of the opposite phenomenon represented by the component.

A difference in the variance between Opera singers and Jazz singers was possible to observe on these three dimensions (Figure 4.6). The Opera vibrato, represented by the blue circles, were grouped more tightly around positive scores on the first principal component—denoting high determinism and line length according to the loadings in Figure 4.5—while the Jazz singers were more widely spread on the three dimensions. In the figure, we can also observe the presence of many outlier notes from both Opera and Jazz singing. A particular separation of these outliers was found in Component 3, where some notes had $\text{SampEn} \approx 8$ in some cases.

Table 4.3: Comparison of vibrato notes from opera singing and jazz singing using the first three principal components.

Scores	Opera	Jazz	p
Component 1 (40%)	0.62 (-0.08 to 1.42)	-0.01 (-1.04 to 0.63)	< 0.001*
Component 2 (16%)	-0.29 (-0.87 to 0.23)	0.13 (-0.43 to 0.96)	< 0.001*
Component 3 (14%)	-0.03 (-0.48 to 0.42)	-0.03 (-0.62 to 0.49)	0.72

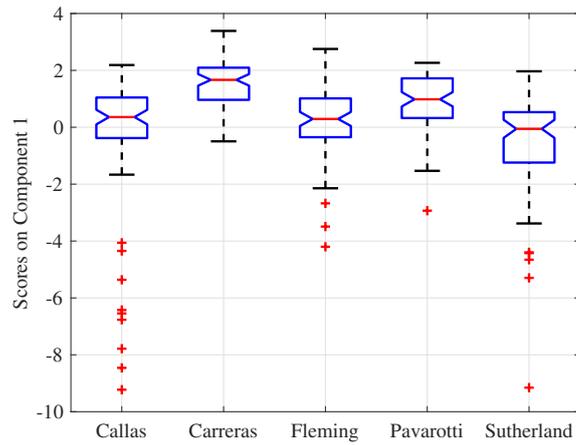
Notes: Data reported as median (interquartile range).

*Wilcoxon rank sum test hypothesis of equal medians rejected at $p < 0.05$ significance.

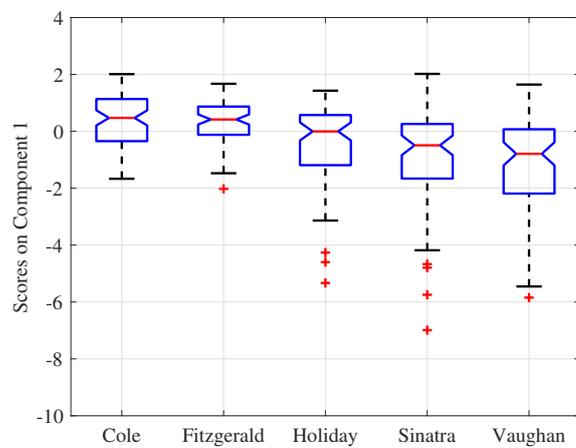
Statistical comparison of the score medians from Opera notes and Jazz notes showed differences in Components 1 & 2 (Table 4.3). Recalling the patterns described in Figure 4.5, this means that the vibrato notes from Opera singing and Jazz singing were not equal in terms of complexity/determinism (Component 1) or rate and extent (Component 2). However, differences in scores on these components reflect the percentage of data that they represent. The scores on Component 3 showed no difference between the two groups.

4.3.3 First Principal Component

The variance explained by the first principal component was particularly relevant to this experiment and complexity differences across the notes and



(a) Opera



(b) Jazz

Figure 4.7: Differences across singers within Opera and Jazz on the first principal component.

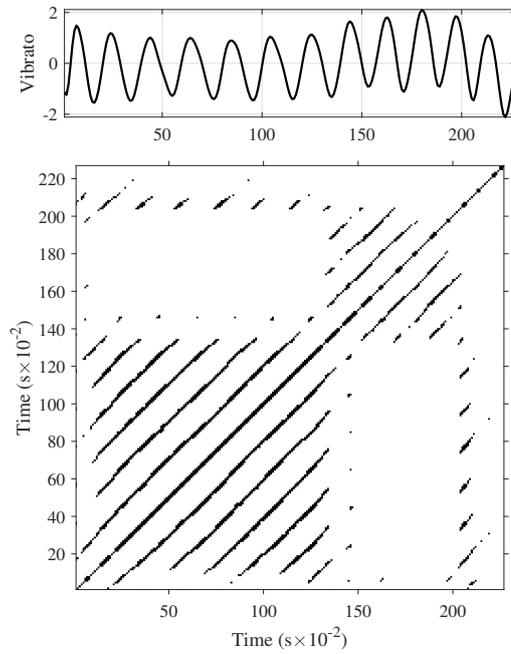
singers. As alluded to above, high values of jitter, shimmer, and sample entropy are measures that indicate irregularity or complexity, whereas high values of determinism and L in recurrence suggest the opposite: High periodicity and predictability. What was not so evident about this relationship was that low rate would also be related to low entropy, as indicated in the loadings.

From the singers in the Opera category, José Carrerás scored the highest on this component (Figure 4.7a), while Sutherland had the notes with lower median value, and Callas had high percentile values but a large number of outlier notes with low scores. From Jazz singing (Figure 4.7b), Nat King Cole was the only singer that had no low outlier notes in vibrato, and he also scored high on this component (75th percentile around 2). The singers Billie Holiday, Frank Sinatra, and Sarah Vaughan had similar results between them within the Jazz category, which were similar to Sutherland's scores from the Opera category.

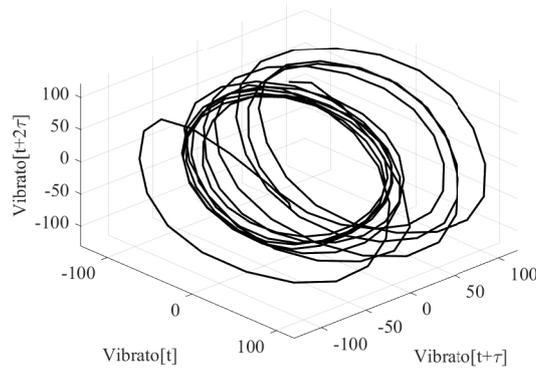
4.3.4 Recurrence Plots

To illustrate the implications of the results described above, the following section presents notes with high scores on the first principal component. Figure 4.8 shows the dynamics from one of Carrera's notes (note 291) with higher value (score = 3.27) on the first principal component. The horizontal and vertical axes of the recurrence plot (Figure 4.8a) correspond to the time vector sample (as described above, with $t_s = 10$ ms), and the autoscaled (unitless) underlying time series is also shown on the upper graph. The plot in Figure 4.8b shows the attractor reconstruction in the phase space, which helps to visualise geometrically the shape of the vibrato. In this case, the axes of the three-dimensional space represent the extent measured in cents from the $\text{vibrato}[t]$ time series.

The long diagonal lines in the recurrence plot seemed to agree with the high determinism and line length represented by the high score on Component 1. This graph also shows that the vibrato note, despite oscillating steadily around the same fundamental frequency up to around $t = 140$, at



(a) Waveform (top) and recurrence plot (bottom)



(b) Phase space

Figure 4.8: Opera vibrato with high score on Component 1 from Carreras (note 291).

that point the note experiences an offset in f_o that would cause a slight change in the perceived pitch of the note. This transition is illustrated by the white bands—in recurrence plots, “abrupt changes in the dynamics as well as extreme events cause white areas or bands” [108]—and also the large number of orbits in the phase space that appear to oscillate outside the main dense orbit.

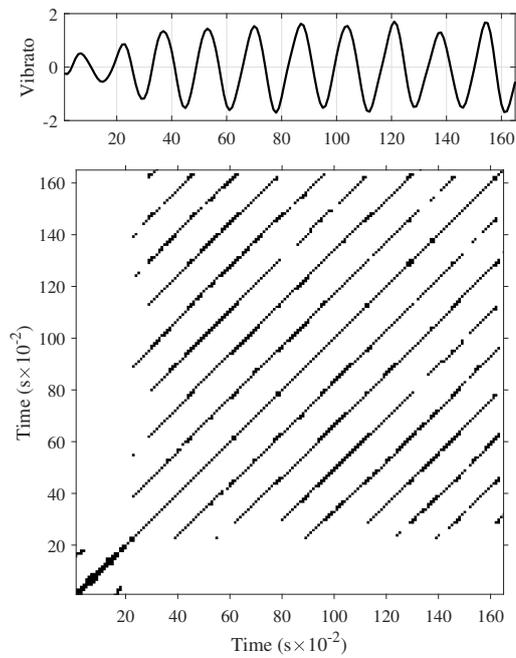
In Figure 4.9 it is shown a recurrence plot and phase space from one of the highest scores (note 431) from Fitzgerald on Component 1. This note illustrates the characteristic of some Jazz vibrato tones to increase the extent slowly. In the recurrence plot of this note, the lines of recurrence only appear once the vibrato extent stabilises and gets attracted to the limit cycle (around sample 25 of the time vector). This is also illustrated in the phase space, where it is possible to see how the note spirals out from the centre and gets attracted into a regular orbit in the space.

4.3.5 Deterministic Controls

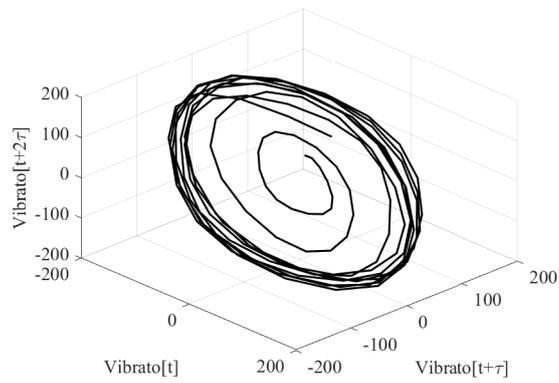
Table 4.4: Comparison of two notes with high score on Component 1 from Opera (note 291), Jazz (note 431), and other highly deterministic signals adjusted to oscillate within standard vibrato parameters.

Characteristic	Opera note	Jazz note	Sine wave	Nonlinear system
Cycles (n)	13	11	7	11
Rate (Hz)	5.6	6.4	5	6.3
Extent (cents)	156	254	121	112
Jitter (%)	7.3	4	0	2.7
Shimmer (%)	11	23	0	25
Sample entropy	0.43	0.94	0	0.37
Determinism	0.97	0.92	1	0.99
Line length	10	8.13	70	42

Table 4.4 shows a comparison of the notes with high Component 1 score described in the last subsection. For this, the original metrics were used (before component transformation) of the notes. And to put in perspective the periodicity/complexity of these notes, the table also shows values ob-

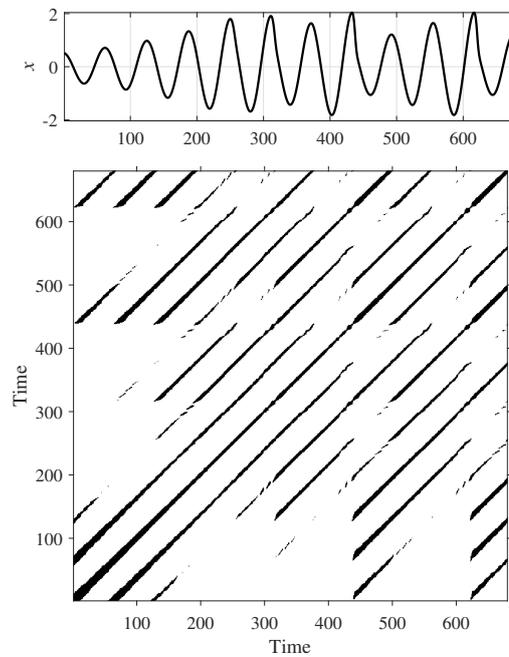


(a) Waveform (top) and recurrence plot (bottom)

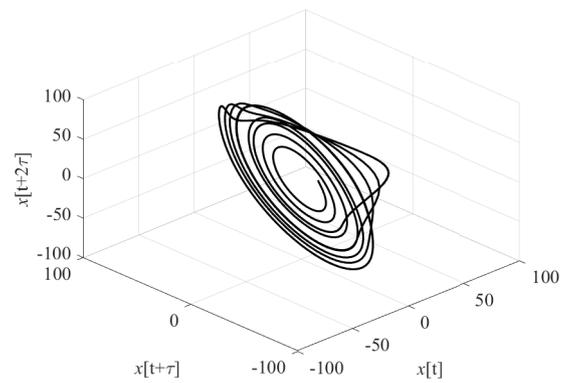


(b) Phase space

Figure 4.9: Jazz vibrato with high score on Component 1 from Fitzgerald (note 431).



(a) Waveform (top) and recurrence plot (bottom)



(b) Phase space

Figure 4.10: Dynamics from one of the states from the Rössler system with $a = 0.1$, $b = 0.1$, and $c = 12$.

tained from a periodic sine wave and a mathematical model with nonlinear dynamics 4.10.

Table 4.4 shows the results from notes with high scores on Component 1 described above (notes 291 & 431), and two highly deterministic signals: a sine wave, and a nonlinear system state, the dynamics of which are illustrated in Figure 4.10. For this comparison, the original metrics were used (before component transformation) of the notes.

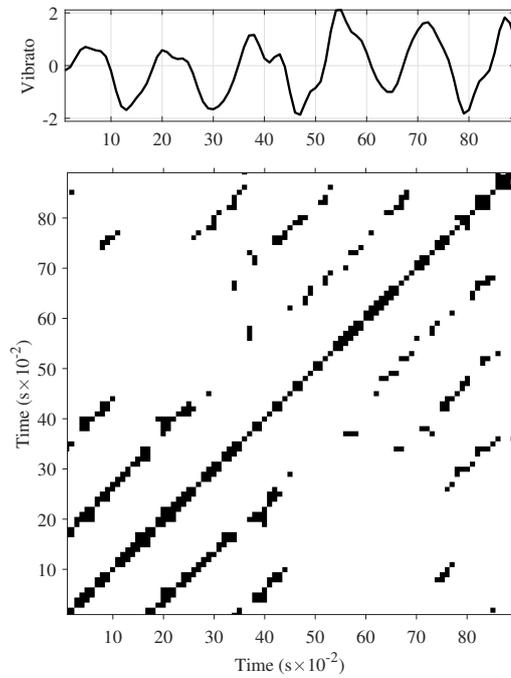
These vibrato notes, although being notes with high scores on Component 1, had regularity values far from those of a sinusoidal wave—particularly L, which was around seven and nine times higher than the Opera and Jazz samples, respectively. However, Table 4.4 shows that the nonlinear system and Jazz note had similar jitter and shimmer values, but different sample entropy and determinism. Similar comparisons with these deterministic signals can be observed with the values presented in Table 4.2 from all the singers in the study.

4.3.6 Low Determinism

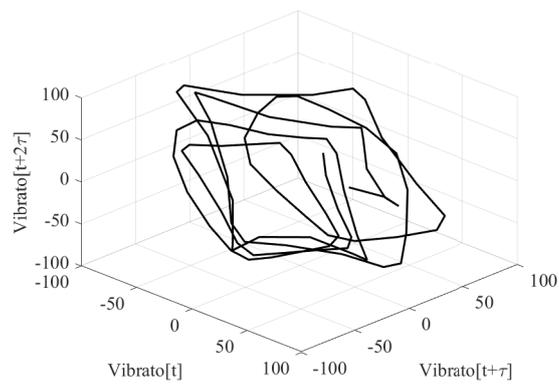
Notes with low scores on Component 1, representing low determinism and high entropy, had different observable shapes than those with positive values. For instance, in Figure 4.11 vibrato note from Maria Callas is presented which represents one of the outliers on Component 1. This note had values of $\text{SampEn} = 8.3$, $\text{Determinism} = 0.74$, and $L = 3.4$, which suggest high complexity. This can be observed in the recurrence plot—having few and short diagonal lines—and also in the phase space, which shows an almost square shape and large distance between each of the trajectory orbits.

4.4 Discussion

In this study, vocal vibrato was analysed from 780 notes to determine the extent to which nonlinear metrics can characterise relevant features of vibrato. Specifically, how these metrics relate to established vibrato parameters was



(a) Waveform (top) and recurrence plot (bottom)



(b) Phase space

Figure 4.11: Opera vibrato with low score on Component 1 from Callas (note 86).

analysed alongside illustrative differences between vibrato notes in Opera and Jazz singing. We found that the principal component, explaining 40% of variability, had positive weights of Determinism and L; while having negative weights of Rate, Shimmer, and Sample Entropy. We also found that using this component the vibrato tones from Opera compared to Jazz singing were statistically different, and it was possible to spotlight singers and notes with high periodicity.

4.4.1 Regularity Metrics

Nonlinear metrics were confirmed to be relevant features of vibrato regularity, validated by the comparison and similarities with jitter and shimmer. Although the relationship between the metrics utilised was not simple or easily reduced using principal components, we note that this is possibly given the subtle but important difference between them—each metric may be contributing information about a particular characteristic of complexity. For instance, jitter and shimmer are known for representing independently the variability of frequency and amplitude of vibrato, respectively. Whereas sample entropy, determinism, and line length are metrics that encapsulate both aspects of regularity from quite different approaches based on their definitions and calculations.

A contribution from this analysis is that we confirmed that geometrical representation of the form and shape of vibrato and their quantifiable metrics yield results similar to the established regularity metrics. This means that the values from nonlinear dynamics from recurrence, illustrated using phase spaces and recurrence plots, can provide useful and reliable information about vibrato regularity. As Korn described, “A remarkable feature of the phase space is its ability to represent a complex behaviour in a geometric and therefore comprehensible form” [114]. As for recurrence plots, we saw that they were also able to describe slight changes in f_0 transitions and intonation across the note, which are not captured by jitter and shimmer. As such, in phase spaces and recurrence plots, it is possible to assess qualitatively individual regularity of vibrato in ways that quantitative and averaging methods

do not.

Another observed strength of this analysis in assessing regularity is its sensitivity to nonlinearities—as “recurrence quantification analysis stands out for being able to detect nonlinear inter-relations between systems [149]—which are not captured by jitter and shimmer (being calculated from rate and extent) but are likely important acoustic attributes: *periodicity within periodicity* [11]. “Superimposed upon the periodic vibrato cycles are rapid, minute pitch fluctuations which probably modify the tonal quality. This periodicity within periodicity, often exhibiting interference, will throw some light upon the physiological basis of the vibrato”, noted Seashore. This was particularly pronounced in some vibrato notes from Maria Callas. These small nuances observed in the data were captured by sample entropy, determinism, and line length—but not by jitter and shimmer, as vibrato cycles were defined to be Extent > 12 cents. This was based on the work by Nix [150], in which the authors discussed that some participants could discriminate notes as containing vibrato with as low as 12 cents. Some notes from Maria Callas had cycles within cycles with less than 12 cents of amplitude and therefore were not included in the extent and rate analysis. However, sample entropy, phase spaces, and recurrence plots did capture this information. This could be another reason why the metrics could not be fully explained using fewer dimensions.

Regarding the implications of cycles within cycles, we note that this phenomenon is important and should be considered for regularity quantification. It could be argued that skill should be needed to produce such peculiar shapes, like the one in Figure 4.11 from Maria Callas, in a healthy way at will. Most vibrato notes from Callas were indeed highly periodic, but her outlier notes were also able to express a different, characteristic type of sound. This Callas sound has been considered unconventional and it is often thought to give it a fragility that makes it special. As Ira Siff wrote in a review about a 1951 recording of *I Vespri Siciliani* from Maria: “Accepted wisdom tells us that Callas possessed, even early on, a flawed voice, unattractive by conventional standards [...] Yet listen to her entrance in this performance and

one encounters a rich, spinning sound, ravishing by any standard, capable of delicate dynamic nuance.”

We can speculate that this flexibility of vibrato usage provides sound diversity and a wider spectrum of expression capabilities according to each note—perhaps analogue to a wide vocal range in a singer, where being able to produce a wide range provides a singer flexibility and freedom. We could contrast this intra-variability to listening to Carreras. His consistency across notes was high, and his singing has an easy and smooth listening where all the notes have a similar weight and give a sense of flow across a whole piece. A similar feel is drawn from listening to Pavarotti, whose notes scored high periodicity and were within the standards associated with good opera singing. In short, the analysis in the present study is based on purely acoustic and objective measures, and the perceptual significance of the results is still unknown. Given the renowned status of singers like Callas and Sutherland, it could be speculated that high complexity can be psychoacoustically pleasant and characteristic of particular singers.

4.4.2 Vibrato Usage

Differences between vibrato characteristics and singers between Jazz and Opera were observed by the visualization of scores on the principal components. This showed a trend in which it was possible to differentiate the singers by genre, suggesting that singers in Opera in general produced vibratos with (1) more determinism and less regularity and (2) lower extent and rate than Jazz singers. These findings were also supported by the statistical differences between the scores, and the trend of lower jitter and shimmer in Opera singers agrees with results from previous studies [124].

Historically, Jazz singing is a type of singing loosely derived from Opera singers. In which the first Jazz singers were trained by Classically trained music, starting with similar techniques until progressively adapting their sound according to their genre. This could be why some of the singers from Opera and Jazz still scored similarly on the scale from component one. The comparison across singers seemed relevant given that, even within their own genre

styles, the technique and characteristic style of each singer can be varied and often crossover genres. As explained in the methods, however, in this study we did not consider another important aspect of vibrato which is duration within a note. As such the results reported here only reflect differences between vibrato when it is indeed employed.

Published research on the application of nonlinear science in vocal vibrato is limited. However, our results from differences in vibrato dynamics agree with those from a compelling study analysing vibrato regularity using jitter and shimmer recently published by Becker and Watson [99]. Here the authors compared singers from belt and legit styles and their use of vibrato in terms of vibrato rate, vibrato extent, and cycle to cycle perturbation (jitter-local and shimmer-local), and also reported that “different styles require modification of phonatory quality and control of vibrato.” Although using jitter and shimmer the authors did not find significant differences between the groups, for future studies it could be questioned if other metrics sensitive to nonlinearities might yield valuable findings.

4.4.3 Limitations

There are several limitations to the current study. Firstly the dataset is limited to 10 singers and 780 vibrato tones. The music material being analysed also differed across singers and styles and was not considered within the analysis. Future studies with more extensive data could not only confirm the findings of the current study but also consider variables introduced by the musical context, such as emotional intention of note, pitch and dynamics. Additionally, the fact that vocal signals were isolated from music recordings instead of clean, controlled vibrato recordings was another limitation. The use of iZotope RX for voice isolation has been documented in the work by Gerdes and Siedenburg [145]. Using a *lead-to-accompaniment ratio* the authors calculated the accuracy of the source separation function of Music Rebalance from stereo music signals by comparing and validating their results with a data set of clean stem vocals. Although it was found that this function also yielded clean and isolated signals in our case, some artifacts

were still present and could have added noise to the calculations. Additionally, a more rigorous definition of straight pitch segments and the manual / visual selection of notes must be explored, as this may greatly influence the results.

The sampling frequency obtained from Praat in the selection of vibrato notes might also introduce a limitation. Although theoretically this should be sufficient to sample vibrato frequencies, it is hypothesised that higher frequencies would play a role in the detailed analysis of vibrato. Since the regularity metrics can be sensitive to very small nuances, increasing the resolution of the signal could prove valuable to reduce noise in the calculations and yield much better representations in the phase space and recurrence plots.

4.4.4 Future Work

Many aspects regarding the significance of regularity are still unanswered. As hinted above, the most prominent one has to be the perceptual correlates with regularity. Previous studies have suggested that the higher the periodicity, the better the quality of a vibrato [17]. However, given the large capabilities of the voice system and our complex psychoacoustic system, it would seem difficult to simplify it as such. The complexities of quantifying regularity as well as measuring its perception make this a compelling challenge for future research. Likewise, in this study we assessed vibrato differences between singers and genres—using a similar method, future studies can focus on vibrato differences of additional aspects/labels such as relative fundamental frequency (and singer’s *tessitura*), emotion, duration, vowel, or phonation type. Related to predictability, these analyses could also help to identify unexpected events, which can occur in music performance for at least two reasons: Musician’s control (or momentary lack of control) of the instrument, or expression of remarkable musical events (harmonic, melodic, rhythmic, etc.). Another aspect related to regularity is the deterministic vs. stochastic nature of vibrato signals, for which additional metrics and analyses would be needed to add to the discussion.

Based on the findings of this study, looking to the future of research on vibrato regularity, identifying the characteristic of interest will be paramount to choosing the most relevant calculation. If, for instance, a study is interested in showing differences of regularity strictly between extent and rate variability, then jitter and shimmer would be highly revealing. If focused on quantifying vibrato regularity using only one variable but more comprehensively capturing its complexity, choosing between SampEn or L should be considered. SampEn proved to be easiest to implement and with parameter dependence of r previously reported [113]. L, whilst being slightly more difficult to obtain and more sensitive to embedding parameters, provides additional information from recurrence plots. It is worth noting that L was the only metric that presented normality, (sample entropy showed to be sensitive to outliers) and also the one that more closely assimilated the results from the first component, in which L had a high positive weight. Lastly, measurements from recurrence plots have been proven to be robust on non-stationary and short data, which is often characteristic of vibrato tones. Additionally, quantifying the Correlation Dimension of vibrato tones was considered in this study, but as reported in Kantz and Schreiber (2004) [100] regarding data requirements: “The correlation dimension is a tool to *quantify* self-similarity when it is *known* to be present. When applied for this purpose it provides a safe and stable algorithm. The concept is less suited and has to be used with much more caution if self-similarity still has to be *established*, that is, if it is uncertain that the data are low dimensional deterministic data at all.”

4.5 Conclusion

In summary, this chapter shows that nonlinear dynamic metrics applied to vibrato tones provide a valuable tool for quantifying and observing regularity beyond the main four established parameters of rate, extent, jitter, and shimmer. Significant differences between the vibrato tones produced by Opera singers compared to Jazz singers were observed, indicating that Opera singers had higher consistency in the periodicity or determinism of their notes, and

also that Jazz singers had higher rate and extent. This could be representative of the specific musical data set used in this study. The analysis also identified singer-specific characteristics based on outliers identified through PCA. The results of this study highlight the potential for more accurate descriptions of regularity, which could be eventually correlated to the perceptual significance and the underlying biological mechanisms responsible for vibrato production.

Chapter 5

Vibrato Synchronicity in Blended Duet Singing

This chapter focuses on the use of recurrence based analysis for the study of vibrato interactions. As was described in the previous chapter, recurrence is a powerful tool to understand how a single vibrato signal recurs to a state in an abstract multidimensional space. This chapter, which presents the main experiment of this thesis, describes how complexity quantification from this analysis can be used to compare the recurrence from coupled pairs of vibrato signals. This is a bivariate extension of the previous work, to further understand the complex and dynamic interactions that can take place during blended duet singing. Part of this research has been presented at an international conference [151] and published in Acosta Martinez and Daffern (2025) [152].

This chapter begins with an introduction to vibrato, but now presented in the context of vibrato modifications and adaptations during blended singing. As such, the first part of this chapter recaps the literature and current evidence of synchronisation effects and potential relationships between vibrato and choral blend before the methods of the experiment are presented.

5.1 Introduction

As described earlier in this thesis, vibrato in the singing voice is a modulation in pitch, intensity, and timbre [14], [153] and singers can modify or alter it depending on musical, physiological, emotional, cultural, and stylistic factors [18], [57], [99], [122], [154]. Vibrato provides singers with a tool with which to dynamically express a unique quality in their voice according to the needs and context of the music. In the context of group singing, choral practitioners and choral directors strive for a choral blend, characterised by a homogeneous, cohesive sound [37], [50], [155], [156]. It has been speculated and anecdotally reported that vibrato affects blend in choral singers, with some choral directors actively recommending a reduction of vibrato in their singers [38], [39], [157]. Parallel to this notion, there is evidence that the interactions occurring during choral singing can lead to a coupling and entrainment of vibrato [62]. This is supported by the large field of synchronisation research [4], [158], [159], which highlights the natural inclination of spontaneous synchrony to occur during group activities [47], [94], [160], and how this might enhance our well being when blending together [161]–[164].

The amount of previous research assessing the relationship between vibrato and choral blend is not extensive [58]–[60], [67], [69], [71], [73], [74], [165]. Moreover, mathematical methods well-suited for short and non-stationary vibrato signals have not previously been utilised in the study of vibrato synchronicity. Although research has provided evidence of vibrato synchronisation in choral singing [53], [62], [63], its magnitude and probability have not been comprehensively studied, so the relationship with choral blend is still not well-understood. Additionally, other intrinsic factors that could influence an individual singer's ability to synchronise have not been assessed, such as singing skill or vibrato quality and regularity.

The aim of this study was to assess the vibrato modifications and vibrato interactions that can arise in group singing as compared to solo singing. This study tested the hypothesis that when choral singers aim to blend to another singer with vibrato, a degree of natural synchronisation will tend to occur. It was also hypothesised that synchronisation can be related to the choral

singer’s regularity of vibrato.

To test this hypothesis, choral singing participants were recorded singing solo while producing long notes with vibrato, as well as while blending to another pre-recorded sung vibrato tone used as stimulus. This stimulus was pre-recorded to isolate, and more effectively study, one-way synchronisation in a controlled and repeatable environment across participants. This study focused on the well-known vibrato parameters of extent and rate [2], [16]. Vibrato regularity in terms of rate and extent, were measured using the established metrics of jitter and shimmer [3], which reflect the average variability across the vibrato cycles. Lastly, to assess synchronisation a bivariate extension of previously validated vibrato analysis is introduced [119] from Nonlinear Dynamical Systems [86], [96], [100]: A new metric applied to vibrato—derived from Recurrence Quantification Analysis [108], [166], [167]—which might be a valuable surrogate for how well singers are able to dynamically adapt to each other.

5.2 Background

Previous research has presented evidence of the phenomenon of vibrato synchronisation. In the study by Duncan et al. [62], the authors used a published recording from world-class singers to assess vibrato interactions in duet singing. To our knowledge, this was one of the first papers to view multiple vibrato tones through a more dynamical lens, comparing their coupled oscillations to those commonly modelled and researched in complex systems. The authors illustrated examples of synchronisation by plotting the waveform of the two vibrato signals, as well as rate and extent values for numerical analyses.

In the paper by Titze et al. 2002 [20], the researchers presented a mathematical model of vibrato using differential equations and two observations particularly interesting to this study: The first being the fact that “training reduces the variability (increases regularity) in vibrato rate and extent, may be explained by increased reflex gains.” And the second, “that the natural

singer’s vibrato can be entrained by an outside stimulus [15], has not been a target of investigation [...] It is conceivable that an auditory stimulus from another singer could override the central tremor oscillators and pull the natural vibrato slightly away from its natural resonance, but this awaits further study.”

In another paper, the pilot study by Daffern [53] assessed vibrato synchronisation using a longitudinal method, where an SATB quartet from a Master’s program in Choir Singing was recorded over a period of 10 weeks. This design was advantageous given that it shows the long-term adaptations that could occur in vibrato modifications after dedicated practice with the same choir members. It is unknown whether practice over several weeks and months would lead to synchronisation at a higher level than a more “spontaneous” synchronisation. Results showed a linear correlation between the time series from the SATB members in some of the music passages. However, it was not reported whether time series auto-correlation could have influenced these results.

A thorough literature review exploring the role of vibrato in choral singing and its diverse applications, methodological approaches, and thematic research categories has previously been reported in the form of a systematic review [6]. The qualitative analysis presented in the review is relevant given that the methods and thematic category of this study are associated with “Lx / Multi-Track Recordings” and “Adjustment: Intrapersonal Adaptations and Vibrato Matching / Interaction: Interpersonal Relationship of Singing Voices”, respectively.

5.3 Methods

This study was approved by the Data Protection and Physical Sciences Ethics Committee of the University of York (reference number Martinez20240402). All participants provided written informed consent to take part.

To study the vibrato adaptation of individual singers towards synchronisation, individual choral singers were asked to blend with a pre-recorded

reference singer. This method of pre-recording a repeatable stimulus has previously been described as a useful approach commonly used in choral research [6]. As mentioned in the cited review, pre-recording helps in isolating the reference signal from the experimental recordings of participants and provides repeatability and control for assessing common behaviour across multiple participants. The detailed processes and experimental protocol are described below and summarised in Figure 5.1.

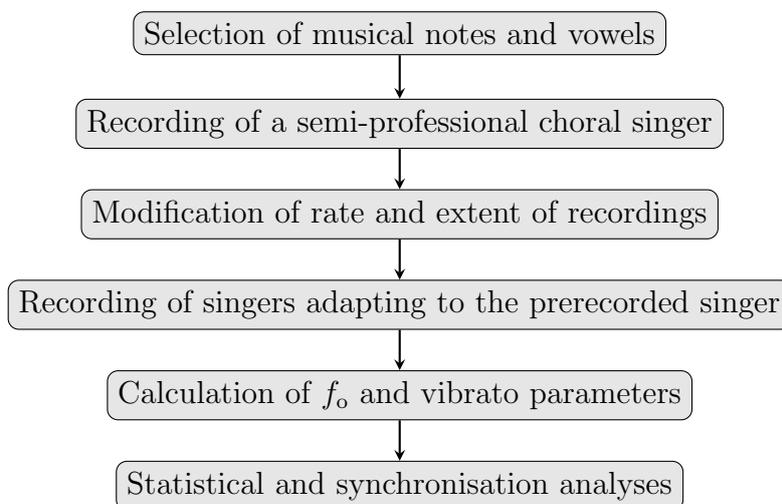


Figure 5.1: Overview of the experimental processes in this study.

5.3.1 Musical Material

Isolated notes were held for approximately 5 to 7 seconds on the vowel /i/ and /a/ on pitches of an E major triad reflecting a comfortable soprano range (E_4 to E_5). Table 5.1 shows information about the pitch, frequency, and vowels used in this study. Vowels /i/ and /a/ were chosen to diversify the stimuli that singers were exposed to during the experiment. However, no differentiation was made regarding vowel production in our analysis.

Table 5.1: Musical notes and vowels used as singing material.

/a/ & /i/	
Pitch	Frequency (Hz)
E ₄	329.62
G ₄ [#]	415.30
B ₄	493.88
E ₅	659.25

5.3.2 Pre-Recorded Singer

This study used a repeatable design for the experiment, which involved selecting a predefined singing stimulus. This allowed us to isolate one of the voice signals and simplify the design by having all of the participants singing and therefore adapting to a fixed and known stimulus. Once the musical notes and vowels were selected, a semi-professional singer was recruited, trained in opera and ensemble singing with 30 years of experience in choral singing.

The characteristics of the pre-recorded singer are shown in Table 5.2. This singer was invited to the study due to her years of experience in choral singing as well as high level education in classical voice. This ensured that the singer was experienced in both high quality vibrato production as well as having a choral tone adequate for the blending and matching task. Since vibrato is one of the main interests in this study, the singer was invited to actively introduce vibrato in all of the recorded notes.

Room and Equipment

The recording of the pre-recorded singer took place in the anechoic chamber of the AudioLab, University of York. This ensured that the recording was free from any acoustic attributes from room reflections so that when the stimulus was played back to the singers it contained only the clean voice signal (background noise level = 43.5 dBA). A single Neumann U87 Ai microphone was connected to a Fireface UCX II audio interface and a MacBook Air running Logic Pro X for the recordings.

All the audio was recorded and analysed at a sampling frequency of 48 kHz. The microphone was set in omnidirectional mode and placed one meter in front of the singer. A reference note from the singer was recorded and measured with an SPL meter (C-weighted) in order to be able to calibrate the pre-recorded signal to the same SPL during playback to participants.

Procedure

After a short period of warm up, the singer was recorded singing long, sustained notes using her “choral tone” and with vibrato of around 7 second long as described in Table 5.1. A total of 8 vibrato notes (2 vowels \times 4 musical notes) were recorded for the study.

Sound Level

The singer produced the notes at a *mezzo-piano* dynamic. As will be described below, the sound level was calibrated and measured during playback to the participants (at the participant’s position). The SPL values from the recorded notes during playback ranged from 70 dBC to 83 dBC.

5.3.3 Vibrato Modification

The recorded notes were edited using the Flex Pitch tool in Logic Pro X. This tool has a *vibrato* function to reduce or increment the extent of vibrato on a given note in terms of percentage. A hundred percent vibrato with this function represents the original, unedited extent, and changes above or below this percent value increase or reduce extent, respectively. Regarding rate, the notes were stretched or compressed in length with Flex Pitch, which would automatically increase or reduce the rate of vibrato.

Five sets of vibrato notes were created and extracted from Logic:

1. Unaltered set of eight vibrato notes
2. Eight notes with 90% of rate

3. Eight notes with 110% of rate
4. Eight notes with 70% of extent
5. Eight notes with 130% of extent.

This provided a total of 40 notes that were separated into 5 different categories of High Rate, Low Rate, High Extent, Low Extent, Unaltered (see Figure 5.2). These categories were then treated as different vibrato conditions during the playback to the choral singers. This was done to study how the singers would try to adapt to vibrato notes with different parameters of rate and extent. This could be particularly interesting in cases where the natural vibrato from the pre-recorded singer had similar rate and extent values to the natural vibrato of the participants. In this case, it would be difficult to know if singers were adapting to the reference or just happened to have similar vibrato.

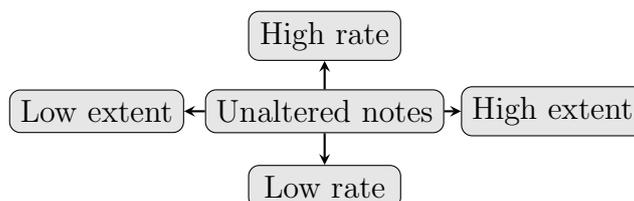


Figure 5.2: Five modulation modifications edited with Flex Pitch in Logic Pro X used as different vibrato conditions.

5.3.4 Participants

Invitation to participate in the study was made to singers that meet the following criteria:

- Had at least two years of choral experience or group singing experience
- Alto, mezzo soprano, or soprano voice
- Comfortable singing in the E_4 to E_5 vocal range.

Table 5.2: Descriptive characteristics of the prerecorded singer and the participants in the study.

Characteristic	Prerecorded singer	Participants
N	1	23
Age (years)	42	28 (11)*
Biological sex	Female	Female
Voice type (alto mezzo-soprano soprano)	Mezzo-soprano	3 4 16
Education level (undergrad masters PhD)	PhD	8 8 6
Known hearing loss	No	0
Ethnicity (white British European Asian white and Asian)	White British	15 1 6 1

Note: *Mean (standard deviation).

An exclusion criteria of the experiment was presenting known hearing loss.

Before the experiments, singers signed a consent form and were given a £10 gift card as a thank you for their participation. Twenty-three choral singers were recruited, with characteristics shown in Table 5.2. Since this study was interested in blending notes in unison, the study was restricted to voice types that typically sing in the range of the reference stimulus. A large number of the singers that took part in the experiment were students at the University of York and/or part of the choirs of the University.

Room and Equipment

The experiment took place in the recording studio at the AudioLab, University of York. Recordings were made using REAPER software and Focusrite hardware (Dante controlled interface and preamps). As mentioned above, f_s was set to 48 kHz for all recordings. The studio has a dedicated live room, where both the participant and a Genelec 8040B speaker were placed 150 cm apart during the experimental session (see Figure 5.3).

A speaker—as opposed to headphones—was chosen to deliver the audio of the stimulus to more closely replicate a choral singers natural environment: They would usually hear and match the sound coming from other singers in the room and headphones may have an impact on their perception of their own voice caused by the airborne feedback [55], [168]. The speaker was then used to give cues to the participant as well as providing the playback system for the pre-recorded singer. The Genelec was calibrated to reproduce the



Figure 5.3: Experimental setup and participant during recording and blending to the prerecorded singer (reproduced from the Genelec speaker).

voice from the pre-recorded singer at the same SPL that was measured in the anechoic chamber recording. The base of the speaker was placed at a height of around 160 cm. At the participant's position in the room, the SPL from the pre-recorded singer stimuli—reproduced from the speaker—ranged from around 70 dBC to 85 dBC. The participants and the speaker faced at an angle of 60 degrees relative to each other's position.

Because the pre-recorded singer was reproduced from the Genelec in the room, isolation of the choral singer voice was critical during recording. For this, a DPA 4066 Omnidirectional Headset Mic was used, and Laryngograph Lx with a preamp box attached to the singer. The Lx signal was recorded at line level straight from the preamp box with fixed gain control. The Focusrite preamp for the DPA microphone was set to a predefined value across all participants (peaking at around -12 dBFS or -10 dBFS for the high E₅ depending on each participant) to allow for sound level calculations using a calibrated tone.

Procedure

After giving a general overview of the experiment and signing the consent form, participants were given 5 minutes to warm up. A Nord Stage 2 keyboard in the studio was made available for the participants to go through particular exercises of their preference (or pre-recorded piano vocal exercises for those who preferred it). As such, there was variability in terms of the warm up time and routine of each singer—some singers chose to skip or shorten the warm up session, detailing that they have had other daily singing activities and felt vocally ready. Afterwards, a short practice take of the four musical notes in the study was performed to give the participant further opportunity to get used to the musical notes and task of the experiment.

Table 5.3 shows a timeline and sequence of recordings in the experimental session. The first part consisted of recording the baseline vibrato characteristics from singers in their solo singing voice. Similar to the procedure described for the pre-recorded singer, the choral singers were cued to sustain long notes with vibrato of around 7 seconds. To assist with this, the singers

were given the musical note (from the Nord keyboard played through the speaker in the live room) each time before starting phonating. Likewise, a click was played after the 7 seconds of note recording to cue them to stop phonation.

The solo singing recording took place first (using the musical notes described in Table 5.1) on /a/, then /i/). Three takes of each note were recorded, providing a total of 24 recorded notes. This was performed in an ascending order of the scale (but individually, pausing before and after each note, as described).

The next part (Blend Singing) involved the vibrato notes from the pre-recorded singer. Participants were asked to “blend” with the pre-recorded singer they heard coming from the speaker (singing same note and vowel). The exact wording of the task elicitation was to: “Sing and blend with the singer coming from the speaker”. The singers were asked to blend as soon as they heard the stimulus (and had taken a deep breath if necessary, to be able to sustain the long note) for the duration of the note.

The 40 notes from the five Blended Singing conditions (see Figure 5.2) from the pre-recorded singer were presented in a randomised sequence. Three takes were recorded. Considering the Solo Singing condition notes, this provided a total of 144 notes from each participant. Three takes with 24 notes per condition was found to be a good balance between sufficient data within condition to allow for statistical testing but without making the experiment too long to cause fatigue to participants. Overall, the whole experiment took around 50 minutes per singer. Audio recordings with the DPA microphone took place across most of the experimental session, which allowed to additionally capture qualitative data from the participants regarding their experience and comments during the experiment.

Table 5.3: Timeline and sequence of the experimental session.

1.1 Baseline singing	1.2 Baseline singing	2. Blend singing				
/a/ vowel	/i/ vowel	Randomised /a/ & /i/ vowels				
Ascending scale	Ascending scale	Randomised musical notes and blending conditions				
4 notes \times 3 takes	4 notes \times 3 takes	2 vowels \times 4 notes \times 5 conditions \times 3 takes				
	Solo	Unaltered	Low rate	High rate	Low extent	High extent
	24 notes	24 notes	24 notes	24 notes	24 notes	24 notes

5.3.5 Data Analysis

Vibrato Analysis

The recorded notes (DPA and Lx signals) were exported as individual wav files and imported into MATLAB for subsequent analyses. This was done in Reaper using regions, ensuring that only the steady-state part of the vibrato notes was captured—in order to avoid noise from the transitional period in the beginning or ending of the note.

For the vibrato analysis—of both the pre-recorded singer notes and the participant notes—the first step was to calculate the fundamental frequency, which was done using the Pitch Estimation Filter method [169] in MATLAB.

A bespoke vibrato function in MATLAB was used to obtain the vibrato parameters from each note. The function was based on rate and extent, which reflect the speed and depth of modulation in the fundamental frequency. It calculated the mean rate and extent of f_o cycles (defined from previous research with thresholds of Rate ≥ 2 Hz, and Extent ≥ 12 cents [146]) as well as jitter and shimmer [3] to study the vibrato regularity associated to rate and extent, calculated as described in Chapter 3.

Statistical Tests

For the eight notes from the pre-recorded singer, the vibrato parameters are reported using the median and interquartile range given the low number of notes—non-normality was assessed by inspection and histogram visualization of the data.

From the experimental sessions, a total of 720 notes were recorded (144

notes per choral singer). Statistical analysis between conditions was made using a one-way Analysis of Variance (ANOVA) as well as Dunnett's test [170] for post-hoc analysis when appropriate. Additionally, variable correlations were assessed using Pearson's correlation coefficients with significance at $p < 0.05$.

Recurrence

Recurrence plots and recurrence quantification analysis were used [171], [172] to detect phase synchronisation of vibrato tones from the pre-recorded singer and participants. Recurrence plots are well suited for short and non-stationary time series and have previously been applied to vibrato signals to assess auto-similarity and complexity [119].

Mathematically, recurrence plots are defined as

$$\mathbf{R}_{i,j}^{m,\epsilon} = \Theta(\epsilon - \|\vec{x}_i - \vec{x}_j\|) \quad (5.1)$$

where Θ is the Heaviside function, ϵ is a distance threshold, and x_i is a trajectory in an m -dimensional phase space. Since this study is interested in paired vibrato time series, we can define two trajectories as

$$\vec{x}(t_1) = \vec{x}_i = (u_i, u_{i+\tau}, \dots, u_{i+(m-1)\tau}), \quad t_1 = i\Delta t \quad (5.2)$$

$$\vec{y}(t_1) = \vec{y}_i = (v_i, v_{i+\tau}, \dots, v_{i+(m-1)\tau}), \quad t_1 = i\Delta t \quad (5.3)$$

where u_k v_k are the two time series with sampling frequency Δt , and τ and m are embedding parameters for reconstruction of the phase space trajectories [173].

With each vector, the resulting recurrence plot is based on a recurrence matrix:

$$\mathbf{R}_{i,j}^\epsilon = \begin{cases} 1 : \vec{x}_i \approx \vec{x}_j \\ 0 : \vec{x}_i \not\approx \vec{x}_j \end{cases} \quad i, j = 1, \dots, N \quad (5.4)$$

which shows when similar states $\vec{x}_i \approx \vec{x}_j$ occur up to the predefined distance ϵ .

As described in previous vibrato analysis [119], recurrence quantification analysis provides various metrics related to the dynamics of the time series based on recurrence. The present study focused on the recurrence rate and phase synchronisation, calculated as the Correlation Probability of Recurrence (CPR) [174].

The τ -recurrence rate is calculated as

$$\hat{p}(\epsilon, \tau) = RR_\tau(\epsilon) = \frac{1}{N - \tau} \sum_{i=1}^{N-\tau} \mathbf{R}_{i,i+\tau}^\epsilon \quad (5.5)$$

and also corresponds to the “probability $\hat{p}(\epsilon, \tau)$ that the system recurs to the ϵ -neighbourhood of a former point x_i of the trajectory after τ time steps.” [108]. Phase synchronisation can be then quantified by comparing a normalised-to-zero mean and standard-deviation-of-one version of $\hat{p}(\epsilon, \tau)$ from both systems as

$$CPR = \langle \bar{p}^x(\epsilon, \tau) \bar{p}^y(\epsilon, \tau) \rangle. \quad (5.6)$$

For this index, “If both systems are in PS [phase synchronisation], the probability of recurrence will be maximal at the same time and $CPR \approx 1$. On the other hand, if the systems are not in PS, the maxima of the probability of recurrence will not occur simultaneously. Then we observe a drift and hence expect low values of CPR” [108].

The creation of recurrence plots and recurrence quantification analysis was performed using the MATLAB CRP Toolbox [175]. This analysis required the experimental vibrato wave from the participant and the reference vibrato from the stimulus, which were multitrack-synced in time and exported as wav files from Reaper. Reconstruction parameters were set to $\tau = 4$, $m = 3$, and $\epsilon = 0.7$. The embedding dimension was defined based on false neighbour estimation [100] (which ensured that three dimensions were sufficient for the phase space visualisation and embedding), and the time delay was set approximately to correspond to a quarter of the vibrato period.

Lastly, a bivariate version of recurrence plots was also used, Cross Re-

currence Plots, on the Lx and DPA signals. This helped to assess the signal similarity of the signals and the successful vocal isolation of the close microphone, and to validate the results obtained from them.

The time series from the microphone and Lx recordings showed the same f_0 information for our analysis. Figure 5.4 shows an example, plotted using the signals from both the microphone and the Lx using a Cross Recurrence Plot—the red line on the upper plot corresponding to the Lx signal, which can be barely seen under the DPA signal in black. It is possible to see how the line of identity is present in the main diagonal of the recurrence plot, which shows that the trajectories from the DPA and Lx are the same by means of the recurrence analysis used in this study. Given the similitude between the signals, only the DPA recordings were used for the analyses, and kept the Lx signals for cases of clipping or high noise in the DPA recordings, which did not occur throughout the study.

5.4 Results

5.4.1 Pre-Recorded Singer Vibrato

Table 5.4: Vibrato parameters of the prerecorded singer notes.

Metric	Unaltered	Low rate	High rate	Low extent	High extent
Rate (Hz)	5.7 (0.4)	5.2 (0.4)	6.3 (0.6)	5.7 (0.5)	5.8 (0.5)
Extent (cents)	78.0 (17.5)	79.0 (18.0)	77.5 (17.7)	56.5 (12.5)	101.5 (22.0)
Jitter (%)	16.3 (3.8)	16.2 (2.5)	15.9 (3.5)	16.3 (3.9)	16.1 (3.5)
Shimmer (%)	35.8 (3.7)	37.4 (5.1)	36.0 (5.0)	35.6 (3.0)	36.2 (4.0)

Note: Median (interquartile range) values from the eight notes in each condition.

Results from the pre-recorded singer (Table 5.4) show that the vibrato modifications applied to the notes were accurate according to the vibrato parameters. For instance, the rate mean went down from 5.7 Hz to 5.2 Hz in the original recording and the Low Rate condition, respectively. Likewise, the extent went from 78 cents to 101.5 cents in the original recording and High Extent condition. These modifications of rate and extent did not seem

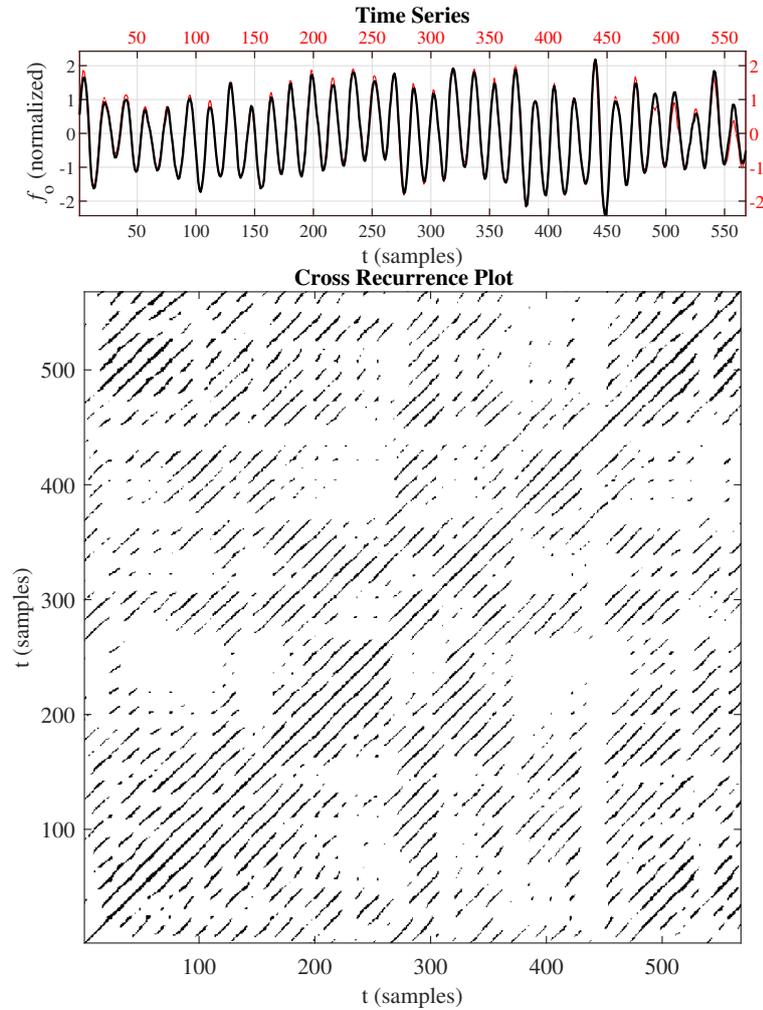


Figure 5.4: Cross Recurrence Plot between the Lx and DPA microphone signals recorded from a participant to assess their similarity. The upper graph represents the underlying time series from the Lx (black) and the DPA (red). The diagonal Line of Identity indicates both signals contain the same frequency information.

to have much of an impact in jitter and shimmer.

5.4.2 Participants' Vibrato

Table 5.5: Vibrato parameters from the participants in the Solo and Blending conditions.

Metric	Solo	Blending				<i>p</i>	
		Unaltered	Low rate	High rate	Low extent		High extent
Rate (Hz)	5.3 (0.8)	5.5 (0.8)	5.4 (0.7)	5.5 (0.8)	5.3 (0.8)	5.6 (0.9)	< 0.001*
Extent (cent)	47.8 (29.4)	52.2 (28.9)	52.4 (28.6)	52.0 (30.5)	49.0 (28.8)	52.8 (29.1)	0.017*
Jitter (%)	30.3 (13.6)	33.3 (15.0)	33.5 (14.7)	33.6 (15.4)	33.4 (15.3)	33.7 (15.7)	< 0.001*
Shimmer (%)	41.9 (12.0)	44.8 (12.4)	45.9 (11.2)	45.5 (11.9)	45.1 (12.3)	44.9 (11.9)	< 0.001*

Notes: Data reported as mean (standard deviation).

*Analysis of variance test hypothesis of equal medians rejected at $p < 0.05$ significance.

Results in Table 5.5 show the values of vibrato parameters averaged from the participants. These values are separated in the six different experimental conditions. Solo singing corresponds to the only condition where the singers phonated without listening to the pre-recorded stimuli. Unaltered, Low Rate, High Rate, Low Extent, and High Extent, correspond to the parameters the singers produced while listening and blending to the recorded singer' examples. The full data from vibrato parameters for each of the 23 participants has been published separately in Mendely Data [176].

Table 5.6: Rate post-hoc comparison of the participant vibrato notes from the Blending and Solo singing conditions.

Group	Control	(Hz)	
		Difference	<i>p</i>
Unaltered	Solo	0.16	0.004*
Low rate	Solo	0.11	0.06*
High rate	Solo	0.16	0.004*
Low extent	Solo	0.00	1
High extent	Solo	0.28	< 0.001*

*Dunnett's test significance at $p < 0.05$.

Singers produced vibrato notes with significantly different parameters of rate, extent, jitter, and shimmer across the singing conditions (Table 5.5). In

Table 5.7: Extent post-hoc comparison of the participant vibrato notes from the Blending and Solo singing conditions.

Group	Control	(cents)	
		Difference	<i>p</i>
Unaltered	Solo	4.4	0.05
Low rate	Solo	4.6	0.03*
High rate	Solo	4.2	0.07
Low extent	Solo	1.2	0.94
High extent	Solo	4.9	0.02*

*Dunnett's test significance at $p < 0.05$.

Table 5.8: Jitter post-hoc comparison of the participant vibrato notes from the Blending and Solo Singing conditions.

Group	Control	(%)	
		Difference	<i>p</i>
Unaltered	Solo	2.97	0.004*
Low rate	Solo	3.19	0.002*
High rate	Solo	3.26	0.001*
Low extent	Solo	3.06	0.003*
High extent	Solo	3.35	< 0.001*

*Dunnett's test significance at $p < 0.05$.

Table 5.9: Shimmer post-hoc comparison of the participant vibrato notes from the Blending and Solo singing conditions.

Group	Control	(%)	
		Difference	<i>p</i>
Unaltered	Solo	2.88	< 0.001*
Low rate	Solo	3.96	< 0.001*
High rate	Solo	3.45	< 0.001*
Low extent	Solo	3.20	< 0.001*
High extent	Solo	2.96	< 0.001*

*Dunnett's test significance at $p < 0.05$.

general, the rate from the participants in the Solo condition was on average lower than the unaltered mean rate produced by the pre-recorded singer (Table 5.4). This was also the case for extent. Regarding regularity, the participants in the study had on average higher jitter and shimmer (solo and blending conditions) than the pre-recorded singer.

Post-hoc analyses revealed that increases in rate were significant between the solo and blending conditions except for Low Extent (Figure 5.6). This means that, on average, the participants produced faster vibrato while listening to the pre-recorded singer in the unaltered, low rate, high rate, and high extent conditions. The pre-recorded singer vibrato had higher rate (except the Low Rate condition) than the rate from the participants in all the conditions. As for the extent post-hoc analysis (Figure 5.7), participants also experienced an increase in their vibrato while blending to the stimulus, but this was only significant for Low Rate and High Extent. Similarly, the extent from the pre-recorded stimulus had higher extent (except the Low Extent condition) than the extent from participants in all conditions.

Although this study found a significant difference in jitter between the vibrato conditions as shown in Table 5.5, post-hoc analysis showed no significant differences across the groups according to the Dunnett's test (Figure 5.8). However, shimmer did present a significant increase in all of the blending conditions compared to Solo (Figure 5.9). This means that singing along to the pre-recorded stimulus increased the variability of vibrato in terms of its amplitude (extent). This was true even though the shimmer from the pre-recorded singer was lower relative to the participants'.

5.4.3 Vibrato Synchronisation

Table 5.10 shows the CPR obtained from the pre-recorded singer and participant vibrato signals. These results reflect directly the degree to which participants synchronised their vibrato with the pre-recorded singers. For these results, individual CPR are obtained from each singer, as opposed to the averaged data as with the vibrato parameters. Nevertheless, vertical and horizontal means are also presented, which correspond to the average CPR

Table 5.10: Mean phase synchronisation coefficients calculated from the time series of participants blending to the prerecorded singer.

Participant	CPR					Mean
	Unaltered	Low rate	High rate	Low extent	High extent	
1	0.394	0.239	0.124	0.356	0.413	0.305
2	0.407	0.242	0.041	0.409	0.341	0.288
3	0.445	0.293	0.054	0.434	0.311	0.307
4	0.250	0.270	-0.016	0.133	0.174	0.162
5	0.135	0.325	-0.024	0.106	0.093	0.127
6	0.052	0.016	0.127	0.113	0.046	0.071
7	-0.023	-0.012	0.107	0.016	0.040	0.026
8	-0.041	0.127	-0.103	-0.067	-0.042	-0.025
9	-0.049	0.040	-0.096	-0.039	-0.031	-0.035
10	0.287	0.526	0.005	0.303	0.345	0.293
11	0.122	0.341	0.077	0.042	0.137	0.144
12	0.056	-0.011	0.017	0.067	-0.031	0.020
13	-0.022	0.041	-0.014	-0.039	0.174	0.028
14	0.281	0.525	-0.013	0.261	0.286	0.268
15	-0.009	0.181	-0.063	0.008	0.087	0.041
16	0.451	0.049	0.244	0.488	0.488	0.344
17	0.383	0.063	0.245	0.451	0.299	0.288
18	0.110	0.051	0.172	0.110	0.029	0.094
19	0.132	0.172	0.117	0.109	0.151	0.136
20	0.217	0.245	-0.016	0.208	0.099	0.151
21	0.115	0.357	-0.050	0.059	0.135	0.123
22	0.157	0.096	0.157	0.163	0.139	0.142
23	0.139	0.090	0.059	0.195	0.187	0.134
Mean	0.173	0.185	0.050	0.169	0.168	

per condition and per singer, respectively.

The CPR is a coefficient that has been previously reported to range from 0 to 1, which reflects the probability of phase synchronisation. Participants were able to reach high levels of vibrato synchronisation of up to mean CPR = 0.344 (values from column 7, Table 5.10). These values represent the mean from 144 notes per participant across the six singing conditions.

Observing individual notes, some participants were able to synchronise to even higher levels of CPR. For instance, Figure 5.5 shows a graphical representation of sync determination from a single note with CPR = 0.82 from Participant 16 (the highest CPR scored participant in the study). In this figure, the top graph shows the underlying time series from both the pre-recorded singer (black line) and the participant (red line) vibrato. The bottom graph shows the RR_τ for both signals, used to calculate CPR. For this recurrence graph: “During phase synchronisation, RR_τ should have high probabilities at the same τ [lag] values; thus, the shape of RR_τ should be very similar” [177]. From the individual vibrato parameters, reported separately, it is noted that this participant had a constant rate of 6.0 Hz in all conditions (except for the High Extent condition for which rate = 5.9 Hz), which is close the pre-recorded singer values. Additionally, and potentially of greater relevance to synchronisation, this participant had very similar Jitter values to the pre-recorded singer.

On the other hand, for lower levels of synchronisation (Figure 5.6), the RR_τ shows different peaks of recurrence at differing delays (or lag) that do not overlap. This makes it easier to observe the phase information of the signals, which might be not evident from the time series representation. [177]

Although the CPR coefficient has been described as ranging from 0 to 1, some of the participants in our study produced negative values of CPR. An example of this is shown in Figure 5.7. From the RR_τ plot, it is possible to see that a negative value seems to be an indication of very different distribution of probability of recurrence.

Regarding vertical CPR means, (Table 5.10, last row), participants achieved higher synchronisation in two conditions: Low Rate and Unaltered. On

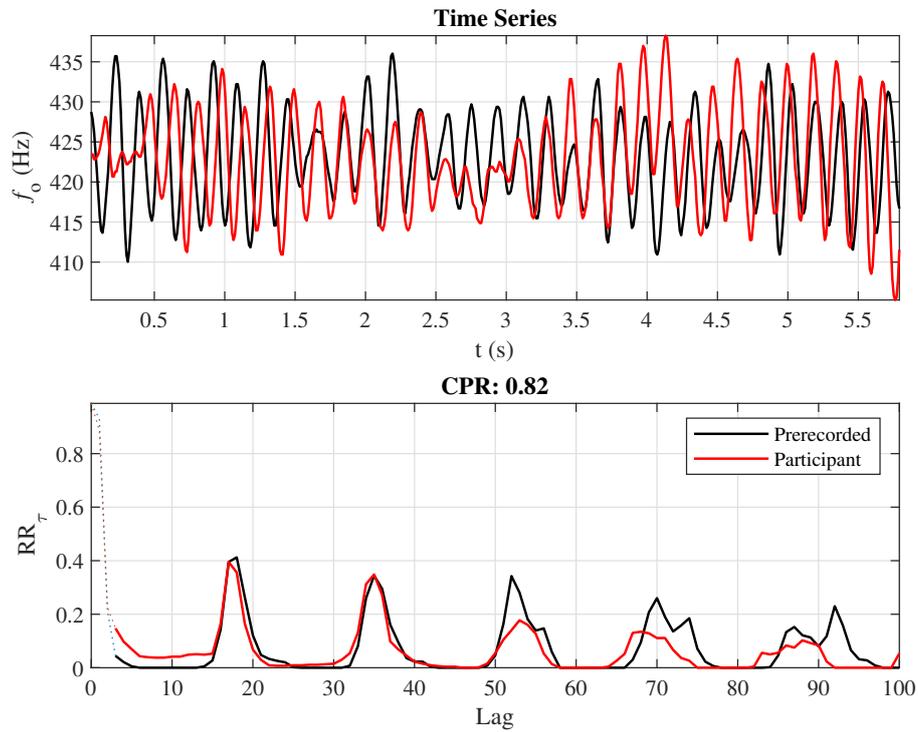


Figure 5.5: Example of paired vibrato notes with higher probability of phase synchronisation between the prerecorded singer and the participant time series.

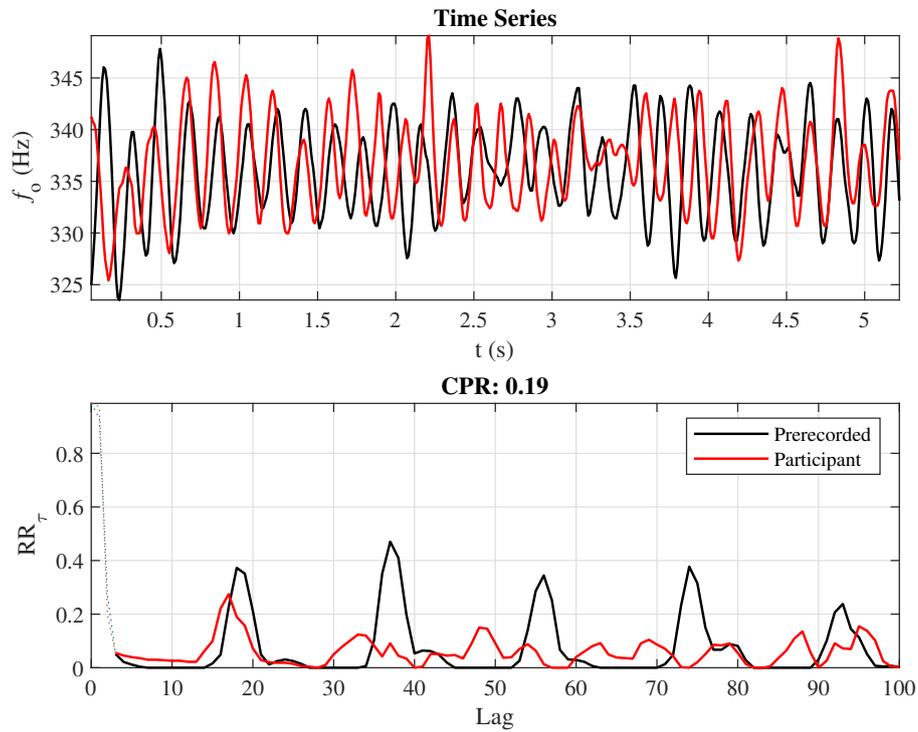


Figure 5.6: Example of paired vibrato notes with lower probability of phase synchronisation between the prerecorded singer and the participant time series.

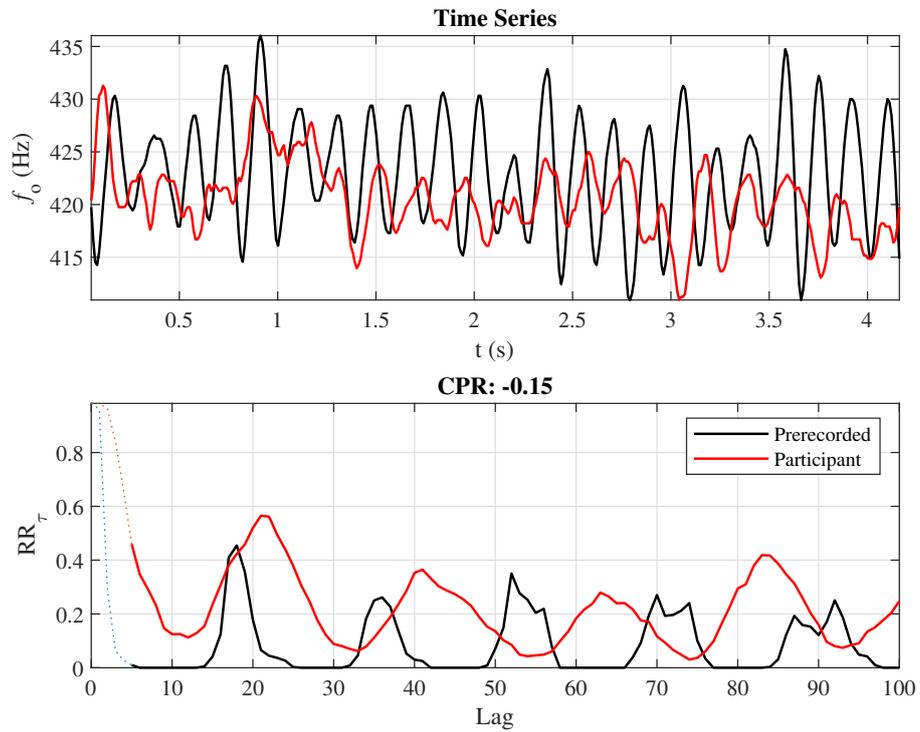


Figure 5.7: Example of paired vibrato notes with negative probability of phase synchronisation between the prerecorded singer and the participant time series.

the other hand, High Rate was the condition in which singers produced the lowest CPR values.

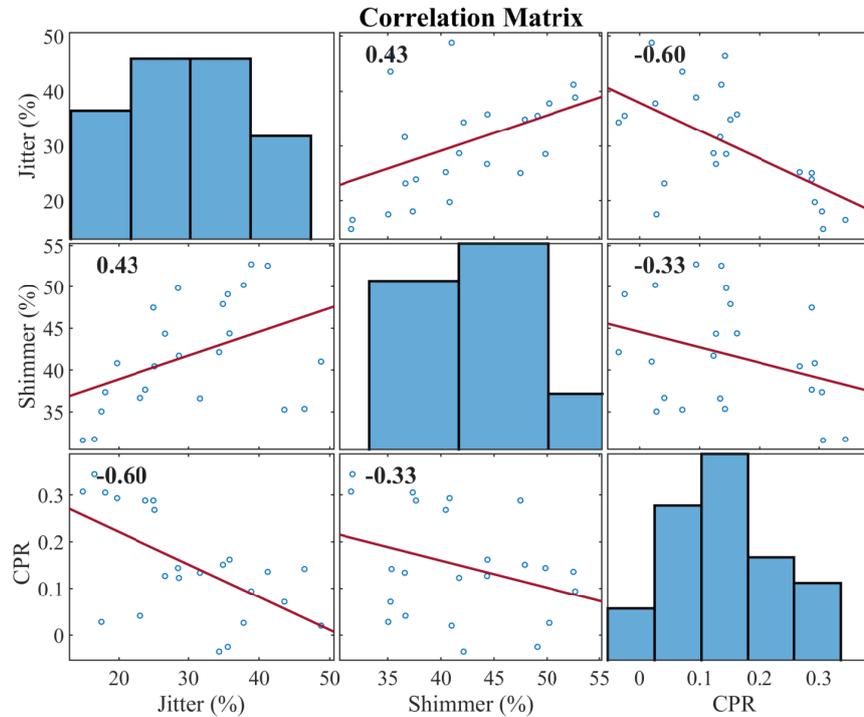


Figure 5.8: Correlation analysis of jitter (Solo condition), shimmer (Solo condition), and mean CPR (Blending conditions) data from the 23 participants, 24 notes per condition, in the study.

There was a statistically significant negative correlation ($p = 0.002$) between jitter and CPR from participants singing in the solo condition vs blending condition (Figure 5.8). This means that, the variability of vibrato modulation while singing Solo was associated with the degree of synchronisation that participants reached in the blending condition. Results showed no correlation between shimmer (extent variability) and CPR. However, as expected, there was a significant correlation between the jitter and shimmer ($p = 0.04$). In Figure 5.8, the diagonal panels show a histogram of the data from the 23 participants for each of the metrics.

Table 5.11 shows the mean SPL (L_C) for all participants in the six singing

conditions. The results from the individual values show that all singers experienced a reduction of SPL between the solo and blending conditions.

Singers in the study in average, produced lower sound level values in the blending condition relative to their solo singing (see Table 5.11). As expected, singers produced different levels of L_C inter-participant.

Table 5.11: Mean sound level calculated from the DPA close microphone notes across the different singing conditions (Solo and Blending) in the study.

Participant	L_C (dB)					
	Solo	Unaltered	Low rate	High rate	Low extent	High extent
1	92	89	88	89	89	88
2	91	90	90	90	90	91
3	86	83	83	83	83	83
4	89	86	86	86	86	86
5	91	87	87	88	88	88
6	89	86	86	86	86	87
7	89	84	84	84	84	84
8	90	88	87	88	87	88
9	89	83	82	83	83	82
10	84	83	83	83	83	83
11	89	82	81	83	83	82
12	84	84	83	84	84	84
13	91	83	83	83	83	83
14	90	87	87	87	88	88
15	91	85	85	85	84	85
16	89	79	79	80	79	79
17	83	79	79	80	80	80
18	85	83	83	84	84	84
19	85	79	78	78	79	78
20	88	80	79	79	79	80
21	77	76	76	76	76	76
22	88	81	81	81	82	81
23	78	77	77	77	77	77

5.5 Discussion

In this study we found that participants altered the parameters of vibrato when blending with the pre-recorded stimuli as compared with their solo singing. Specifically, on average they increased their rate (up to 0.3 Hz in the High Extent condition), which brought them closer to the higher values from the pre-recorded singer. This was true for all the blending conditions, with the exception of the Low Extent condition, which might be speculated that the low extent of the stimulus (mean 56.5 cents) was below a threshold to be more difficult to be perceived and therefore blended. (Although lower values have been previously reported as the perception threshold for expert listeners [146], this might be different in a blended singing context). Alternatively, it could have been that the extent was very “close” to their own so that they did not try to make any adaptations. Another significant vibrato adaptation was that participants increased their jitter and shimmer (Blending compared to Solo).

Another finding was that the ability to synchronise was highly individualised across participants. This made singers stand out in terms of synchronisation, which was easier to observe using the single metric of CPR, compared to the vibrato parameters. Moreover, there was a significant negative correlation between jitter in the Solo condition (which is the variation of the “speed” of the vibrato undulation) and the mean CPR across all blending conditions. This is tightly related to the finding regarding CPR individuality—the phenomenon of reaching high synchronisation with another vibrato was correlated with the vibrato regularity (surrogate for vibrato quality) in the solo voice.

Regarding the absolute values of rate and extent in our study, we can note that most singers produced a slower and smaller amplitude than, for instance, values from Operatic or Jazz vibratos reported previously [98], [119]. The commonly measured values for rate are from around 4 Hz to 6 Hz (but can go up to 8 Hz), with slower speeds often referred to as a wobble, and faster as a tremolo. As for extent, 100 cents to 200 cents have been cited to be common in classically trained singers [2].

The baseline values of rate, extent, jitter, and shimmer in solo singing shown in the Mendeley dataset [176], allow comparison of the metrics from the individual participants to that of the stimuli, however, only provide a general picture of the signals. For instance, the mean rate from Participants 16 & 17 are very similar and rather constant across blending conditions. However, vibrato from Participant 16 had lower jitter (range of 16.5% to 19.1%) and more similar to the pre-recorded stimuli jitter, compared to Participant 17 (range of 22.1% to 25.2%). Similarly, Participant 16 scored a higher level of synchronisation. Although it might be tempting to use rate to assess the modifications (or similarity) of vibrato, we can note that rate and extent are only based on averages of the notes, and equal rates are necessary but not sufficient for phase synchronisation.

In previous work [119], we explained this aspect in detail, showing that the commonly used parameters are averages that do not capture the nuanced dynamics that vibrato signals contain. For instance, if one were to analyse the rate of the first half-second of a note, and find a value of 5 Hz, and for the next half-second a value of 6.6 Hz. The mean rate would still be 5.8 Hz. However, this value does not reflect any information regarding its dynamics. Similarly, the difference or distance, between rate values from paired notes might not provide the full picture of the relationship and dynamics between them. This was the main topic previously studied, where it was shown that complexity metrics are a useful tool to incorporate the dynamics and nonlinear properties of vibrato. Similar to vibrato, synchronisation is a dynamical process. (See for instance Figure 1A in Fell and Axmacher, 2011 [178].) This is a strength of this study given that CPR, by using time delays and point-to-point recurrence, integrates the dynamic nuances across the entire notes.

Titze et al. [20] defined the natural resonance as in the range of 5 Hz to 6 Hz. In our study, the mean rate stimuli (Table 5.4) ranged from 5.2 Hz to 6.3 Hz, the high end being above the natural resonance proposed by Titze. Our results suggest that singers vibrato can be entrained by an external auditory stimulus, but within the bound extent and rate that the oscillator

can function at—the High Rate condition was the one that produced lower synchronisation, suggesting that singers found it more difficult to go higher than the natural resonance.

Our results also indicate that a possible aspect for the different degrees of vibrato synchronisation within singers might be the skill or quality of vibrato itself. This is supported by the data, where the singers with higher synchronisation rate were the ones with more periodic and regular vibrato. If a singer does not have extensive experience in producing vibrato, it might be less probable that they will be able to synchronise blending to another singer, which might be more of an unconscious behaviour and thus difficult task.

Informal discussion and anecdotal reporting of participants was recorded as part of the experiment providing valuable qualitative data to supplement the quantitative analysis. Although not formally studied or measured here, the experience and observations reported by participants as they took part indicated a potential relationship between (1) the subjective interest in the juxtaposition of vibrato and blending that the participants conveyed (often along with reported experience in both solo and choral endeavours); and (2) the CPR levels that they produced.

For instance, Participant 15 described her experience in solo singing and choir singing: “Blending is one of the most important things you got to be earful on and mindful of. [...] I’ve done so many different things, and being a soloist is probably easier than being in a choir, a vocal group, or an *a cappella* group. Because being a soloist you can do what you want, when you want, if you have a good accompanist he or she will follow—you got freedom. You don’t have that in a choir or a vocal group.”

Regarding vibrato, Participant 14 noted “I think she [the pre-recorded singer] has a natural vibrato that I don’t have. [...] When I detect vibrato I find it more difficult to blend, I feel we have a different quality of sound. If there was no vibrato I’d be more comfortable, just because I would relax. I would say: That’s the note she’s singing, we need to blend! And we are both kind of pure.”

The concept of a “pure sound” as an ideal was reported frequently by a number of participants when discussing vibrato and choral tone. This might be a reference to the psychoacoustic term “pure tone”—which is the sound from a sine wave with only one oscillating frequency, no harmonics—although the human voice is a more complex wave, containing harmonics and filtering by the vocal tract. One participant in our experiment, when prompted about this, speculated that, for her, purity of sound relates to the tone containing only one single frequency (f_o) and nothing else, hence purity—as opposed to a note with vibrato, which is a sequential combination of fluctuating frequencies. In this sense a pure sound was synonymous with a straight tone deviating as little as possible from the f_o . It is also compared to the sound of a tuning fork, with no fluctuation in frequency.

Participant 14 went on: “But if I detect vibrato, I feel I should be singing with vibrato as well. And [...] when you start trying you’re not necessarily relaxing. [...] If I was in a choir would I try and match somebody’s vibrato that’s next to me—I’m not sure I would! I’d probably want to [...] neutralise the vibrato.” This participant received one of the highest synchronisation scores in our study (mean CPR = 0.268).

Participants in this study were not asked to try to “match” perfectly the vibrato from the other singer. Rather, they were instructed to “match” and “blend” with the other singer as a whole. This was to assess whether they would use synchronisation as a strategy to blend. In dynamical systems, synchronisation is characterised in three main forms: Phase synchronisation, generalised synchronisation, and complete synchronisation [158], [174], [179]–[182]. As its name suggests, phase synchronisation considers the phase ϕ information, whereas generalised synchronisation is more relevant to the amplitude of the signals, such that for instance the amplitude of one signal is a function of the other. Complete synchronisation happens when both the amplitude and the phase of the signals coincide. In this study, we focused on phase synchronisation as it has been reported to be the first to appear after a weak coupling of the signals necessary for synchronisation (generalised synchronisation and complete synchronisation can only be attained once phase

synchronisation has been reached). One strength of this study is that, to our knowledge, this is the first experiment to implement nonlinear time series analysis applied to the study of multiple, paired vibratos.

Although different to phase synchronisation, a flat tone can be seen as a strategy that singers might use to blend and synchronise at a different level. This is given that a flat tone in a vibrato time series involves less variation of the fundamental frequency. This would lead to an increase in the recurrence rate, as the fluctuations become smaller than the predefined threshold for recurrence analysis. This might be a mathematical way to conceptualise why intuitively choral practitioners sometimes advocate for a straight tone in order to improve blending. It is the equivalent of syncing by staying within a state, rather than recurring to a state.

As such, having a flat tone can then be regarded as another form of synchronisation. Vibrato is a tool that allows singers to express their own uniqueness, and by removing this uniqueness, probability of fundamental frequency synchronisation increases. By adding vibrato in group singing, singers may or may not synchronise depending on various aspects. But by removing vibrato, this uncertainty gets reduced and a type of static synchrony is achieved by locking on a single f_o to optimise blend.

Regarding vibrato sync, Participant 16 reported: “I am listening to her tuning: It’s weird to try and match up her exact vibrato. Usually, because she has quite a lot on, what I would do in a more choral sense is I would back it with something less. [...] I would match her tone exactly and her tuning exactly. [...] I wouldn’t try to match the vibrato [...] because I can also hear phase issues. It’s not phasing but you can hear different harmonics going in and out and it’s affecting the overall timbre.” It might be plausible that this ability to perceive phase information and harmonics could have influenced the highest phase synchronisation scores that this participant produced.

As mentioned, participants were not explicitly instructed to match vibrato, but only to blend and match *the other singer*. As such, the different strategies that singers would unconsciously (or consciously) try to use to increase blend may vary from singer to singer: For instance, try to synchro-

nise vibrato, or try to “neutralise” it with a straight tone. The description provided by this participant also urges the inclusion of perceptual studies in future research. We can speculate, that for multiple vibrato tones containing any type of entrainment or coupling, potentially “if executed correctly, vibrato draws no attention to itself, but rather enhances the quality of the sound” [20], analogue to what Titze described about solo vibrato tones.

This study only considered the blending of duo pairs, whilst group singing would usually involve multiple singers which may change the approach of the singers to blending. By adding vibrato, a couple of singers, as was the case in our exercise, might be able to synchronise to each other if the coupling between them is high enough. But statistics and complexity systems would point out that the higher the number of singers trying to synchronise, the lower the probability (and difficulty to analyse). This is again due to the complex nature of the voice and multiple singers. The higher the number of participants singing together, probably the higher the diversity. This might also contribute to why choral directors often advocate for a straight tone. If they have 8 singers trying to sing in unison and expressing their uniqueness in the form of a complex vibrato, the probability of synchronisation is still unknown—although extrapolating from our results, we could hypothesise that in theory some degree could be possible. On the other hand, having a straight tone could allow, at least temporarily, reduced individuality and therefore blend into a single note, where the group can merge into a sound larger than the individual voices.

In high numbers, probably choral practitioners have a good argument in the sense of trying to reduce vibrato in big choirs, where a cohesive sound of straight tone will blend better—this may be likely explained with the concept of auditory roughness (many singers singing with pronounced, but non-synchronised vibrato) although this is out of the scope of this study. However, in other types of group singing, like duet singing, or even chamber choirs, vibrato, and synchronisation of that vibrato might be viable in achieving blend. It is not uncommon to hear a clear expression of vibrato in such cases, particularly opera chorus singers singing together in duets. This

is also true for large opera choruses singing with vibrato all the time, as their aesthetic goal is not so much for blending but more about being heard over an orchestra and multiple vibrato tones might be beneficial in this case. This is speculation that awaits future research.

In contemporary music, the studio recording technique of vocal doubling has historically allowed singers to record (and potentially adapt/blend) multi-tracks with multiple vibrato waves. Although pronouncedly different than traditional choral aesthetics, this phenomenon might be relevant to the study and understanding of vibrato dynamics and synchronisation. As an example, in 1974's record *Fly to the Rainbow* [183], the singer Klaus Meine delivers a double vocal performance on the track *Fly People Fly*, where it is possible to hear a multi-tracked vibrato-rich harmony. Another stylistic example of vibrato harmonies has been applied by doubling the vocals only in specific sections of a song—to emphasise or add depth to the vocals dynamically across the piece. In the vocal recording of the song *Children of the Sea* by the artist Black Sabbath [184], the choruses of the song are accentuated by multiple, multi-tracked vibrato notes. Another example, and perhaps an archetypal model of this sound, is the singer Layne Staley, who was known for his extensive use of multi-tracking combined with a natural vibrato tone. A particular instance can be found in the song *Rotten Apple* from 1994's record *Jar of Flies* [185], where the whole vocal recording feels like a large layer of waves moving together, due to the superposition of several doubling vocals and harmonics with vibrato. This artist is also known for the vocal harmonies between the singer and the guitarist, for which the guitarist described: “Two voices make one. [... For] the blend, we were always singing together.” This duet is interesting as it presents one of the possible scenarios that can occur during blended singing: Staley usually sang with vibrato, while the lower-voice guitarist, Cantrell, sang with a straight tone—as exemplified in the song *Down in a Hole* from their record *Dirt* [186]—which creates their distinguishable sound. When discussing singer vs. guitar in a contemporary music band, Freddie Mercury and Brian May from Queen can also be highlighted: Mercury's vocal vibrato has been subject of previous research [85], which was often followed / accompanied by May's guitar vibrato

during solos in their music. Lastly, as a final exemplar, Mariah Carey in her monumental record *Emotions* (1991) [187] demonstrates the sound of vibrato in both solo singing and backing vocals in the female voice. For instance, the very last note of *Can't Let Go* highlights this phenomenon, starting in F \sharp 6 before being harmonised by (1) low-frequency backing vocals and (2) a high B6 that transitions to C \sharp 7—all notes adorned with Carey's world-class vibrato. However, the presence and relevance of vibrato synchronisation and blending in this context of contemporary music also awaits future research.

One key difference from this type of vibrato superposition and the potential to blend is that studio singers tend to be more used to singing and tracking over headphones. In our study, Participant 20, who reported experience in choral singing as well as jazz solo singing, described: "I'm so much more used to when I am being recorded having headphones on—but if you were with another singer in the room, it wouldn't be coming to you through headphones, so it is more like having somebody there."

Further questions regarding the role of vibrato in blended singing remain unanswered—especially the impact of vibrato synchronisation on the auditory perception of blend. As mentioned in the introduction section, detecting and quantifying vibrato synchronisation from an acoustical analysis perspective is only the first step towards a better understanding of its influence over blend. Blend can be defined as a psychoacoustical aspect. As such, the relevance of phase synchronisation of vibrato in its complex forms—for instance, in phase and out of phase—still needs to be addressed in future work focusing on the perceptual and psychoacoustic assessment. Additionally, with the analysis presented, it cannot be determined whether a phenomenon of "synchronisation establishment period" exists (such that vibrato synchronisation would require a few cycles to being established, due to auditory and neuronal feedback occurring within the singer).

5.5.1 Limitations

One limitation of the study was that the notes that were recorded from the participants were exported and processed manually by the first author. This

might have introduced human error, although great care was taken regarding the selection for analysis to include only the steady state of the note. Another limitation is that blending can be affected by other factors such as pitch, vowel quality, and formants, which were not measured (but controlled) in this study. Notably, amplitude vibrato (which can assume rates of twice the frequency vibrato [188]) could have influenced the participants' vibrato characteristics in the study. Additionally, this study found that singers can adapt their vibrato rate and extent, but the mechanism for how they do it (whereas consciously or unconsciously) is still unknown. Another limitation is that this experiment included only high voice, female choral singers. As such, these results are only representative of this particular cohort.

The regularity of the vibrato from the pre-recorded singer was an important aspect. In order to provide a repeatable, and good quality model stimuli for participants to adapt to, we recruited a choral singer with years of experience in operatic singing and choral singing, and high level of education/teaching/academic knowledge of vibrato. Thus, the pre-recorded vibrato stimuli presented high regularity, according to analysis of individual notes. However, given the complex nature of vibrato, the stimuli did not present perfect periodicity. This is, however, something that was embraced as part of the ecological design as vibrato is naturally dynamical and complex. Future studies could consider integrating synthesised stimuli using sine waves or voice synthesis with perfect periodicity to compare against the more chaotic-like human resonances. The addition of multiple stimuli singers would also be beneficial, as opposed to altered conditions of vibrato, as well as analysis of variation as a function of musical pitch.

Another potential limitation is that the pre-recorded singer's vibrato rate is slightly on the high side of values reported in the literature. If vibrato can indeed be explained by a neuronal feedback-reflex-model [20], different individuals will be "wired" differently, and perhaps for some participants these high modulation rates were outside the achievable value range.

The selection of a speaker was reported by some participants to be difficult to hear in the room, and a number of them reported being more used to

singing along to pre-recordings and tracks via headphones. The methodology chosen to assess vocal aspects in choral singing deserves attention and can influence the results. For instance, the difference between speakers, headphones, or live recording environments has not been fully established. In this study, the speaker was chosen due to the unwanted influence of bone-conducted self-feedback.

Although sound level was not the main interest of the study, the mean SPL (L_C) for all participants in the six singing conditions is reported. A limitation of the study is that the sound level reported for the pre-recorded singer in the Methods section cannot be compared to the ones from participants. This is given that the vocal recording from the pre-recorded singer was made using a studio microphone to provide a cleaner and higher quality recording, which was placed 1 meter from the singer, as well as the sound level calibration. Therefore, the SPL calculation that we present for the pre-recorded singer was taken at a different position than the ones from the participants, which were estimated at a close-mouth position. Nevertheless, the results from Table 5.11 are relevant to understand relative changes in sound level intra-participant between the solo and blending conditions. All singers showed a reduction of sound level in the Blending conditions relative to Solo, so a relationship between SPL and vibrato production might be possible. In other words, it may be that their results are the outcome of lowering SPL rather than blending, but causality is not possible to derive from the experiment. Related to this, additional studies can incorporate analysis of the Self-to-Other ratio: “Difference in decibels between the sound levels of Self and Other, as experienced by a given singer” [42].

5.5.2 Future Directions

Four main areas of future work have been identified from the findings of this study: To study the influence of the playback systems and the environments in which singers are better able to adapt, for instance, speakers versus headphones; to add a level of complexity by assessing two-way synchronisation between a pair of singers in a live-recording setting; to scale up the study and

analysis to multiple singers singing together to understand blending of larger groups; and lastly, to assess the relevance and implication of synchronisation using through perceptual testing.

An important finding in this study is that higher vibrato synchronisation metrics in response to blending were correlated with the intrinsic characteristics of vibrato measured when the participants sang solo. This means that having a regular vibrato increased the probability of a singer to be able to sync to the pre-recorded singer. This study treated solo singing and blended singing as having different mechanisms of production and goals. However, the experience and training of individual singers across the genres of choral and solo singing may be fundamental to how singers can adapt to each other in a blending scenario.

5.6 Conclusion

This study analysed vibrato signals from high voice choral singers in two different conditions of solo singing and blending singing. We found that singers do adapt their vibrato when blending to a pre-recorded solo stimulus, with evidence of phase synchronisation for most participants. The extent of adaptation seems to be influenced by the parameters of rate, extent, and jitter—synchronisation was higher the closer the participant’s baseline vibrato was to the stimulus’ rate and jitter, and there was a negative correlation between jitter and synchronisation. Whether synchronisation of vibrato tones between singers is a conscious or unconscious strategy to blend is still unknown to singing research, however we have presented further evidence of the presence of this phenomenon. This research is relevant for pedagogical practices across both solo singing and group singing and supports a need for future research to understand levels of entrainment that take place in group singing activities. This experiment utilises vibrato as a proxy for contributing to the wider understanding of the human side of sync and its potential benefits and applications in music.

Chapter 6

Vibrato in Blended Live Singing Duets

This chapter expands on the work presented in the previous chapter, which analysed vibrato of singers blending to a fixed repeatable stimulus, pre-recorded and artificially adapted. Once it was established that synchronisation can occur during multitracking blended vocal recordings, the next step of this project was to understand if these vibrato adaptations and synchronisation would also be present during the live interactions of choral singers in duet blended singing. Therefore, the experiment presented in this chapter assesses the live adaptations that can be produced by duet pairs of singers when they are performing and blending to each other live.

6.1 Introduction

The research question asked here relates to the hypothesis of this thesis that singers are able to synchronise their vibrato tones during choral blended singing. This question was tested by recording singing duet signals in a live blending context. It was hypothesised that singers singing in a live environment would be able to produce higher synchronisation values as compared to the pre-recording.

The last chapter showed that singers do adapt and modify their vibrato parameters of rate, extent, jitter, and shimmer when singing along to a pre-recorded stimulus. This was studied by comparing the vibrato metrics under a Solo singing condition and Blending conditions. It was also shown that high levels of synchronisation were reached between the vibrato from singing participants and the vibrato from a pre-recorded singer.

This chapter studies synchronisation introducing a two-way adaption environment, where not only one singer is trying to adapt and blend to a static stimulus, but rather both singers blend simultaneously. The hypothesis for the expected synchronisation values to be higher is based on the idea that two adaptive systems would be better than a single one.

6.2 Methods

This study was approved by the Data Protection and Physical Sciences Ethics Committee of the University of York (reference number Martinez20240402). All participants provided written informed consent to take part.

Participants that took part in the experiment described in Chapter 5 were invited to take part in this follow up experiment. They could complete this second phase on the same day as part one or on a later day as long as part one was completed first.

6.2.1 Duets

Since in this study the main aim was to record duets blending together whilst still maintaining control and isolation to be able to compare the results with the previous phase, we first recruited a singer that was able to participate in all of the experiments.

Singer Λ

To be able to replicate the methodology from the previous experiment to some degree, the singer that participated as the reference, pre-recorded singer was recruited for this second phase of this experiment. This reference singer, defined here as Singer Λ for clarity, participated in this next part of the experiment by singing along with each of the participants of the study but now in a live blending scenario. This singer had never sung with any of the participants before.

Participants

Once Singer Λ was recruited, invitation to participate in the study was made to the singers that had participated in the previous study, with the same following inclusion criteria:

- Had participated in the previous experiment
- Had at least two years of choral experience or group singing experience
- Alto, mezzo soprano, or soprano voice
- Comfortable singing in the E_4 to E_5 vocal range.

Before the experiments, singers signed a consent form and were given a £10 gift card as a thank you for their participation. Thirteen choral singers (along with Singer Λ) participated in the experiment, with characteristics shown in Table 6.1.

As such, each of the participants sang with Singer Λ during this experiment and was assigned a duet number defined in Table 6.2.

6.2.2 Musical Material

The music material used in this part of the experiment was kept consistent with the previous experiment and the same notes were used (Table 5.1). The only difference in this experiment was that 4 takes per note were recorded

Table 6.1: Descriptive characteristics of Singer Λ and the participants that conformed the duets in the study.

Characteristic	Singer Λ	Participants
N	1	13
Age (years)	42	18 to 61*
Biological sex	Female	Female
Known hearing loss	No	0
Ethnicity (white British European Asian white and Asian)	White British	8 1 3 1

Note: *Range.

Table 6.2: Assignment of duet number based on the singer labels from the experiment in Chapter 5, Table 5.10. (Singer Λ took part in all duets.)

Duet (N)	Participant (N)
1	2
2	3
3	6
4	7
5	8
6	10
7	11
8	14
9	15
10	16
11	19
12	22
13	23

(instead of 3 from the previous part). This was given that this experiment was much shorter compared to the previous one, and an extra take to have a total of 32 notes per duet was plausible without fatiguing the singers and would improve the statistical power of the study, with more data points per singer. Isolated notes were also held for a duration of 5 to 7 seconds on vowels /a/ and /i/.

6.2.3 Room and Equipment

The experiments took place in the Live Room and the Music Studios in the AudioLab, University of York. Recordings were made using REAPER software and Focusrite hardware. Sampling frequency was set to 48 kHz for all recordings. Both participants were placed 150 cm apart during the experimental session. A Genelec 8040B speaker was placed in front of the singers and was used to reproduce the notes as cues before each note.

Both singers wore a DPA microphone and an Lx attached to a preamp box, as described for the previous experiment (Chapter 5, Section 5.3.4).

6.2.4 Procedure

After giving a general overview of the experiment and signing the consent form, participants were given 5 minutes to warm up. A Nord Stage 2 keyboard in the studio was made available for the participants.

A randomised sequence of notes was presented through the speaker, and the participants were instructed to start singing and blending as soon as they heard the cue. Additionally, a click was played to cue the participants to stop singing after around 7 seconds, until they heard the next note cue. A total of 32 notes per singer were recorded.

6.2.5 Data Analysis

Vibrato

The vibrato analysis of the recorded notes was made following the same procedure described in Chapter 5. As such, this study also used the metrics of rate, extent, jitter, and shimmer of vibrato. The parameters were calculated from the vocal recordings from both Singer A and the choral participants.

In order to understand the amount of vibrato that participants and Singer A produced for this experiment, a new vibrato metric was defined as:

$$\text{Vibrato coefficient (unitless)} = \frac{\text{Cycles}}{\text{Length} \times \text{Rate}} \left(\frac{1}{\lambda \times (1/\lambda)} \right). \quad (6.1)$$

This is useful given that it reflects the amount of vibrato that was added across the whole length of a note. The units of *cycles/length* (1/second) are similar to that of the Rate (Hz), except that Hz is a standardised unit that reflects the amount of cycles or events in one single second and is derived from the period or wavelength λ . Dividing *cycles/length* by the rate provides a unitless coefficient defined here as the coefficient of vibrato.

The vibrato coefficient as utilised here is useful because it incorporates all of the cycles across a note. This provides information regarding vibrato density independently from the duration of the note. As such, this metric should in theory have values ranging from 0 to 1, or

$$0 \lesssim \text{Vibrato coefficient} \lesssim 1. \quad (6.2)$$

A value of 0 would be an indication that the algorithm did not detect any vibrato cycles within the note. And a value of 1 would be for instance found on a note with vibrato that was added across the whole length of a note.

To illustrate this, we can consider first the term *cycles/length* of the equation and two vibrato notes with a rate of 5 Hz but different lengths and dynamics: If a singing note is one-second long and contains 5 vibrato

cycles, $cycles/length$ would be 5 (which in this case would be equal to rate = 5 Hz as the length is 1 second). But if another singing note is two-seconds long and contains 5 vibrato cycles which were produced only during the second half of the note (so the first half being a straight tone), then $cycles/length = 5/2 = 2.5$, whilst vibrato rate is still 5 Hz. In other words, if a note has vibrato for the whole note, then $cycles/length \approx rate$. So diving this expression by the rate provides a normalised coefficient that captures a ratio of how much vibrato was detected across the note.

As such, this coefficient would complement the other vibrato metrics to understand how much vibrato the note contains. This new metric was necessary to introduce given that some duets sang with less vibrato. This was different to the previous experiment, which was biased towards more vibrato production by the pre-recorded vibrato notes with vibrato present consistently throughout notes.

Recurrence

The analysis of recurrence and the CPR coefficient was performed using the same procedure described in Chapter 5. The main difference is that in the previous analysis, the CPR was calculated between the participant signal and the pre-recorded signal. For this experiment, CPR is calculated from the signals of Singer A and the participant, singing together simultaneously.

6.2.6 Statistical Tests

Normality test were performed on the metrics from the 13 singers using the Anderson-Darling test in MATLAB. As the test failed to provide evidence against normality, the mean values are used. Additionally, a *t*-test analysis was performed between the values obtained from the participants singing live and those obtained from the previous experiment when they sang in their solo condition. This is important as it illustrates the degree of adaptation that the singers experienced singing live when adapting to another singer as compared to their solo voice.

6.3 Results

6.3.1 Vibrato Metrics

Table 6.3: Rate live adaptations.

Duet	Singer Λ (Hz)		Participant (Hz)	
	Rate	Δ to Pre-recording	Rate	Δ to Solo
1	5.5	-0.2	5.5	-0.3
2	5.5	-0.2	5.4	-0.5
3	5.5	-0.2	5.5	0.5
4	5.5	-0.2	6.8	0.6
5	5.0	-0.7	4.3	-0.1
6	5.7	0.0	5.8	0.2
7	5.5	-0.2	4.7	-0.1
8	5.3	-0.4	5.4	0.0
9	5.5	-0.2	5.0	0.0
10	5.5	-0.2	5.1	-0.9
11	5.3	-0.4	5.1	-0.3
12	5.4	-0.3	5.2	0.2
13	5.1	-0.6	5.1	-0.3
Mean	5.4		5.3	

Table 6.3 shows the rate values obtained from both the Singer Λ and the participants. Additionally, the columns showing the delta values indicate a difference to (1) the pre-recording values from Singer Λ and (2) the Solo values for the participant. These Δ to Pre-recording values indicate that the Singer Λ experienced a reduction of her rate when singing and blending live to the participants. The last row of Table 6.3 shows that the mean rate for Singer Λ was 5.4 Hz, and the pre-recording value obtained in the previous experiment was Rate = 5.7 Hz.

Table 6.3 shows that the changes in rate from the participants were increases in some cases and decreases in others. For instance, Duet 3, Duet 4, Duet 6, and Duet 12 showed an increase of rate relative to their solo voice. Regarding absolute values, the participant from Duet 5 had the lowest rate (4.3 Hz), for which the Singer Λ responded with a reduction of 0.7 Hz (max

reduction across participants) in her rate compared to the pre-recording values.

Table 6.4: Extent live adaptations.

Duet	Singer Λ (cents)		Participant (cents)	
	Extent	Δ to Pre-recording	Extent	Δ to Solo
1	33	-45	34	-13
2	39	-39	64	-50
3	37	-41	27	5
4	38	-40	40	9
5	30	-48	23	-4
6	51	-27	68	-3
7	34	-44	55	-23
8	35	-43	29	-10
9	46	-32	79	1
10	34	-44	31	-45
11	33	-45	21	-13
12	37	-41	27	4
13	26	-52	26	-1
Mean	36		40	

Table 6.4 shows the extent values obtained from the Duets. Similar to rate, Singer Λ showed a consistent reduction of extent across all participants. The magnitude of this reduction was high, ranging from -32 cents and -52 cents of reduction. The mean value from the pre-recording notes was 78 cents, whereas the mean value during live blending was 36 cents (last row of the Table). The results from the participants in the duets show that increases and decreases of extent were produced. This varied dramatically between individuals, and the Δ to Solo ranged from -50 cent decrease to 9 cent increase.

Results from the jitter and shimmer (Tables 6.5 and 6.6) show that Singer Λ produced vibrato with less regularity singing in the live blending condition as compared to the values from the pre-recorded vibrato. The mean percentage values were 31% for jitter, and 43% for shimmer. As for the participants in the duets, it is worth noting that jitter was the only metric that was sta-

Table 6.5: Jitter live adaptations.

Duet	Jitter	Singer Λ (%)		Participant (%)	
		Δ to Pre-recording	Jitter	Jitter	Δ to Solo
1	30	14		35	11
2	32	16		23	8
3	29	13		42	-1
4	31	15		34	-4
5	39	23		43	7
6	29	13		23	3
7	31	15		33	4
8	26	10		36	11
9	24	8		28	5
10	35	19		40	24
11	30	14		49	7
12	27	11		51	4
13	39	23		38	7
Mean	31			36*	

*Hypothesis of t -test equal means (blending vs solo) rejected at $p \leq 0.05$.

Table 6.6: Shimmer live adaptations.

Duet	Shimmer	Singer Λ (%)		Participant (%)	
		Δ to Pre-recording	Shimmer	Shimmer	Δ to Solo
1	42	6		45	7
2	42	6		35	3
3	42	6		43	8
4	45	9		55	5
5	42	6		43	-6
6	47	11		41	0
7	44	8		53	3
8	44	8		45	4
9	41	5		40	3
10	44	8		39	7
11	43	7		37	-16
12	43	7		47	12
13	43	7		42	6
Mean	43			43	

tistically different from the solo values. This means that the average value that singers produced in live blending, was statistically different from that of their solo voice. This difference corresponds to an increase of jitter, given that all participants showed an increase of jitter relative to solo (Δ to Solo), with the exception of Duet 3 and 4.

Table 6.7: Vibrato coefficient live adaptations.

Duet	Singer Λ		Participant	
	Coefficient	Δ to Pre-recording	Coefficient	Δ to Solo
1	0.832	-0.151	0.813	-0.078
2	0.850	-0.132	0.822	-0.171
3	0.854	-0.129	0.672	0.139
4	0.866	-0.116	0.897	0.058
5	0.696	-0.286	0.582	-0.162
6	0.916	-0.067	0.901	-0.015
7	0.815	-0.168	0.920	-0.059
8	0.843	-0.139	0.786	-0.100
9	0.894	-0.089	0.913	0.002
10	0.810	-0.173	0.651	-0.277
11	0.868	-0.115	0.537	-0.108
12	0.853	-0.129	0.573	0.070
13	0.697	-0.286	0.464	-0.171
Mean	0.830		0.733	

Table 6.7 shows the obtained vibrato coefficients from the singing duets. Results show that Singer Λ sang with less vibrato during live blending as compared to the pre-recording (Δ to Pre-recording values). In one case, Duet 13, her vibrato coefficient went as low as 0.697, which means that on average, the notes contained vibrato waves only in around two thirds of the entire notes. This also corresponds to a reduction, or difference, of -0.286 from her solo voice during pre-recording, which had a mean vibrato coefficient of 0.985 (presence of vibrato for almost all the duration of the note).

The participants, on average, produced notes with less vibrato than Singer Λ (mean of 0.733 versus mean of 0.830, correspondingly). The participant that presented with the greatest decrease of vibrato coefficient compared to

Solo singing was from Duet 10. This corresponded to a vibrato coefficient decrease of -0.277 (from 0.928 in Solo to 0.651 in Live Blending).

6.3.2 Synchronisation

Table 6.8: Phase synchronisation in the live context and the unedited stimuli values for comparison.

Duet	Unedited pre-recording CPR	Live CPR	p
1	0.407	0.483	
2	0.445	0.470	
3	0.052	0.207	
4	-0.023	0.101	
5	-0.041	0.139	
6	0.287	0.440	
7	0.122	0.187	
8	0.281	0.464	
9	-0.009	0.190	
10	0.451	0.278	
11	0.132	0.253	
12	0.157	0.170	
13	0.139	0.421	
Mean	0.184	0.293	0.004*

*Hypothesis of t -test equal means rejected at $p \leq 0.05$.

Table 6.8 shows the phase synchronisation coefficients obtained from the notes of Singer A and the participants blending together. Additionally, the table presents the CPR values that the corresponding singers produced during the experiment described in Chapter 5. As such, these two columns described the degree of synchronisation between the two singers in the duet but one produced in a live singing scenario, and the other one using the pre-recording.

Results show that all singers produced higher levels of synchronisation in the Live singing blending condition compared to blending to the pre-recorded notes. (The increase in mean CPR was statistically significant as shown in the last column and row of Table 6.8.) This was true with the exception of Duet

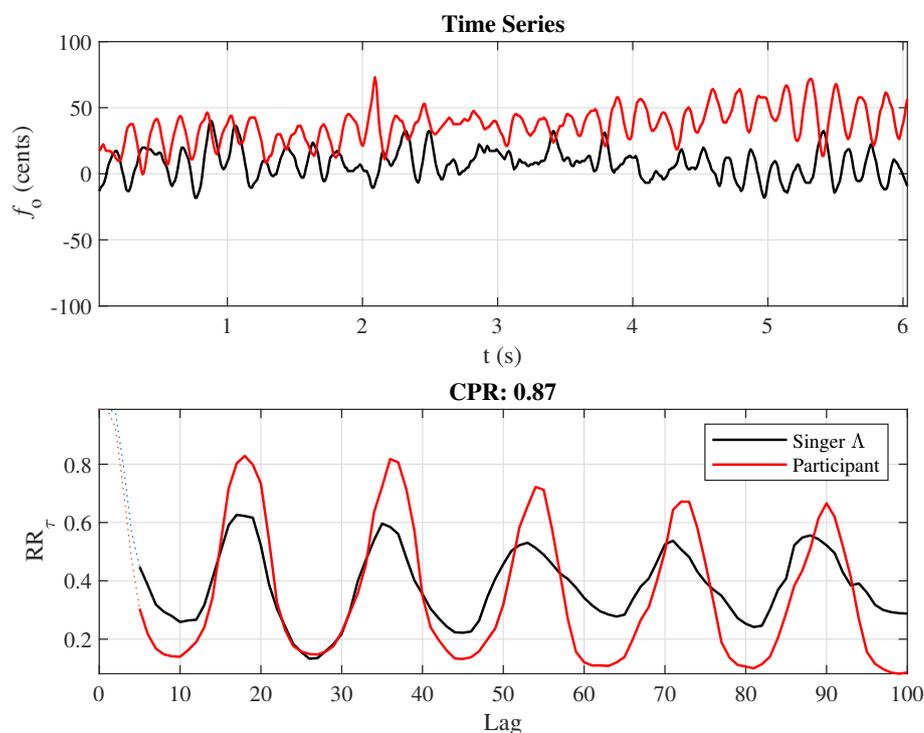


Figure 6.1: High sync with pitch drifting.

10, which actually presented an almost twofold reduction of synchronisation during Live blending. It's worth nothing that Duet 10 also presented the highest decrease of vibrato coefficient as described for Table 6.7.

From the values obtained from CPR, a number of special cases and scenarios are valuable to illustrate using the graphic visualisation of RR_τ .

High Sync and Pitch Drifting

The paired vibrato notes shown in Fig 6.1 correspond to the highest synchronisation value obtained in this experiment ($CPR = 0.87$). It is also noticeable that the mean pitch trajectory from the participant increased towards the end of the note, whilst the vibrato from Singer A kept oscillating around 0 cents.

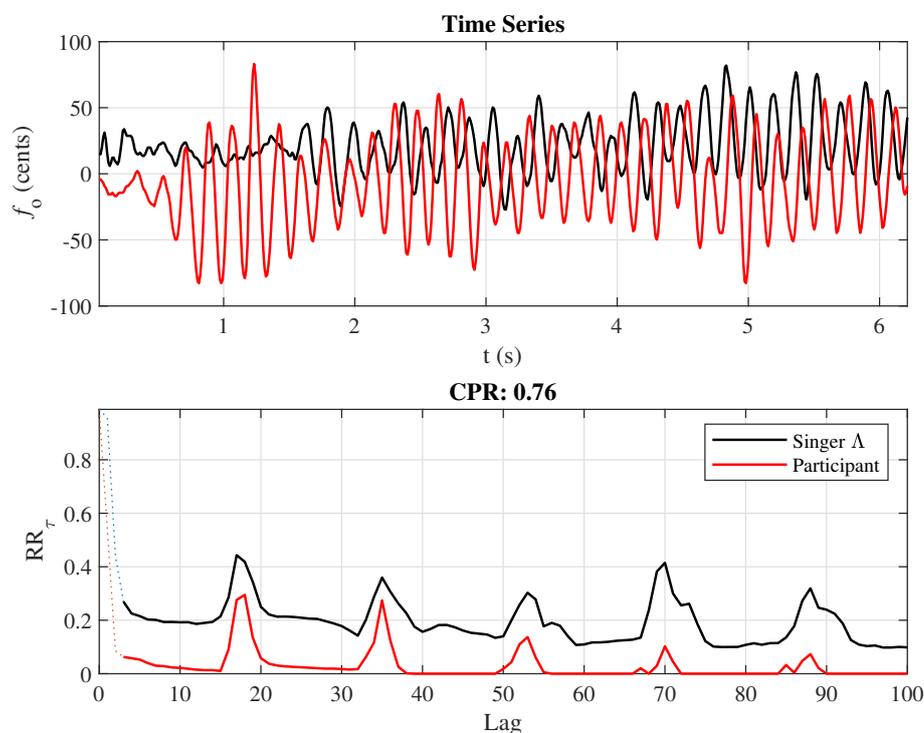


Figure 6.2: A singer leading vibrato production.

Leading Vibrato Singer

Figure 6.2 shows a pair of vibrato tones from Duet 2. This Duet also produced high levels of synchronisation. And this particular note shows a phenomenon relevant to vibrato dynamics, where it is possible to see that the participant starts the vibrato production, and Singer A joins only after a couple of seconds. It is also noticeable that right at the moment when Singer A starts to add vibrato, both signals present almost complete synchronisation, but only for two or three cycles.

Duet 10

Duet 10 presents an interesting case that is worth highlighting. The participant from Duet 10 corresponds to Participant 16 in the previous study of multitracked blending, who was also the singer that achieved the highest

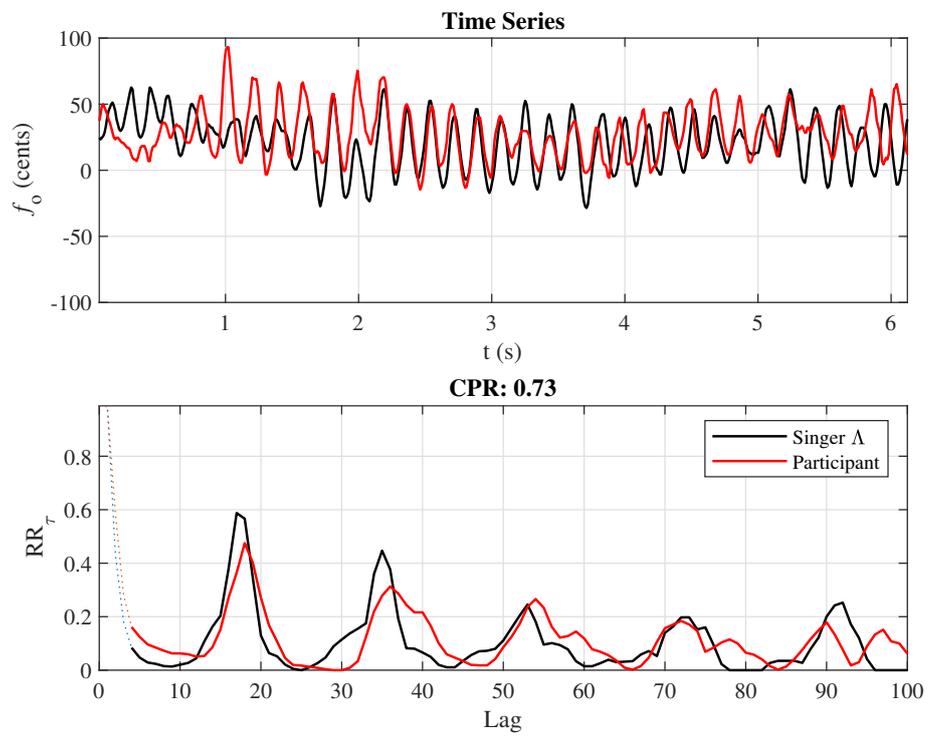


Figure 6.3: High sync blending.

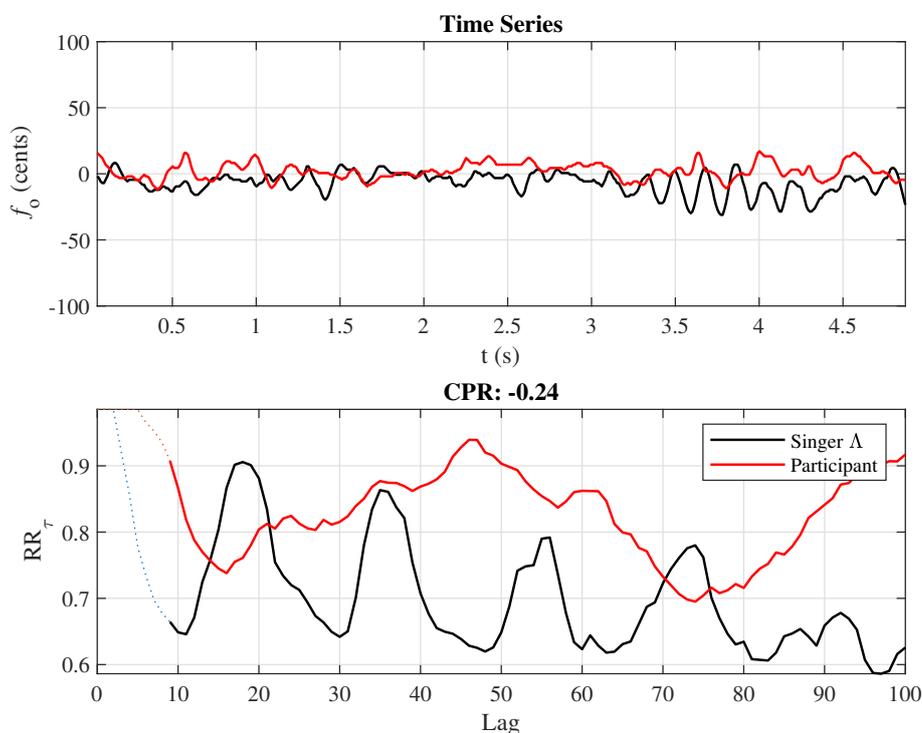


Figure 6.4: Both singers singing with a straighter tone.

mean synchronisation in the whole study. As such, from the previous study, it was hypothesised that this participant has a high ability to synchronise. However, the results in the present experiment of Live Blending show Duet 10 was the only duet that experienced a decrease of CPR as compared to the multitracked blending (Table 6.8).

Regardless of this average value presenting a decrease of CPR, looking at the individual notes, it was found that this singer/pair still produced notes with high levels of synchronisation (Figure 5.5) of up to $\text{CPR} = 0.73$. However, for other individual notes, some vibrato dynamics were found, with lower and even negative CPR values.

Figure 6.4 shows a pair of notes with a CPR of 0.24. It is possible to observe that this note contained almost no vibrato at all. And since $p(\epsilon, \tau)$ is normalised-to-zero mean and standard-deviation-of-one independently, the RR_τ look pronouncedly different to each other. This agrees with the results

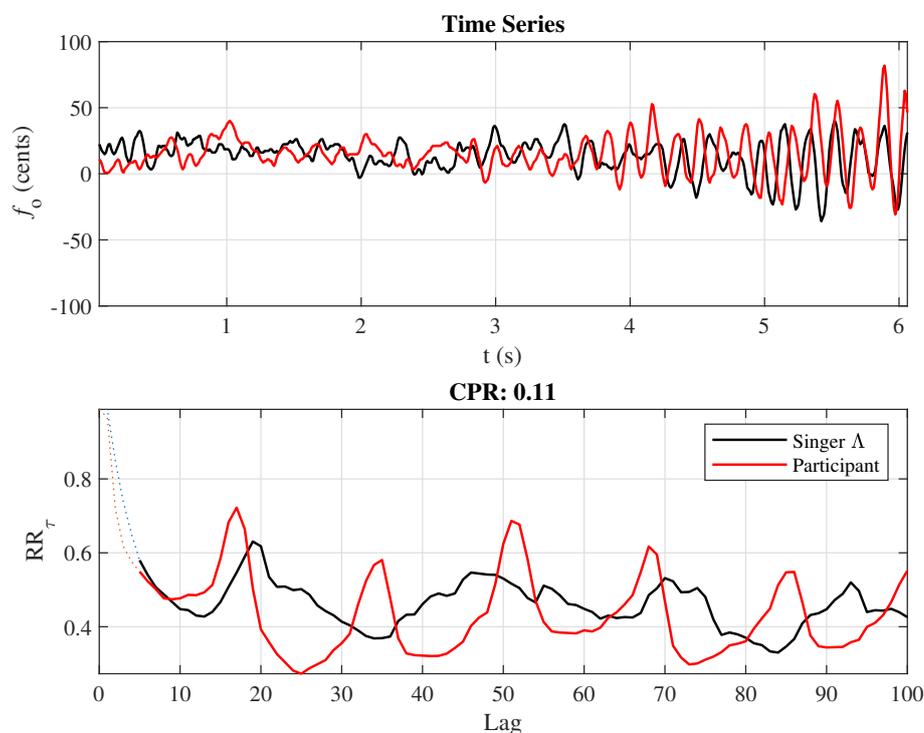


Figure 6.5: Both singers adding late vibrato.

from Table 6.7, that show that the participant from Duet 10 sang with less vibrato than in her solo voice resulting in a great reduction in her vibrato coefficient.

Another example is shown in Table 6.5. In this case, both singers sing with vibrato, but only until the end of the note, for around 2 seconds. In terms of vibrato coefficient, this would be equivalent of a value of 0.3, given that vibrato was added only for around one third of the whole note. This note also presented a low CPR value (0.11).

Figure 6.6 illustrates another possible blending scenario, where only one singer (Singer Λ) in Duet 10 sang with vibrato, but the participant maintained a straight tone. The RR_τ shows an offset in the participant signal given that a straight tone would increase recurrence as the oscillations become smaller than ϵ ; however, that did not increase CPR. This phenomenon is interesting to study using a phase space (Figure 6.7), and plotting both

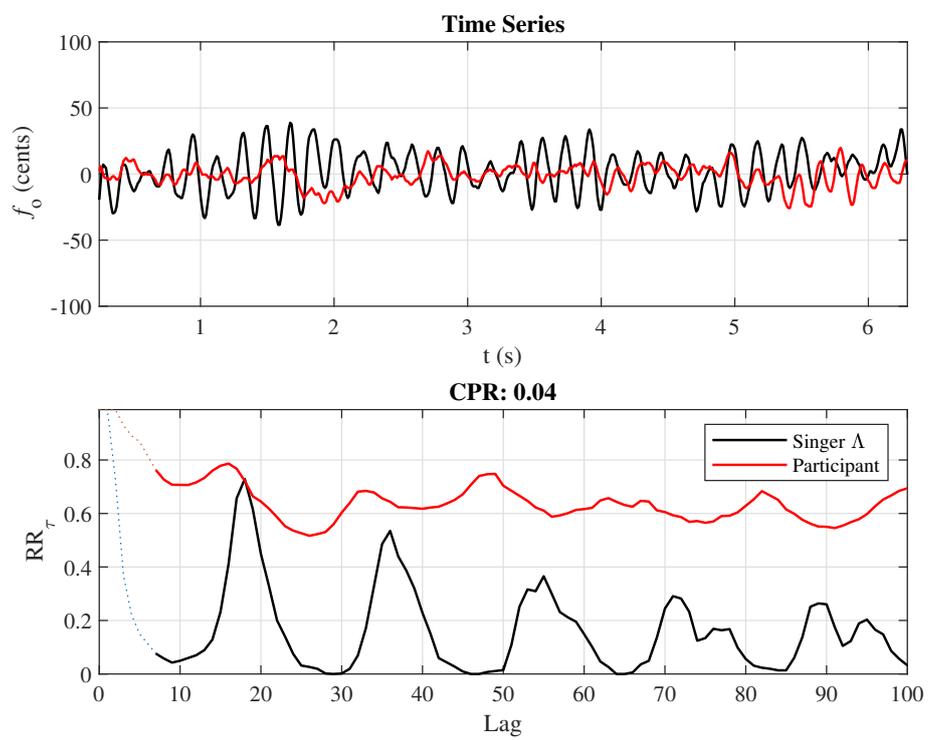


Figure 6.6: Participant singing with a straight tone whilst Singer A added vibrato.

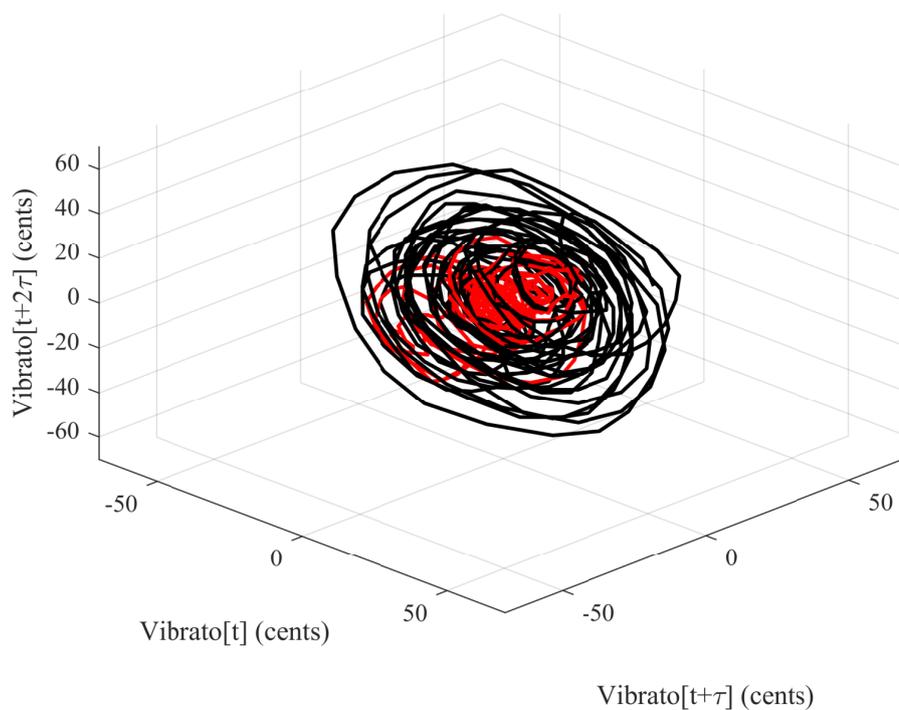


Figure 6.7: Attractor from a straighter tone engulfed by an attractor with higher extent, depicted on the same phase space.

trajectories on it, from which recurrence plots and $RR\tau$ is derived from. It is worth mentioning that from the previous experiment, this participant mentioned during informal qualitative reporting that in a choral scenario, she would “back off vibrato with something less”. In the phase space, it is possible to observe how the straight note from the participant stays in the centre of the attractor from Singer Λ . In the next section, the possible implications and perspectives from these results are discussed.

6.4 Discussion

The study presented in this chapter addressed whether any adaptations or synchronisations in vibrato waves would be observable when singers sing and blend with each other live. We found that the participants and Singer Λ

did present differences in the vibrato parameters of their voice as compared to the values obtained from solo singing. Specifically, a reduction of rate and extent, along with an increase of the regularity metrics of jitter and shimmer. From these only the jitter increase from participants blending live versus blending to a pre-recording were statistically significant.

Another finding of the study was that phase synchronisation coefficients were higher during the live blending of the duets in a two-way adaptation scenario, as compared to blending to a pre-recording. This agrees with the hypothesis of the study, which was that singers were going to be able to blend and adapt better whilst singing together in real time.

This study also found that the new metric, presented as the vibrato coefficient, was useful to determine the amount or “density” of vibrato that the notes contain. By utilising this metric, this study found that some singers sang with less vibrato during live blending than when singing on their Solo voice, but others actually increased their vibrato. Additionally, in absolute terms, the vibrato coefficients that participants produced ranged from 0.464 to 0.920. This is important as a value of 0.464 would indicate that vibrato was added only for half of the duration of the notes.

The rate of vibrato produced tells an interesting story that is best analysed as individual cases. Firstly, the mean values obtained from both the participants and Singer A were within the expected values commonly reported in the literature. The similarity between the rate values between each duet suggest that singers adapt in an attempt to improve blend. For instance, for Duet 3, both singers had a rate of 5.5 Hz. But the interesting point is that Singer A *reduced* her vibrato rate by -0.2 Hz, while the participant *increased* her vibrato rate by 0.5 Hz. That magnitude of change observed from the participant is not trivial, and although our experiment cannot prove causality, it can be speculated that the rates converged into that mid-land value of 5.5 Hz. Duets 1 and 2 present a similar case, where the singers were close to that rate value of 5.5 Hz, although in this case the participants experienced a decrease of rate relative to their solo rate.

The mechanistic reason for the extent values from Singer A presenting

dramatic reductions is difficult to unpick. It might be that lower levels of extent are better for blend in a choral context. Such low extent values are different than, for instance, those found in operatic singing, where it would not be uncommon to find values of 200 cents [2], [119].

As this study is interested in the blending strategies that singers would utilise when singing together. This poses questions regarding vibrato production and perceptual blending. A set of possible scenarios was illustrated using the waveforms and derived RR_τ plots. The concept of a “leader” in group singing has already been explored and described in the study by Daffern 2017 [53]. The work reported in this chapter presents additional evidence of this phenomenon in live blended singing. These results provide further evidence of vibrato synchronisation in a live setting context; and perspectives are discussed in the next and final chapter of this thesis.

6.5 Conclusion

This chapter shows that singers blending together in real time do adapt and modify their vibrato. The rate of vibrato produced between the singers was similar between the duets. A number of possible scenarios that can occur during blended singing was presented. Future studies can identify how the different dynamics of vibrato and synchronisation values relate perceptually to blend.

Chapter 7

Conclusion

The main hypothesis that was considered throughout this thesis is as follows:

Singers adapt their vibrato when aiming to achieve blend with another singer, often resulting in synchronisation as observed through the application of nonlinear dynamics and recurrence analysis.

This thesis presents novel contributions to the field of voice science, nonlinear dynamics, and synchronisation. The implications give insights into understanding mathematical methods of dynamical systems but also understanding coordinating behaviours in humans, specifically in music and singing. The overall conclusions derived from the chapters in this thesis do suggest that vibrato signals contain variations in fundamental frequency that can be highly complex and are rarely periodic. Results from the analysis of nonlinear metrics and comparison with other deterministic signals revealed that vibrato signals can present high irregularity, and rich complexity. Moreover, the results from Chapter 5 show that metrics of recurrence can also be used to mathematically assess synchronisation between pairs of vibrato signals, with results showing that singers can synchronise with each other when blending together.

The methods used throughout this thesis provide insights into the further development and establishment of understanding vibrato regularity in voice

science. In this work, close microphones and Lx recording devices were used to extract the fundamental frequency of voice signals in a multi-voice setting.

It is important to point out that there are limitations to the algorithm used for the estimation of fundamental frequency. These limitations become particularly relevant for studies that aim to address, for instance, intonation and pitch accuracy in the singing voice. In the analysis of vibrato, it could be mentioned that relative changes in the fundamental frequency are more relevant than the absolute values obtained from the algorithms. But future studies can explore the role and accuracy of different algorithms for the vibrato analyses described in this thesis.

Likewise, another key aspect that deserves attention in future research is the development of robust and systematic methods for vibrato tones selection. In this work, it is imperative to note that there was a human factor in the vibrato tone identification and selection. This reduces the potential for reproducibility, and a caveat of our results is that they might change depending on this aspect.

As described in Chapter 2, frequency modulation is only one of three aspects that conform vibrato in the singing voice. As such, future studies focusing on the amplitude modulation and timbre characteristics of superimposed vibrato tones can provide further insights.

In conclusion, the results from this thesis show that singers can adapt and blend to each other and this can be observed through vibrato synchronisation. This brings insights into the mechanistic reasons of how group activities bring a sense of connection between participants.

7.1 Further Work

The conclusions from this thesis also offer potential research avenues to be explored after the presentation of our results.

- The choice of window and selection criterion for *steady vibrato* as defined in this thesis can heavily influence the analysis results, partic-

ularly those from nonlinear dynamics. This thesis therefore has the limitation that the steady vibrato phonation selection was performed manually and has not been defined reproducibly, which presents an area of improvement for future research.

- Regarding vibrato complexity, the perceptual significance is still unknown and awaits for future research.
- The work of synchronisation can be extended to a higher number of participants, to further explore the potential of group singing and the wellbeing benefits that might come from interconnectedness of multiple participants. This would involve more challenging mathematical methods, such as graph theory or connectivity metrics measured between pairs after the extraction of synchronisation metrics between individual nodes.
- Vibrato could be described as a frequency modulation concept, which could open up way to a more rigorous mathematical modelling—and potentially using a dynamical system that integrates the plausible nonlinear interaction between the power source, glottal source, vocal tract filter, and neural mechanisms associated with vibrato production.

7.2 Final Remarks

The purpose of this thesis was to explore the potential usage of complexity metrics applied to vibrato, and to test the value that might come from analysing vibrato synchronisation in this area. It is acknowledged that the conclusions of this thesis have only considered a small part of a wider phenomenon that vibrato synchronisation can produce, and this thesis has explained into further areas and aspects that can be explored in future research.

Appendix A

Systematic Review Data

This appendix shows the complete records and data of articles that were reviewed for a systematic review of vibrato in group singing, (Tables A.1, A.2, and A.3). The tables provide characteristic information about the papers, and the categories and methodologies employed by them.

Table A.1: Methodology approaches of studies included in the review.

Study	Synthesis	Multi-track		Stereo/Binaural	Lx/Contact
		Prerecording	Live		
Coleman 1994				×	×
Culling & Summerfield 1995	×				
Daffern 2017			×		×
Dai & Dixon 2019		×			
Dromey <i>et al.</i> 2003		×			
Duncan <i>et al.</i> 2000				×	
Erickson <i>et al.</i> 2012	×				
Goodwin 1980		×			
Jers & Ternström 2004			×		
Le Beux <i>et al.</i> 2011	×				
Mann 2014				×	
McAdams 1989	×				
Reid <i>et al.</i> 2007		×			
Rossing <i>et al.</i> 1987		×		×	
Sacerdote 1957				×	
Ternström <i>et al.</i> 1988	×				
Ternström & Sundberg 1988	×	×		×	×
Total	6	6	2	6	3

Table A.2: Objective category of each study.

Study	Adjustment	Identification	Interaction	Intonation	Style
Coleman 1994					×
Culling & Summerfield 1995		×			
Daffern 2017			×		
Dai & Dixon 2019				×	
Dromey <i>et al.</i> 2003	×				
Duncan <i>et al.</i> 2000			×		
Erickson <i>et al.</i> 2012		×			
Goodwin 1980					×
Jers & Ternström 2004			×		
Le Beux <i>et al.</i> 2011				×	
Mann 2014					×
McAdams 1989		×			
Reid <i>et al.</i> 2007					×
Rossing <i>et al.</i> 1987					×
Sacerdote 1957					×
Ternström <i>et al.</i> 1988					×
Ternström & Sundberg 1988				×	
Total	1	3	3	3	7

Table A.3: List of assessed records and exclusion criteria.

Study	No vibrato	No choral singing	No scientific paper	Other
Electronic databases				
Berghs <i>et al.</i> 2013 [76]		×		
Reddy & Subramanian 2015 [24]		×		
Howard 2018 [189]		×		
Takasu 1997 [190]		×		
Marchini 2014 [191]	×	×		
Wooding & Nix 2016 [146]		×		
Dai & Dixon 2019 [192]	×			
Nix 2015 [25]		×		
Maher & Beauchamp 1994 [193]	×	×		
Key journals				
Ravi <i>et al.</i> 2019 [194]	×			
Kirsh <i>et al.</i> 2013 [162]	×			
Howard <i>et al.</i> [75]	×			
Nix <i>et al.</i> 2016 [150]		×		
Duvvuru & Erickson 2016 [165]		×		
Guzman <i>et al.</i> 2012 [19]		×		
Ternström 1991 [49]				×
D'Amario <i>et al.</i> 2020 [78]	×			
Sacerdote 1956 [195]				×
Brown 1993 [50]				×
Dai & Dixon 2019 [196]	×			
Grell <i>et al.</i> 2009 [197]	×			
Mauch <i>et al.</i> 2014 [73]	×			

Table A.3: (cont.)

Study	No vibrato	No choral singing	No scientific paper	Other
Other sources				
Sublett 2019 [157]			×	
Olson 2008 [198]			×	
Nix 2014 [36]			×	
Galante 2011 [39]			×	
Pórolniczak & Kramarczyk 2013 [52]		×		
Wyatt 1967 [37]			×	
Sundberg 1988 [199]			×	
Katok 2021 [40]			×	
Titze <i>et al.</i> 1994 [15]		×		
Titze <i>et al.</i> 2002 [20]		×		
Carter <i>et al.</i> 2010 [122]			×	
Culling & Darwin 1993 [200]	×	×		
Dejonckere <i>et al.</i> 1990 [81]				×
Reference lists				
Sundberg & Lindqvist 1973 [201]	×	×		
Rossing <i>et al.</i> 1986 [74]	×			
Howard 2007 [65]	×			
Howard <i>et al.</i> 2013 [51]		×		
Troup 1981 [202]	×	×		
King & Horii 1993 [203]				
Culling & Darwin 1993 [204]	×			
Demany & Semal 1990 [205]		×		
Marin & McAdams 1996 [206]		×		
McCoy 2011 [41]			×	
Ekhholm 2000 [155]	×			
Ford 2003 [207]	×			
Howard <i>et al.</i> 2012 [121]		×		
Daugherty 1999 [208]	×			
Daugherty 2003 [209]	×			
Daugherty <i>et al.</i> 2013 [210]	×			
Hsiao <i>et al.</i> 1994 [211]		×		
Sundberg <i>et al.</i> 1998 [212]		×		
Rasch 1978 [213]		×		
Gardner <i>et al.</i> 1989 [214]		×		
Carlyon 1991 [215]		×		
Carlyon 1992 [216]		×		
Carlyon 1994 [217]		×		
Darwin & Culling 1990 [218]			×	
Darwin 1981 [219]	×			
Chalikia & Bregman 1989 [220]	×			
Chalikia & Bregman 1993 [221]	×			
Wilson <i>et al.</i> 1990 [222]		×		
Cohen & Chen 1992 [223]	×	×		
Darwin <i>et al.</i> 1994 [224]		×		
Moore <i>et al.</i> 1985 [225]	×	×		
Zwicker 1984 [226]	×	×		
Assmann & Summerfield 1990 [227]	×	×		
Gardner & Darwin 1986 [228]		×		
Total	28	35	15	1

Appendix B

Vibrato Dynamics

In this appendix, a preliminary study on synthesis of vibrato waves is presented. For the synthesis, we focused on the use of a well known dynamical system that produces signals similar to sine waves but which dynamics are easy to bifurcate (or control) with the use of parameters. These results were presented as part of the following reference from an abstract in conference proceedings [229].

B.1 Introduction

Dynamics is a term that describes the evolution and variations over time of certain state variables that can be defined in terms of differential equations. A general system of differential equations can be defined as:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{B.1}$$

such that

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_m \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \tag{B.2}$$

where overdots denote differentiation with respect to t ($\dot{x}_m \equiv dx_m/dt$); $\mathbf{x} = [x_1, x_2, \dots, x_m]^T \in \mathbb{R}^m$ denotes the state vector of the system; and the coefficients $a_{m,n}$ with $(m, n) = [1, 2, \dots]$ in matrix \mathbf{A} denote the parameters of the corresponding state. For nonlinear systems, this is commonly expressed as:

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, \dots, x_m, t) \\ \dot{x}_2 &= f_2(x_1, x_2, \dots, x_m, t) \\ &\vdots \\ \dot{x}_m &= f_m(x_1, x_2, \dots, x_m, t) \end{aligned}, \quad (\text{B.3})$$

where f_1, f_2, \dots, f_m are functions defined in terms of system's state and the independent variable of time t . Thus, nonlinear systems can be defined using the following compact notation:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \quad (\text{B.4})$$

where $\mathbf{x} = (x_1, x_2, \dots, x_m)$ and $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x}))$.

B.2 Methods

B.2.1 Rössler System

This is a system in \mathbb{R}^3 published by Otto Rössler [230] and described as follows:

$$\begin{aligned} \dot{x} &= -y - z \\ \dot{y} &= x + ay \\ \dot{z} &= b + z(x - c) \end{aligned} \quad (\text{B.5})$$

where $(x, y, z) \in R^3$ are the states, and $(a, b, c) \in R^3$ the parameters of the system. Originally, the model was published using the following values of $a = 0.2$, $b = 0.2$, and $c = 5.7$; for which the model generates chaotic

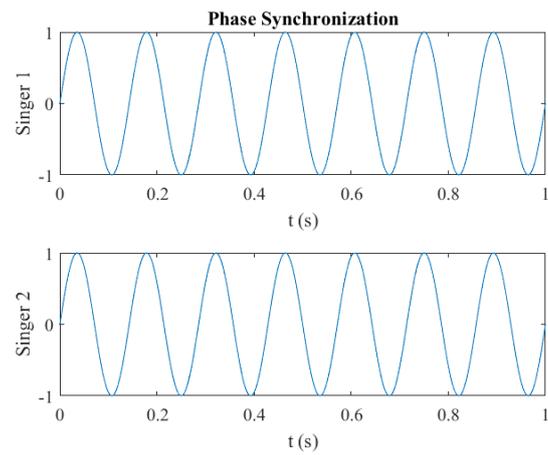
behaviour.

For this study, the derivation of the signals used to create f_o time series, we used MATLAB and the Runge–Kutta method. In this way, a numerical (discrete) solution of the system is created, in which the “speed” at which system evolves can be controlled and regulated by the time step t_s of the algorithm. This was exploited to adjust the speed of the oscillation of the states to be in accordance with those from the oscillatory frequencies of vibrato rate. As for the extent, the amplitude of the system states were normalised to be bounded between 1 and -1 (semitones). Although this is a three-dimensional system, only the x -state was used for the vibrato creation.

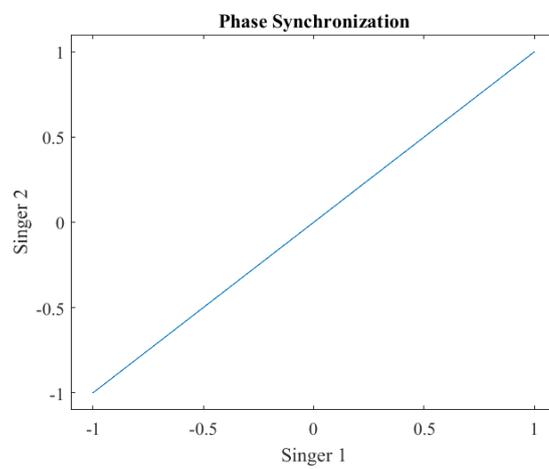
B.3 Results

Figures B.1, B.2, B.3 & B.4 show examples of vibrato signals that were able to be generated from the sine waves with perfect periodicity and the dynamical system. These vibrato synthesis signals are shown in their normal waveform time series form, and also using phase spaces, to illustrate a graphical representation of their dynamics.

Figures B.5 & B.6 show that it was also possible to derive multiple vibrato signals, and to be able to represent them in a four-dimensional plot to illustrate their interactions.

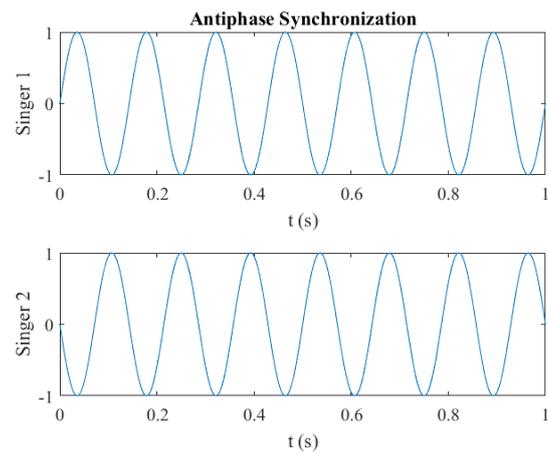


(a) Time series

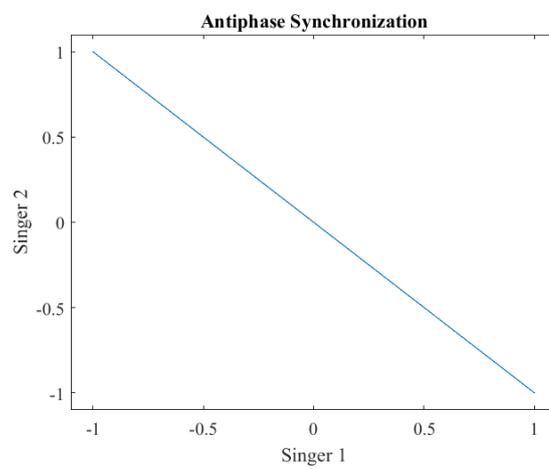


(b) Phase space

Figure B.1: In-phase synchronisation.

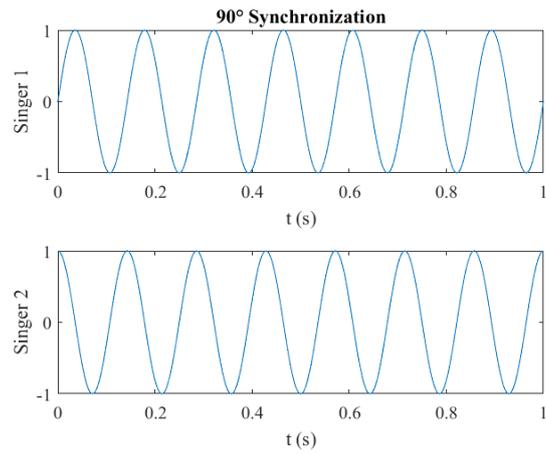


(a) Time series

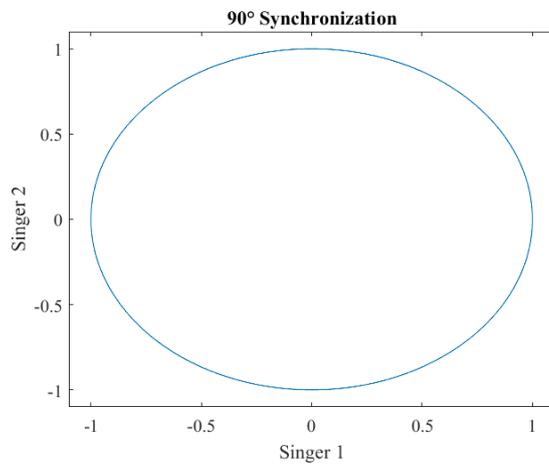


(b) Phase space

Figure B.2: Anti-phase synchronisation.

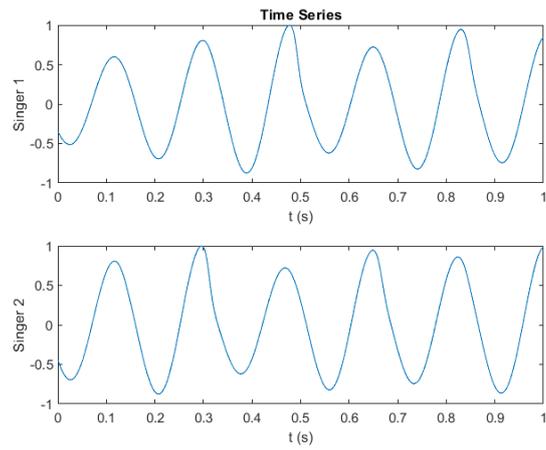


(a) Time series

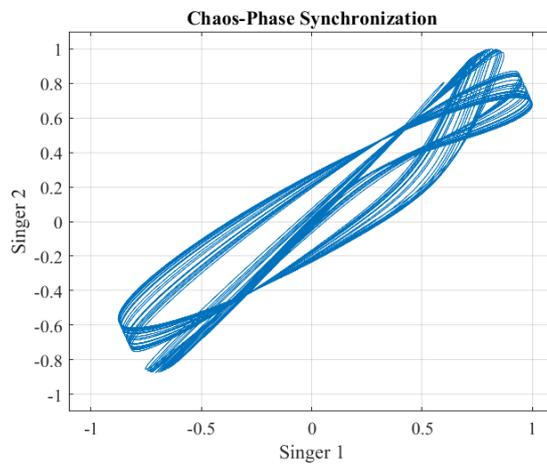


(b) Phase space

Figure B.3: Ninety-degree synchronisation.



(a) Time series



(b) Phase space

Figure B.4: Chaos synchronisation.

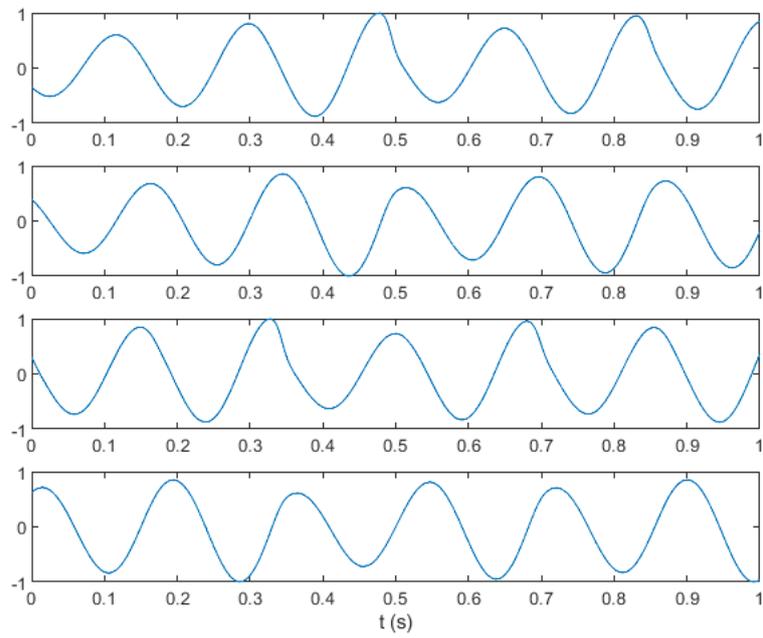


Figure B.5: Time series of four synthesised vibrato waves.

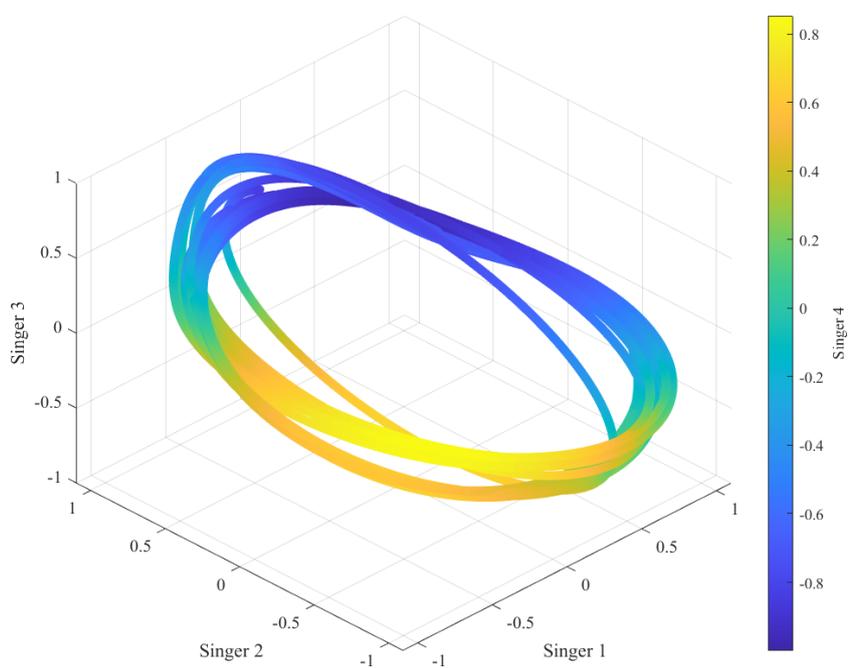


Figure B.6: Four-dimensional vibrato dynamics: The Tesseract vibrato.

Appendix C

Experimental Supplementary Material

Participant Information Sheet: Blended Aesthetics in Group Singing

Background

The University of York would like to invite you to take part in the following PhD research project. Please read the following information before agreeing to participate in the study.

For any queries regarding this study please contact the principle investigator:

Gerardo Acosta Martínez

gerardo.acosta@york.ac.uk

07479344386

This study has been approved by the University of York Physical Sciences Ethics Committee

What is the purpose of the study?

The study is designed to investigate the effects of vibrato in duet singing under different singing conditions. To do this we will record your voice with a set of microphones and electrodes while singing in either (1) solo singing condition (2) acoustic stimulus condition, or (3) singing partner condition.

How do I take part?

If you agree to participate you will be invited to a recording session at the AudioLab and asked to sing a number of tasks while your voice is being recorded. For this, you will have to wear a laryngograph (electrode) on the neck, and a close microphone (DPA) to record your voice. The task will consist of simple musical scales sung on vowels, material that will be available for you beforehand and so that you can be familiar with it before the session begins if necessary. If you agree, photos will be taken during the singing session for reporting purposes.

Why have I been invited to take part?

You have been invited to take part because you meet the criteria for the experiment: Have experience in choral singing of at least two years, unimpaired hearing, are 18 years of age or older, and a strong musical background or studies in classical music.

Do I have to take part?

No, participation is optional. If you do decide to take part, you will be given a copy of this information sheet for your records and will be asked to complete a participant consent form. If you change your mind at any point during the study, you will be able to withdraw your participation without having to provide a reason. Please note, if you withdraw from the project after data has been collected the data may still be used.

On what basis will you process my data?

Under the General Data Protection Regulation (GDPR), the University has to identify a legal basis for processing personal data and, where appropriate, an additional condition for processing special category data.

In line with our charter which states that we advance learning and knowledge by teaching and research, the University processes personal data for research purposes under Article 6 (1) (e) of the GDPR:

Processing is necessary for the performance of a task carried out in the public interest

Research will only be undertaken where ethical approval has been obtained, where there is a clear public interest and where appropriate safeguards have been put in place to protect data.

In line with ethical expectations and in order to comply with common law duty of confidentiality, we will seek your consent to participate where appropriate. This consent will not, however, be our legal basis for processing your data under the GDPR.

How will you use my data?

Data will be processed for the purposes outlined in this notice.

Will you share my data with 3rd parties?

No identifiable or personal data collected will be shared with 3rd parties.

How will you keep my data secure?

The University will put in place appropriate technical and organisational measures to protect your personal data. For the purposes of this project we will keep the participant information form with your signature on it in a locked office at the University of York and no copies will be made. Access will be restricted to the project team only. The voice recordings will be kept on the secure Google Drive of the University of York and be accessible to the main researcher and supervisor.

Will you transfer my data internationally?

Yes. All data will be stored on a secure University of York Google Drive, and may therefore be stored on international servers.

Will I be identified in any research outputs?

Your data will always be kept anonymous unless you give express permission for identifiable data (e.g. photographs or video) to be used in research publications. You may be identifiable.

How long will you keep my data?

Data will be retained in line with legal requirements or where there is a business need. Retention timeframes will be determined in line with the University's Records Retention Schedule. Please note, once published to the internet, the University will not be able to control how long recordings are retained and publicly accessible.

What rights do I have in relation to my data?

Under the GDPR, you have a general right of access to your data, a right to rectification, erasure, restriction, objection or portability. You also have a right to withdrawal. Please note, not all rights apply where data is processed purely for research purposes. For further information see, <https://www.york.ac.uk/records-management/general-dataprotectionregulation/individualsrights/>.

Questions or concerns

If you have any questions about this participant information sheet or concerns about how your data is being processed, please contact Gerardo Acosta Martinez at gerardo.acosta@york.ac.uk in the first instance. If you are still dissatisfied, please contact the University's Data Protection Officer at dataprotection@york.ac.uk.

Right to complain

If you are unhappy with the way in which the University has handled your personal data, you have a right to complain to the Information Commissioner's Office. For information on reporting a concern to the Information Commissioner's Office, see www.ico.org.uk/concerns.

Participant ID:

Participant Consent Form: Blended Aesthetics in Group Singing

This study has been approved by the Physical Sciences Ethics Committee with reference Martinez20240402

Participants should complete the following items themselves

- | | Please circle either
YES or NO |
|---|------------------------------------|
| 1. I confirm that I have read and understood the information sheet entitled “Blended Aesthetics in Group Singing” | YES / NO |
| 2. I confirm that I have been given the opportunity to ask any questions I may have about my participation in the project and that these questions have been answered to my satisfaction. | YES / NO |
| 3. I understand that I am free to withdraw from this study: <ul style="list-style-type: none"> ● At any time without having to give a reason, but not after the study is published ● Any data previously collected will be stored, analysed, and the results published in academic journals | YES / NO
YES / NO |
| 4. I understand that I can discuss the study with a researcher at any time if I wish. | YES / NO |
| 5. I understand that my name will not be linked with the research materials, and I will not be identified in the report or reports that result from the research. | YES / NO |
| 6. I understand that recordings of my voice and my anonymised responses may be used for scientific research purposes, and published at conferences, in journals or online. | YES / NO |
| 7. I give permission for members of the research team to have access to the data collected as part of this study, including audio recordings and anonymised related information that I provide. | YES / NO |
| 8. If I have any questions or concerns about the research, I know I can contact the main investigator at the AudioLab and gerardo.acosta@york.ac.uk. | YES / NO |
| 9. I give consent to the data collected from me to be used in future research. | YES / NO |
| 10. I agree to take part in the study. | YES / NO |
| 11. I agree for photos to be taken of my participation and used in publications and media related to the project. (I understand that I can say no to this question and still take part in the study.) | YES / NO |
| 12. I have received a £10 gift voucher as a thank you for taking part. | YES / NO |

13. Participant:

Name:

Signature

Date.....

14. Investigator:

I have explained the study to the above participant and he/she has indicated her/his willingness to take part.

Name:

Signature

Date.....

Abbreviations

ANOVA	Analysis of Variance
CPR	Correlation Probability of Recurrence
PCA	Principal Component Analysis
SPL	Sound Pressure Level

Bibliography

- [1] C. E. Seashore, “The natural history of the vibrato,” in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 17, National Academy of Sciences, 1931, pp. 623–626.
- [2] J. Sundberg, “Acoustic and psychoacoustic aspects of vocal vibrato,” *STL-QPSR*, vol. 35, pp. 45–68, 1994.
- [3] Y. Horii, “Frequency modulation characteristics of sustained /a/ sung in vocal vibrato,” *Journal of Speech, Language, and Hearing Research*, vol. 32, no. 4, pp. 829–836, 1989.
- [4] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A universal concept in nonlinear science*. Cambridge University Press, 2003.
- [5] G. Buzsáki, *Rhythms of the brain*. Oxford University Press, 2006.
- [6] G. Acosta Martínez and H. Daffern, “The role of vibrato in group singing: A systematic review,” *Journal of Voice*, vol. 39, no. 3, 847.e21–847.e39, 2025. DOI: 10.1016/j.jvoice.2022.11.004.
- [7] J. Potter, *Vocal authority: Singing style and ideology*. Cambridge University Press, 2006.
- [8] C. J. Butte, Y. Zhang, H. Song, and J. J. Jiang, “Perturbation and nonlinear dynamic analysis of different singing styles,” *Journal of Voice*, vol. 23, no. 6, pp. 647–652, 2009.
- [9] J. Neubauer, M. Edgerton, and H. Herzel, “Nonlinear phenomena in contemporary vocal music,” *Journal of Voice*, vol. 18, no. 1, pp. 1–12, 2004.

- [10] K. R. Scherer, J. Sundberg, B. Fantini, S. Trznadel, and F. Eyben, “The expression of emotion in the singing voice: Acoustic patterns in vocal performance,” *The Journal of the Acoustical Society of America*, vol. 142, no. 4, pp. 1805–1815, 2017.
- [11] C. E. Seashore, Ed., *Studies in the psychology of music. The vibrato*. University of Iowa, 1932, vol. 1.
- [12] H. Daffern and D. M. Howard, “Voice source comparison between modern singers of early music and opera,” *Logopedics Phoniatrics Vocology*, vol. 35, no. 2, pp. 68–73, 2010.
- [13] C. Dromey, L. Reese, and J. A. Hopkin, “Laryngeal-level amplitude modulation in vibrato,” *Journal of Voice*, vol. 23, no. 2, pp. 156–163, 2009.
- [14] P. H. Dejonckere, M. Hirano, and J. Sundberg, *Vibrato*. Singular Publishing Group, 1995.
- [15] 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*, vol. 8, no. 3, pp. 215–223, 1994.
- [16] J. Sundberg, “Perceptual aspects of singing,” *Journal of Voice*, vol. 8, no. 2, pp. 106–122, 1994.
- [17] J. A. Diaz and H. B. Rothman, “Acoustical comparison between samples of good and poor vibrato in singers,” *Journal of Voice*, vol. 17, no. 2, pp. 179–184, 2003.
- [18] S. Anand, J. M. Wingate, B. Smith, and R. Shrivastav, “Acoustic parameters critical for an appropriate vibrato,” *Journal of Voice*, vol. 26, no. 6, pp. 820e19–820e25, 2012.
- [19] M. A. Guzman, J. Dowdall, A. D. Rubin, *et al.*, “Influence of emotional expression, loudness, and gender on the acoustic parameters of vibrato in classical singers,” *Journal of Voice*, vol. 26, no. 5, pp. 675e5–675e11, 2012.

- [20] I. R. Titze, B. Story, M. Smith, and R. Long, “A reflex resonance model of vocal vibrato,” *Journal of the Acoustical Society of America*, vol. 111, no. 5, pp. 2272–2282, 2002.
- [21] Y. Horii, “Acoustic analysis of vocal vibrato: A theoretical interpretation of data,” *Journal of Voice*, vol. 3, no. 1, pp. 36–43, 1989.
- [22] Y. Meron and K. Hirose, “Synthesis of vibrato singing,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, IEEE, vol. 2, 2000, pp. 745–748.
- [23] N. Amir, O. Michaeli, and O. Amir, “Acoustic and perceptual assessment of vibrato quality of singing students,” *Biomedical Signal Processing and Control*, vol. 1, no. 2, pp. 144–150, 2006.
- [24] A. A. Reddy and U. Subramanian, “Singers’ and nonsingers’ perception of vocal vibrato,” *Journal of Voice*, vol. 29, no. 5, pp. 603–610, 2015.
- [25] J. Nix, “Listener preferences for vibrato rate and extent in synthesized vocal samples,” in *Proceedings of Meetings on Acoustics*, Acoustical Society of America, vol. 23, 2015, p. 035 002.
- [26] H. Cuesta, E. Gomez Gutierrez, A. Martorell Domínguez, and F. Loáiciga, “Analysis of intonation in unison choir singing,” in *Proceedings of the 15th International Conference on Music Perception and Cognition*, 2018.
- [27] S. Nandamudi and R. C. Scherer, “Airflow vibrato: Dependence on pitch and loudness,” *Journal of Voice*, vol. 33, no. 6, pp. 815–830, 2019.
- [28] 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*, vol. 18, no. 2, pp. 216–230, 2004.
- [29] J. D. Glasner and J. Nix, “Perception of vibrato by professional voice teachers: A pilot study,” *Journal of Voice*, 2023.

- [30] 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*, vol. 43, no. 5, pp. 675–685, 2015.
- [31] C. d’Alessandro and M. Castellengo, “The pitch of short-duration vibrato tones,” *Journal of the Acoustical Society of America*, vol. 95, no. 3, pp. 1617–1630, 1994.
- [32] P. Desain, H. Honing, R. Aarts, and R. Timmers, “Rhythmic aspects of vibrato,” in *Rhythm perception and production*. Swets & Zeitlinger, 1999.
- [33] M. L. Erickson, “Can inexperienced listeners hear who is flat? The role of timbre and vibrato,” *Journal of Voice*, vol. 30, no. 5, pp. 638e9–638e20, 2016.
- [34] C. Dromey, S. O. Holmes, J. A. Hopkin, and K. Tanner, “The effects of emotional expression on vibrato,” *Journal of Voice*, vol. 29, no. 2, pp. 170–181, 2015.
- [35] H. Kondo, S. Kondoh, and S. Fujii, “Perceived vibrato and the singing power ratio explain overall evaluations in opera singing,” *Frontiers in Psychology*, 2025.
- [36] J. Nix, “Shaken, not stirred: Practical ideas for addressing vibrato and nonvibrato singing in the studio and the choral rehearsal,” *Journal of Singing*, vol. 70, no. 4, pp. 411–418, 2014.
- [37] L. Wyatt, “Factors related to choral blend: Tone quality, vibrato, intonation,” *The Choral Journal*, vol. 8, no. 2, pp. 7–9, 1967.
- [38] G. Walker, “Good vibrations: Vibrato, science, and the choral singer,” *The Choral Journal*, vol. 47, no. 6, pp. 36–46, 2006.
- [39] B. Galante, “Vibrato and choral acoustics: Common voice science issues for the choral conductor,” *The Choral Journal*, vol. 51, no. 7, pp. 67–78, 2011.
- [40] D. Katok, “Healthy minimization of vibrato: An exploration of straight tone,” *The Choral Journal*, vol. 62, no. 4, pp. 8–19, 2021.

- [41] S. McCoy, “The choir issue, part 1,” *Journal of Singing*, vol. 67, no. 3, pp. 297–301, 2011.
- [42] S. Ternström, “Choir acoustics - an overview of scientific research published to date,” *TMH-QPSR*, vol. 43, no. 1, pp. 1–8, 2002.
- [43] S. D’Amario, H. Daffern, and F. Bailes, “A new method of onset and offset detection in ensemble singing,” *Logopedics Phoniatics Vocology*, vol. 44, no. 4, pp. 143–158, 2019.
- [44] C. Seibert, “Using performance sociograms to investigate inter - performer relationships in music ensembles,” in *Together in music: Coordination, expression, participation*. Oxford University Press, 2021, pp. 155–161.
- [45] G. Kearney, H. Daffern, L. Thresh, H. Omodudu, C. Armstrong, and J. Brereton, “Design of an interactive virtual reality system for ensemble singing,” in *Interactive Audio Systems Symposium*, York, UK, 2016.
- [46] H. Daffern and S. D’Amario, “Understanding expressive ensemble singing through acoustics,” in *Together in music: Coordination, expression, participation*. Oxford University Press, 2021, pp. 129–138.
- [47] V. Müller and U. Lindenberger, “Cardiac and respiratory patterns synchronize between persons during choir singing,” *PloS ONE*, vol. 6, no. 9, e24893, 2011.
- [48] H. Jers and S. Ternström, “Intonation analysis of a multi-channel choir recording,” *TMH-QPSR*, vol. 47, no. 1, pp. 1–6, 2005.
- [49] S. Ternström, “Physical and acoustic factors that interact with the singer to produce the choral sound,” *Journal of Voice*, vol. 5, no. 2, pp. 128–143, 1991.
- [50] L. R. Brown, “Effects of vibrato and vowel matching on choral blend: A correlation of spectral analyses and perceptual evaluations,” in *126th Meeting*, vol. 94, Acoustical Society of America, 1993, p. 1860.

- [51] D. M. Howard, H. Daffern, and J. Brereton, “Four-part choral synthesis system for investigating intonation in a cappella choral singing,” *Logopedics Phoniatrics Vocology*, vol. 38, no. 3, pp. 135–142, 2013.
- [52] E. Póhrolniczak and M. Kramarczyk, “Analysis of the signal of singing using the vibrato parameter in the context of choir singers,” *Journal of Electronic Science and Technology*, vol. 11, no. 4, pp. 417–423, 2013.
- [53] H. Daffern, “Blend in singing ensemble performance: Vibrato production in a vocal quartet,” *Journal of Voice*, vol. 31, no. 3, pp. 385e23–385e29, 2017.
- [54] S. D’Amario and H. Daffern, “Using electrolaryngography and electroglottography to assess the singing voice: A systematic review,” *Psychomusicology: Music, Mind, and Brain*, vol. 27, no. 4, pp. 229–243, 2017.
- [55] S. Ternström and J. Sundberg, “Intonation precision of choir singers,” *Journal of the Acoustical Society of America*, vol. 84, no. 1, pp. 59–69, 1988.
- [56] G. G. Sacerdote, “Researches on the singing voice,” *Acustica*, vol. 7, no. 2, pp. 61–68, 1957.
- [57] C. Dromey, N. Carter, and A. Hopkin, “Vibrato rate adjustment,” *Journal of Voice*, vol. 17, no. 2, pp. 168–178, 2003.
- [58] M. L. Erickson and C. S. Gaskill, “Can listeners hear how many singers are singing? The effect of listener’s experience, vibrato, onset, and formant frequency on the perception of number of simultaneous singers,” *Journal of Voice*, vol. 26, no. 6, pp. 817e1–817e13, 2012.
- [59] S. McAdams, “Segregation of concurrent sounds. I: Effects of frequency modulation coherence,” *The Journal of the Acoustical Society of America*, vol. 86, no. 6, pp. 2148–2159, 1989.
- [60] J. F. Culling and Q. Summerfield, “The role of frequency modulation in the perceptual segregation of concurrent vowels,” *The Journal of the Acoustical Society of America*, vol. 98, no. 2, pp. 837–846, 1995.

- [61] J. M. Chowning, “The synthesis of complex audio spectra by means of frequency modulation,” *Journal of the Audio Engineering Society*, vol. 21, no. 7, pp. 526–534, 1973.
- [62] M. Duncan, C. Williams, and G. Troup, “Vibrato frequency and phase lock in operatic duet quality,” in *Acoustics Australia*, vol. 28, 2000, pp. 5–9.
- [63] H. Jers and S. Ternström, “Intonation analysis of a multi-channel choir recording,” in *Joint Baltic-Nordic Acoustics Meeting*, Åland, Finland, 2004.
- [64] M. Mauch, K. Frieler, and S. Dixon, “Intonation in unaccompanied singing: Accuracy, drift, and a model of reference pitch memory,” *The Journal of the Acoustical Society of America*, vol. 136, no. 1, pp. 401–411, 2014.
- [65] D. M. Howard, “Intonation drift in a capella soprano, alto, tenor, bass quartet singing with key modulation,” *Journal of Voice*, vol. 21, no. 3, pp. 300–315, 2007.
- [66] S. Le Beux, L. Feugère, and C. d’Alessandro, “Chorus digitalis: Experiment in chironomic choir singing,” in *Annual Conference of the International Speech Communication Association*, ISCA, Firenze, Italy, 2011, pp. 2005–2008.
- [67] J. Dai and S. Dixon, “Intonation trajectories within tones in unaccompanied soprano, alto, tenor, bass quartet singing,” *The Journal of the Acoustical Society of America*, vol. 146, no. 2, pp. 1005–1014, 2019.
- [68] S. Ternström, A. Friberg, and J. Sundberg, “Synthesizing choir singing,” *Journal of Voice*, vol. 1, no. 4, pp. 332–335, 1988.
- [69] A. W. Goodwin, “An acoustical study of individual voices in choral blend,” *Journal of Research in Music Education*, vol. 28, no. 2, pp. 119–128, 1980.

- [70] T. D. Rossing, J. Sundberg, and S. Ternström, “Acoustic comparison of soprano solo and choir singing,” *The Journal of the Acoustical Society of America*, vol. 82, no. 3, pp. 830–836, 1987.
- [71] R. F. Coleman, “Acoustic and physiologic factors in duet singing: A pilot study,” *Journal of Voice*, vol. 8, no. 3, pp. 202–206, 1994.
- [72] K. L. Reid, P. Davis, J. Oates, *et al.*, “The acoustic characteristics of professional opera singers performing in chorus versus solo mode,” *Journal of Voice*, vol. 21, no. 1, pp. 35–45, 2007.
- [73] L. M. Mann, “Effects of solo and choral singing modes on vibrato rate, extent, and duration exhibited by undergraduate female singers,” *International Journal of Research in Choral Singing*, vol. 5, no. 1, pp. 26–38, 2014.
- [74] T. D. Rossing, J. Sundberg, and S. Ternström, “Acoustic comparison of voice use in solo and choir singing,” *The Journal of the Acoustical Society of America*, vol. 79, no. 6, pp. 1975–1981, 1986.
- [75] D. M. Howard, G. F. Welch, E. Himonides, and M. Owens, “The developing female chorister voice: Case-study evidence of musical development,” *Journal of Voice*, vol. 33, no. 4, pp. 516–525, 2019.
- [76] G. Berghs, N. Creylman, M. Avaux, W. Decoster, and F. de Jong, “A lifetime of professional singing: Voice parameters and age in the Netherlands Radio Choir,” *Logopedics Phoniatics Vocology*, vol. 38, no. 2, pp. 59–63, 2013.
- [77] S. P. Hinkley, “Effects of vibrato and pitch-varied vocal models on acoustic measures of high school and undergraduate singers’ vocal performance,” *International Journal of Research in Choral Singing*, vol. 9, pp. 1–21, 2021.
- [78] S. D’Amario, D. M. Howard, H. Daffern, and N. Pennill, “A longitudinal study of intonation in an a cappella singing quintet,” *Journal of Voice*, vol. 34, no. 1, pp. 159e13–159e27, 2020.
- [79] J. R. Wolpaw and E. W. Wolpaw, Eds., *Brain-computer interfaces: Principles and practice*. New York, NY: Oxford University Press, 2012.

- [80] B. Hagerman and J. Sundberg, “Fundamental frequency adjustment in barbershop singing,” *STL-QPSR*, vol. 21, no. 1, pp. 028–042, 1980.
- [81] P. H. Dejonckere, R. Miller, and G. Wieneke, “Vibrato in duet singers,” in *Proceedings CoMeT Congress*, Cairo, 1990.
- [82] H. B. Rothman and A. A. Arroyo, “Acoustic variability in vibrato and its perceptual significance,” *Journal of Voice*, vol. 1, no. 2, pp. 123–141, 1987.
- [83] A. T. Winfree, *The geometry of biological time*, 2nd ed. Springer, 2000.
- [84] S. H. Strogatz, *Sync: The emerging science of spontaneous order*. Hyperion, 2003.
- [85] C. T. Herbst, S. Hertegard, D. Zangger-Borch, and P. Å. Lindestad, “Freddie Mercury — acoustic analysis of speaking fundamental frequency, vibrato, and subharmonics,” *Logopedics Phoniatics Vocology*, vol. 42, no. 1, pp. 29–38, 2017.
- [86] H. Herzel, D. Berry, I. Titze, and I. Steinecke, “Nonlinear dynamics of the voice: Signal analysis and biomechanical modeling,” *Chaos*, vol. 5, no. 1, pp. 30–34, 1995.
- [87] L. Matassini, R. Hegger, H. Kantz, and C. Manfredi, “Analysis of vocal disorders in a feature space,” *Medical Engineering & Physics*, vol. 22, no. 6, pp. 413–418, 2000.
- [88] H. Herzel, “Bifurcations and chaos in voice signals,” *Applied Mechanics Reviews*, vol. 46, no. 7, pp. 399–413, 1993.
- [89] A. Anikin, “The perceptual effects of manipulating nonlinear phenomena in synthetic nonverbal vocalizations,” *Bioacoustics*, vol. 29, no. 2, pp. 226–247, 2020.
- [90] M. Little, P. Mcsharry, S. Roberts, D. Costello, and I. Moroz, “Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection,” *Nature Precedings*, pp. 1–35, 2007.

- [91] J. G. Švec, H. K. Schutte, and D. G. Miller, “A subharmonic vibratory pattern in normal vocal folds,” *Journal of Speech, Language, and Hearing Research*, vol. 39, no. 1, pp. 135–143, 1996.
- [92] H. Daffern, D. Camlin, H. Egermann, *et al.*, “Exploring the potential of virtual reality technology to investigate the health and well being benefits of group singing,” *International Journal of Performance Arts and Digital Media*, vol. 15, no. 1, pp. 1–22, 2019.
- [93] J. Sundberg, *The science of the singing voice*. Northern Illinois University Press, 1987.
- [94] V. Müller, J. Saenger, and U. Lindenberger, “Intra and inter-brain synchronization during musical improvisation on the guitar,” *PLoS ONE*, vol. 8, no. 9, e73852, 2013.
- [95] D. A. Camlin, H. Daffern, and K. Zeserson, “Group singing as a resource for the development of a healthy public: A study of adult group singing,” *Humanities and Social Sciences Communications*, vol. 7, no. 1, pp. 1–15, 2020.
- [96] S. H. Strogatz, *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*, 2nd ed. Westview Press, 2015.
- [97] G. Acosta Martínez and H. Daffern, “A pilot study of vocal vibrato incorporating nonlinear time series analysis,” in *Proceedings of the Stockholm Music Acoustic Conference*, S. D’Amario, S. Ternström, and A. Friberg, Eds., Stockholm, Sweden, 2023, pp. 182–189. DOI: 10.30746/978-91-8040-865-3.
- [98] H. Daffern, “Distinguishing characteristics of vocal techniques in the specialist performance of early music,” Ph.D. dissertation, University of York, 2008.
- [99] 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*, 2022.

- [100] H. Kantz and T. Schreiber, *Nonlinear time series analysis*, 2nd ed. Cambridge University Press, 2004.
- [101] C. T. Herbst, H. Herzel, J. G. Švec, M. T. Wyman, and W. T. Fitch, “Visualization of system dynamics using phasegrams,” *Journal of the Royal Society Interface*, vol. 10, no. 85, p. 20130288, 2013.
- [102] W. C. Costa, F. M. Assis, B. G. A. Neto, S. C. Costa, and V. J. D. Vieira, “Pathological voice assessment by recurrence quantification analysis,” in *ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living*, IEEE, 2012, pp. 1–6.
- [103] A. Selamtzis and S. Ternström, “Analysis of vibratory states in phonation using spectral features of the electroglottographic signal,” *The Journal of the Acoustical Society of America*, vol. 136, no. 5, pp. 2773–2783, 2014.
- [104] C. Fabris, W. De Colle, and G. Sparacino, “Voice disorders assessed by (cross-) Sample Entropy of electroglottogram and microphone signals,” *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 920–926, 2013.
- [105] P. Tranchant, E. Scholler, and C. Palmer, “Endogenous rhythms influence musicians’ and non-musicians’ interpersonal synchrony,” *Scientific Reports*, vol. 12, no. 1, p. 12973, 2022.
- [106] C. Palmer and A. P. Demos, “Are we in time? How predictive coding and dynamical systems explain musical synchrony,” *Current Directions in Psychological Science*, vol. 31, no. 2, pp. 147–153, 2022.
- [107] A. Ravignani and P. Norton, “Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations,” *Journal of Language Evolution*, vol. 2, pp. 4–19, 2017.
- [108] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, “Recurrence plots for the analysis of complex systems,” *Physics Reports*, vol. 438, no. 5-6, pp. 237–329, 2007.

- [109] N. Marwan, “How to avoid potential pitfalls in recurrence plot based data analysis,” *International Journal of Bifurcation and Chaos*, vol. 21, no. 4, pp. 1003–1017, 2011.
- [110] A. Delgado-Bonal and A. Marshak, “Approximate entropy and sample entropy: A comprehensive tutorial,” *Entropy*, vol. 21, no. 6, p. 541, 2019.
- [111] S. Ternström, D. Johansson, and A. Selamtzis, “FonaDyn—A system for real-time analysis of the electroglottogram, over the voice range,” *SoftwareX*, vol. 7, pp. 74–80, 2018.
- [112] P. Boersma, “Praat, a system for doing phonetics by computer,” *Glott International*, vol. 5, no. 9, pp. 341–345, 2001.
- [113] J. S. Richman and J. R. Moorman, “Physiological time-series analysis using approximate entropy and sample entropy,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, no. 6, H2039–H2049, 2000.
- [114] P. Faure and H. Korn, “Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation,” *Comptes Rendus de l’Académie des Sciences, Sciences de la Vie*, vol. 324, no. 9, pp. 773–793, 2001.
- [115] J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, and N. Stergiou, “The appropriate use of approximate entropy and sample entropy with short data sets,” *Annals of Biomedical Engineering*, vol. 41, no. 2, pp. 349–365, 2013.
- [116] I. T. Jolliffe and J. Cadima, “Principal component analysis: A review and recent developments,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150202, 2016.
- [117] R. A. Lester-Smith and B. H. Story, “The effects of physiological adjustments on the perceptual and acoustical characteristics of vibrato as a model of vocal tremor,” *The Journal of the Acoustical Society of America*, vol. 140, no. 5, pp. 3827–3833, 2016.

- [118] D. Abur, R. A. Lester-Smith, A. Daliri, A. A. Lupiani, F. H. Guenther, and C. E. Stepp, “Sensorimotor adaptation of voice fundamental frequency in Parkinson’s disease,” *PloS One*, vol. 13, no. 1, e0191839, 2018.
- [119] G. Acosta Martínez and H. Daffern, “Complexity of vocal vibrato in opera and jazz recordings: Insights from entropy and recurrence analyses,” *Journal of Voice*, 2023. DOI: 10.1016/j.jvoice.2023.11.020.
- [120] J. Sundberg, “Perception of singing,” in *The Psychology of Music*. Academic Press, 1982.
- [121] 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*, vol. 7, pp. 58–64, 2012.
- [122] N. Carter, J. A. Hopkin, and C. Dromey, “Volitional control of vibrato in trained singers,” *Journal of Singing*, vol. 67, no. 1, pp. 9–17, 2010.
- [123] 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 *Acoustics 2012*, Nantes, France, 2012.
- [124] 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*, vol. 29, no. 4, 517e1–517e9, 2015.
- [125] J. LoVetri, “Contemporary commercial music,” *Journal of Voice*, no. 3, pp. 260–262, 2008.
- [126] A. Ponchielli, *Ponchielli: La Gioconda, Op. 9* [Album recorded by Orchestra del Teatro alla Scala di Milano]. Warner Classics, 2023.
- [127] M. Callas, *Maria Callas legendary recordings* [Album]. Profil, 2014.
- [128] J. Carreras, *The very best of José Carreras* [Album]. EMI Records, 2003.

- [129] R. Fleming, *Renée Fleming* [Album]. Decca Music Group Limited, 2000.
- [130] L. Pavarotti, *The big Luciano* [Album]. Replay Music, 2010.
- [131] G. Puccini, *Puccini: Turandot* [Album recorded by London Philharmonic Orchestra]. Decca Music Group Limited, 1973.
- [132] G. Verdi, *Verdi: La Triavata* [Album recorded by National Philharmonic Orchestra]. Decca Music Group Limited, 1980.
- [133] N. K. Cole, *Unforgettable* [Album]. Capitol Records, 1954.
- [134] —, *The Nat King Cole story* [Album]. Capitol Records, 1991.
- [135] E. Fitzgerald and L. Armstrong, *Porgy & Bess* [Album]. UMG Recordings, 1958.
- [136] E. Fitzgerald, *Clap hands, here comes Charlie!* [Album]. UMG Recordings, 1989.
- [137] —, *Ella Fitzgerald sings the Rodgers and Hart song book* [Album]. UMG Recordings, 1956.
- [138] B. Holiday, *Solitude* [Album]. UMG Recordings, 1956.
- [139] —, *Billie Holiday* [Album]. UMG Recordings, 1957.
- [140] F. Sinatra, *In the wee small hours* [Album]. Capitol Records, 1955.
- [141] —, *Strangers in the night* [Album]. Reprise Records, 1966.
- [142] S. Vaughan, *Vaughan and violins* [Album]. UMG Recordings, 1958.
- [143] —, *After hours* [Album]. Sony Music Entertainment, 1955.
- [144] —, *The magic of Sarah Vaughan* [Album]. UMG Records, 1959.
- [145] K. Gerdes and K. Siedenburger, “Lead-vocal level in recordings of popular music 1946–2020,” *JASA Express Letters*, vol. 3, no. 4, 2023.
- [146] R. Wooding and J. Nix, “Perception of non-vibrato sung tones: A pilot study,” *Journal of Voice*, vol. 30, no. 6, 762e15–762e21, 2016.
- [147] M. Herzog, G. Francis, and A. Clarke, *Understanding statistics and experimental design: How to not lie with statistics*. Springer, 2019.

- [148] G. Acosta and H. Daffern, *Vibrato from Opera and Jazz singers*, 1. Mendeley Data, 2023. DOI: 10.17632/36kctyh64.1.
- [149] P. G. Rodrigues, R. Attux, G. Castellano, and D. C. Soriano, “Space-time recurrences for functional connectivity evaluation and feature extraction in motor imagery brain-computer interfaces,” *Medical & Biological Engineering & Computing*, vol. 57, no. 8, pp. 1709–1725, 2019.
- [150] J. Nix, N. Perna, K. James, and S. Allen, “Vibrato rate and extent in college music majors: A multicenter study,” *Journal of Voice*, vol. 30, no. 6, 756e31–756e41, 2016.
- [151] G. Acosta Martínez and H. Daffern, “Analysis of vibrato variability and adaptability in a duet singing experiment: A pilot study,” in *Abstracts Book of the 15th Pan-European Voice Conference*, Santander, Spain, 2024, pp. 87–88.
- [152] —, “Vibrato synchronicity in sustained notes from female singers blending to a prerecorded singer,” *Journal of Voice*, 2025. DOI: 10.1016/j.jvoice.2025.09.006.
- [153] C. E. Seashore, “The psychology of music. VI. The vibrato: (1) What is it?” *Music Educators Journal*, vol. 23, no. 4, pp. 30–33, 1937.
- [154] 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*, vol. 151, no. 6, pp. 4131–4139, 2022.
- [155] E. Ekholm, “The effect of singing mode and seating arrangement on choral blend and overall choral sound,” *Journal of Research in Music Education*, vol. 48, no. 2, pp. 123–135, 2000.
- [156] J. N. Killian and L. Basinger, “Perception of choral blend among choral, instrumental, and nonmusic majors using the continuous response digital interface,” *Journal of Research in Music Education*, vol. 55, no. 4, pp. 313–325, 2007.

- [157] V. Sublett, “Vibrato or nonvibrato in solo and choral singing: Is there room for both?” *Journal of Singing*, vol. 65, no. 5, pp. 539–544, 2009.
- [158] A. Pikovsky, M. Rosenblum, and J. Kurths, “Phase synchronization in regular and chaotic systems,” *International Journal of Bifurcation and Chaos*, vol. 10, no. 10, pp. 2291–2305, 2000.
- [159] S. H. Strogatz, “From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators,” *Physica D: Nonlinear Phenomena*, vol. 143, pp. 1–20, 2000.
- [160] S. D’Amario and F. Bailes, “Ensemble timing and synchronization,” in *Together in music: Coordination, expression, participation*. Oxford University Press, 2021, pp. 139–147.
- [161] N. Sebanz, H. Bekkering, and G. Knoblich, “Joint action: Bodies and minds moving together,” *Trends in Cognitive Sciences*, vol. 10, no. 2, pp. 70–76, 2006.
- [162] E. R. Kirsh, E. van Leer, H. J. Phero, C. Xie, and S. Khosla, “Factors associated with singers’ perceptions of choral singing well-being,” *Journal of Voice*, vol. 27, no. 6, 786e25–786e32, 2013.
- [163] I. Konvalinka, N. Sebanz, and G. Knoblich, “The role of reciprocity in dynamic interpersonal coordination of physiological rhythms,” *Cognition*, vol. 230, p. 105 307, 2023.
- [164] A. Zamm, S. Debener, and N. Sebanz, “The spontaneous emergence of rhythmic coordination in turn taking,” *Scientific Reports*, vol. 13, no. 1, p. 3259, 2023.
- [165] S. Duvvuru and M. Erickson, “The effect of timbre, pitch, and vibrato on vocal pitch-matching accuracy,” *Journal of Voice*, vol. 30, no. 3, 378e1–378e12, 2016.
- [166] J. P. Eckmann, S. O. Kamphorst, and D. Ruelle, “Recurrence plots of dynamical systems,” *Europhysics Letters*, vol. 4, no. 9, pp. 973–977, 1987.

- [167] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, “Recurrence-plot-based measures of complexity and their application to heart-rate-variability data,” *Physical Review E*, vol. 66, p. 026 702, 2002.
- [168] P. Cairns, T. Rudzki, J. Cooper, *et al.*, “Singer and audience evaluations of a networked immersive audio concert,” *Journal of the Audio Engineering Society*, vol. 72, pp. 467–478, 2024. DOI: 10.17743/jaes.2022.0145.
- [169] S. Gonzalez and M. Brookes, “A pitch estimation filter robust to high levels of noise,” in *European Signal Processing Conference*, IEEE, 2011, pp. 451–455.
- [170] C. W. Dunnett, “A multiple comparison procedure for comparing several treatments with a control,” *Journal of the American Statistical Association*, vol. 50, no. 272, pp. 1096–1121, 1955.
- [171] J. P. Zbilut, A. Giuliani, and C. L. Webber Jr, “Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification,” *Physics Letters A*, vol. 246, pp. 122–128, 1998.
- [172] N. Marwan, M. Thiel, and N. R. Nowaczyk, “Cross recurrence plot based synchronization of time series,” *Nonlinear Processes in Geophysics*, vol. 9, no. 3, pp. 325–331, 2002.
- [173] F. Takens, “Detecting strange attractors in turbulence,” in *Dynamical Systems and Turbulence*, Berlin: Springer, 1981, pp. 366–381.
- [174] M. C. Romano, M. Thiel, J. Kurths, I. Z. Kiss, and J. Hudson, “Detection of synchronization for non-phase-coherent and non-stationary data,” *Europhysics Letters*, vol. 71, no. 3, pp. 466–472, 2005.
- [175] N. Marwan, *CRP Toolbox*, version 5.29, 2024.
- [176] G. Acosta and H. Daffern, *Individual vibrato parameters from high choral singers blending to a pre-recorded singer*, 1. Mendeley Data, 2025. DOI: 10.17632/kvj6vrd6rf.1.

- [177] A. M. Nkomidio, E. K. Ngamga, B. R. N. Nbandjo, J. Kurths, and N. Marwan, “Recurrence-based synchronization analysis of weakly coupled bursting neurons under external ELF fields,” *Entropy*, vol. 24, no. 2, p. 235, 2022.
- [178] J. Fell and N. Axmacher, “The role of phase synchronization in memory processes,” *Nature Reviews*, vol. 12, no. 2, pp. 105–118, 2011.
- [179] V. Afraimovich, A. Cordonet, and N. Rulkov, “Generalized synchronization of chaos in noninvertible maps,” *Physical Review E*, vol. 66, no. 1, p. 016 208, 2002.
- [180] G. V. Osipov, B. Hu, C. Zhou, M. V. Ivanchenko, and J. Kurths, “Three types of transitions to phase synchronization in coupled chaotic oscillators,” *Physical Review Letters*, vol. 91, no. 2, p. 024 101, 2003.
- [181] L. M. Pecora and T. L. Carroll, “Synchronization in chaotic systems,” *Physical Review Letters*, vol. 64, no. 8, pp. 821–824, 1990.
- [182] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, “From phase to lag synchronization in coupled chaotic oscillators,” *Physical Review Letters*, vol. 78, no. 22, pp. 4193–4196, 1997.
- [183] Scorpions, *Fly to the rainbow* [Album]. RCA, 1974.
- [184] Black Sabbath, *Heaven and hell* [Album]. Vertigo, 1980.
- [185] Alice In Chains, *Jar of flies* [Album]. Columbia, 1994.
- [186] ———, *Dirt* [Album]. Columbia, 1992.
- [187] M. Carey, *Emotions* [Album]. Columbia, 1991.
- [188] Y. Horii and K. Hata, “A note on phase relationships between frequency and amplitude modulations in vocal vibrato,” *Folia Phoniatrica et Logopaedica*, vol. 40, no. 6, pp. 303–311, 1988.
- [189] D. M. Howard, “The vocal tract organ: A new musical instrument using 3-D printed vocal tracts,” *Journal of Voice*, vol. 32, no. 6, pp. 660–667, 2018.
- [190] M. Takasu, “Vocalness of the female singing voice,” *Journal of the Acoustical Society of Japan*, vol. 18, no. 1, pp. 19–28, 1997.

- [191] M. Marchini, R. Ramirez, P. Papiotis, and E. Maestre, “The sense of ensemble: A machine learning approach to expressive performance modelling in string quartets,” *Journal of New Music Research*, vol. 43, no. 3, pp. 303–317, 2014.
- [192] J. Dai and S. Dixon, “Understanding intonation trajectories and patterns of vocal notes,” in *International Conference on Multimedia Modeling*, Springer, 2019, pp. 243–253.
- [193] R. C. Maher and J. W. Beauchamp, “Fundamental frequency estimation of musical signals using a two-way mismatch procedure,” *The Journal of the Acoustical Society of America*, vol. 95, no. 4, pp. 2254–2263, 1994.
- [194] S. K. Ravi, S. Shabnam, K. S. George, and T. Saraswathi, “Acoustic and aerodynamic characteristics of choral singers,” *Journal of Voice*, vol. 33, no. 5, 803e1–803e5, 2019.
- [195] G. G. Sacerdote, “Researches on the singing voice,” in *51th Meeting*, vol. 28, Acoustical Society of America, 1956, pp. 789–790.
- [196] J. Dai and S. Dixon, “Singing together: Pitch accuracy and interaction in unaccompanied unison and duet singing,” *The Journal of the Acoustical Society of America*, vol. 145, no. 2, pp. 663–675, 2019.
- [197] A. Grell, J. Sundberg, S. Ternström, M. Ptok, and E. Altenmüller, “Rapid pitch correction in choir singers,” *The Journal of the Acoustical Society of America*, vol. 126, no. 1, pp. 407–413, 2009.
- [198] M. Olson, “Vibrato vs. nonvibrato: The solo singer in the collegiate choral ensemble,” *Journal of Singing*, vol. 64, no. 5, pp. 561–564, 2008.
- [199] J. Sundberg, “Synthesizing singing,” in *115th Meeting*, vol. 83, Acoustical Society of America, 1988, S30–S31.
- [200] J. F. Culling and C. Darwin, “Perceptual separation of simultaneous vowels: Within and across-formant grouping by F0,” *Journal of the Acoustical Society of America*, vol. 93, no. 6, pp. 3454–3467, 1993.
- [201] J. Sundberg and J. Lindqvist, “Musical octaves and pitch,” *Journal of the Acoustical Society of America*, vol. 54, no. 4, pp. 922–929, 1973.

- [202] G. J. Troup, “The physics of the singing voice: Measurements on the voices and vocal apparatus of trained singers, and their interpretation,” *Physics Reports*, vol. 74, no. 5, pp. 379–401, 1981.
- [203] J. B. King and Y. Horii, “Vocal matching of frequency modulation in synthesized vowels,” *Journal of Voice*, vol. 7, no. 2, pp. 151–159, 1993.
- [204] J. F. Culling and C. Darwin, “The role of timbre in the segregation of simultaneous voices with intersecting F0 contours,” *Perception & Psychophysics*, vol. 54, no. 3, pp. 303–309, 1993.
- [205] L. Demany and C. Semal, “The effect of vibrato on the recognition of masked vowels,” *Perception & Psychophysics*, vol. 48, no. 5, pp. 436–444, 1990.
- [206] C. M. Marin and S. McAdams, “The role of auditory beats induced by frequency modulation and polyperiodicity in the perception of spectrally embedded complex target sounds,” *The Journal of the Acoustical Society of America*, vol. 100, no. 3, pp. 1736–1753, 1996.
- [207] J. K. Ford, “Preferences for strong or weak singer’s formant resonance in choral tone quality,” *International Journal of Research in Choral Singing*, vol. 1, no. 1, pp. 29–47, 2003.
- [208] J. F. Daugherty, “Spacing, formation, and choral sound: Preferences and perceptions of auditors and choristers,” *Journal of Research in Music Education*, vol. 47, no. 3, pp. 224–238, 1999.
- [209] ———, “Choir spacing and formation: Choral sound preferences in random, synergistic, and gender-specific chamber choir placements,” *International Journal of Research in Choral Singing*, vol. 1, no. 1, pp. 48–59, 2003.
- [210] J. F. Daugherty, J. N. Manternach, and M. C. Brunkan, “Acoustic and perceptual measures of SATB choir performances on two types of portable choral riser units in three singer-spacing conditions,” *International Journal of Music Education*, vol. 31, no. 3, pp. 359–375, 2013.

- [211] 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*, vol. 8, no. 3, pp. 224–229, 1994.
- [212] J. Sundberg, M. N. Thörnvik, and A. M. Söderström, “Age and voice quality in professional singers,” *Logopedics Phoniatrics Vocology*, no. 4, pp. 169–176, 1998.
- [213] R. A. Rasch, “The perception of simultaneous notes such as in polyphonic music,” *Acustica*, vol. 40, no. 1, pp. 21–33, 1978.
- [214] R. B. Gardner, S. A. Gaskill, and C. Darwin, “Perceptual grouping of formants with static and dynamic differences in fundamental frequency,” *The Journal of the Acoustical Society of America*, vol. 85, no. 3, pp. 1329–1337, 1989.
- [215] R. P. Carlyon, “Discriminating between coherent and incoherent frequency modulation of complex tones,” *The Journal of the Acoustical Society of America*, vol. 89, no. 1, pp. 329–340, 1991.
- [216] ———, “The psychophysics of concurrent sound segregation,” *Philosophical Transactions of the Royal Society of London*, no. 1278, pp. 347–355, 1992.
- [217] ———, “Further evidence against an across-frequency mechanism specific to the detection of frequency modulation incoherence between resolved frequency components,” *The Journal of the Acoustical Society of America*, vol. 95, no. 2, pp. 949–961, 1994.
- [218] C. Darwin and J. F. Culling, “Speech perception seen through the ear,” *Speech Communication*, vol. 9, pp. 469–475, 1990.
- [219] C. J. Darwin, “Perceptual grouping of speech components differing in fundamental frequency and onset-time,” *The Quarterly Journal of Experimental Psychology*, vol. 33, no. 2, pp. 185–207, 1981.
- [220] M. H. Chalikia and A. S. Bregman, “The perceptual segregation of simultaneous auditory signals: Pulse train segregation and vowel segregation,” *Perception & Psychophysics*, vol. 46, no. 5, pp. 487–496, 1989.

- [221] ———, “The perceptual segregation of simultaneous vowels with harmonic, shifted, or random components,” *Perception & Psychophysics*, vol. 53, no. 2, pp. 125–133, 1993.
- [222] A. S. Wilson, J. W. Hall, and J. H. Grose, “Detection of frequency modulation in the presence of a second FM tone,” *The Journal of the Acoustical Society of America*, vol. 88, no. 3, pp. 1333–1338, 1990.
- [223] M. F. Cohen and X. Chen, “Dynamic frequency change among stimulus components: Effects of coherence on detectability,” *The Journal of the Acoustical Society of America*, vol. 92, no. 2, pp. 766–772, 1992.
- [224] C. Darwin, V. Ciocca, and G. Sandell, “Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic,” *The Journal of the Acoustical Society of America*, vol. 95, no. 5, pp. 2631–2636, 1994.
- [225] B. C. Moore, B. R. Glasberg, and R. W. Peters, “Relative dominance of individual partials in determining the pitch of complex tones,” *The Journal of the Acoustical Society of America*, vol. 77, no. 5, pp. 1853–1860, 1985.
- [226] U. T. Zwicker, “Auditory recognition of diotic and dichotic vowel pairs,” *Speech Communication*, vol. 3, no. 4, pp. 265–277, 1984.
- [227] P. F. Assmann and Q. Summerfield, “Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies,” *The Journal of the Acoustical Society of America*, vol. 88, no. 2, pp. 680–697, 1990.
- [228] R. B. Gardner and C. Darwin, “Grouping of vowel harmonics by frequency modulation: Absence of effects on phonemic categorization,” *Perception & Psychophysics*, vol. 40, no. 3, pp. 183–187, 1986.
- [229] G. Acosta Martínez and H. Daffern, “Vibrato dynamics in group singing: Understanding vibrato behaviour through dynamical systems theory,” in *Abstracts of the 14th Pan-European Voice Conference*, Tallinn, Estonia, 2022, p. 46.

- [230] O. Rössler, “An equation for continuous chaos,” *Physics Letters A*, vol. 57, no. 5, pp. 397–398, 1976.