Cetacean citations: rhythmic variability in the composition and recomposition of humpback whale song

Alexander Mitchell South

A thesis submitted for the degree of PhD
at the
Royal Conservatoire of Scotland
&
University of St Andrews

2024

Full metadata for this item is available in
St Andrews Research Repository
at:
https://research-repository.st-andrews.ac.uk/

Identifier to use to cite or link to this thesis:

DOI: https://doi.org/10.17630/sta/851

This item is protected by original copyright

CETACEAN CITATIONS

Rhythmic Variability in the Composition and Recomposition of Humpback Whale Song

Alexander Mitchell South

This thesis is submitted in partial fulfilment
for the degree of PhD at the

University of St Andrews
Royal Conservatoire of Scotland

15 August 2023

General Abstract

In this interdisciplinary thesis I attend to the rhythmic variability of the phrases whose repetition is such a distinctive feature of humpback whale song, listening and responding as bioacoustician, musician, and zoömusicologist. I developed methods to visualize and measure individual distinctiveness and rhythmic precision in shared song phrases, and assessed thirteen hours of humpback song from ten singers, collected off Mo'orea, French Polynesia, September-November 2019. Using multiple regression and multivariate distance techniques, I found that individual singers sing shared phrases with their own individually distinctive rhythms but with equal levels of rhythmic precision, across a wide range of phrase variants. As a musician I recomposed with transcriptions and recordings of humpback song phrases, the 'cetacean citations' of my title, producing a portfolio of six works in collaboration with other musicians. In a reflective zoömusicological analysis of my compositional processes, I examined how composers might avoid both an 'anthropodenial' that fails to recognize the similarities between humans and other animals, and a 'naïve anthropomorphism' that fails to recognize the differences. Further, I challenged the presumption that there are no ethical questions involved in the musical use of other-than-human audio recordings via an analysis of their inevitable 'objectification' in contemporary composition. Drawing on Adorno's critique of instrumental rationality and Plumwood's critique of anthropocentrism, I elaborated the concept of an aesthetic rationality that treats other-than-human sounds with respect, through a set of strategies for composers to produce a non-anthropocentric multispecies music that neither enacts nor embodies human exceptionalism. Finally, I proposed the notion of 'multispecies heterophony' to describe the overlapping nonhierarchical sounding of difference. This compositional form was jointly inspired by the 'asynchronous chorus' of the collective singing of humpback whales, and the human dynamics of large group improvisation; I suggest that it offers a promising musical model for ecological thinking.

Candidate's declaration

I, Alexander Mitchell South, do hereby certify that this thesis, submitted for the degree of PhD, includes:

- A thesis which is approximately 57,700 words in length
- A portfolio comprised of 6 works

has been produced by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews and the Royal Conservatoire of Scotland in September 2018.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Signature of candidate Date 15 August 2023

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Signature of supervisor Date 15 August 2023

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Alexander Mitchell South, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Title and Abstract

I agree to the title and abstract being published

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Alexander Mitchell South, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Signature of candidate Date 15 August 2023

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

No embargo on underpinning research data or digital outputs.

General Acknowledgements

I consider myself extremely fortunate to have spent much of the last four and a half years listening to and re-composing the sounds made by humpback whales, learning and thinking about cetaceans and their culture, and carrying out bioacoustical research into the rhythms of humpback song. My navigation along shifting scholarly currents would not have been possible without the expert, patient and wise guidance of my supervisors Emily Doolittle (Royal Conservatoire of Scotland), Luke Rendell (University of St Andrews) and Ellen Garland (University of St Andrews). Emily, Luke and Ellen: thank you for your generosity and encouragement, and for providing so much detailed feedback on the writing of this thesis. I have also been much helped by the members of my progress review panels at both my institutions: Stephen Broad and Bethany Whiteside (Royal Conservatoire of Scotland), Ines Jentzsch, Andy Gardner and Anne Magurran (University of St Andrews).

Along the way I have been inspired by a number of artistic collaborators, some of whom were involved in the creation and performance of my portfolio works (as detailed in §4.2): Katherine Wren, Sonia Cromarty and Alice Rickards (Sequoia), Jan Foote; also Frances Barker, Sally Day and Nicola Long (Scottish Clarinet Quartet). Others played important roles in further pieces and performances: Lesley Harrison, Aya Kobayashi, Becky Milne, Anne Rankin, Jer Reid, Emma Roche, Glasgow Improvisers Orchestra, Nichola Scrutton, Oliver Searle, Steve Smart, members of St Andrews Chamber Orchestra, Paul Stone and Solene Weinachter.

My fieldwork around the coasts of the beautiful mountainous island of Mo'orea, French Polynesia, was led by Ellen Garland and made possible by funding from the RCS Research Degrees Committee and the Russell Trust (University of St Andrews). In Mo'orea, the collection of whale song recordings was made possible by M. Michael Poole and Richmond Frank ('Arii'). The whole wonderful experience was shared with and much enlivened by fellow PhD student Natalie Sinclair. Of course it was also in Mo'orea that I was able to encounter the hauntingly familiar yet unfamiliar song of those humpback singers whose voices I have quoted directly in my portfolio works and whose patterns of vocalizations I have subjected to statistical

analysis, and here I wish to acknowledge them for providing me with so much musical, scientific and philosophical inspiration since the start of my PhD.

It has turned out that many people are interested in the connections between whale song and human music, and I have had productive and helpful conversations with friends and colleagues including but not limited to Robert Churchill, Franca Eichenberger, David Fennessy, Jan Foote, Norman Gray, Iraklis Ioannidis, Bianca Keeley, Eugene de Klerk, Annie Lewandowski, Alistair MacDonald, Jennifer Martin, Sara Niksic, Éadin O'Mahony, Bryony Pierce, members of RASCAL, David Rothenberg, Susan Stuart (who also kindly read and commented on Chapter 4), Şebnem Susam-Saraeva and Hollis Taylor. I thank Ana Eguiguren and Taylor Hersh for reading and commenting on Chapter 3, and Amy Cutler for providing me with the title to my PhD. It was a particular privilege to be able to talk to Katy Payne, whose passionate writing on both music and humpback song have been inspirational throughout my PhD.

I also want to thank those who have helped me to take my research out of the academy, whether into schools, onto the concert stage, or for public broadcast: Jessica Argo, Raymond MacDonald and Gerry Rossi ('GIOfest'), Connie Edwards (CBC), Howie Firth ('Orkney International Science Festival'), Fiona Robertson and Pete Stollery ('Sound Festival'), Monika Smekot ('Make a Splash'), Alison Stirling ('Human Threads'), and Bede Williams (University of St Andrews).

Unswerving support has come throughout from Mum, Robert, Fiona, and other members of my family. Finally, a leviathan debt is owed to Oliver Rundell for having indulged and encouraged my enthusiasms for all things cetacean, for coping with the peaks and troughs of my research mood, and for not complaining about my almost-constant distractedness over the past several months of our shared life. Olly also read the whole of this thesis and helped me to improve the expression of my ideas.

I would like to dedicate this thesis to a friend who died too soon to see it, but who would have fully appreciated and robustly critiqued all of its aspects. Matthew Neale (1972–2017): zoologist, musician, philosopher.

Ethics Statement

This study was approved by the University of St Andrews School of Biology Ethics Committee. Song recording in French Polynesia was conducted under permits issued to M. Michael Poole by the Ministry of the Environment, French Polynesia.

Funding

This PhD was supported by a St Leonard's Interdisciplinary Scholarship funded by the University of St Andrews and Royal Conservatoire of Scotland (RCS), with additional funding for fieldwork and conferences provided by the RCS Research Degrees Committee, the Russell Trust and the Ken Murray Fund.

Digital Outputs Access Statement

Digital outputs underpinning this thesis are available at osf.io/rg6bx

Table of Contents

Table of Acronyms

Acronym	**Full Expression**
IOI	Inter-Onset Interval
NSU	Number of Song Units
PD	Phrase Duration
PV	Phrase Variant
SI	Supplementary Information
SNR	Signal to Noise Ratio
WRV	Within-individual Rhythmic Variability

Portfolio Contents

My portfolio contains three humpback whale song recordings selected because of their importance to this thesis, alongside six original works of music listed below. For each of these works, scores (where appropriate) are included alongside other performance materials and audio and video recordings of performances (in bold in the tables below).

Humpback Whale Song

1.	Singer 150918 Audio Recordings (WAV, MP3)	'Singer150918.wav' 'Singer150918.mp3'
2.	Singer 190919 Audio Recordings (WAV, MP3)	'Singer190919.wav' 'Singer190919.mp3'
3.	Singer 191026 Audio Recordings (WAV, MP3)	'Singer191026.wav' 'Singer191026.mp3'

'(im)possible gestures'

1.	Score	'impossible gestures score.pdf'
2.	Performance Notes	'impossible gestures performance notes.pdf'
3.	Programme Note	'impossible gestures programme note.pdf'
4.	Audio Recording (WAV)	**'impossible gestures SCQ live Glasgow 2019.wav'**
5.	Audio Recording (MP3)	**'impossible gestures SCQ live Glasgow 2019.mp3'**
6.	Video Recording (MP4)	**'impossible gestures Phillip Island 2019.mp4'**

‘The Path of the Unseen Whale’

1.	Score, Performance Instructions, Programme Note	‘The Path of the Unseen Whale Score, Performance Instructions and Programme Note.pdf’
2.	Ableton Live 10 Project File	‘The-Path-of-the-Unseen-Whale.als’
3.	Fixed Media Audio Files (WAV)	‘150918-003-silence1-stereo.wav’ ‘150918-003-silence1-stage1-stereo.wav’ ‘150918-003-silence2-stage1-stereo.wav’ ‘150918-003-silence3-stage1-stereo.wav’
4.	Audio Recording (WAV)	**‘The Path of the Unseen Whale Studio Alex South.wav’**
5.	Audio Recording (MP3)	**‘The Path of the Unseen Whale Studio Alex South.mp3’**

‘Underwater Rain (Farehau Humpback, 26.10.2019)’

1.	Audio Recording (WAV)	**‘Underwater Rain.wav’**
2.	Audio Recording (MP3)	**‘Underwater Rain.mp3’**
3.	Programme Note	‘Underwater Rain Programme Note.pdf’
4.	Preparatory Score	‘Underwater Rain Bassline.pdf’

‘Entanglement’ (Alex South with Katherine Wren)

1.	Audio Recording (WAV)	**‘Entanglement Studio Alex South Katherine Wren.wav’**
2.	Audio Recording (MP3)	**‘Entanglement Studio Alex South Katherine Wren.mp3’**
3.	Programme Note	‘Entanglement Website Text.docx’
4.	Preparatory Scores	‘Farehau Humpback 'bassline' transcribed viola.pdf’ ‘Farehau Humpback 'bassline' transcribed bass clarinet.pdf’

‘Whale, Bow, Echo’ (Alex South with Sequoia)

1.	Score	‘Whale Bow Echo Score.pdf’
2.	Fixed Media Track (WAV)	‘Whale Bow Echo Fixed Media Track.wav’
3.	Projections (M4V)	‘Whale Bow Echo Projections with Fixed Media Track.m4v
4.	Performance Instructions	‘Whale Bow Echo Performance Instructions.pdf’
5.	Programme Note	‘Whale Bow Echo Programme Note.pdf’
6.	Audio Recording (WAV)	**‘Whale Bow Echo Studio Sequoia.wav’**
7.	Audio Recording (MP3)	**‘Whale Bow Echo Studio Sequoia.mp3’**
8.	Video Recording (M4V)	**‘Whale Bow Echo Live Sequoia Glasgow 2022.m4v’**

‘Submergence’ (Jan Foote, realization of live part by Alex South)

1.	Score and Part	‘Submergence Score with Solo Realization.pdf’ ‘Submergence Live Clarinet Part.pdf’
2.	Fixed Media Track (WAV)	‘Submergence Fixed Media Track.wav’
3.	Performance Notes	‘Submergence Performance Notes.pdf’
4.	Programme Note	‘Submergence Programme Note.pdf’
5.	Audio Recording (WAV)	**‘Submergence with Solo Realization Studio Alex South.wav’**
6.	Audio Recording (MP3)	**‘Submergence with Solo Realization Studio Alex South.mp3’**

Chapter 1 Whale Song and Whale Music

The main goals of this chapter are to provide sufficient material on humpback whales and their ‘song’ on the one hand (§1.2), and humanly-created ‘whale music’ on the other (§1.3), to enable the original music and research presented in the remainder of this thesis to be situated and evaluated. I preface this material with a description of the unusually broad cross-disciplinary scope of my project. Before that, I mention some relevant facts from my own multidisciplinary background. For the reader more centred in the arts, the humanities, or the sciences, I have provided a Glossary towards the end of this thesis (p. 245) to explain my usage of general terms whose meanings vary according to discipline. Technical bioacoustical and music theoretic terms regarding humpback song structure, individual variability, and rhythm are defined in Chapter 2 (Table 2.1).

1.1 Multiple Perspectives

I am a musician whose performance practice in recent years has circled around the areas of free improvisation, experimental music, and contemporary classical music (The One Ensemble Orchestra, 2009; Hair, 2013; Sound of Yell, 2014, 2020; The One Ensemble, 2017; Glasgow Improvisers Orchestra and Maggie Nichols, 2019), and has included collaborations with dancers, poets and visual artists, and empirical musicologists (Bailey *et al.*, 2014, 2017). My initial training on my instruments (clarinet, bass clarinet) was within the conventional systems, ensembles, and institutions of Western classical music, including postgraduate studies at the Royal Conservatoire of Scotland. Before this, I had studied the natural sciences at the University of Cambridge, gaining a PhD and doing post-doctoral work in atmospheric chemistry (Povey *et al.*, 1998; South *et al.*, 1998; Mohamed-Tahrin *et al.*, 2001). For many years, then, I was immersed in the empirical and quantitative methods of an observational science. Following conservatoire musical training, I studied philosophy at the University of

Glasgow, where I wrote my undergraduate dissertation on the French philosopher Maurice Merleau-Ponty, and later taught on courses on Consciousness, Kant and Phenomenology (South, 2011). Here I came to value the rigorous first-personal approach to areas including mind and embodiment offered by hermeneutic phenomenology, with its focus on thick description and self-understanding, and first came into contact with Critical Theory. I also began to question the default scientific naturalism that I had generally accepted until then, and became more aware of the assumptions underlying many of my beliefs about the external world, other humans, and myself. Finally, it was here that I started to learn more about the perceptual worlds, cognitive powers and social lives of other-than-human animals, which primed me for the research I describe in this thesis.

Methodology

These studies, carried out partly at a conservatoire for the performing arts, partly at a university, combine practice-as-research, bioacoustics, and zoömusicology (Glossary). My Portfolio, containing six musical works influenced in various ways by humpback whale song, is one component of my attempt to answer the question "How can I compose so as to maintain fidelity both to the songs of the humpback whale, and to the singers themselves?" Another component is found in Chapter 4, where I use concepts and analytical tools drawn from zoömusicology, Critical Theory and critical environmental ethics in order to critique the construction, interpretation and musical use of representations of other-than-human animal sounds, and to develop new strategies to produce multispecies music that neither enacts nor embodies human exceptionalism. The pieces in this Portfolio were written at the same time as I was carrying out the bioacoustical research into the variability of humpback song rhythms described in Chapters 2 and 3. For this research I followed standard biological principles of data collection and analysis, for example using statistical methods to assess the validity of scientific hypotheses. This methodology manifests an epistemology that is robustly scientific: the methods detailed in Chapters 2 and 3 have enabled me to make claims that are open to replication by any other investigator who accesses the same data with the same tools.

Epistemology, Ontology, Ethics

Despite these different disciplinary approaches, I contend that all my analyses begin from experience, and so can be brought together in an epistemology that philosopher Bruno Latour has referred to as an "additive empiricism" that is "interested in objective facts and grounded claims," but also seeks to "complicate, to specify, and, whenever possible, to slow down and, above all, hesitate so as to multiply the voices that can be heard" (Foreword to Despret, 2016, p. ix). And although I have not used zoömusicological or philosophical methodologies to attempt to answer scientific questions (or vice versa), I believe that this additive empiricism unifies my bioacoustical analyses of song rhythms in Chapters 2 and 3 with my zoömusicological-philosophical critique of representation in Chapter 4.

Throughout my research I have been concerned to avoid misrepresentation: both in the interests of scientific accuracy (Chapters 2 and 3), and in the interests of the humpback singers themselves (Chapter 4). And although it is the latter that is more obviously connected with the ethical question of how humans treat other animals, there are also clear ethical questions involved in their scientific study. My own fieldwork, which involved making recordings from a small boat in the proximity of humpback whales, required ethical approval from the Animal Welfare and Ethics Committee at the University of St Andrews. However, for the scientist there is also an ethical concern for truth in the pursuit of objective knowledge.

My turn to ethics and recommendations for praxis naturally follows from my adoption of the perspectives of Critical Theory and Plumwood's environmental ethics, which fuse empirical analysis and philosophical argumentation with a concern for emancipation (Bohman *et al.*, 2021). It is also consistent with philosopher Baptiste Morizot's proposal for an approach to the close observation and study of other animals that resonates with my own thinking on fidelity and critical anthropomorphism. “The neo-naturalist is a field naturalist who practises his art without forgetting that he is an animal, without forgetting that those he investigates are much more than inert matter...” (Morizot, 2021, p. 114). Morizot describes his neo-naturalist as freed from reductive naturalist ontologies, thereby restoring agency to other-than-human animals; additionally he argues for a movement from a detached scientific to a participatory ethical attitude: “The question of knowledge is not… that of disembodied truth, but that of the best cohabitation between living beings, in shared territories...” (2021, p. 115). Morizot’s writing situates him in an emergent and interdisciplinary ‘multispecies studies’, an umbrella term bringing together researchers in the social sciences and the humanities with artistic practitioners “united by a common interest in better understanding what is at stake—ethically, politically, epistemologically—for different forms of life caught up in diverse relationships of knowing and living together” (van Dooren *et al.*, 2016, p. 5). Consistent with biologist von Uexküll’s theory of *Umwelten* (Uexküll, 2010), the dichotomy between the simplistic realism of “a singular world ‘out there’ awaiting description” and a total relativism in which every subject has their own world is rejected (van Dooren *et al.*, 2016, p. 12). Rather, what exists is shaped by “multispecies worldings,” dynamic material-semiotic processes in which particular ways of knowing are “never innocent… but, rather, are situated, embodied, and historical practices” (van Dooren *et al.*, 2016, p. 12). “It matters what matters we use to think other matters with…” (Haraway, 2016, p. 12). According to this understanding, epistemology cannot be distinguished from ontology. Furthermore, as van Dooren *et al.* argue, our recognition that we are involved in making worlds implies that we bear an accountability “for how and what they become” (2016, p. 15), i.e., ethical issues are at the heart of all our relations within the more-than-human world. They are also central to our human *relatings* of these relations, whether verbal or otherwise. It matters what musics we use to music other musics.

1.2 Humpback Whale Song from the Musical Standpoint

Here I introduce humpback whale song from the viewpoints of a human musician interested in creating music with it and a zoömusicologist interested in its similarities and differences to human music. In Chapters 2 and 3 I discuss its possible functions for the whales themselves. In this section, I focus on the 'what' and the 'how' rather than the 'why', but I begin with some evolutionary history.

Humpback whales belong to the taxon of mysticetes (baleen whales) that diverged from the odontocetes (toothed whales) c.36 million years ago (McGowen *et al.*, 2009). Together, mysticetes and odontocetes are known as cetaceans, and include approximately 90 species of whales, dolphins and porpoises. Evolving in an environment that placed severe constraints on vision and olfaction, all cetacean species have highly developed systems of acoustic communication. Odontocetes, including the well-studied bottlenose dolphin (*Tursiops* spp.), orca (*Orcinus orca*) and sperm whale (*Physeter macrocephalus*), use echolocation to detect and pursue their prey, sometimes to extreme depths, and communicate with each other via diverse systems of clicks and calls. But they do not possess what biologists call 'song'. Definitions vary by taxon, and the dichotomy between calls and song may have fuzzy edges (Clark and Gagnon, 2022, p. 22), but for cetaceans song refers to the repetitive vocalizations predominantly produced in the breeding season and so far described in around half the known species of mysticetes (Clark and Garland, 2022). Documented singers have been restricted to males in humpback (Glockner, 1983), blue (*Balaenoptera musculus*, Oleson *et al.*, 2007), and fin (*Balaenoptera physalus*, Croll *et al.*, 2002) whales, and only for the bowhead (*Balaena mysticetus*) has it been proposed that the females may also sing (Tervo *et al.*, 2011).[1] The parallels with better-known avian songsters makes it likely that mysticete song is a reproductive

---

[1] Among odontocetes and mysticetes, both males and females produce calls or 'social sounds', but in this thesis I am only concerned with song.

advertisement display that has evolved through inter- and/or intra-sexual selection (Tyack, 2022), but its adaptive functions have been hard to establish with certainty due to the challenges of studying such long-lived and inaccessible animals. Such an evolutionary explanation is also not the only possible answer to a question such as "Why does a humpback whale sing?"[2]

1.2.1 Song Traditions

In the humpback, the song is shared by all singers in a population and progressively changes during the breeding season (Payne *et al.*, 1983; Cerchio *et al.*, 2001) and over much longer timescales (Payne and Payne, 1985). This pattern of song evolution was early suspected (Payne *et al.*, 1983) to require a capacity for vocal learning that is rare among mammals (Janik and Slater, 1997) but has been demonstrated in the bottlenose dolphin (Herman, 2002) and orca (Abramson *et al.*, 2018). The existence of changing humpback song traditions has led to comparisons with the dynamics of oral transmission in human music from within biology (Payne *et al.*, 1983; Guinee and Payne, 1988; Payne, 1995, 2000), zoömusicology (Mâche, 1992, p. 152; Martinelli, 1999), and ethnomusicology (Sorce Keller, 2012). Observations in the South Pacific of rapid population-wide replacement of one song type with another (Noad *et al.*, 2000; Garland *et al.*, 2011) have sharpened such comparisons (Doolittle, 2007a; Rothenberg, 2008b; Martinelli, 2009; Whitehead and Rendell, 2015). For example, the spatiotemporal patterns of these replacement events are inexplicable without vocal learning (Janik, 2014), pointing to humpback whale song as a key example of other-than-human animal culture (Whitehead and Rendell, 2015; Whiten, 2019; Garland and McGregor, 2020). Agent-based modelling studies have suggested that learning biases such as for conformity and novelty may play a role in constraining this cultural transmission (Mcloughlin *et al.*, 2018; Zandberg *et al.*, 2021). This is analogous to trends in human musical features or styles, where

---

[2] From within biology, expansions of Tinbergen's (1963) classic framework of 'four whys' (mechanism, ontogeny, phylogeny and evolutionary function) have proposed additional levels of explanation indexing animal awareness (Burghardt, 1997; Bateson and Laland, 2013) and culture (Fitch, 2018).

musicological accounts have been supported by the modelling of sampling traditions in popular music (Youngblood, 2019) and harmonic dissonance in Western classical music (Nakamura and Kaneko, 2019). There are further parallels with human culture, such as between the cyclical rise and fall of complexity of South Pacific humpback song (Allen *et al.*, 2018, 2022) and changes in instrumental complexity in human popular music (Percino *et al.*, 2014). Finally, the way in which humpback song structures continually accumulate modifications, possibly driven by individuals seeking to increase their attractiveness to potential mates, has raised the question of whether humpback song traditions should count as an example of cumulative cultural evolution (Garland *et al.*, 2021; Sinclair *et al.*, 2022), believed by many researchers to be a uniquely human characteristic (Dean *et al.*, 2014).

1.2.2 Song Structure and Rhythm

Although the discoveries of song evolution and revolution have been partially responsible for provoking comparisons between humpback and human singers, biologists and musicologists alike have also been struck by similarities in structural features, in particular the multilevel organization of humpback song (Figure 1.1), unique among mysticetes. The nomenclature of ‘song’, ‘theme’, ‘phrase’, and ‘unit’ (sometimes ‘song unit’) was established by Roger Payne and Scott McVay (1971) in the publication that first provided evidence for this multilevel organization, and although later work on classification has provided valuable advice on how to resolve ambiguities in the delineation of phrases (Cholewiak *et al.*, 2013), the basic description used by most current researchers remains unchanged.[3]

---

[3] Roger Payne and McVay use musical terminology throughout, e.g., describing one theme as “quicker in tempo and less rhythmic than the rest of the song… [in another] the arpeggios become more elaborate with each repetition of this phrase, until, by the end of the theme, the component notes become more sustained” (1971, p. 593). In a spoken presentation ‘Songs of the Humpback Whale: Faith, Structure and Listening in the Twentieth Century’ to the online discussion group ‘Research in Animal Songs in Culture, Art and Literature’, Rachel Mundy suggested that such descriptions may owe much to forms of analysis studied by Katy Payne at Cornell University with musicologist Donald Grout, and that “her role in the discovery of humpback whale song is part of a more complex story about gender, ethical responsibility, and women in the sciences”. This presentation, given 5 November 2020, forms the basis of a chapter in Mundy’s forthcoming monograph (Mundy, forthcoming).

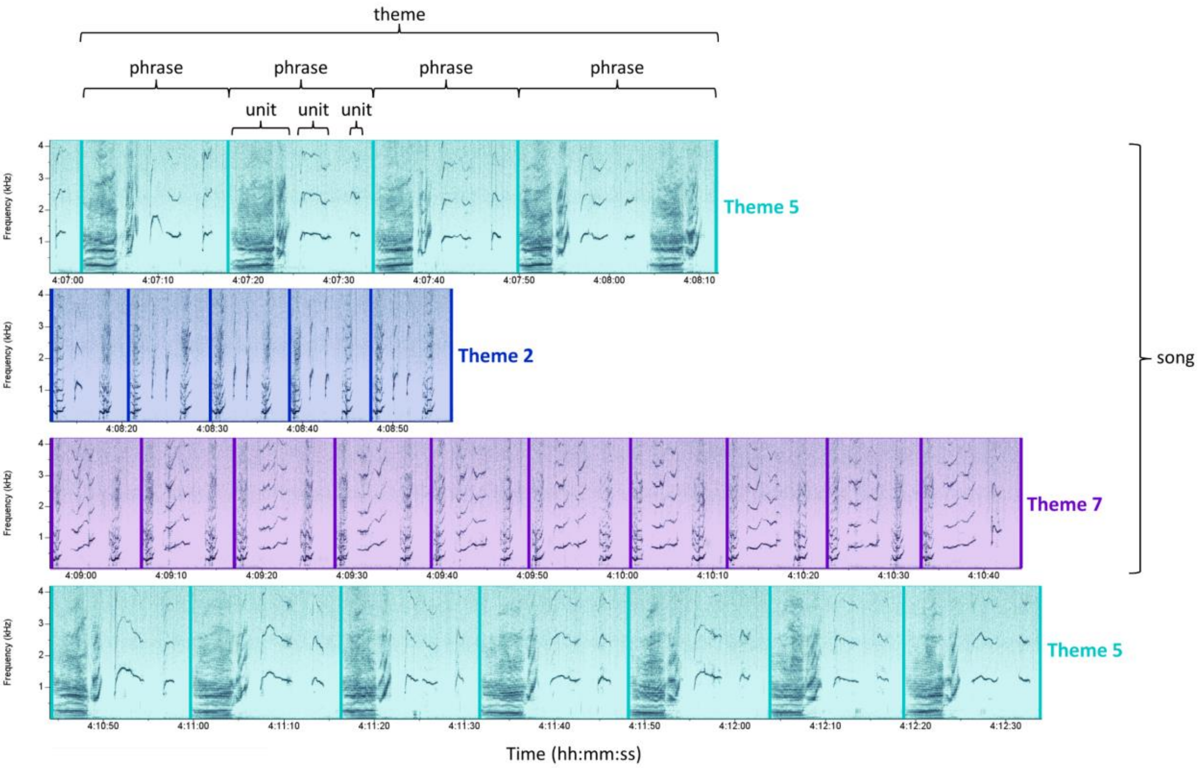


Figure 1.1 Structure of humpback whale song, recorded 25 September 2019 off Mo'orea, French Polynesia. Individual sounds are called units, and arranged in short patterns known as phrases, here separated by vertical coloured bars. Phrases repeat with variation to form themes, which follow each other without a break, and a sequence of themes (here: 5-2-7) is known as the song.

Working upwards from the unit, "the shortest sound continuous to our ears when heard in 'real time'" (Payne and McVay, 1971, p. 590), a phrase consists of a fixed pattern of units. Phrases are repeated with variation, to make up a theme, and a sequence of themes is termed the song. The order of themes within a sequence is stereotyped, though this too can vary and individual themes are frequently dropped (Payne *et al.*, 1983, p. 17). Finally, songs are repeated without break to comprise a 'song session'. Other terminology occasionally used is the 'subunit', (Payne and McVay, 1971, p. 591) to describe fine-scale variability within a unit, and 'subphrase' (Payne *et al.*, 1983, p. 17) to describe repeating patterns on a scale below the phrase. Falling between two themes, 'transitional phrases' contain song units or subphrases from both (Payne *et al.*, 1983, p. 49). Recently 'motif' has been used to describe any repeating pattern of units within a theme (Lamoni *et al.*, 2023), matching the 'motivic pattern' identified as a cross-cultural regularity in human music (Savage *et al.*, 2015). Themes are differentiated into static, shifting, and unpatterned, according to whether their phrases are nearly identical, subject to progressive modification within a theme, or "have no clear organization" (Payne and Payne, 1985, p. 99). These different levels span a wide range of timescales, with typical units lasting 1–2 seconds (and inter-unit silences of similar timing), phrases 10–20 seconds, themes a minute or two, and songs 5–15 minutes. Within a theme, repetitions of a specific phrase tend to be rather consistent in duration (Thompson, 1981), but the number of repetitions varies widely (Payne *et al.*, 1983). The repetition of song can continue for many hours of each day during the breeding season, with the longest recorded song session of c.22 hours documented in the West Indies (Winn and Winn, 1978).

From the perspectives of composition and zoömusicology it is interesting that the nested timescales are familiar to human audiences for many forms of music. In an extensive comparison between humpback whale song and human music, Roger Payne noted the similarity in duration of song phrases and commented that "[t]he length of humpback whale songs falls between the length of a modern ballad and a movement of a symphony" (Payne, 1995, p. 146). The "rhythmic monotony" (Thompson, 1981, p. 17) of phrase repetitions in static themes establishes a temporal grid that Handel *et al.* (2012) have compared with musical metre. This aligns much humpback song with the distinctive and memorable cyclic rhythms driving so many forms of dance and popular musics (Toussaint, 2020, p. 11), even if the humpback 'tempo' (rate of events) is slower than is typical for human musicians. It is the degree of rhythmic consistency of phrase repetitions that will concern me in Chapters 2 and 3 of this thesis.

1.2.3 Song Sounds and Vocal Production

The majority of humpback song units fall within human vocal range, with those whose pitches are too low or high to be sung mainly capable of being played on musical instruments. Although harmonics have been recorded up to 24 kHz (Au *et al.*, 2006), most acoustic energy falls within 30–5000 Hz (Girola *et al.*, 2019). This is comparable with the pitch range (A0–C8) of a standard grand piano (fundamental frequencies 27.5–4186 Hz when tuned to A4 = 440 Hz), and exceeds that of my own instruments (the combined fundamental frequency range for clarinet and bass clarinet spans c.58–c.2100 Hz). The smoothly changing frequency contours of song units evident in Figure 1.1 are very characteristic of humpback sounds, distinguishing them from the sets of more or less discrete pitches more commonly found in human musics. Song sounds, especially the low frequency units, tend to be loud, with Girola *et al.* (2019) reporting source levels of 138–187 dB re 1 μPa at 1 m across nine unit types. These levels are higher than a typical jet airliner at 10 m heard in air, when the necessary conversion for the different transmission medium has been performed (National Research Council (US) Committee on Low-Frequency Sound and Marine Mammals, 1994). Due to the high density of water compared with air, sound travels faster and further underwater, and typical humpback song “may be audible over thousands of square kilometres in favourable conditions” (Payne, 2000, p. 136).

Although details are lacking of the sound production mechanisms *in vivo*, dissections performed on stranded whales have allowed biologists to specify the anatomical structures responsible for the wide variety of humpback song units. The humpback sings underwater, typically head down and at a depth of c.15–25 m (Au *et al.*, 2006), surfacing every 5–20 minutes to breathe (Chu, 1988). Under these circumstances, sound is produced with the blowholes closed against water ingress, and is thought to be generated by the internal movement of air through three valve-like structures (Cazau *et al.*, 2013) including the U-fold located in the larynx and homologous to the vocal folds found in other mammals (Reidenberg and Laitman, 2007). With the aid of a theoretical model (Adam *et al.*, 2013), Cazau *et al.* (2013) proposed that recirculation of air between the lungs and the laryngeal sac, an elastic structure found in many primates including the chimpanzee (Nishimura *et al.*, 2007; Nishimura, 2020), coupled with three different configurations of the valves and consequently different resonating cavities, permits three different classes of sounds to be produced: (1) low frequency pulsative sounds, (2) mid-frequency harmonic sounds modulated by formants, (3) high-frequency harmonic sounds without formants. Cazau *et al.* (2013) further argued that sounds with different qualities could be generated during both ingressive and egressive airflow. However, subsequent anatomical investigation of the U-fold has revealed a dorsal-ventral asymmetry that makes it unlikely that sounds can be produced during ingressive flow to the lungs (Damien *et al.*, 2019), which could explain the silences between humpback song units.

Humpback whale science has tended to use onomatopoeic terms for different unit types (e.g., Garland *et al.*, 2017), and I give some examples here to indicate the variety of sounds represented in the spectrograms above (Figure 1.1): pulsative long moans, pulsative ascending cries, and whistles (Theme 5), modulated moans and whistles (Theme 2), and modulated moans and long ascending cries (Theme 7). Such terms provide more information on timbre than pitch or duration; a more complete idea of the timbral range produced may be gained by listening to the song or examining the spectrograms and unit names given in Appendix B. To the human ear, the transition from one theme to another is apparent in part through the changes in ‘sonic palette’ produced by the typical restriction of song unit types to specific themes. This qualitative assessment is supported by an analysis of conditional source entropy showing that specific unit types predict which theme the whale is singing (Handel *et al.*, 2012). Exceptions would be expected where adjacent themes share subphrases, the so-called ‘rhyming’ character sometimes found in more complex humpback song structures (Guinee and Payne, 1988) and displayed by the shared modulated moans of Themes 2 and 7 (Figure 1.1). In non-rhyming song, Handel *et al*. (2012) point out that theme-specific unit types differentiate humpback song composition from typical human musical practice, where the same notes are re-used from one piece or phrase to another, although it is true that contrasts between pieces or between sections of pieces are sometimes achieved through the use of pitch sets (whether or not a key structure is formally in place). The fact that this sonic palette is also continuously morphing during song evolution, with changes in frequency, contour, duration and timbre, further differentiates humpback vocalizations from the oral traditions of human song in which the composition of new works and evolution of style, at least within the lifetime of individual performers, has not typically been accompanied by changes in the basic elements of composition (notes, scales, rhythms).

1.2.4 Solo and Group Singing

So far, I have considered song sounds and structures without considering social or behavioural context. But for cultural transmission to occur, humpback singers must be monitoring the song produced by other singers, and it is possible that singing may mediate male-male interactions. On the breeding grounds, singers are generally solo (Winn and Winn, 1978; Darling *et al.*, 2006) and "relatively stationary" (Frankel *et al.*, 1995, p. 1135). Found several kilometres apart (Tyack, 1981), singers tend to be more widely-spaced than non-singers (Frankel *et al.*, 1995). This suggests that singing functions to maintain spacing between them, with the mean distance of c.5 km implying that singers are close enough to hear each other. This is supported by Payne and Payne's observation that "[c]ircumstances in which only one singing whale is audible on the offshore banks are relatively rare" (1985, p. 105). However, this separation does not mean that males do not interact: in an analysis of observations of 167 singers off Maui, Hawai'i, Darling *et al.* (2006) found that a large majority of singers (89%) sang until approached by non-singing males, at which point they ceased singing. Such approaches were usually followed by non-agonistic interactions. Interpreting their results, Darling *et al.* speculated that individual differences in song offer a "mutual recognition system" (2006, p. 1093) that allowed singers to judge their degree of association with others, though this has been judged implausible on the grounds of song convergence within a population (Herman, 2017, p. 1805). At present, however, the level of individuality of song is not known, and recent detailed work on South Pacific song is beginning to illustrate that there are significant between-individual differences in phrase repertoires (Murray *et al.*, 2018; Lamoni *et al.*, 2023). In Chapter 3 I investigate between-individual variability in the rhythms of shared song phrases.

For a human musician, this picture of widely-spaced solo singers within earshot of each other immediately raises the question of whether this might lead to song synchronization or other direct vocal interaction. After all, duo and other group singing is known to take place in other mammals, including white-hand gibbons (*Hylobates lar,* Terleph *et al.*, 2017), and a wide variety of bird species, including happy wrens (*Pheugopedius felix,* Templeton *et al.*, 2013) and pied butcherbirds (*Cracticus nigrogularis,* Taylor, 2017b). However, thus far there have been very few studies bearing on humpback group song, and early post-hoc analyses of recordings containing the voices of multiple singers from a 19-year dataset revealed no evidence for synchronization: "The songs overlap randomly and with typical internal variability, as if each singer were oblivious of the specific progress of the others around it and was locked into its own routine" (Payne and Payne, 1985, p. 105). For example, a solo whale joined by other singers did not change its song presentation, and in a further recording featuring three singers song durations varied widely. A more recent and targeted evaluation of simultaneous singing of humpbacks off Isla Socorro, México (Cholewiak *et al.*, 2018) also did not find compelling evidence for song synchronization. Assessed as the tendency for singers to overlap phrase types, only two out of eight pairs of singers showed significantly more overlap than expected by chance, one significantly less. However, a majority of singers (11/12) were found to modify their song presentation (the rate of switching between phrase types) when a second singer became audible, which offers a potential mechanism for the mediation of intrasexual interactions. Payne and Payne suggested that there are more interesting possibilities for group song than simple singing in unison, and that "[o]ccasional portions of whale choruses can be described in terms of musical effects like antiphony, fugue, or pedalpoint" (1985, p. 110), but for now it appears safer to conclude that the collective singing of humpback whales is better described as an "asynchronous chorus" (Herman, 2017), or as heterophony (Sorce Keller, 2012, p. 172). In Chapter 4 I shall put the latter concept to work in characterizing the multispecies musicking that occurs when human musicians use field recordings of whale song in their compositions, and I turn now to this practice.

1.3 Whale Music and Activism[4]

The recent histories of human scientific and musical engagements with the sounds of cetaceans are inextricably entangled with those of Western anti-whaling activism and modern environmental movements (Rothenberg, 2008b; Burnett, 2012; Ritts, 2017), with the work of Katy and Roger Payne forming a focal point. It took their biological and musical skills and training to recognize that the long-breathed vocalizations of humpback whales fitted the biological description of song and had much in common with the structures of human music (Rothenberg, 2008b; Mundy, forthcoming), and their research carried out individually and together (Payne and McVay, 1971; Guinee *et al.*, 1983; Payne and Guinee, 1983; Payne *et al.*, 1983; Payne and Payne, 1985; Guinee and Payne, 1988; Payne, 1995, 2000) is still regarded as foundational. The nature recording *Songs of the Humpback Whale* (Payne, 1970) introduced the sounds of the humpback whale to the general public shortly before the publication of Roger Payne and Scott McVay's (1971) article in the widely-read journal *Science*. In §1.3.2 I provide a summary account of the impact that the field recordings had on contemporary musicians and activism, but I begin by describing some of the past and present human existential threats to cetaceans.

---

[4] In this section I draw from previously published material (South, 2022).

1.3.1 Whales have Voices but do not have a Voice

Despite the growing awareness of the diversity of cetacean vocalizations (at the level of species, population and the individual), contemporary industrial human societies have until recently denied cetaceans any form of voice in the sense of representation or recognition of their interests, rights or agency. The working of the "anthropological machine" (Agamben, 2004) has ensured that whales (among other nonverbal species) have been positioned outside of moral consideration, valued only as means to human ends. Many species were serially hunted during the nineteenth and twentieth centuries to near-extinction; the oils derived from blubber and spermaceti being highly prized as lubricants and for any number of other purposes from lighting to the manufacture of soap and margarine (Burnett, 2012). Despite the moratorium on the commercial hunting of large whales declared in 1982 by the International Whaling Commission, Iceland, Norway and Japan have continued such industrial whaling. Dolphins and smaller whales are under threat both from the deliberate hunting that kills around 100,000 cetaceans per year (Altherr and Hodgins, 2018) and as the accidental bycatch from human fishing practices that Michael Moore refers to as "whaling by default" (Moore, 2014) and is estimated to kill over 300,000 cetaceans every year (Read *et al.*, 2006). Large whales too are vulnerable to entanglement in human fishing gear, which can cause extreme suffering and a prolonged death (Moore and van der Hoop, 2012). Annual mortality is extremely hard to estimate, but recent research is beginning to reveal the startling extent of the danger: in North American coastal waters the *majority* of adequately examined blue, fin and humpback whales are marked with the scars of entanglement with fishing gear (Ramp *et al.*, 2021). Other existential threats include ship strikes, plastic pollution, and perhaps most invidious of all in the context of this thesis, the increasing levels of underwater anthropogenic noise linked with human economic and military activity (McDonald *et al.*, 2006) that have been shown to cause increased stress levels (Rolland *et al.*, 2012), changes to foraging and communication behaviour (Parks *et al.*, 2007) and most likely mass strandings (Simonis *et al.*, 2020). This final threat demonstrates an awful irony: that in their own domain the functionally complex, potentially expressive and arguably articulate vocalizations of cetaceans are being increasingly obscured (Clark *et al.*, 2009) by the mechanical noises made by human container vessels, sonar, and air guns used for oil and gas exploration.

1.3.2 A Soundtrack for Activism

Songs of the Humpback Whale (Payne, 1970), which remains the best-selling nature recording of all time, was released with an explicit plea to consider the plight of whales and rapidly became "the soundtrack to the 'Save the Whales' campaign" (Burnett, 2012, p. 629). Played on the floor of the House of Congress in the US and at meetings of the International Whaling Commission (Burnett, 2012, pp. 630–631), and receiving radio airtime in the US and Japan (Rothenberg, 2008b), the album is widely credited with being a key resource for activism in the period leading up to the International Whaling Commission's 1982 moratorium (Rothenberg, 2008b; Burnett, 2012; Ritts, 2017). Furthermore, alongside the identification of the sounds made by humpbacks as song (Payne and McVay, 1971), the LP helped to drive what Burnett labels the "dramatic transformation of attitudes toward cetaceans in the period 1960–1975" that took place in the Western world, from "industrial commodity of dwindling importance" (Burnett, 2012, p. 521) to intelligent and self-aware agents with complex inner lives. Ongoing research into cetacean cognition, communication, sociality and culture (Whitehead and Rendell, 2015) continues to support demands that humans ought to widen the sphere of the "morally considerable" (Gruen, 2021) and radically reconsider the ethical position of cetaceans (White, 2007; Marino and Frohoff, 2011; Wichert and Nussbaum, 2017, 2019).

However, my focus here is not on the *Songs of the Humpback Whale* field recordings, but on the mediating role between humpback singers and human listeners played by music creators. Since the album was released, dozens of musicians directly incorporated recordings of humpback song into their musical works (Appendix A) in such ways as to allow the song to be heard by the sceptical listener as having music-like qualities, taking it beyond "haunting mewls and honks" or "wheezing bleats" (Burnett, 2012, pp. 629–631) into what I term the realm of the 'aesthetically considerable'. The link between aesthetics and ethics was stated starkly by Roger Payne: "Human musicians like Judy Collins and Paul Winter began to include the sounds of whales in their performances and compositions. Killing and eating a whale now became killing and eating a 'musician'" (Payne, 2013). Some of these performances and compositions reached very large audiences, hence enhancing the potential for awareness-changing and transformative encounters between humans and humpback song. For example, Judy Collins' album *Whales and Nightingales* (1970b) sold over 500,000 copies in the USA within a year of its release (RIAA, no date). Alan Hovhaness's 'And God Created Great Whales' (1970) has been recorded by three major orchestras. Kate Bush's 'Moving' (1978) charted No.1 as single in Japan (Bush, no date a), was performed at the Tokyo Music Festival and broadcast on Japanese TV to a reputed audience of 35 million people (Horner, 2014); the album from which it was taken sold over 1 million copies in the UK (Bush, no date b). In this way, contemporary human musical culture has sampled, interpreted and amplified the reach of the works of an other-than-human vocal culture vastly older than our own. In the next section I examine the musical techniques used in this process, and what they might reveal about our relationship with other-than-human song and its singers.

1.3.3 A Hierarchy of Voices

An obvious starting point for analysis is polyphony, the 'many-voicedness' of music. Multipart musical practices and works can be described in terms of how they combine different melodic lines (Jordania, 2015). There may be a single dominant voice served by accompanying voices, whether these tend to move together ('homophony') or remain in a steady state ('drone polyphony'); alternatively, the music might display a non-hierarchical play amongst equals (‘contrapuntal polyphony’). Sometimes the voices alternate in time in call and response (‘antiphony’). In the cases of antiphony and contrapuntal polyphony it is useful to distinguish between imitative and non-imitative versions. Finally, the term ‘heterophony’ aptly describes humpback whale song, in which multiple individuals simultaneously produce minor variations of the same material without temporal synchronization (§1.2.4).

The human works of music which incorporate the voices of humpback whales using field recordings display a wide range of such textures. Here I consider principally the dominant interplay (or its absence) found between human and cetacean voices, rather than the polyphony occurring among the human musical parts (which may be vocal or instrumental). In a perhaps unlikely pairing, both Bush and Hovhaness make use of antiphony. In the former, short excerpts from *Songs of the Humpback Whale* (Payne, 1970) effectively frame Bush’s three-minute love song for a “moving stranger,” without any transparent musical relationship between human and whale. In the latter, rather longer sections of whale song are heard partly on their own, partly accompanied by aleatoric (free) string writing said by the composer to represent “waves in a vast ocean sky” (Philharmonia Orchestra and Amos, 1989). During the orchestral sections of the piece, there are short but memorable passages in which portamenti (slides) for trombones and violins imitate whale sounds (Figure 4.7). An imitative and responsorial antiphony is found at the start of Paul Winter’s ‘Lullaby from the Great Mother Whale for the Baby Seal Pups’ (1980), in which a humpback phrase whose intervals are readily fitted to conventional Western harmonies seems to initiate close melodic imitation, both whale and human lines underpinned by a bass drone and tonal guitar accompaniment. A similar technique is found in ‘Ocean Dream’ (Winter, 1978), in which “Winter sings a melody that heads up a tritone, the inverse of the whale’s own interval” (Rothenberg, 2008b, p. 42). Towards the end of a later version of this track (Winter and Halley, 1987), the vocal is replaced with soprano saxophone, with its almost uncanny timbral similarity to the humpback. Elsewhere in Winter’s pieces the sense of a human musical responsivity is maintained even where the interweaving of parts is non-imitative (using homophony and contrapuntal polyphony). Collins takes a very different approach in her performance of the nineteenth-century Scottish whaling song ‘Farewell to Tarwathie’ (1970a), in which her solo voice is juxtaposed throughout against that of a solo humpback. This unusual contrastive variant of contrapuntal polyphony (or even mashup) seems to put both performers on an equal footing: the attention of the listener is free to move between two very different kinds of sonic material. Here it is clear that, as in Bush’s pop song, “this animal music comes from a whole different harmonic world” (Rothenberg, 2008b, p. 28), although as Michal Grover Friedlander points out, “[e]ach timbre carries with it qualities of the lament” (2020, p. 35).

1.3.4 Between Naïve Anthropomorphism and Anthropodenial

A number of the composers of the early works incorporating field recordings are known to have been motivated by the goal of bringing the inhumanities of contemporary whaling practices to the awareness of the general public (Rothenberg, 2008b). However, some of these works appear to misrepresent cetaceans either in a naïvely anthropomorphic way, in which the other animal is refused its own characteristics, or through a failure to recognize those characteristics that it shares with humans, analogous to primatologist Frans de Waals’s notion of ‘anthropodenial’ (1999). I have argued elsewhere (South, 2022) that both of these attitudes are aligned with human exceptionalism. There are parallels between my analysis of whale music and that of Şebnem Susam-Saraeva (forthcoming), who from the post-colonial perspective of translation studies argues that since the 1970s musical responses to whale song have displayed “an oscillation between two highly anthropocentric positions: neo-colonialist exoticization (‘They’re so different from us! And therefore intriguing’) or a narcissistic benevolence (‘They’re just like us! Therefore, they deserve protection’).” In both cases, as for the Indigenous inhabitants of colonized nations, the whales themselves may have been silenced.

Alina Feldman has also drawn attention to the exoticization provoked by the presentation of the field recordings:

> “The culture at the time was left to spectacularize the whale sounds as new aural tapestries to get high with. An anthropocentric framework positioned the whales as something external to humanity, placed outside the social body, which made it easier to perpetuate fantasies of the exotic” (Feldman, 2021).

Like anthropodenial, exoticization operates to divide whales from humans, whereas narcissism, like naïve anthropomorphism, compels whales to be like us. To describe the latter phenomenon I will occasionally borrow the term ‘domestication’ from translation studies.

1.3.5 Difference without Distance

In my article (South, 2022), I drew on work on anthropomorphism in ethology (Burghardt, 1991; de Waal, 1999), zoömusicology (Taylor, 2017b) and animal ethics (Karlsson, 2012) to argue that musicians working with other-than-human animal sounds can overcome human exceptionalism through a ‘critical anthropomorphism’, which recognizes continuity between humans and other animals whilst avoiding the projection of human music or musicality into those sounds. I also claimed that adopting an attitude of critical anthropomorphism in re-compositional practices is consistent with an aesthetics of “difference without distance,” perhaps enabling an ethics of “proximity without indifference” (South, 2022). The distance I discussed was the temporal distance effected between human music and field recordings of humpback song. Here I summarize the treatment of difference in three more recent works of whale music, in which distance is minimized by the temporal interweaving of humpback and human soloists.

In Emily Doolittle’s ‘Social Sounds from Whales at Night’ (2007b) difference is evident from the start, where we hear processed other-than-human sounds against melodic gestures from the human vocal or instrumental soloist. It will turn out that these human gestures are mimetic of the humpback song phrase that they seem to summon into existence, a mimesis led not by human harmony and notions of timbral purity but by a fuller account of whale sound. In the central duet between humpback recording and human, differences are welcomed: the composer's instruction is to "Mirror whale song as closely as possible (including 'out of tune' notes). Of course it won't be possible to mimic it perfectly: relish the little clashes between your version and the whale's!" (Doolittle, 2007b, p. 4).

There are certain formal parallels between Annie Lewandowski's appropriately expansive 'Cetus: Life after Life' (2018) and Collins's 'Farewell to Tarwathie' (1970a): humpback song and church carillon sound more or less continuously throughout and retain their acoustic individualities through their very different timbres and registers. Similarities between humpback and human music are also present, though, located at the level of song structure and change. Details of theme and phrase in the humpback song determine the phrases of the human performer, who also at one place models the process of vocal learning that occurs in humpback populations through improvising musical gestures in response to the changes occurring in the field recording. Here the influence of Katy Payne's research and her collaboration with the composer is evident.

Finally, I turn to David Rothenberg's 'Never Satisfied' (2008a), a piece described by its human performer as "my favourite live unedited duet of me playing along with humpback whales" (Rothenberg, 2008b, p. 248). This remarkable performance is fully analysed by Rothenberg elsewhere (Rothenberg, 2008c), and is notable not just for the apparent interaction between clarinettist and whale. From my perspective it is also striking that the distance between humpbacks and human is apparently minimized both temporally and spatially: the recording gives us a real sense that the human player is situated in the same *space* as the whales. Of course, this is not really the case: Rothenberg remained on a boat during the performance with an underwater speaker relaying his sound into the ocean, but nonetheless the perceived proximity has a powerful effect. In terms of difference, Rothenberg's clarinet sound is kept timbrally distinct from humpback cries through either deliberate processing or a quirk of the recording that makes it sound as though the instrument has been played through a narrow-band filter. His material ranges widely across imitative and non-imitative gestures, and the overall impression is as of two experienced free jazz improvisers dialoguing and co-creating, spinning off each other's musical ideas. Crucially, the whale remains fully cetacean throughout, neither domesticated nor exoticized.

1.4 Thesis Overview

In my thesis, I ask the question of how we can faithfully represent the varying structures and rhythms of the song of another species, where the representations take the forms of visualizations, statistical summaries of quantitative measures, and musical scores and performances. I approached this topic from multiple perspectives, bringing with me the tools of different disciplines. With the ears of a musician I listened to and marvelled at the rich and varied sounds made by humpbacks in their songs, structured in some ways so similarly to the human musics I have loved and performed for much of my life, in other ways so very differently. With my ears open, I played with field recordings, where 'playing' signifies both improvising on my instruments in mimetic, complementary and contrastive ways, and creatively experimenting with selected sections as transcriptions or audio clips, reorganizing, transforming or layering them in order to re-compose new multispecies works (Portfolio).

As a natural scientist I learnt to think these sounds through the concepts of bioacoustics, formalizing the categories with which I heard the elements of song, and using existing techniques to represent long stretches of humpback song in symbolic form. I developed further representations and statistical methods used in the study of sounds both human and other-than-human in order to visualize and quantify the variability in the production of the rhythms of those short patterns of sounds known as phrases (Chapter 2).

This variability exists both within the repetitions of phrases produced by an individual singer, where, as a human musician, I might talk about rhythmic precision or consistency, and as differences between one singer and another in the production of a phrase they both possess in their repertoires (individually distinctive rhythms). Here it is important to remember that the humpback whale singers of a specific population share the same song, though individual singers often introduce their own variations. When considering the differences among individuals, I examined the discrepancies in the level of rhythmic precision with which each singer repeats its phrases, and the recurrent or systematic differences that give rise to individually distinctive rhythms. In human music, these systematic differences might allow us to recognize individual performers, or to distinguish among genres or styles. For example, the same jazz standard might be played with different degrees of swing, or indeed with none at all. I found that the humpback singers in my dataset sang with equal rhythmic precision, but with their own individually distinctive rhythms (Chapter 3).

The pieces in my Portfolio explore structural and rhythmic aspects of humpback whale song, but only as my research progressed did I come to realize that the focus of my compositional inquiry was as much ethical as musical: I sought to work respectfully with these songs in ways that did not misrepresent them or their singers. In Chapter 4 I consider my pieces through the concepts of zoömusicology and Critical Theory, approaching the question of representation and misrepresentation through the lens of *fidelity*. In both my compositional and bioacoustical work I was concerned with being faithful to the structures and rhythms of the humpback whale songs which I was listening to, playing with and analysing. I wanted to cite them accurately. This meant becoming more aware of my own musical presuppositions and embodied responses arising from my personal musical enculturation and, possibly, species membership. I wanted to avoid the inaccurate projection of human musical features into humpback song (‘naïve anthropomorphism’), but at the same time I wanted to avoid being blind to the similarities that *do* exist between human music and humpback song (‘anthropodenial’). During my research, I had become increasingly aware that the song recordings in my dataset were not only representations of song, but also represented, in the sense of ‘standing for’, the individual singers who had been recorded. Convinced that humpback whales, like human beings, are the subjects of lives which they themselves help to shape, I began to feel the ethical call to be faithful to them as individual actors, in other words to ensure that my work respected their dignity. Not only did I want to cite them accurately, in my music I wanted to ensure that my ‘cetacean citations’ appropriately and respectfully acknowledged their original sources. I was struck by the realization that with the manipulations of transcriptions and audio clips in my compositional work I ran the risk of objectifying humpback song and potentially failing to respect individual singers through treating field recordings instrumentally, as a mere resource. I wanted to challenge the presumption that there were no ethical issues involved in working with other-than-human audio recordings. Here I found critical theorist Theodor W. Adorno’s critique of the workings of contemporary industrial societies helpful. Adorno identified the dominant mode of reason as an ‘instrumental rationality’ that objectified all entities and treated them as mere means to ends. He argued that art offered the potential for resistance, offering audiences non-objectifying experiences. As a musicologist, he also recognized that composition itself necessarily involves objectification and manipulation of its materials, giving rise to tensions within music between instrumental rationality and an “aesthetic rationality [that] wants to make good on the damage done by nature-dominating rationality” (Adorno, 2002, p. 289). As described in Chapter 4, I proposed that in my compositional context a development of this aesthetic rationality would treat its materials, i.e., the sounds of other animals, with the respect owed to anything possessing intrinsic value. I drew on Val

Plumwood's analysis of anthropocentrism to propose a set of strategies for composers who wish to work with the sounds of other animals without re-enacting objectification or producing works that might be read as disrespecting the producers of these sounds. These strategies for the musical representation of the sounds of other animals are analogous to my development of scientific representations of humpback song rhythms described in Chapter 2. Finally, turning to the pieces in my own Portfolio, I found a common musical form that mirrors the organization of the collective yet asynchronous singing of humpback whales. I hope that this 'multispecies heterophony' can serve as a model for an ecological thinking that recognizes the intrinsic value of individual other animals and of the entangled more-than-human worldings on our damaged planet.

This page intentionally left blank.

Chapter 2 Visualizing and Quantifying Rhythmic Variability in Humpback Whale Song

Abstract. In this chapter, I report on the development of visualizations and quantitative methods for assessing rhythmic variability in the repetitive heterochronous rhythms typical of humpback whale song phrases but also present in human music and the vocalizations of other animals. I compare and evaluate a range of visualizations previously used in bioacoustics and empirical musicology. To evaluate quantitative methods, I used Monte Carlo simulations of humpback song rhythms, modelling intra- and inter-individual rhythmic variability as a function of 'IOI jitter', 'Phrase Variant Diversity', and systematic rhythmic differences in phrase rhythm templates. These showed that (1) Lempel-Ziv Complexity is a promising indirect measure of IOI jitter, and (2) mean pairwise chronotonic distance is approximately linearly dependent on IOI jitter and independent of systematic rhythmic differences, and in combination with the Mantel test is sensitive to typical systematic differences in the presence of discrepancies in IOI jitter. In an exploratory application of the analysis to sample song sessions from two French Polynesian humpback whales, some song themes and one shared phrase variant exhibited inter-individual variability exceeding intra-individual variability, indicating that these two singers produced individually distinctive rhythms.

2.1 Introduction

2.1.1 Repetitive Rhythms

From the insistent cooing of woodpigeons to the opening dactyls of Mozart's Symphony No.40, repetitive (cyclic) rhythms are ubiquitous features of the communicative and expressive systems of living organisms. The mere repetition of a sound stimulus can lead to it being perceived as music by human listeners (Deutsch *et al.*, 2011; Simchy-Gross and Margulis, 2018), and the prevalence of repetition – of motifs, phrases, sections, and whole pieces – is one factor tending to distinguish human music from language (Fitch, 2006; Huron, 2007; Margulis, 2014). Repetition in music enhances its comprehensibility (Clayton, 2000; Margulis, 2014), learnability (Huron, 2007, p. 229), and the predictability that strongly influences the musical pleasure experienced by the listener (Meyer, 1956; Huron, 2007; Margulis, 2013, 2014; Cheung *et al.*, 2019). The repetition of short rhythms, defined as patterns of inter-onset intervals (IOIs) (Table 2.1), is one of the few cross-cultural musical features identified by Savage *et al.* (2015). The effects of such rhythmic repetition are perhaps most obvious when music entrains the movements of our bodies, with such entrainment facilitating social synchrony in shared or participatory music and dance performances (Turino, 2009; Tarr *et al.*, 2016). Indeed, the social coordination and subsequent social bonding made possible via the predictability of repetitive rhythms may even have been a key driver in the evolution of human musicality itself (Bispham, 2006; Tomlinson, 2015; Savage *et al.*, 2021).

Turning from sound patterns and behaviours to their underlying mechanisms, the capacity for periodic motor pattern generation is practically universal in the animal kingdom, and that for audiomotor entrainment – at least to the signals of conspecifics – is also widespread (Ravignani *et al.*, 2014; Kotz *et al.*, 2018). Hence it is unsurprising to discover repetitive rhythms in the acoustic productions of many insects, anurans (frogs and toads), birds and mammals. These may be simple as the metronomic croaks of the African painted reed frog (*Hyperolius* *marmoratus,* Jennions *et al.*, 1995) or the barks of a California sea lion (*Zalophus californianus,* Ravignani and Madison, 2017), or contain multiple durations such as displayed by the collared dove (*Streptopelia decaocto,* ten Cate *et al.*, 2002) or bowhead whale (*Balaena mysticetus,* Johnson *et al.*, 2015). In nonhuman vocalizations, where the components of rhythm cognition may have arisen through convergent evolution (Ravignani *et al.*, 2014; Kotz *et al.*, 2018), we need to look beyond social bonding for plausible evolutionary explanations for the presence of repetitive rhythms. An obvious benefit for any communication system operating in a noisy environment is that repetition increases signal redundancy, helping to get the message across (Bradbury and Vehrencamp, 2011, p. 305). In nonhuman animals, signals can carry information on species, sex, individual identity, motivation and quality (Catchpole and Slater, 2008, p. 5), and consequently play a very important role in sexual selection where mates may be chosen partly on the basis of their song or calls. Here, it is not just repetitiveness that counts but also the consistency of the signal, because this can be related to aspects of singer quality including size, as found in leopard seals (*Hydrurga leptonyx,* Rogers, 2017), or age, as for field crickets (*Gryllus bimaculatus,* Simmons and Zuk, 1992). Intriguingly, evolutionary modelling has shown that a dependence of rhythmic consistency on singer quality tends to lead to the fixation of more regularly alternating rhythms in a population (van den Broek and Todd, 2009), as these are more discriminable with respect to production errors than less structured forms (van den Broek and Todd, 2003).

*Table 2.1 Glossary of terms relating to the variability of rhythm and other aspects of temporal structure related to human music and humpback whale (*Megaptera novaeangliae*) song.*

Term	**Definition**
Anisochrony, Mean	The degree of departure from isochrony. For a sequence of IOIs, the mean proportional deviation from the mean IOI (§2.4.2).
Beat	A temporal grid inferred from a given timeseries such that there are sound events at most points of the grid. Not all timeseries possess a beat.
Chronotonic Function	IOI as a function of onset time (Toussaint, 2006).
Common Fast Pulse	For a given timeseries, the coarsest temporal grid on which the onsets of all events in that timeseries may be placed (Thul and Toussaint, 2008b). For a timeseries with a beat, the pulse may coincide with the beat, or be a subdivision of the beat.
Heterochronous Rhythm	Rhythm characterized by >1 IOI value, in contrast with Isochronous Rhythm (see below).
Inter-Individual Variability	Between-individual variability in shared rhythm performance.
Intra-Individual Variability	Within-individual variability in rhythm performance.
Inter-Onset Interval (IOI)	The duration between the starts of successive sound events.
IOI Jitter	Rhythmic variability in the duration between two successive sound events, simulated using repeated random sampling.
IOI Vector ([IOI])	Sequence of IOI values. E.g., [2.5, 5.5, x.x, ….]
Isochronous Rhythm	Rhythm characterized by a single IOI value.
Microtiming or Expressive Timing	Small deviations from notated, customary or theorized rhythms, occurring both within and among individuals.
Phrase	A short sequence of humpback song units, typically repeated to form a theme.
Phrase Type	Within a humpback song theme, phrases whose structure can be described in generalized form, e.g., ‘two modulated moans with intervening units’. Different general arrangements of song units, e.g., ‘three modulated moans with intervening units’, are classified as new phrase types (Cholewiak *et al.*, 2013). In this thesis phrase types are distinguished by uppercase letter appended to theme number (e.g., 2A, 2B).

Term	**Definition**
Phrase Variant (PV)	Within a humpback song phrase type, phrases possessing the same unit sequence (Murray *et al.*, 2018), e.g., ‘modulated moan – high cry – modulated moan’. Different phrase variants within a phrase type are related through minor changes, such as the addition, deletion or substitution of a few song units (e.g., ‘modulated moan – high cry – high cry – modulated moan’, whilst key structural features remain the same. In this thesis PV names are distinguished by the addition of lowercase letters or other symbols to phrase type names (e.g., 2Aa, 2Ab).
Phrase Variant Diversity (PVD)	Defined here as the balance of phrase variants within a theme (§2.4.2).
Rhythm	A temporally patterned configuration of sound events (Toussaint, 2020), represented herein as an IOI vector.
Song	A sequence of humpback song themes, tending to be sung in the same order, although some may be omitted.
Song Unit (also termed Unit)	The shortest element of humpback song continuous to the human ear when listened to in “real time” (Payne and McVay, 1971).
Systematic Rhythmic Differences	Recurrent or habitual differences in the production of shared rhythms that distinguish individuals or groups.
Tempo	Rate (or pace) of sound events (McAuley, 2010).
Theme	A sequence of similar humpback song phrases. In this thesis themes were numbered (1, 2, 3, …) in the order they were encountered during analysis.
Timeseries	A series of sound events.
Within-individual Rhythmic Variability (WRV, also termed IOI Jitter for Monte Carlo simulations)	Non-recurrent variability in the repetition of a rhythm. WRV is the converse of rhythmic consistency (rhythmic precision).

2.1.2 Rhythmic Variability

The repetition of rhythms by living systems is inevitably accompanied by variability. In the study of human music performance, such expressive timing or microtiming variability has been usefully differentiated into systematic or recurrent, and unsystematic, residual or random components (Kvifte, 2007; Davies *et al.*, 2013; Hellmer and Madison, 2015). In this thesis, I refer to the latter (unsystematic, residual or random) as unsystematic variability, IOI jitter, or Within-individual Rhythmic Variability (WRV) (Table 2.1).[5] Systematic departures from notated, customary, or theorized rhythms may occur as idiosyncratic differences among performers or groups of performers of the same work of music, or at a more general level, such as of a preferred swing ratio in jazz (Frane, 2017; Corcoran and Frieler, 2021) or the anticipations of the third and fourth semiquavers in Samba (Naveda *et al.*, 2011). Systematic rhythmic differences can therefore differentiate between performance styles of individuals or groups and arguably should be thought of as defining rhythmic norms rather than being deviations from them (Kvifte, 2007; London, 2012, p. 179). In contrast, unsystematic variability describes the variability within an individual's performance, which arises from several causes including neural noise (Glaze and Troyer, 2012). Both forms of variability have recently come under scrutiny for their contribution to the groove experience, defined as the "pleasurable urge to move [one's] body in synchrony with music" (Senn *et al.*, 2018). Many studies agree that lower unsystematic variability (higher rhythmic consistency) leads to higher groove ratings (Davies *et al.*, 2013; Frühauf *et al.*, 2013; Datseris *et al.*, 2019), although other factors may be more important (Senn *et al.*, 2018).

---

[5] Where I am referring to intra-individual rhythmic variability for repetitions of specific PVs, as in Chapter 3, I use WRV. For the Monte Carlo simulations employed in §2.5, I use IOI jitter to describe the random timing variability introduced at IOI-level, and unsystematic variability to cover both IOI jitter and Phrase Variant Diversity modelled by the random selection of phrase variants within a theme.

In bioacoustics, between- or inter-individual variability of an acoustic feature X is often tested against within- or intra-individual variability. The former must exceed the latter for it to be possible for a given receiver to reliably discriminate between individuals (Robisson *et al.*, 1993). The inter-/intra-individual distinction partially maps onto that made above between systematic difference and unsystematic variability, as intra-individual variability may be equated with unsystematic variability σ_X. However, inter-individual variability can result from (1) systematic differences in X between individuals ($X1$, $X2$, etc.), which in the case of rhythms could comprise a form of rhythmic signature, and/or (2) discrepancies in intra-individual variabilities in X (σ_{X1}, σ_{X2}, etc.), e.g., different levels of consistency in repeating a rhythmic signature. The proximate effects of rhythmic variability on nonhuman subjects are hard to gauge, but from the long-range (ultimate) point of view of signal evolution both forms of inter-individual variability might signal or serve as a cue for motivation and quality, allowing the singer to be assessed as a potential mate or rival (Riebel and Slater, 2003). Studies of the role played by temporal features in nonhuman animal vocalizations have tended to focus on call duration or call rate rather than rhythm *per se*. For example, in the humpback whale, Chu (1988) proposed that systematic differences in dive times might signal physical condition. In other species, playback experiments have confirmed that systematic differences can function as cues, for example the call rates of male blue petrels (*Halobaena caerulea*) predict female interest (Gémard *et al.*, 2021). In the case of differences in intra-individual variability, the consistency of temporal features has been shown to correlate with increased reproductive success in hyraxes (*Procavia capensis*, Demartsev *et al.*, 2022) and chestnut-sided warblers (*Dendroica pensylvanica*, Byers, 2007), although no correlation between consistency and female preference was found for African painted reed frogs (Jennions *et al.*, 1995) and field crickets (Simmons and Zuk, 1992). A further function of inter-individual or inter-group rhythmic differences, as mentioned above for human musicians, is that they can contribute to signalling identity (Robisson *et al.*, 1993; Prögler, 1995; Rendell and Whitehead, 2003b; Antunes *et al.*, 2011; Garcia *et al.*, 2012; Budka and Osiejuk, 2014; Gero *et al.*, 2016; Frane, 2017; Mathevon *et al.*, 2017; Osiejuk *et al.*, 2019; Dodson, 2020; Corcoran and Frieler, 2021). Finally, for animals such as humans, humpback whales and songbirds that possess vocal production learning and song traditions, systematic rhythmic differences (whether intended or accidental) provide new forms that may be taken up by others and thus contribute to song evolution (Payne *et al.*, 1983; Payne and Payne, 1985; Garland *et al.*, 2011, 2017; Sorce Keller, 2012; Williams *et al.*, 2013; Savage, 2019; Otter *et al.*, 2020).

2.1.3 The Rhythms of Humpback Whale Song

Among cetaceans, the humpback whale stands out for the complexity of its hierarchical song structure: single elements (song units) are repeated in combination with others to make up phrases, which repeat (with some variation) to make up a theme (Payne and McVay, 1971). Where variation includes trends in song unit form, number or duration, themes are described as "shifting" as opposed to "static" (Payne and Payne, 1985). A sequence of themes is termed a song, and the population-wide song is typically repeated for hours at a time by adult males during the breeding season. Repetition thus occurs on multiple levels: here I focus on that of the phrase, characterized as rhythmically monotonous (Payne and McVay, 1971; Thompson, 1981) and described by Cholewiak *et al.* (2013) as the "salient element of repetition." Phrase durations are more consistent than those of themes or the song (Thompson, 1981; Frumhoff, 1983; Payne *et al.*, 1983; Cerchio *et al.*, 2001), and such consistency can span multiple phrase variants (PVs) (Frumhoff, 1983; Handel *et al.*, 2012). Thus far there has been very limited research into phrase rhythms analysed as IOI patterns, although the independence of the durations of song units and inter-unit silences within phrases has been demonstrated for a restricted dataset (Handel *et al.*, 2009). In the single study of humpback song that focused on the intervals between song units (Schneider and Mercado III, 2019), a hierarchical classification of song was omitted. In this initial exploration of rhythm, Schneider and Mercado III showed that successive sections of humpback song had contrasting temporal structure falling into categories derived from Ravignani *et al.* (2014): periodic (including static or shifting isochronous or heterochronous patterns), quasiperiodic and aperiodic. However, they did not quantify the rhythmic consistency of the repetitive phrases that typically serve as the basis for comparisons of song structures (Cholewiak *et al.*, 2013).

In this chapter, I describe the development of methods to fill this gap in the analysis of phrase rhythms within the context of a classical coding of humpback song structure. The methods were tested on simulated phrase rhythms and sample recordings of humpback song from French Polynesia. My objectives were to provide visualizations and measures of the rhythmic variability present in sets of humpback song phrases, including heterochronous phrase rhythms containing varying numbers of song units. Visualizations are important to reveal patterns, while quantitative measures are needed to establish the robustness of these patterns. Tools were drawn both from bioacoustics and, where assumptions of human musicality could be avoided, from empirical musicology. For example, an important criterion for a good representation was that it avoid the anthropocentric assumption that humpback song units would either be organized to an underlying regular beat or possess a common fast pulse. To allow microtiming variability to be captured, representations should also avoid numerical approximations or categorizing of IOI values, such as occurs in conventional musical notations. In addition, visualizations should maintain, and measurements of variability be sensitive to, the IOI ordering within successive repetitions of a phrase. Ideally, visualizations should be easy to interpret for scientists and musicians alike, and allow large amounts of data to be presented in a compact form. Visualizations leading directly to an associated measure of variability could be advantageous for comprehensibility. I also considered measurements of intra-individual rhythmic variability insensitive to the base rhythms to be preferable, as they would allow comparison and aggregation across different PVs. I begin with accounts of the humpback song collection, coding, and song unit timing extraction (§2.2). In subsequent sections I describe and present results for methods focused on visualizing rhythmic structure and variability (§2.3), quantifying within-individual variability (§2.4), and comparing within- and between-individual variabilities (§2.5). I evaluate the methods and provide a general discussion in §2.6.

2.2 Data collection, Song coding, and IOI calculation

In this section I describe the collection of humpback whale song data and its coding, i.e., sequencing of song units in symbolic form into phrases and themes (§2.2.1). This is followed by a description of the semi-automated threshold detection process used to extract accurate IOI timeseries from the song audio (§2.2.2). Phrase start information from coding allowed the IOI timeseries to be divided into ordered IOI vectors $[\mathrm{IOI}_1, \mathrm{IOI}_2, \mathrm{IOI}_3, \ldots, \mathrm{IOI}_n]$ corresponding to phrase rhythms (Figure 2.1). Readers less interested in the details of this procedure can skip to the timeseries itself (Figure 2.4), or to the broad description of the parallel process for a performance of ‘Auld Lang Syne’ (§2.2.3).

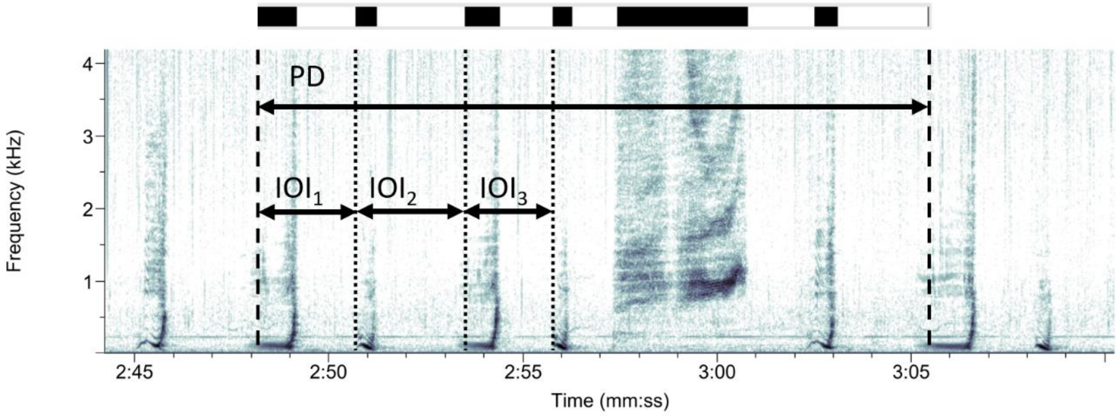


Figure 2.1 Spectrogram (Raven Pro 1.6, Hann window, 2048 samples, 75% overlap) of a stretch of song from Singer 190924, with the phrase duration (PD) and first three IOIs of an example phrase variant (PV 4A°) indicated. Song unit onsets from threshold detection (§2.2.2) are marked with vertical lines. Above the song units, a ‘phrase raster’ plot shows the song units (black rectangles) and inter-unit silent intervals (white rectangle). Similar representations from multiple phrases are combined to visualize phrase rhythm variability (e.g., Figure 2.9).

2.2.1 Humpback Song Data Collection and Coding

Songs were recorded off the east coast of Mo’orea, French Polynesia using two OceanInstruments SoundTrap ST300STD recorders tethered on a single cable to the ocean floor at depths of c. 27 m and c. 28 m in ~30 m of water, approximately 800 m from the shore and 150 m outside the reef, at 17°33’ S, 149°46’ W. The SoundTrap hydrophones (bandwidth 0.02–60 kHz ±3 dB) were configured to operate on high gain with a sampling frequency of 24 kHz. Audio files were stored locally in WAV format and transferred to computer after retrieval. The lower device recorded continuously from 11–28 September 2019, while the upper recorder ran on a duty cycle of 30 mins/2 hours from 11 September – 17 November 2019, resulting in over 800 hours of stored audio.

For this thesis I coded c.13 hours of song down to the unit level following previous studies (Garland *et al.*, 2011, 2012; Garland, Noad, *et al.*, 2013; Garland *et al.*, 2017) (Table 3.1). Coded song for this chapter (Table 2.2) consisted of two 30-minute continuous sessions selected on the basis of signal to noise ratio. Each song unit was classified to song unit types (following previous analyses), repeating general arrangements of units were classified as phrase types (denoted with uppercase A, B, etc.) with particular sequences within these labelled as phrase variants (PVs) to capture fine-scale variability (A°, A’, Aa, Ab, A-, etc.). Repeated phrases were grouped into themes (labelled 1, 2, 3, etc.). Phrases were delineated following Cholewiak *et al.* (2013), keeping consecutive similar units together and minimizing incomplete phrases at the ends of sequences of similar phrases. These guidelines were combined with the rule that the maximum silence in a repeating pattern of song unit types defines the end of a phrase. Where rules conflicted, this maximum silence rule was given priority. Naming conventions for PVs depended on the kind of variation that was present in each theme, but major differences in structure or song unit types were recognized by the creation of new phrase types. For spectrograms and codings of phrase types and common variants see Appendix B.

Table 2.2 Song sessions used in the exploratory analysis reported in this chapter comprised two 30-minute samples from the full coded dataset (Table 3.1). The number of song cycles was established by examining the theme sequence and counting a new cycle each time any theme repeated. For example, theme sequence ***3****412756****3****12****1****75634****1*** *would contain three cycles.*

Recording Date (yyyy.mm.dd)	Singer ID	Recorder	Coded Duration	# Song Cycles	Theme							
					1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
2019.09.19	190919	lower	0h30	3	●	●	●	●	●	●	●	
2019.09.24	190924	lower	0h30	5	●	●	●	●	●	●	●	

Both recordings used here (Singers 190919 and 190924) included seven different themes, with Theme 3 accounting for around a third of the song duration (Table 2.3). For most themes, the majority of phrases fell into one or two PVs (Table 2.4).

Table 2.3 Total number of phrases, units and duration of each theme for sample recordings.

	Singer 190919				**Singer 190924**			
Theme	N	#Phrases	#Units	Time (s)	N	#Phrases	#Units	Time (s)
1	4	25	131	292	1	7	51	66
2	4	9	46	163	5	23	71	212
3	4	39	268	677	5	28	197	487
4	2	8	54	146	4	20	119	348
5	3	7	35	127	5	13	70	228
6	3	7	38	144	5	5	33	129
7	4	15	57	182	5	16	75	255
TOTAL	24	110	629	1731	30	112	616	1725

Table 2.4 Total phrases and prevalence of common PVs (for spectrograms see Appendix B), giving frequency of occurrence in theme. E.g., For Singer 190919 20% of Theme 1 phrases were PV 1Ba5. PVs occurring only once in the data are excluded from PV counts. Shared PVs are those sung by both singers.

	Singer 190919			**Singer 190924**			
Theme	Total #Phrases	Common PVs (frequency, %)	#PVs	Total #Phrases	Common PVs (frequency, %)	#PVs	#Shared PVs
1	25	1Ba5 (20)	7	7	1Ba6 (71)	3	2
2	9	2Aa (78)	2	23	2Ba (91)	2	0
3	39	3A° (41)	5	28	3A° (57)	5	2
4	8	4C° (25) 4Da-4 (25)	2	20	4A° (30)	4	0
5	7	5Aa (29) 5Aa’ (29) 5Aa- (29)	4	13	5Aa (38) 5Aa’ (38)	3	3
6	7	6A° (43)	2	5	6A° (40) 6A’+ (40)	2	1
7	15	7B° (53)	4	16	7A° (56)	3	3
TOTAL	110		26	112		22	11

2.2.2 Extracting Song Unit Timing

I wrote custom code in MATLAB version 9.10.0 (R2021a) (The MathWorks Inc., 2021) to extract song unit onsets and other acoustic features from the audio recordings. The DC component (estimated as the mean signal level across the whole recording) was subtracted from the original signal, background snapping shrimp sounds were reduced through the use of a Finite Impulse Response bandpass filter (passband 60–5000 Hz, stopband 40–5100 Hz, attenuation 100 dB, Kaiser window), and sharp noise spikes removed by median filtering (5 ms window). The filtered signal amplitude was then squared to convert it into the dimensions of signal power, and smoothed by integrating onto a 25 ms grid.

A double-threshold detector developed from the algorithm described by Pace *et al.* (2010) was then applied to the resampled signal power timeseries to find candidate song units (CSUs): in a first pass through the audio the starting point (onset) of a CSU was set to time points when the Signal to Noise Ratio (SNR) reaches a value greater than the specified threshold value T_{on} (for a minimum duration) and the end point when the SNR subsequently falls below the specified value T_{off}. Using two thresholds rather than one allowed for improved capturing of the intensity profiles of those many song units which tail off or which fade out and back in ('dropout'). The double-threshold detector ($T_{on} = 3.01$ dB, $T_{off} = 1.76$ dB, minimum duration 50 ms) was also run on the time-reversed signal, and the resulting detection curves combined, to estimate better the onset times of song units with initial quiet starts. The use of a double-threshold detector responds to the intuition that if a high signal level (SNR > T_{on}) is detected at a particular time point, it is probable that this corresponds to whale song, and it is then more likely that the signal level at an adjacent time point falling between the threshold levels is caused by a continuation in the song unit than it is caused by random fluctuation in the noise.

The use of SNR thresholds requires estimation of the noise level. This can be obtained because each song unit is surrounded by silence, allowing the signal power curve to be treated to pass through these power minima. The raw signal power curve was smoothed by taking a moving mean (taken across a time window, W1, narrower than the typical inter-unit silence, but wide enough not to be too influenced by noise spikes). The noise level was then estimated by interpolating between the local minima of the resulting smoothed curve in logarithmic space, located using the MATLAB function **islocalmin** (with minimum separation, W2, wider than the duration of typical song units but small enough to be responsive to fluctuations in background noise; minimum prominence 20% of the range of the signal power).[6] Minima

---

[6] The noise estimation method described here incorporates minor alterations made partway through onset detection to deal with a specific problem encountered in the recording of Singer 190924, where boat sounds had caused the

occurring during longer song unit dropouts were at higher signal levels than inter-unit silences and could be automatically excluded. Values of W1 = 0.5 s and W2 = 2.5 s worked well for the sample audio files.

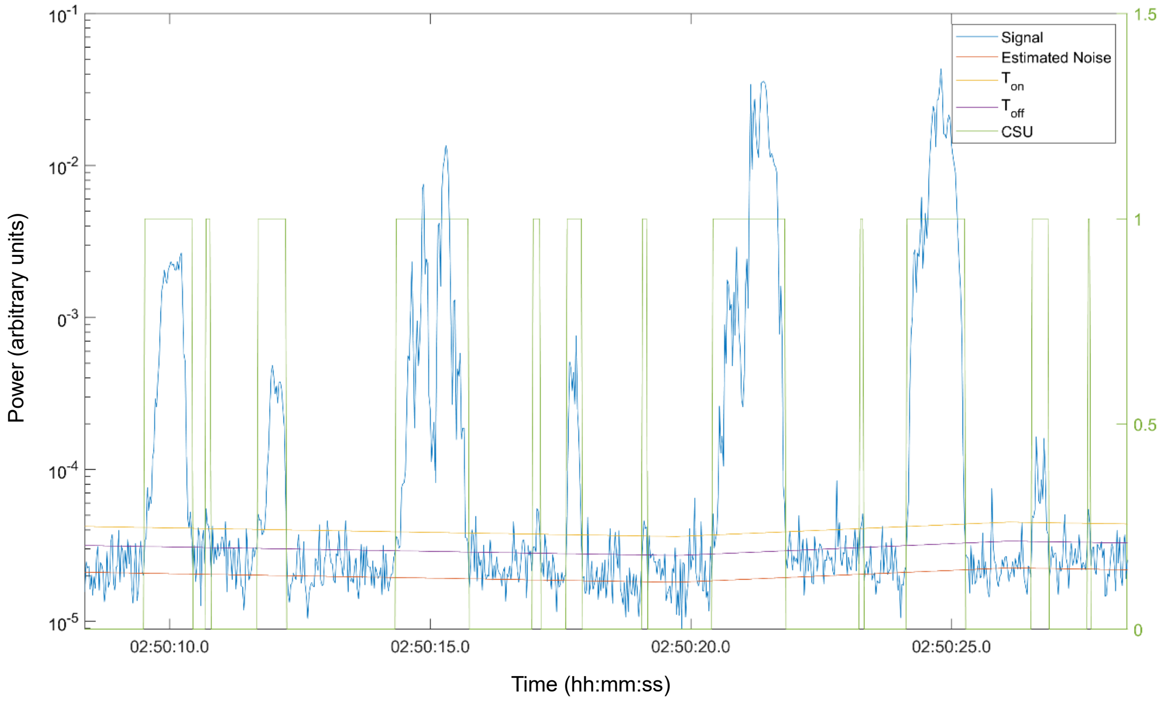


Figure 2.2 Thresholds and noise estimation for Singer 190924. Broad peaks in the filtered signal power curve (blue) correspond to song units, narrower peaks may result from knocks to the detector and other noise spikes. The estimated noise curve (red) successfully tracks inter-song unit minima and the 'Threshold On' curve (yellow) is generally above noise spikes in the inter-song unit silences. The detection curve (green, right-hand axis) outlines Candidate Song Units.

After the initial automatic identification of CSUs (Figure 2.2, green curve), selection tables were written to file and imported into Raven Pro 1.6 (K. Lisa Yang Centre for Conservation Bioacoustics at the Cornell Lab of Ornithology, 2021) to allow the rejection of false positive CSUs corresponding to occasional loud fish sounds, background humpback song, or other noise spikes (Figure 2.3A). This manual inspection also allowed for the identification of

---

background noise to rise rapidly by an order of magnitude. Comparing noise estimation methods for this recording showed that the alterations made no or little difference under normal quiet conditions: for 79/100 CSUs there were no changes to onset timing. The mean absolute change of 12.5 ms was negligible compared with typical IOI variability c.250 ms (Table 2.8).

occasional dropouts in longer song units, where the sound was judged continuous by ear but for which the song unit SNR < T_{off}. Finally, the existence of weak song units heard in the audio and/or spotted through visual inspection but which fell beneath the double-threshold detector's SNR criterion was recorded. This allowed for exclusion from further analysis of the inter-onset interval (IOI) between the preceding and subsequent detected song units. Once all spikes and dropouts had been identified, remaining CSUs were upgraded to song units (Figure 2.3B) and the onset and IOI timeseries calculated (Figure 2.4). Phrase starts were identified during song-coding and matched to CSU onsets, which allowed the calculation of IOI vectors ('phrase rhythms').

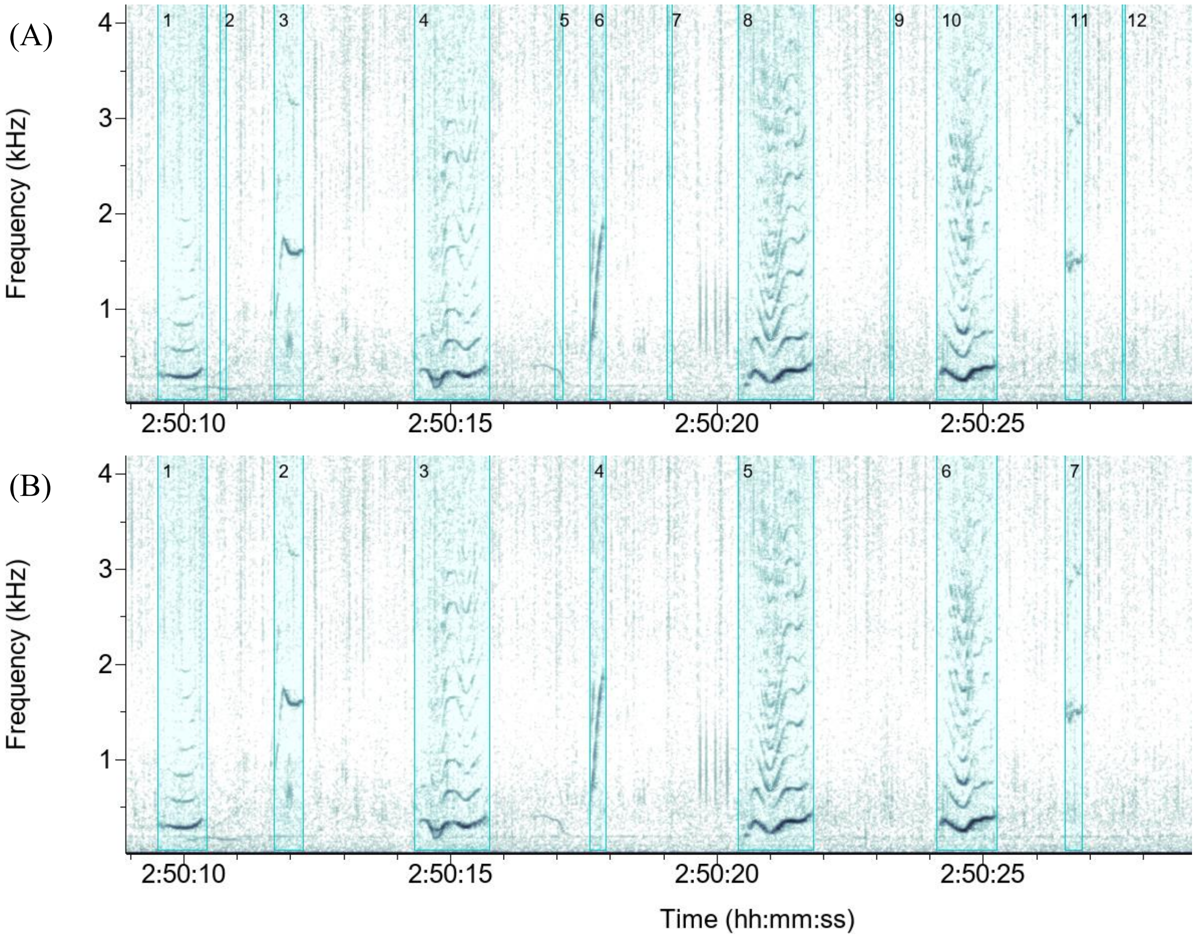


Figure 2.3 Song Unit Timing for Singer 190924. (A) CSUs identified by double-threshold detector, containing false positives resulting from background song (selections 2,5) and spikes in snapping shrimp sounds (selections 7, 9, 12). (B) Song Units remaining after false positives had been manually identified in Raven and removed. Spectrogram settings: Hann window, 2048 samples, 75% overlap.

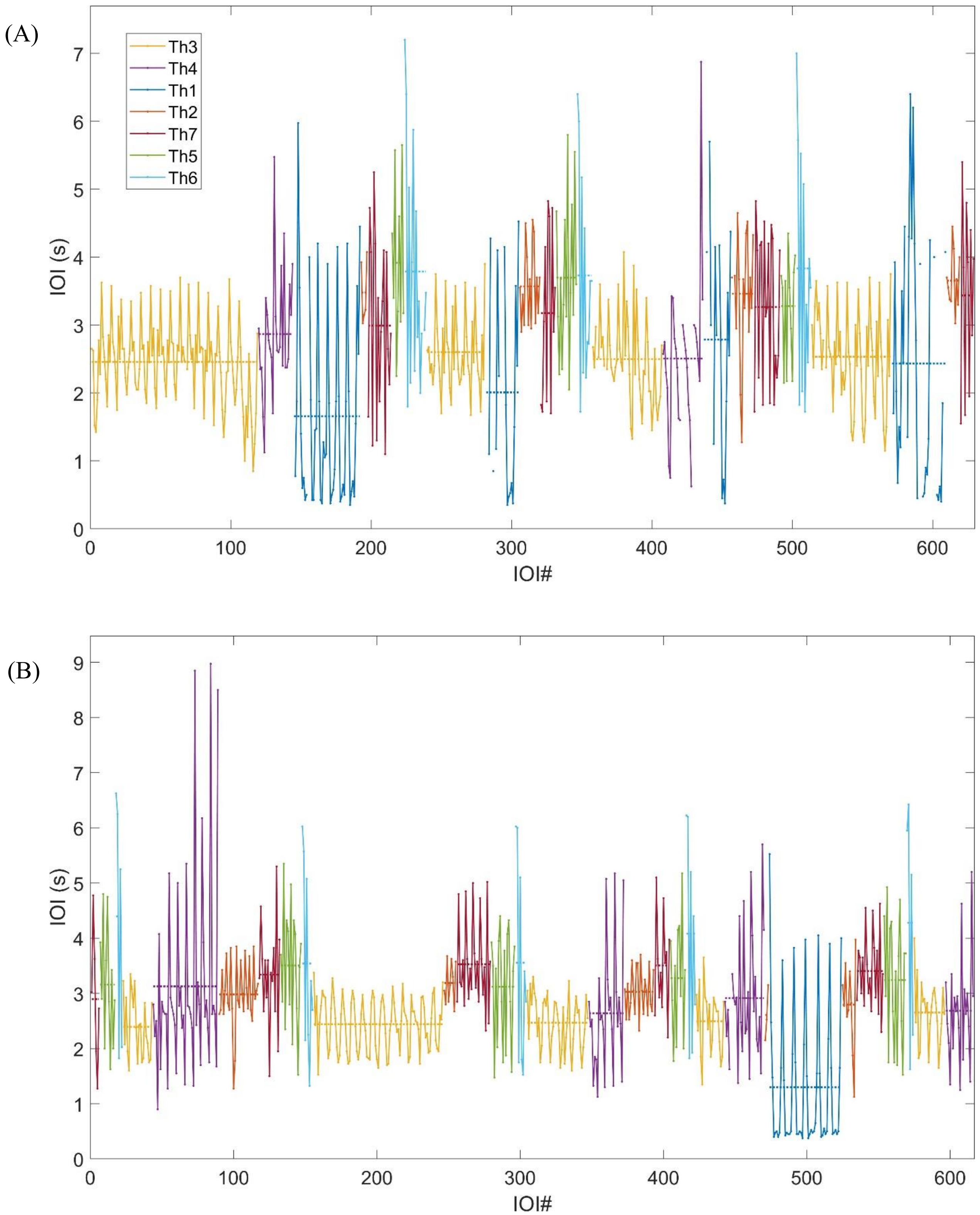


Figure 2.4 IOI timeseries for Singers (A) 190919, (B) 190924 (30 minutes of continuous song). Themes are colour-coded, and dotted lines indicate mean IOI values for each theme. Timeseries plots are continuous within themes except where a song unit has been flagged as not having been picked up by the detector: in these cases the IOI is not plotted, resulting in gaps. Even at this scale it is possible to discern regularities of phrase rhythms within themes and differences among different themes. Substantial inter- and intra-individual variability is also evident.

2.2.3 Human Music Example: 'Auld Lang Syne'

To facilitate understanding of the representations of humpback whale song rhythms and how they compare with some typical human rhythms, I made a recording of 'Auld Lang Syne' (Figure 2.5) by clapping the melody for three verses and refrains to allow for accurate rhythm segmentation without manual validation. The performance was recorded using a Zoom H5 digital recorder and note starts identified using threshold detection (§2.2.2). Song coding into phrases was based purely on rhythm and coincides with lines in the lyrics. The resulting IOI timeseries (Figure 2.6) illustrates the repetitive rhythmic structure of this melody.

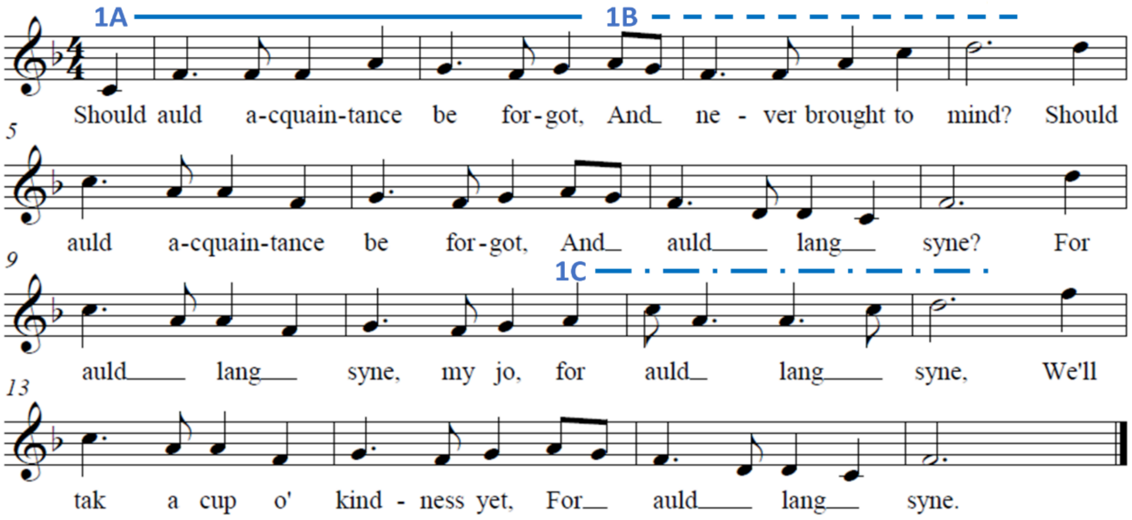


Figure 2.5 Score of 'Auld Lang Syne' used in making recording for comparison with humpback song. The three different phrase rhythms are indicated where they first appear; a single verse and refrain contain 4×1A, 3×1B and 1×1C.

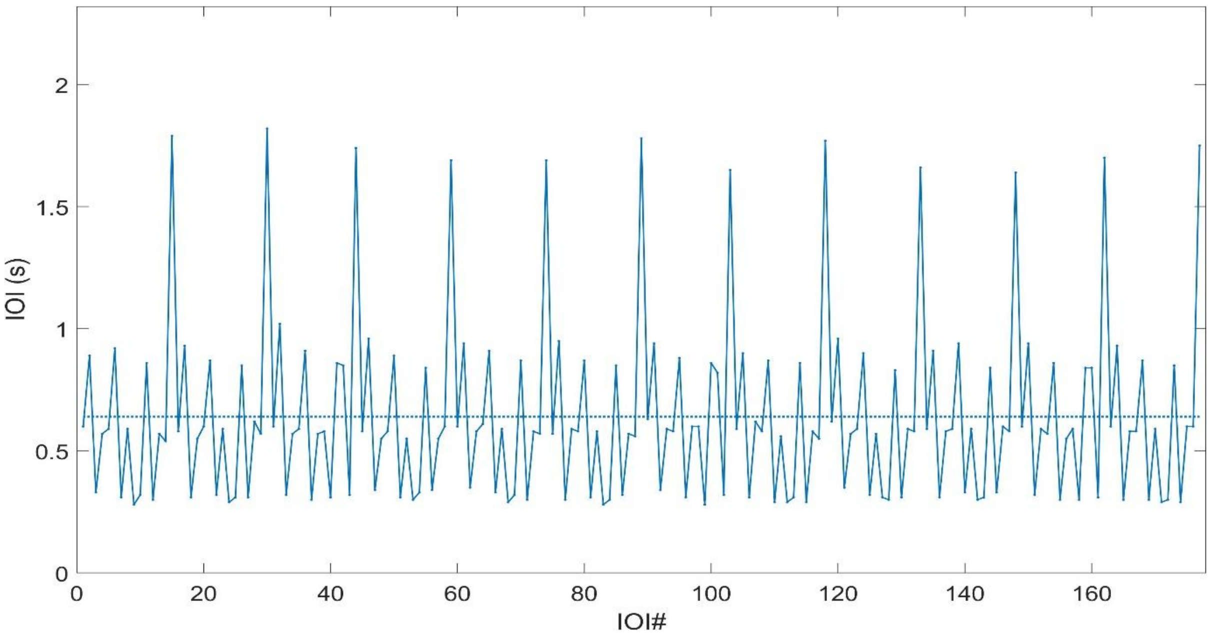


Figure 2.6 IOI timeseries for performance of three verses and refrains of 'Auld Lang Syne'. The dotted horizontal line indicates the mean IOI. Long IOI values occur regularly at the end of every four bar phrase, corresponding to the notated dotted minims (dotted half notes).

2.3 Visualizing the Variability of Repetitive Phrase Rhythms

2.3.1 Methods

The starting point for all visualizations (compared with common musical notation systems in (Table 2.5) was the inter-onset interval (IOI) timeseries (Figure 2.4), divided into vectors representing each phrase [IOI_1, IOI_2, IOI_3,…, IOI_n]. I define a normalized IOI vector [NIOI] for a phrase of duration PD as [IOI]/PD. Alongside the vectors I employed the chronotonic function (Toussaint, 2006), CF (CFN), which takes the value of IOI_i ($NIOI_i$) between onsets relative to the phrase start rO_{i+1} and rO_i (nrO_{i+1} and nrO_i), and for a phrase containing NSU song units is given by:

Equation 2.1

$$CF = \{IOI_i \; if \, rO_i \leq t < rO_{i+1}, for \; 1 \leq i \leq NSU$$

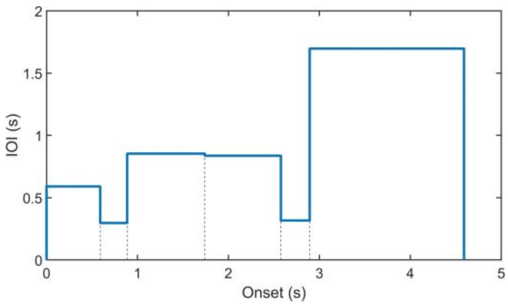
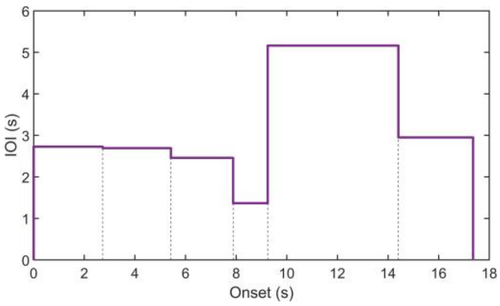
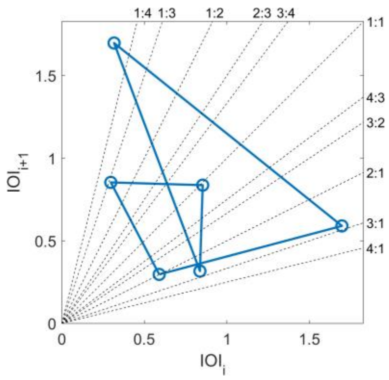
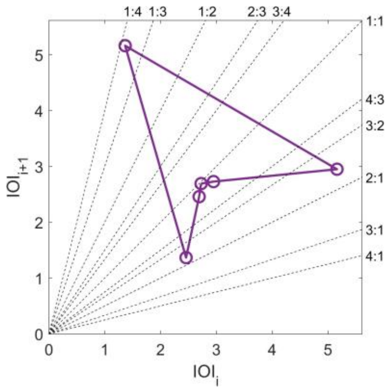
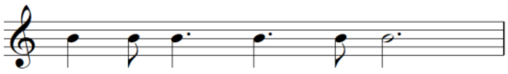
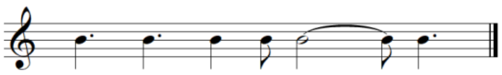
Equation 2.2

$$CFN(t) = \{NIOI_i \; if \, nrO_i \leq t < nrO_{i+1}, for \; 1 \leq i \leq NSU$$

A chronotonic diagram was generated by plotting these functions against time. The phase space plot or 'phase portrait' previously employed in studies of rhythmic variability in humpback whale and thrush nightingale (*Luscinia luscinia*) song (Rothenberg *et al.*, 2014; Schneider and Mercado III, 2019) was created by jointly plotting adjacent IOIs (IOI_{i+1} vs IOI_i) and connecting successive points with lines. Small integer ratios of adjacent points (1:1, 1:2, etc.) were visualized by adding constant gradient diagonals (Ravignani, 2017). For the 'phrase raster plot' (Janney *et al.*, 2016) successive phrases were plotted one per line, with every song unit represented as filled coloured blocks and inter-unit silences in white. I colour-coded human rhythms by loudness and humpback song units by peak frequency, using a logarithmic scale to best characterize the approximately constant Weber fraction ($\Delta f/f$) found for frequency discrimination in cetacean hearing across a wide range of frequencies (Nachtigall *et al.*, 2000). Like a musical score, the raster plot provides information on pitch or loudness, as well as showing the division of the inter-onset interval into sound and silence. Reading vertically down the plot provides a phrase duration timeseries.

Chronotonic diagrams, phase portraits and phrase raster plots provide graphical alternatives to musical notation (Table 2.5), and may be used to depict rhythmic variability in repeated phrases (e.g., Figure 2.11B). In the case of repetitions of single PVs, 'phrase rhythm box plots' were used to display summary statistics (median, interquartile range, outliers) for performed IOIs as a function of song unit position (e.g., Figure 2.11D), following Clayton's (2020) analysis of North Indian rhythm performances.

Table 2.5 Methods of representing rhythmic structure, comparing phrase rhythms from examples of human music and humpback whale song. For representations of the humpback rhythm with the Time Unit Box System and Western musical notation the IOI vector was rounded to the nearest 1s. The Western musical notation for 'Auld Lang Syne' is taken from the score (Figure 2.5) rather than derived from performance, with bar lines removed to allow better comparison with the humpback song. For the phase portraits the IOI vectors were extended to show the full shape of a repeating rhythm. Verticals have been added to the chronotonic diagram to indicate onsets.

	'Auld Lang Syne' **Phrase Rhythm 1C**	**Humpback Singer (190924)** **PV 4A**
IOI vector [IOI] [NIOI]	[0.59 0.30 0.85 0.84 0.32 1.70] [0.129 0.065 0.186 0.182 0.069 0.370]	[2.73 2.69 2.46 1.37 5.16 2.95] [0.157 0.155 0.142 0.079 0.297 0.170]
Time Unit Box System	●□●●□□●□□●●□□□□□	●□□●□□●□●●□□□□●□□
Western Musical Notation		
Phrase Raster Plot (Janney *et al.*, 2016)		
Chronotonic Diagram (Toussaint, 2006)		
Phase Portrait (Ravignani, 2017)		

2.3.2 Results

For readers more familiar with human music than humpback whale song, I begin with ‘Auld Lang Syne’, both in its complete rendition (Figure 2.7) and for a specific phrase rhythm (Figure 2.8). The variability in the multiple performances of the three distinct rhythms in this piece shows up in the chronotonic diagram as three sets of closely overlapping curves, and in the phase portrait as superimposed polygons. The plots for phrase rhythm 1A (Figure 2.8) show more detail.

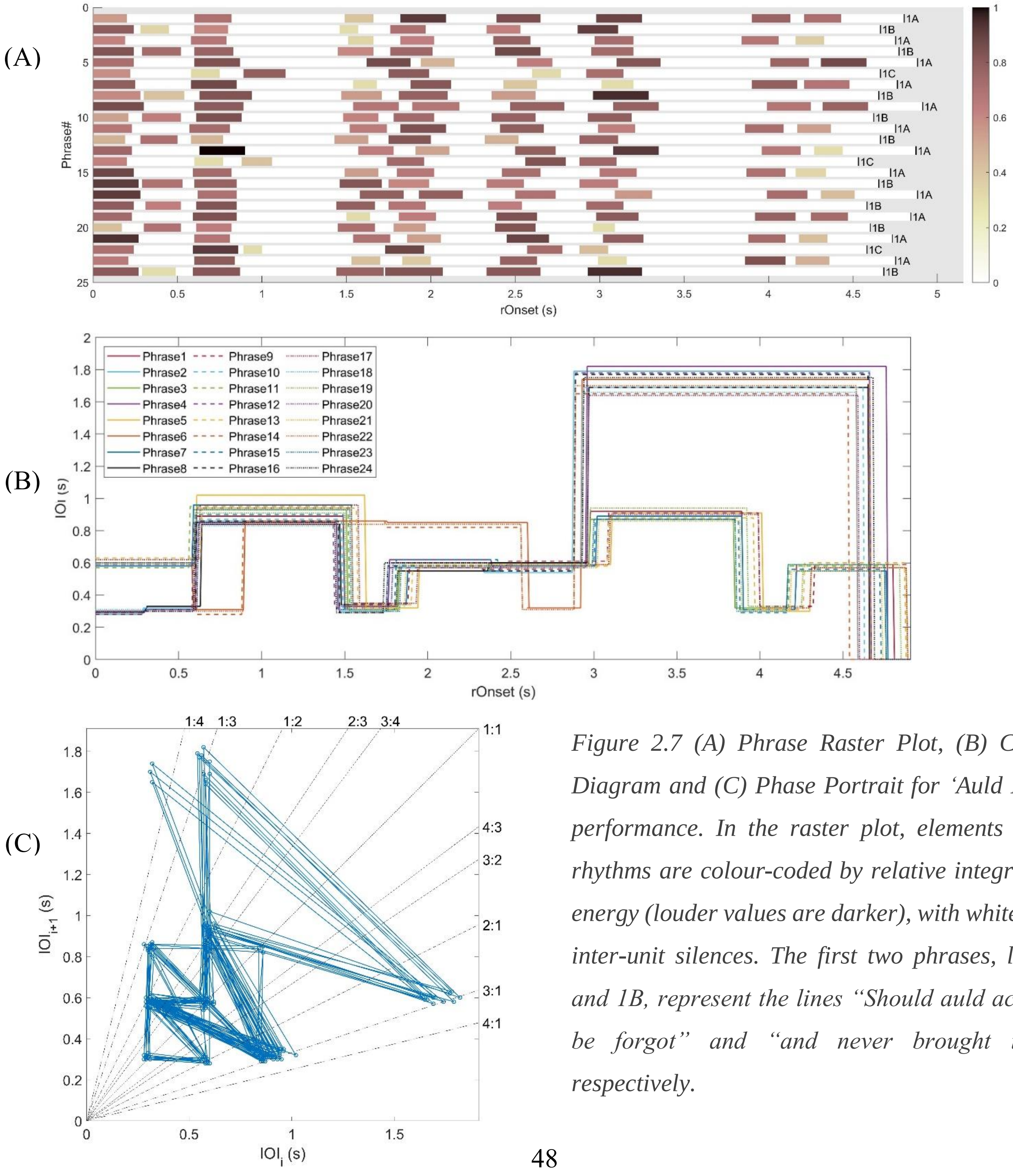


Figure 2.7 (A) Phrase Raster Plot, (B) Chronotonic Diagram and (C) Phase Portrait for ‘Auld Lang Syne’ performance. In the raster plot, elements of clapped rhythms are colour-coded by relative integrated signal energy (louder values are darker), with white spaces for inter-unit silences. The first two phrases, labelled 1A and 1B, represent the lines “Should auld acquaintance be forgot” and “and never brought to mind?” respectively.

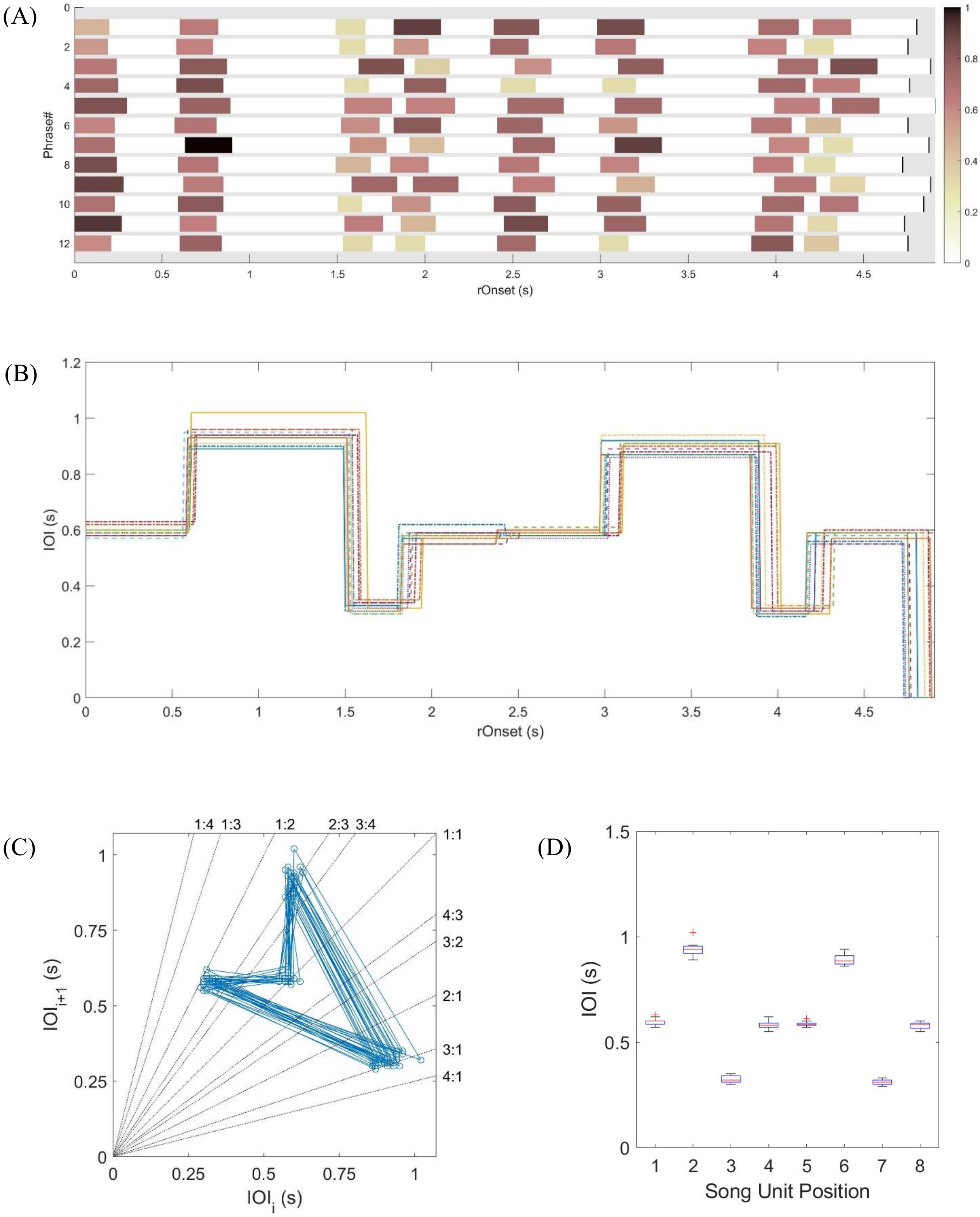


Figure 2.8 (A) Phrase Raster Plot, (B) Chronotonic Diagram, (C) Phase Portrait, (D) Phrase Rhythm Box Plot for 'Auld Lang Syne' performance, Phrase Rhythm 1A.

The following visualizations of humpback song rhythms progress downwards through hierarchical levels, moving from full 30 minute song sessions (Figure 2.9), via themes (Figure 2.10), to the fine-scale PV level for both singers (Figure 2.11,Figure 2.12). The raster plot for the 30 min humpback song session (Figure 2.9) reveals both the repetitiveness and diversity within themes, and the marked differences in phrase rhythms from one theme to another. It also shows trends in IOI and song unit duration. Extracting all Theme 3 phrases from this song session (Figure 2.10) confirms in compact form what may be seen in spectrograms, i.e., that almost all phrases are based on a similar seven-song unit template.

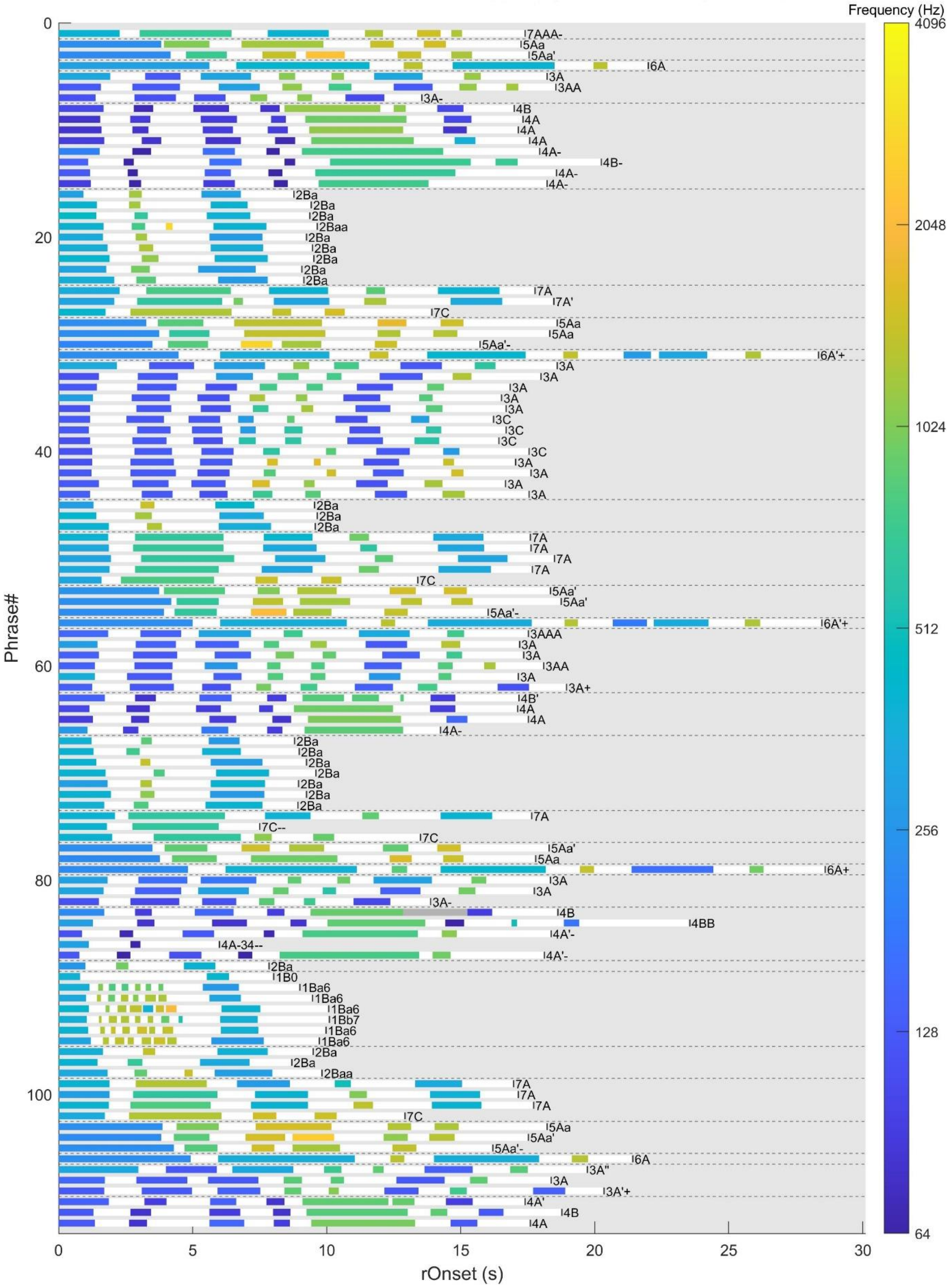


Figure 2.9 Phrase Raster Plot for humpback singer 190924, showing 30 minutes of song (one phrase per line). Dotted horizontal lines indicate theme boundaries.

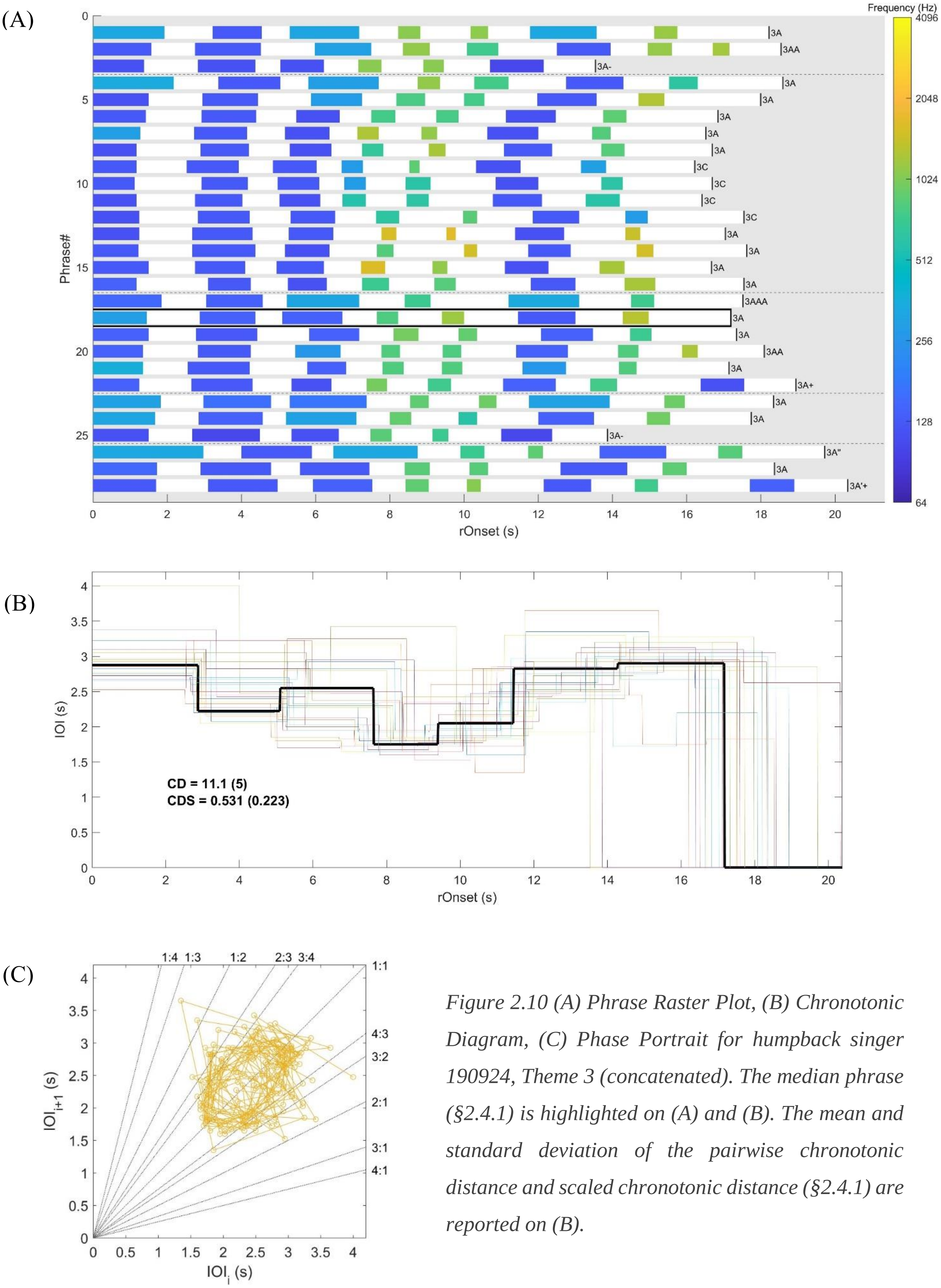


Figure 2.10 (A) Phrase Raster Plot, (B) Chronotonic Diagram, (C) Phase Portrait for humpback singer 190924, Theme 3 (concatenated). The median phrase (§2.4.1) is highlighted on (A) and (B). The mean and standard deviation of the pairwise chronotonic distance and scaled chronotonic distance (§2.4.1) are reported on (B).

When the repetitions of PV 3A° of both singers are compared (Figure 2.11, Figure 2.12), the level of intra-individual variability is clearly higher for Singer 190919: at each song unit position there is more scatter (range of IOI values), which shows up as a broader range of values in the chronotonic diagram and as a more variable phase portrait. In addition, there is a systematic difference between the singers in the phrase rhythm itself, perhaps most obviously illustrated by the phrase rhythm box plot: the second half of the rhythm (song unit positions 4–7) is more uneven for Singer 190919. The average pairwise chronotonic distances reported (Figure 2.11B, Figure 2.12B) and statistical tests assessing these differences support these conclusions (§2.5.4).

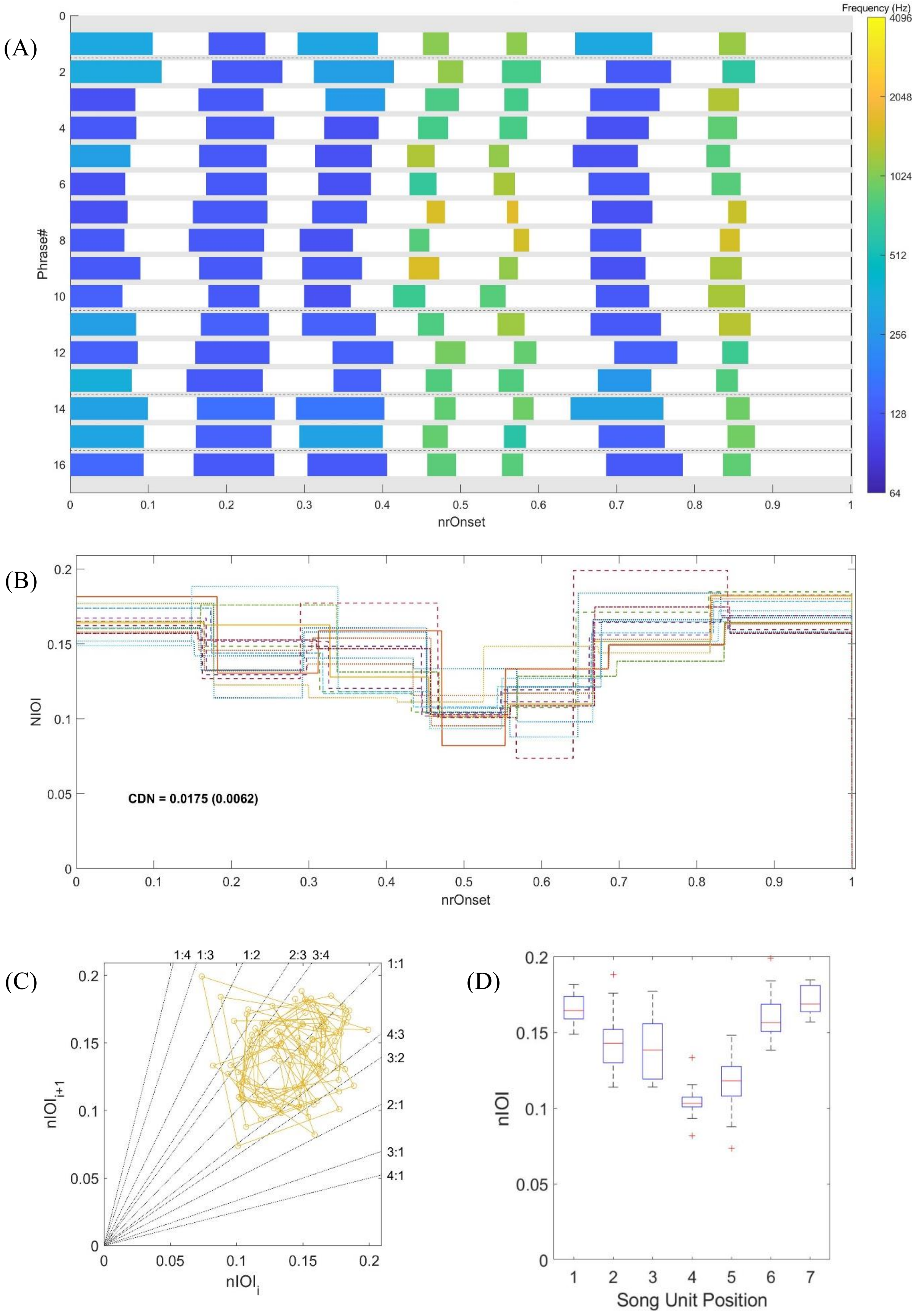


Figure 2.11 (A) Phrase Raster Plot, (B) Chronotonic Diagram, (C) Phase Portrait, (D) Phrase Rhythm Box Plot for humpback singer 190924, Phrase Variant 3A° (concatenated). IOI values are scaled to phrase duration. The mean and standard deviation of the pairwise normalized chronotonic distance (§2.4.1) are reported on (B).

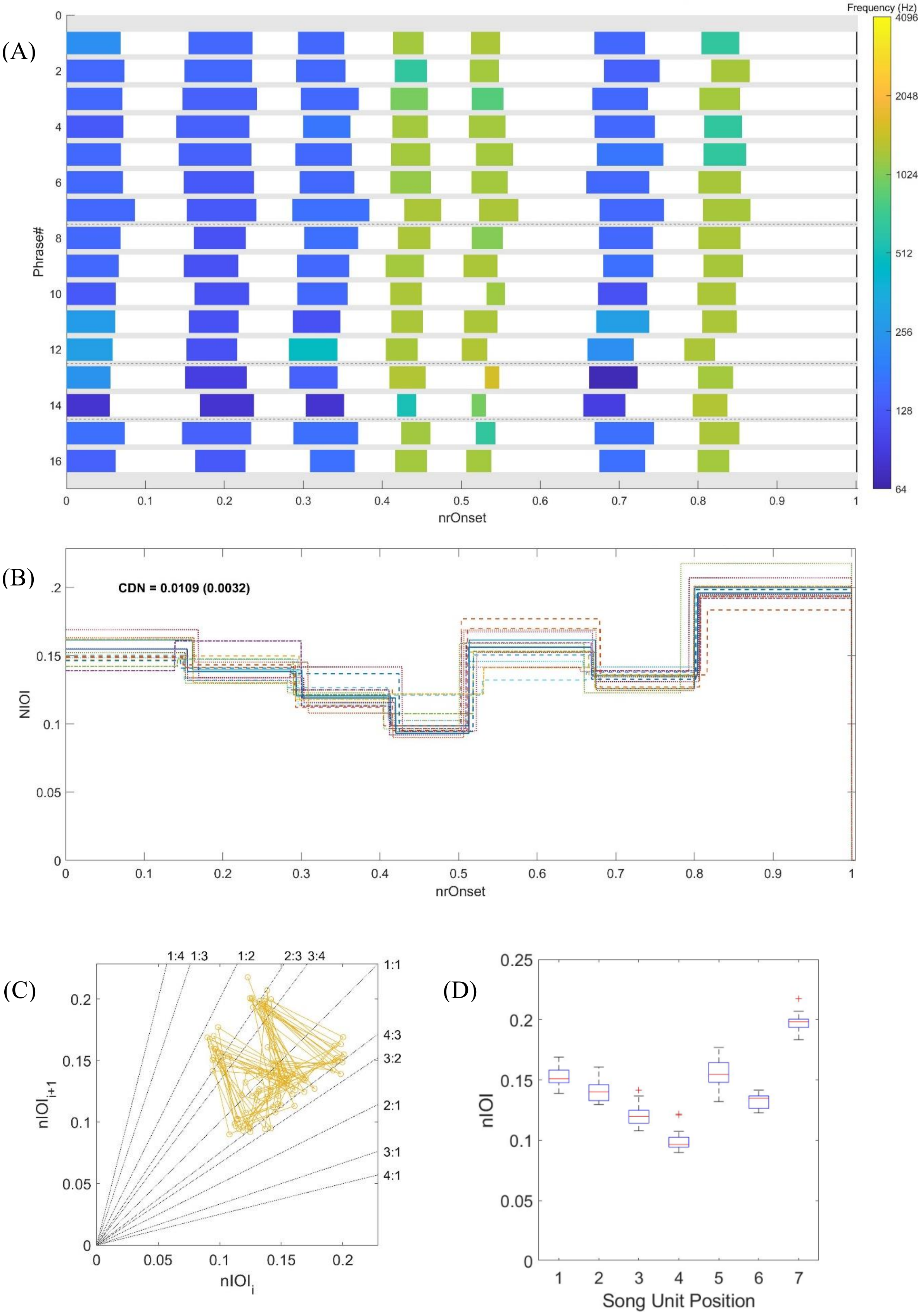


Figure 2.12 (A) Phrase Raster Plot, (B) Chronotonic Diagram, (C) Phase Portrait, (D) Phrase Rhythm Box Plot for humpback singer 190919, Phrase Variant 3A° (concatenated). IOI values are scaled to phrase duration. The mean and standard deviation of the pairwise normalized chronotonic distances (§2.4.1) are reported on (B).

2.4 Quantifying Within-Individual Phrase Rhythm Variability

To quantify rhythmic variability, I drew on research into the closely-related problem of measuring the (dis)similarity among stereotyped cyclic rhythms in human music. In the approach summarized by Toussaint (2020), objective measures of rhythmic (dis)similarity were assessed for their ability to capture subjective perceptual judgements in human listeners (Guastavino *et al.*, 2009; Post and Toussaint, 2011; Toussaint *et al.*, 2011; Beltran *et al.*, 2014; Toussaint and Oh, 2016) or knowledge derived from historical musicology (Díaz-Báñez *et al.*, 2004; Toussaint, 2006; Thul and Toussaint, 2008a; Toussaint *et al.*, 2012). In general, transformation measures such as the Levenshtein distance (Levenshtein, 1966) outperformed assorted music- or information-theoretic measures in matching perceptual judgements or musicological knowledge (Toussaint *et al.*, 2012; Beltran *et al.*, 2014). A 'chronotonic distance' measure based on chronotonic functions (§2.3.1) has been shown to perform well (Díaz-Báñez *et al.*, 2004; Toussaint, 2006; Guastavino *et al.*, 2009) and I test it here as it offers significant advantages over other distance measures. It is computationally much faster than the Levenshtein distance,[7] and unlike the Euclidean distance it can be applied to sets of rhythms containing different numbers of onsets.

Other promising methods developed specifically to quantify microtiming variation in human music performances were unsuitable for their reliance on musical assumptions, such as the existence of a beat and beat subdivisions (Hellmer and Madison, 2015) or a fixed swing ratio (Datseris *et al.*, 2019), that cannot be carried over into the study of the vocalizations of other animals. These studies measured departures from idealized or target rhythms, whether pre-defined by the experimenters or other players (Prögler, 1995; Davies *et al.*, 2013; Frühauf *et al.*, 2013; Hellmer and Madison, 2015), or established via transcriptions of performance (Senn

---

[7] Results not included here indicated that chronotonic and Levenshtein distances are highly-correlated for the humpback whale song rhythms explored in this chapter.

et al., 2016). They therefore also assumed the existence of a template, and the intention to perform it, that we cannot be certain of in humpback whale song.

I begin with measures applied directly to humpback song phrases (§2.4.1). Straightforward statistical measures can be applied to repeated single phrase variants (PVs), but I also seek a measure of rhythmic variability for themes containing mixtures of PVs. I compared direct measures with each other and with indirect measures derived from song themes (§2.4.2).

2.4.1 Direct Measures of Phrase Variability

I initially specify single measures capable of capturing the unsystematic variability in phrase repetitions that can be considered to arise from IOI jitter. To be a true measure of *rhythmic* variability such a measure should be responsive to the order of IOIs in each performed rhythm. This rules out simple aggregate ('zeroth order') quantities such as the standard deviation across all IOIs (*sdIOI*) as an appropriate direct measure, though I include *sdIOI* as an indirect measure in the next section.

Where all repeated rhythms contain the same number of onsets (song units) *NSU*, IOI jitter may be directly assessed through computing the standard deviation of IOI values at each onset (song unit) position *I* across all repetitions, and then taking the mean across all onset (song unit) positions. Where μ_{IOI_I} and σ_{IOI_I} are the mean and sample standard deviation of the IOI values at onset position *I* for all repetitions of a PV *J*, the mean IOI standard deviation is given by:

Equation 2.3

$$\overline{\sigma_{IOI_J}} = \frac{1}{NSU_J} \sum_{I=1}^{NSU_J} \sigma_{IOI_I}$$

If $\overline{\sigma_{IOI}}$ varies with interval size, then the mean IOI coefficient of variation may also be useful:

Equation 2.4

$$\overline{CV_{IOI_J}} = \frac{1}{NSU_J}\sum_{I=1}^{NSU_J}\frac{\sigma_{IOI_I}}{\mu_{IOI_I}}$$

These measures are adequate for capturing the variability in single PVs, where all phrases contain the same number of song units. However, this constraint means that closely-related rhythms created by the addition or deletion of onsets (song units) cannot be compared. To quantify the variability in such quasi-repetitive rhythms I calculated average pairwise chronotonic distances, based on the chronotonic functions (§2.3.1) for each phrase rhythm in the set. Each phrase was compared pairwise to every other through the calculation of the chronotonic distance cd_{ab} between the rhythms in the pair, defined as the integrated area between them when expressed as chronotonic functions (Figure 2.13).

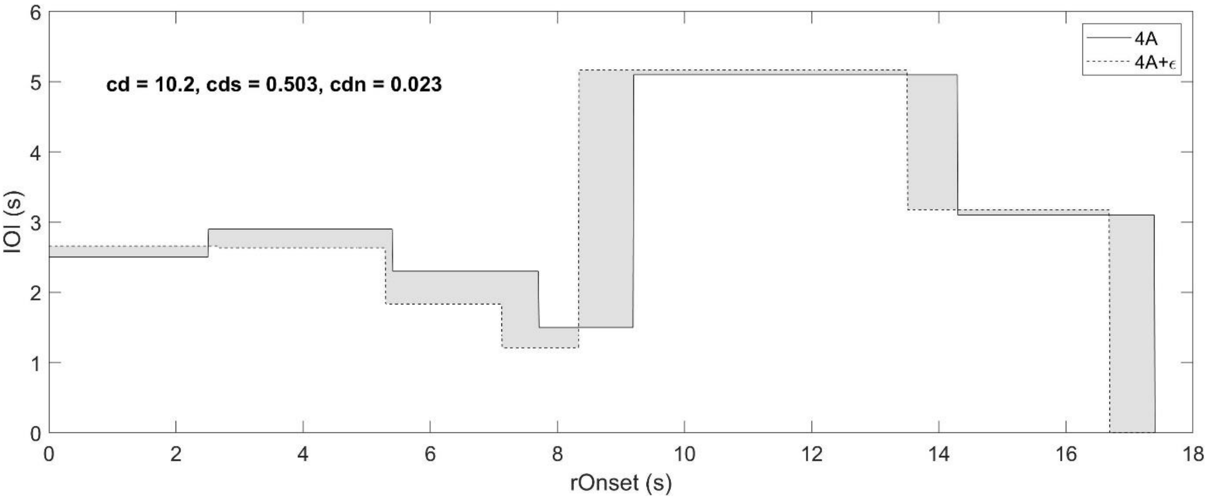


Figure 2.13 Two chronotonic functions based on Phrase Variant 4A (NSU=6), showing typical differences arising from IOI jitter. The chronotonic distance cd between these two rhythms is defined as the total area (shaded) between these functions (Toussaint, 2006). The calculation of scaled (cds) and normalized (cdn) versions is described in the text. Roughly speaking, cds corresponds to the average breadth of the shaded area (considered as a perimeter with length and breadth), cdn to the average proportional difference between the chronotonic functions. Both are less dependent than cd on differences in phrase duration.

For discrete computation of functions on a grid spacing Δt, and for phrase durations PD, this chronotonic distance is given by

Equation 2.5

$$cd_{ab} = \sum_{t=0}^{max(PD_a,PD_b)} |\,CF_a - CF_b|\,\Delta t$$

and the overall intra-individual variability in a set {A} containing N_A phrases by

Equation 2.6

$$\overline{CD_{AA}} = \frac{\sum_{a=1}^{N_A}\sum_{b=1}^{N_A} cd_{ab}}{N_A(N_A - 1)}$$

Setting the denominator of this equation to N(N-1) avoids counting the self-dissimilarity of individual phrases (zero by definition). In general, the area between two functions depends on phrase duration and/or the number of song units in the rhythm, but we may be interested in removing this dependence to better measure IOI jitter, i.e., the breadth of the area when this is considered as a perimeter with length and breadth (Figure 2.13). Simple geometric considerations suggest for pairs of similar quasi-isochronous rhythms the chronotonic distance is described by

Equation 2.7

$$cd \approx (NSU + 1) \times \overline{IOI} \times \overline{\delta IOI}$$

where $\overline{IOI}$ is the mean IOI across both phrases and $\overline{\delta IOI}$ the mean absolute difference between the IOIs from each phrase, averaged across song unit positions. The mean IOI is given by the mean phrase duration divided by the number of song units, and we can write

Equation 2.8

$$cd \approx \overline{PD}\left(\frac{NSU + 1}{NSU}\right)\overline{\delta IOI}$$

Removing the dependency on PD and NSU leads us to a quantity I refer to as the scaled chronotonic distance:

Equation 2.9

$$cds = cd/\left(\overline{PD}\left(\frac{NSU+1}{NSU}\right)\right)$$

For isochronous rhythms in which the only source of variability is IOI jitter, *cds* is a direct measure of this jitter, independent of the phrase duration or number of song units. Modelling results for heterochronous rhythms showed that mean pairwise *cds* is less dependent than *cd* on departures from isochrony for the kinds of rhythms found in humpback song (SI §2.7.6). It is therefore a useful quantity for the comparison of IOI jitter present in different PVs or themes. A final complication arises when comparing phrases from different PVs, if these possess different numbers of song units. We must decide on which values of the PD and NSU to use: to avoid undue weighting being given to shorter phrases I selected the pair providing the largest scaling factor (Equation 2.10), before taking the mean value over all pairwise comparisons to give $\overline{CDS_{AB}}$.

Equation 2.10

$$cds_{ab} = cd_{ab}/max\left(PD_a\left(\frac{NSU_a+1}{NSU_a}\right), PD_b\left(\frac{NSU_b+1}{NSU_b}\right)\right)$$

For sets of phrases that differ in phrase duration (‘tempo’), resulting in proportional scaling of the IOI vector, this scaling can be removed by calculating chronotonic distances *cdn* (and the pairwise mean CDN) between pairs of rhythms scaled to their phrase durations ([NIOI]). For this calculation the NIOI vectors are first converted to normalized chronotonic functions CFN plotted on a proportional time grid.

Equation 2.11

$$cdn_{ab} = \sum_{t=0}^{1} | \, CFN_a - CFN_b \, | \, \Delta t$$

I used $\overline{CDS_{AA}}$ to measure intra-individual rhythmic variability, or $\overline{CDN_{AA}}$ where I wished to remove dependence on phrase duration (meaningful only for closely related PVs). For PVs, these values are alternatives to the mean IOI standard deviation $\overline{\sigma_{IOI}}$ described above. For themes containing a mixture of PVs they measure the combination of IOI jitter and phrase variant diversity (PVD). The chronotonic distance between different PVs can be large compared to that arising from IOI jitter alone (Figure 2.19), and also depends on the precise rhythmic patterns involved (SI §2.7.6).

Before leaving intra-individual variability, I note that mean pairwise distances can be used to select a representative or 'median' phrase: that which has the smallest mean distance to all others in a set of phrases (Garland *et al.*, 2012). An alternative measure of rhythmic variability in this set is then provided by the average distance $\overline{MCD_{AA}}$ to the median (setting the denominator to $N_A - 1$ because one of the phrases the median rhythm itself and the self-dissimilarity to this rhythm is zero by definition).

Equation 2.12

$$\overline{MCD_{AA}} = \frac{\sum_{a=1}^{N_A} mcd_a}{(N_A - 1)}$$

The median rhythm for a mixture of PVs is somewhat analogous to the mean IOI vector for a single PV: both may be described as rhythmic 'stereotypes'. MCD (together with its scaled and normalized versions MCDS, MCDN) is then analogous to the mean IOI standard deviation $\overline{\sigma_{IOI}}$ as an appropriate measure of rhythmic variability.

2.4.2 Indirect Measures of Phrase Variability

The indirect measures described in this section (Table 2.6) do not require prior identification of song phrases, and are calculated for single themes. This means both that they might allow wider datasets to be analysed for the same effort and are not sensitive to ambiguity in coding decisions.

Table 2.6 Indirect (theme-level) measures of phrase rhythm variability.

Measure	**Description**
sdIOI	Standard deviation of IOI timeseries
ANISOS	Mean anisochrony, or proportional deviation from isochrony, scaled to account for dependence on number of song units
NPVI	normalized Pairwise Variability Index
LZC	Lempel-Ziv Complexity

To investigate whether these methods might provide useful indirect measures of phrase rhythm variability I used Monte Carlo techniques, in which simulated phrase rhythms are generated via random sampling. Assumptions are discussed and phrase rhythm templates provided in SI §2.7.3. Simulated themes (concatenated sets of N phrases) were based on observed phrase rhythm templates $[IOI\alpha]_0$, with variability introduced via random sampling for IOI jitter σ_{IOI} and phrase variant diversity PVD (Table 2.7). In the model, as PVD increases from 0 to 1 the simulated theme changes from containing PV α only to a mixture in which each phrase has a 50% probability of belonging to PV α or β. The model was run 1000 times for N=10 across a range of values of σ_{IOI} and PVD, outputting direct and indirect measures of phrase variability. In the case of single PVs, linear regression analyses were then performed on the indirect measures against a pairwise chronotonic distance measure (a direct measure of phrase rhythm variability) to assess how well indirect measures captured phrase rhythm variability.

Table 2.7 Monte Carlo model used to simulate humpback themes for investigating the dependence of indirect measures of phrase variability on IOI jitter ε (drawn from a normal distribution, standard deviation σ_{IOI}) and PV diversity (PVD).

Phrase simulation	$[IOI\alpha]_i = [IOI\alpha]_0 + [\varepsilon]_i$ $[IOI\beta]_i = [IOI\beta]_0 + [\varepsilon]_i$	$[\varepsilon]_i \sim Normal(0, \sigma_{IOI})$
Theme simulation (concatenated set of *N* phrases)	$\{[IOI\alpha]_i$ if $1 \leq i \leq N_\alpha$ $\{[IOI\alpha]_i$ if $N_\alpha < i \leq N, x \leq 0.5$ $\{[IOI\beta]_i$ if $N_\alpha < i \leq N, x > 0.5$ $\left\{[IOI]_j\right\} = permutate\ (\{[IOI]_i\})$	$N_\alpha = round\big((1 - PVD)N\big)$ $x \sim Uniform(0,1)$

The indirect methods include two zeroth order measures that are not responsive to the variability resulting from a re-ordering of IOIs within phrases. The simplest of these is the sample standard deviation of the IOI timeseries, *sdIOI*. Although a function of the rhythmic template, where this is fixed *sdIOI* was expected to be responsive to IOI jitter. I also calculated the mean proportional deviation from isochrony, inspired by a study of cyclic rhythms in North Indian music (Clayton, 2020). Although, like *sdIOI*, this 'mean anisochrony' is a function of the rhythmic template, it is insensitive to changes in tempo because it is calculated as a proportion of the mean IOI. For timeseries duration *DUR* and number of song units *NSU*, the mean anisochrony is given by

Equation 2.13

$$\overline{ANISO} = \frac{1}{NSU}\sum_{i=1}^{NSU}\left|\frac{IOI_i}{\overline{IOI}} - 1\right| = \frac{1}{DUR}\sum_{i=1}^{NSU}\left|IOI_i - \overline{IOI}\right|$$

Its maximum value is dependent on *NSU* (SI §2.7.1),

Equation 2.14

$$max(\overline{ANISO}) = 2 - \frac{2}{NSU}$$

and thus when comparing among themes with different numbers of song units, I used a scaled version

Equation 2.15

$$ANISOS = \overline{ANISO}/max(\overline{ANISO})$$

A first order measure of IOI contrast, comparing adjacent IOIs and hence somewhat responsive to IOI ordering, is the normalized Pairwise Variability Index (*NPVI*) introduced by Grabe and Low (2002) to study speech rhythms. *NPVI* is insensitive to local fluctuations in tempo and takes a numerical value from zero for perfect isochrony to a maximum of 200 for highly contrasting alternate IOIs.

Equation 2.16

$$NPVI = \left(\frac{100}{m-1}\right)\sum_{i=1}^{m-1}\left|\frac{IOI_i - IOI_{i+1}}{(IOI_i + IOI_{i+1})/2}\right|$$

Finally, I included an algorithmic measure, Lempel-Ziv Complexity (*LZC*) (Lempel and Ziv, 1976). *LZC* captures repetition across a wide range of time-scales, in essence counting the number of classes of repetitive rhythms contained in the IOI timeseries. Its value is therefore low for isochronous timeseries and increases with the number of rhythmic substructures, reaching a maximum for a purely random timeseries. For a given rhythmic stereotype, the addition of IOI jitter is likely to move the timeseries towards randomness, hence we might expect *LZC* to perform the function of capturing levels of IOI jitter.

In my application of *LZC* to rhythm, the IOI timeseries was first transformed into a Binary Onset String, e.g., the IOI vector [3333] on a unit grid is written as ‘1001001001001’. I then used the algorithm described in Lempel and Ziv (1976),[8] which parses the string into a smaller number of substrings (‘codewords’). The complexity of the string may then be expressed as the size *c* of the set containing the codewords (the ‘codebook’) (Toussaint, 2020). Here I followed Kershenbaum and Garland (2015) in normalizing with respect to the asymptotic size of the codebook for a random string, given by $L \log d / \log L$ for a string of length L and alphabet size d (Small, 2005, p. 60), resulting in the following expression, where d=2 in the case of a binary string:

Equation 2.17

$$LZC = \frac{c \log L}{L \log d} = \frac{c \log_d L}{L}$$

In general *LZC* is a function of grid size; guided by Monte Carlo simulations (SI §2.7.2) I selected a grid size of 0.5 s for humpback song data, resulting in typical string lengths of 100–300 symbols. For such strings the lower and upper numerical limits of the normalized *LZC* corresponding to isochronous rhythms and random strings can readily be calculated (< 0.2 and c.1). A possible limitation of *LZC* is that its value depends on string length (Hu *et al.*, 2006). To explore whether this is likely to impact my observations, I modelled the dependence of *LZC* on theme duration (SI §2.7.2).

2.4.3 Results

I begin by comparing direct measures of intra-individual phrase rhythm variability in observed song data (Table 2.8). To assess which of the chronotonic distance measures might be most appropriate as a measure of IOI jitter at PV-level, I performed linear regression analyses for all

---

[8] Using MATLAB code from Quang Thai with the ‘exhaustive’ switch, which implements ‘Scheme 1’ from Hu et al (2006).

candidate distances against $\overline{\sigma_{IOI}}$, the most direct measure of IOI jitter (Table 2.9). The coefficients of determination (R^2) suggest that the normalized chronotonic distances best capture IOI jitter, that scaled distances are to be preferred to absolute distances, and that there is no reason to favour the average pairwise distance or the average distance to the median phrase. For all measures shown, both humpback singers displayed similar variabilities when averaged across PVs, although Singer 190924 was more variable for the single shared variant PV 3A° (§2.5.4). I note that the mean coefficient of variation (CV) for these humpback song PVs is approximately twice as high as the value for my performance of 'Auld Lang Syne'.

Table 2.8 Direct measures of PV intra-individual variability ($N \geq 5$).

Singer 190919						**Singer 190924**					
PV	**N**	$\overline{\sigma_{IOI}}$ **(s)**	$\overline{CV_{IOI}}$ **(%)**	$\overline{CD}$ **(s^2)**	$\overline{CDN}$	**PV**	**N**	$\overline{\sigma_{IOI}}$ **(s)**	$\overline{CV_{IOI}}$ **(%)**	$\overline{CD}$ **(s^2)**	$\overline{CDN}$
						1Ba6	5	0.09	7.6	4.6	0.017
2Aa	7	0.25	6.8	8.9	0.018	**2Ba**	21	0.24	7.9	4.2	0.025
3A° **3A'**	16 7	0.17 0.19	6.8 7.9	5.9 5.7	0.011 0.012	**3A°**	16	0.28	11.4	8.3	0.018
						4A°	6	0.16	6.1	5.9	0.014
						5Aa **5Aa'**	5 5	0.24 0.27	6.9 8.6	10.7 11.1	0.025 0.028
7B°	8	0.30	10.1	6.1	0.039	**7A°**	9	0.21	5.8	7.0	0.015
Mean		**0.23**	**7.9**	**6.7**	**0.020**			**0.22**	**7.8**	**7.4**	**0.020**

Table 2.9 Linear regression results for chronotonic distance measures against $\overline{\sigma_{IOI}}$ for selected PVs ($N \geq 3$ for 17 PVs across both singers).

Measure type	**Average pairwise distance**	**Ordinary R^2**	**Average distance to median phrase**	**Ordinary R^2**
Absolute	CD	0.548	MCD	0.517
Normalized to PD	CDN	0.814	MCDN	0.812
Scaled to PD, NSU	CDS	0.635	MCDS	0.605

I next examine whether indirect measures calculated at *theme* level may serve as proxies for the intra-individual variability of phrase rhythms. The results of Monte Carlo simulations for two different PVs show that all measures increased monotonically with increasing IOI jitter (Figure 2.14), but for the more strongly anisochronous PV the use of *sdIOI*, *NPVI* and *ANISOS* as proxies for random IOI jitter seems likely to be less effective than phrase-level CD or CDS. At low levels of jitter (< c.0.2 s) this is because these measures do not respond much to changes in jitter, and at higher levels the wider confidence limits imply that the correlation will be less certain. I also note that the magnitudes of *sdIOI*, *NPVI* and *ANISOS* are highly dependent on PV, making inter-variant comparisons uninformative with respect to IOI jitter. The similar dependence of *LZC* on IOI jitter for both PVs makes this a more promising measure.

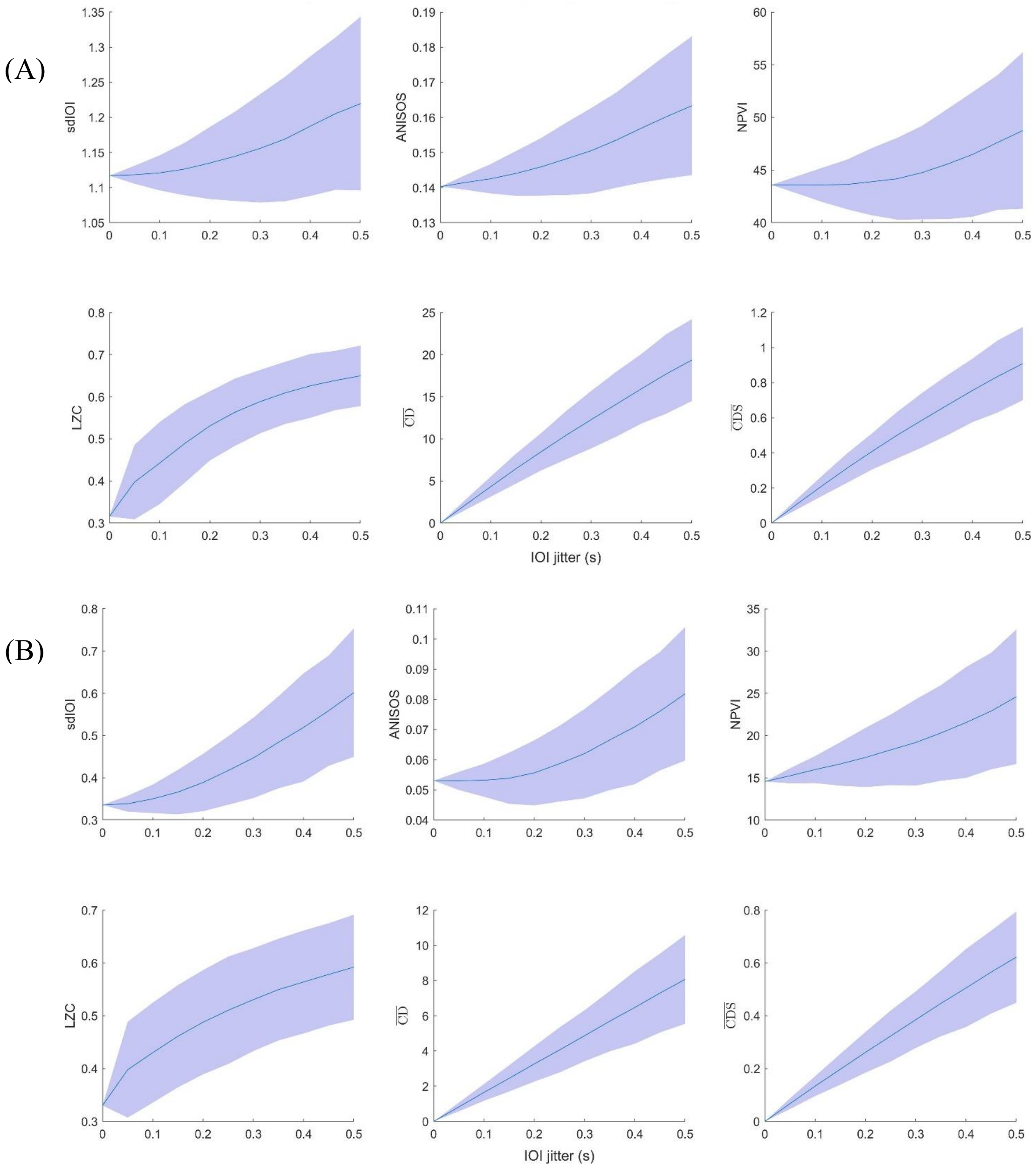


Figure 2.14 Simulations of the dependencies of rhythmic features on IOI jitter, for phrase rhythms based on (A) Phrase Variant 4A, a six unit phrase with pronounced anisochrony, and (B) Phrase Variant 2A-, a three unit phrase closer to isochrony. Theme-level measures (sdIOI, ANISOS, NPVI and LZC) are compared with phrase-level pairwise distance measures (CD, CDS). Mean values and 95% confidence limits are shown for 1000 model runs of a single theme of 10 phrases. LZC grid size is set to 0.5 s. A reduction in the confidence limits could be expected for multiple theme observations. The monotonic increase in sdIOI, ANISOS and NPVI with increasing IOI jitter would be expected for an isochronous phrase rhythm and these simulations show that it also occurs for anisochronous phrase rhythms. The increase in LZC can be explained by considering that increasing IOI jitter leads to increased randomness in the IOI timeseries.

To assess further the suitability of these indirect measures as proxies for the measurement of IOI jitter I performed a linear regression on their values versus a direct measure (CDS) across all model runs. For an isochronous rhythm and PVs closer to isochrony (2A-, 3A), both zeroth order measures and LZC correlate rather well with pairwise CDS (Table 2.10). LZC correlates better across all PVs including the more anisochronous 4A, and should thus perhaps be generally preferred. Conversely, the correlation of NPVI drops off rapidly with increasing anisochrony and is therefore not recommended for the purpose of measuring IOI jitter in humpback song.

Table 2.10 Linear regression performed on simulated (N=1000) theme-level measures versus scaled chronotonic distance (CDS), for four PVs presented in order of increasing anisochrony.

Measure	**Ordinary R^2**			
	Isochronous	**PV 2A-**	**PV 3A**	**PV 4A**
Standard deviation of IOI timeseries (*sdIOI*)	0.984	0.803	0.769	0.456
Mean anisochrony (*ANISOS*)	0.981	0.640	0.758	0.615
Normalized Pairwise Variability Index (*NPVI*)	0.953	0.618	0.570	0.321
Lempel-Ziv Complexity (*LZC*)	0.879	0.707	0.752	0.810

The dependence of *sdIOI*, *ANISOS* and *NPVI* on PV diversity reveals that although all are affected in the same way, the sign of dependence is a function of the PVs modelled (Figure 2.15). *LZC* was found to be relatively insensitive to PV diversity, perhaps because in both cases the pairs of PVs contain much of the same rhythmic structure. CD and CDS increase as PV diversity increase (i.e., as these distances are more likely to measure rhythmic variants), though this increase levels off at high diversities and is not observed in mixtures of closely related variants (results not presented here).

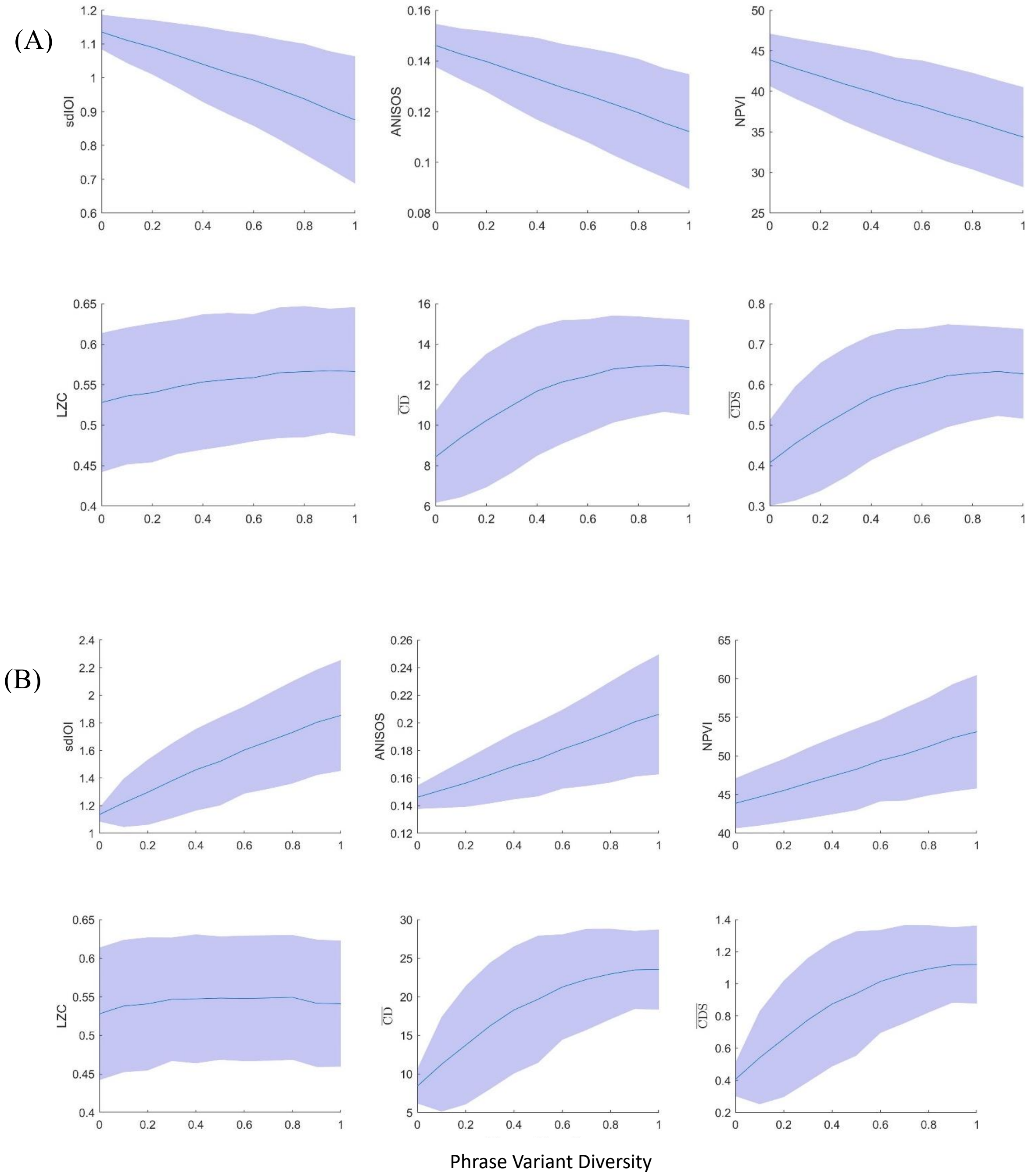


Figure 2.15 Simulations of the dependencies of rhythmic features on Phrase Variant Diversity, for PV 4A combined with variants (A) 4Asplit, (B) 4A-. Theme-level measures (sdIOI, ANISOS, NPVI and LZC) are compared with phrase-level pairwise distance measures (CD,CDS). Mean values and 95% confidence limits are shown for 1000 model runs for a single theme of 10 phrases. LZC grid size is set to 0.5 s, IOI jitter to 0.2 s. A reduction in the confidence limits could be expected for multiple observations of themes. As Phrase Variant Diversity (PVD) increases from 0 to 1 the phrase mixture changes from a theme consisting of PV 4A only to a 50/50 split between 4A and the second variant. The opposite dependencies of sdIOI, ANISOS and NPVI observed for the two different mixtures (A vs B) suggests that no general relationship holds between these features and PVD. On the other hand, in both mixtures LZC is relatively insensitive to PVD, and chronotonic distance increases rapidly with PVD at low levels of PVD before tailing off at high levels.

We can conclude that (1) among the theme-level features LZC has the highest potential as a measure of IOI jitter. However, its dependence on theme duration at short durations restricts its application, and some initial estimate of IOI jitter is required to allow an appropriate grid size to be used (SI §2.7.2), (2) no theme-level feature is able to capture variations in PV diversity, (3) pairwise distance measures are sensitive both to IOI jitter and PV diversity and cannot distinguish between them. However, with phrase coding IOI jitter may be calculated directly as the mean IOI standard deviation, and this information might be combined with theme-level pairwise averages to allow PV diversity to be estimated.

Calculating these theme-level measures for real humpback data (Figure 2.16) showed that *sdIOI*, *NPVI* and *ANISOS* vary widely from one theme to another, but only slightly between the two singers. The close similarity between the patterns of theme dependence of *sdIOI* and *ANISOS* is expected as these are both zeroth order measures of the departure from the mean IOI. Given the sensitivity of *NPVI* to IOI ordering it is perhaps surprising to find that it also tracks these quantities, although this behaviour was observed in the simulations above. These simulations also found *LZC* to be less sensitive to base rhythms and PV diversity than the other theme-level measures so it is not surprising to find that it is less theme-dependent here. However, modelling also demonstrated the problems of comparing *LZC* values across a range of theme durations (SI §2.7.2), and consequently we should not read too much into the box plot summary.

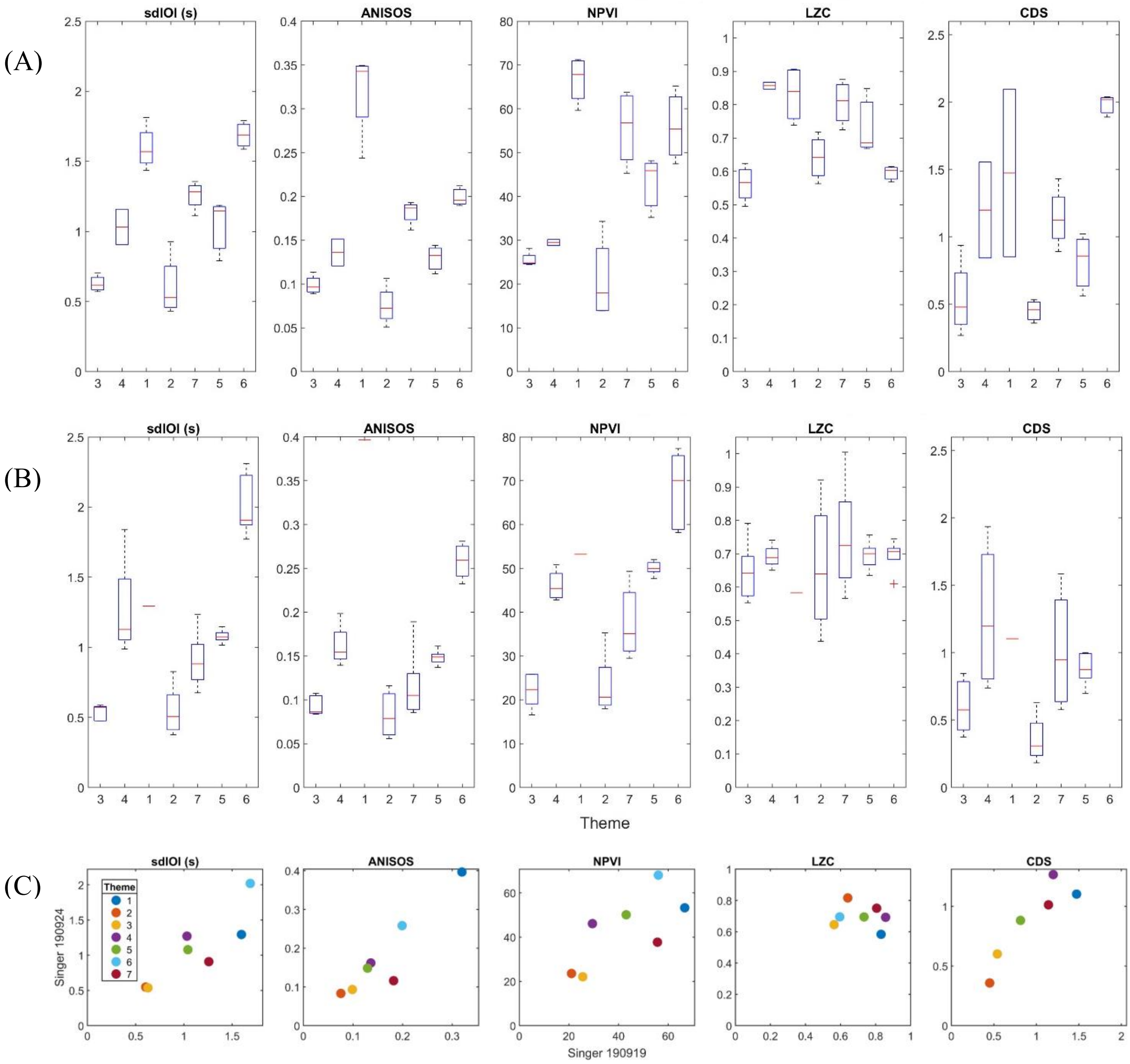


Figure 2.16 Theme-level comparison by Theme (in order of usual occurrence) between Singer (A) 190919 and (B) 190924 for calculations of IOI standard deviation, Scaled Anisochrony, Normalized Pairwise Variability Index, Lempel-Ziv Complexity and pairwise Scaled Chronotonic Distance for each theme, N = 3–5 except for 190919 Th4 (N = 2) and 190924 Th1 (N = 1). For LZC themes shorter than 10 s were excluded. CDS values are mean pairwise distances within each theme. For Singer 190924 Theme 6 consisted of single phrases so CDS could not be calculated. The comparison between singers for mean values of each measure (C) confirms that for all measures except LZC both singers exhibit similar and pronounced inter-theme variability, and that the range of inter-theme variability is greater than inter-singer variability.

2.5 Comparing Between and Within-Individual Phrase Rhythm Variabilities

In §2.4 I described methods for estimating rhythmic variability within the song of an individual whale. For sets of single phrase variants (PVs), I found the normalized chronotonic distance to be an appropriate direct measure of IOI jitter. For PV mixtures, the total variability was measured using scaled chronotonic distance.

Here, I consider rhythmic variability within and *among* humpback singers in themes made up of single PVs and mixtures of PVs. I assess the power of a technique combining pairwise chronotonic distances with the Mantel test to measure whether inter-individual (between individual) variability exceeds intra-individual (within individual) variability. I also examine whether this technique can distinguish between systematic rhythmic differences among individuals *and* discrepancies in IOI jitter, as these sources may play different roles in signal evolution (§2.1.2).

I first discuss Monte Carlo simulations used to test whether systematic rhythmic differences could be detected in the presence of varying levels of IOI jitter, how sensitive the technique was to systematic differences in IOI pattern and phrase duration, and whether IOI jitter could be assessed in the presence of systematic difference (§2.5.1). Next, I describe statistical tests for determining the significance of (1) levels of inter-individual variability compared against intra-individual variability, (2) differing levels of intra-individual variability (§2.5.2). These are followed by the results of applying the combined technique to model runs (§2.5.3) and sample observations (§2.5.4).

2.5.1 Monte Carlo Simulations of Inter-Individual Variability

In general chronotonic distance is a nonlinear function of the rhythmic patterns involved, IOI jitter and systematic difference, so I used Monte Carlo simulations of phrase rhythms from two singers (Figure 2.17, Table 2.11) to assess the ability of chronotonic distance combined with the Mantel test to distinguish among several kinds of variability that arise in humpback whale song. Building on the model described above (§2.4.2), I simulated themes as sets of phrases $\{[IOI]_i\}$ for both singers, based on rhythm templates $[IOI\alpha]_0$ taken from real song. In Singer A, variability consisted only of random variability arising from (a) IOI jitter σ_{IOI}^A and (b) the balance of PVs within a theme ('phrase variant diversity') PVD^A; for Singer B, I introduced systematic inter-individual differences to A, in (a) rhythmic pattern $[\Delta IOI\alpha]^{AB}$ and (b) template phrase duration ('tempo') k^{AB}. Simulated themes comprised sets of 10 phrases, and I ran the model 100 times to allow confidence limits to be estimated (based on the standard deviation of the mean pairwise distance across all model runs). Two sets of base rhythms were used to explore any dependence of the results on the choice of rhythm template.

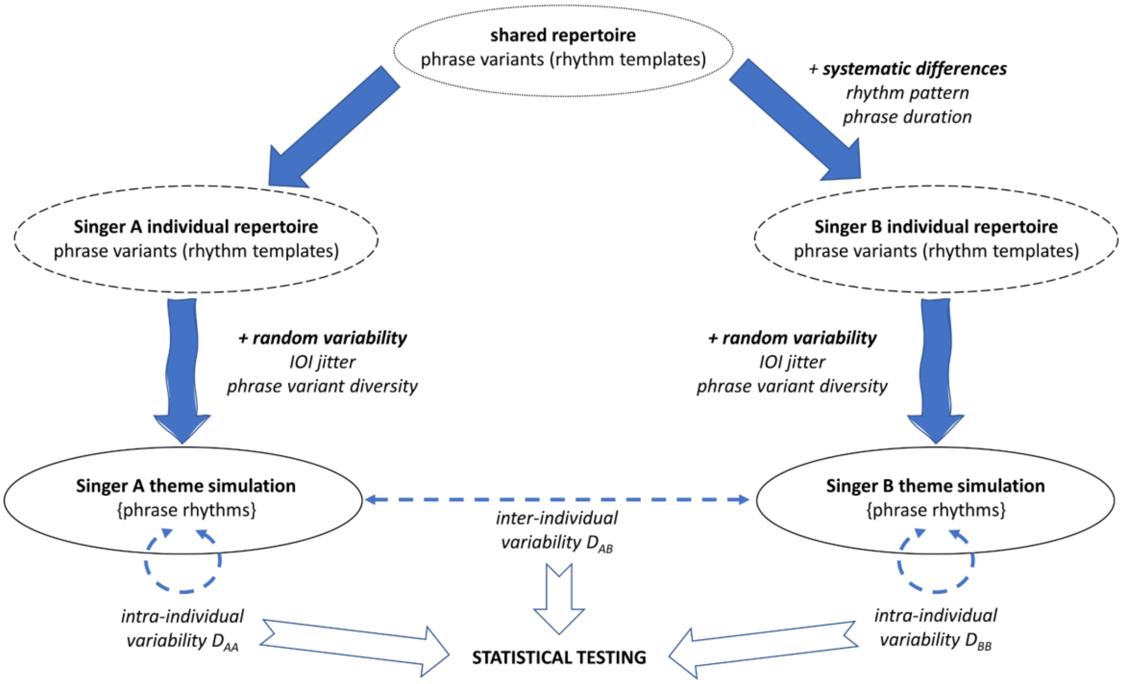


Figure 2.17 Model used to simulate phrase rhythms and assess the ability of distance measures to distinguish among different sources of variability occurring for two singers A and B.

Table 2.11 Monte Carlo model used to assess the ability of the Mantel test applied to pairwise distance measures to distinguish among different sources of variability occurring for two Singers A and B: random IOI jitter ε *drawn from a normal distribution (standard deviation* σ_{IOI}*), random PV diversity (PVD), and systematic differences in rhythmic pattern* $[\Delta IOI]$ *and template phrase duration (*k^{AB}*). Sets of phrase rhythms for each singer* $\{[IOI]_i\}$ *are generated from single or mixtures of phrase templates* $[IOI]_0$ *determined by the level of PVD.*

Shared Repertoire PVs α, β	$[IOI\alpha]_0^A, [IOI\beta]_0^A$	
Singer A Phrase simulation	$[IOI\alpha]_i^A = [IOI\alpha]_0^A + [\varepsilon]_i^A$	$[\varepsilon]_i^A \sim Normal\ (0, \sigma_{IOI}^A)$
Singer A/B Systematic Rhythmic Differences	Rhythm pattern $[\Delta IOI\alpha]^{AB}$ Phrase duration ratio $k^{AB} = PD_0^B / PD_0^A$	
Singer B Phrase simulation	$[IOI\alpha]_i^B = k^{AB}([IOI\alpha]_0^A + [\Delta IOI\alpha]^{AB}) + [\varepsilon]_i^B$	$[\varepsilon]_i^B \sim Normal\ (0, \sigma_{IOI}^B)$
Theme simulation (concatenated set of N phrases)	$\{[IOI\alpha]_i$ if $1 \leq i \leq N_\alpha$ $\{[IOI\alpha]_i$ if $N_\alpha < i \leq N, x \leq 0.5$ $\{[IOI\beta]_i$ if $N_\alpha < i \leq N, x > 0.5$ $\{[IOI]_j\} = permutate\ (\{[IOI]_i\})$	$N_\alpha = round\big((1 - PVD)N\big)$ $x \sim Uniform(0,1)$

I calculated mean pairwise intra-individual distances $\overline{CD_{AA}}$ and $\overline{CD_{BB}}$ from each individual's phrase rhythms (§2.4.1), and $\overline{CD_{AB}}$ across the combined set of their phrase rhythms:

Equation 2.18

$$\overline{CD_{AB}} = \frac{\sum_{a=1}^{N_A} \sum_{b=1}^{N_B} cd_{ab}}{N_A N_B}$$

Typical chronotonic distances resulting from systematic rhythmic differences (Figure 2.18) were comparable with chronotonic distances between PVs (Figure 2.19). I define 'excess' inter-individual distance as the difference between $\overline{CD_{AB}}$ and the weighted mean of the intra-individual differences:

Equation 2.19

$$CD_{AB}^{excess} = \overline{CD_{AB}} - \frac{N_A \cdot \overline{CD_{AA}} + N_B \cdot \overline{CD_{BB}}}{N_A + N_B}$$

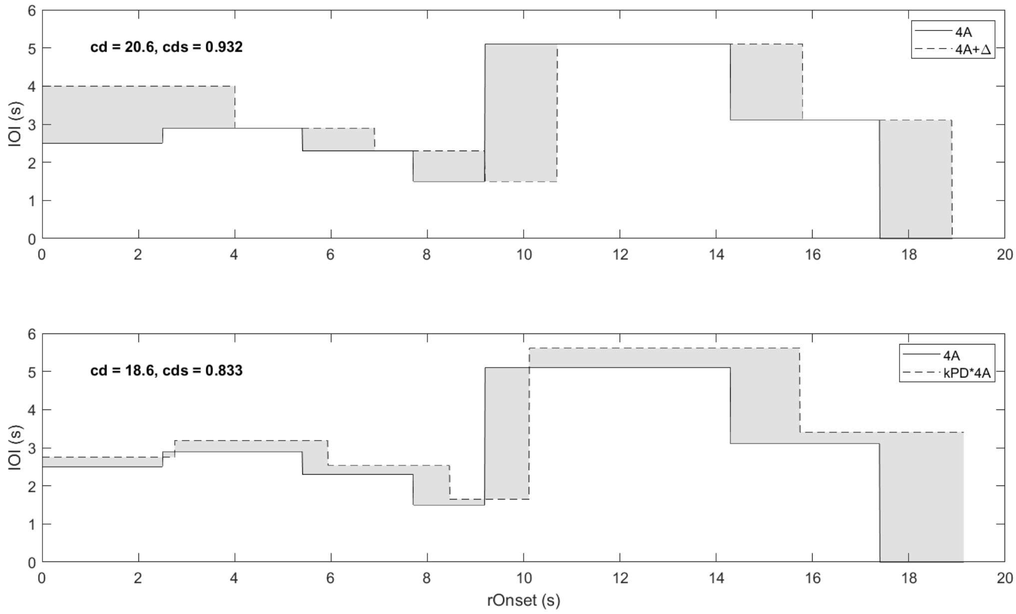


Figure 2.18 Chronotonic functions and calculated distances showing the effect of introducing systematic differences to the base rhythm (PV 4A, solid line). In the upper plot the difference consists in an increase of 1.5 s in IOI_1*, resulting in a shift of all intervals to the right. In the lower plot the difference consists in proportional scaling of the IOI vector corresponding to a 10% increase in phrase duration.*

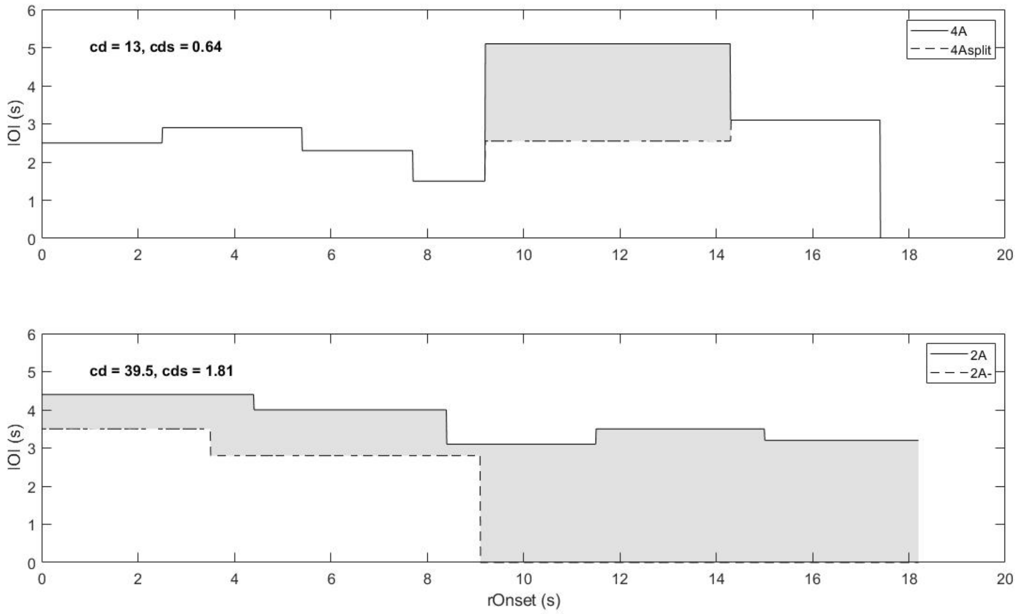


Figure 2.19 Chronotonic functions and calculated distances for PV mixtures (for more details see SI 2.7.3). The 7-song unit phrase variant '4Asplit' shows the effect of splitting the fifth song unit in 4A; the 3-song unit phrase variant 2A- lacks the second subphrase of 5-song unit 2A. Chronotonic distances between the 'high contrast' pair 2A/2A- are around three times greater than between the 'low contrast' pair 4A/4Asplit.

Simulations of single PVs (MODEL RUNS 1–3) and PV mixtures (MODEL RUNS 4–7) addressed specific questions of whether pairwise chronotonic distances combined with the Mantel test were able to discriminate between unsystematic variability and systematic differences among singers, and the sensitivity to systematic difference (Table 2.12). An additional general question was whether these results were dependent on the base rhythms of the PVs. For single PV comparison, I also compared chronotonic with the computationally simpler Euclidean distance and examined which chronotonic measure provided the best relative measure of intra-individual difference (SI §2.7.6).

Table 2.12 Questions addressed by Monte Carlo simulations, with model parameters. Unless otherwise stated IOI jitter = 0.2 s (MODEL RUNS 1–7), PV diversity = 0.3 (MODEL RUNS 4–7). MODEL RUNS 1–3 were additionally carried out for PV 2A, MODEL RUNS 4–7 with phrase mixtures PV 2A/2A–.

RUN	**Question(s)**	**Singer A**	**Singer B**
1	For single PVs, • How do measures of intra-individual variability scale with IOI jitter? • Are different levels of IOI jitter alone sufficient to give rise to significant correlation between the distance and singer difference matrices, potentially confounding attempts to establish whether there are significant systematic rhythmic differences?	PV 4A	PV 4A IOI jitter=0.1–0.4 s
2	For single PVs, can relative intra-individual variability (discrepancies in IOI jitter) be assessed against a background of typical systematic inter-individual differences in (A) rhythmic pattern or (B) template PD?	PV 4A	PV 4A IOI jitter=0.1–0.4 s (A) ΔIOI_1=+1.5 s (B) PD=+10%

3	For single PVs, how sensitive is the Mantel test to systematic differences in (A) rhythmic pattern or (B) template PD?	PV4A	PV 4A (A) ΔIOI_1=0.25–1.5 s (B) PD=-30%–+40%
4	For PV mixtures, with each singer equally diverse, to what degree does the existence of PV diversity change the outcomes of MODEL RUN 1?	PV4A/ 4Asplit	PV 4A/4Asplit IOI jitter=0.1–0.4s
5	For PV mixtures, with each singer equally diverse, to what degree does the existence of PV diversity change the outcomes of MODEL RUN 3?	PV4A/ 4Asplit	PV 4A/4Asplit (A) ΔIOI_1=0.25–1.5 s (B) PD=-30%–+40%
6	For PV mixtures, do different levels of PVD alone give rise to significant correlation between distance and singer difference matrices, potentially confounding attempts to establish whether there are significant systematic rhythmic differences?	PV4A/ 4Asplit	PV 4A/4Asplit PVD=0-1
7	For PV mixtures, is the Mantel test responsive to systematic differences in (A) rhythm pattern, (B) template phrase duration, in the presence of discrepancies in PVD?	PV4A/ 4Asplit	PV 4A/4Asplit PVD=0–1 (A) ΔIOI_1=0.25–1.5 s (B) PD=+10%

2.5.2 Statistical Testing of Inter- and Intra-Individual Variabilities

To assess whether inter-individual variability significantly exceeds intra-individual variabilities, I adapted a procedure previously used to investigate the individuality of the broadband click sequences ("codas") produced by sperm whales (*Physeter macrocephalus*; Antunes *et al.*, 2011; Gero *et al.*, 2016). I calculated the correlation between distance matrices **D** and **F** that contain pairwise comparisons for the phrases sung by all individuals, where **D** contains rhythmic distance values (e.g., chronotonic) and **F** is a binary measure of singer difference ($f_{ij} = 0$ where the two phrases being compared come from the same singer, 1 where they come from a different singer). Intuitively, the greater the positive correlation between **D** and **F** the greater the inter-individual exceeds intra-individual variability. As **D** contains non-independent values (changing any one phrase rhythm will affect the distances between it and all others) the correlation and its significance is estimated following Mantel's procedure (Mantel, 1967). This draws on the insight that if there were no correlation between **D** and **F** (the null hypothesis) then a random permutation of either matrix should be equally likely to lead to a lower or higher value of the correlation. Performing many permutations and recalculating the correlation each time allows a distribution function to be estimated: the p-value of the observed correlation is then given by the proportion of permutations giving a correlation coefficient greater than or equal to that observed (Sokal and Rohlf, 1995, p. 813). Here I used pre-existing code (Glerean *et al.*, 2016),[9] calculating the Pearson correlation coefficient for 10,000 permutations.

To address recent criticism of the Mantel test (Guillot and Rousset, 2013) that highlights the possibility of enhanced Type I error rates (i.e., finding a significant correlation where none exists), I performed a separate study to assess whether the proportion of model runs returned with a p-value below stipulated significance levels exceeded expected levels (SI §2.7.4). In addition, for the sample observations I compared Mantel test results with an inspection of standard errors (*SE*). These were calculated using bootstrapping (Sokal and Rohlf, 1995, p. 823): the observed set of phrase rhythms was resampled with replacement to generate new sets, and $\overline{CD_{AA}}$, $\overline{CD_{BB}}$ and $\overline{CD_{AB}}$ recalculated each time. Standard errors in the means of the original sets of phrase rhythms were then given by the standard deviation in the distribution of mean pairwise distances of the resampled sets. Finally, in order to decide whether intra-individual

---

[9] Code and some description here: https://users.aalto.fi/~eglerean/permutations.html.

variabilities $\overline{CD_{AA}}$ and $\overline{CD_{BB}}$ differed significantly from each other for observed humpback song data, I used Welch's unequal variances *t*-test (Sokal and Rohlf, 1995, p. 404).

2.5.3 Simulation Results and Interpretation

For typical humpback rhythms and variabilities, discrepancies in random variabilities (IOI jitter and PVD) showed up as distinct levels of mean intra-individual variability but did not result in significant correlation between the distance and individual difference matrices (using the Mantel test). On the other hand, systematic inter-individual differences resulted in significant correlations under the Mantel test, but did not affect relative levels of intra-individual variability. Differences in intra-individual variabilities can be tested for using Welch's *t*-test, but were not assessed for simulated data. In the case of themes consisting of highly contrasting PVs, the sensitivity to both systematic difference and discrepancies in random variability was reduced. Additional model runs showed that for single PV comparison, correlations and *p*-values were similar whether based on chronotonic or Euclidean distance (SI §2.7.6).

Comparisons for Single PVs

I initially explored whether and how much inter-individual discrepancies in IOI jitter contributed to inter-individual distances for theme simulations containing single PVs. One singer's jitter was allowed to vary across a wide range, whilst the other's remained fixed, resulting in an approximate linear dependence of intra-individual chronotonic distance on IOI jitter (MODEL RUN 1, Figure 2.20). Inter-individual distances increased with increasing IOI jitter discrepancy, but the excess inter-individual distance (Equation 2.19) was small enough that correlations between distance and singer difference matrices were small and *p*-values large. This shows that on its own a discrepancy in IOI jitter does not show up as significant under the Mantel test.

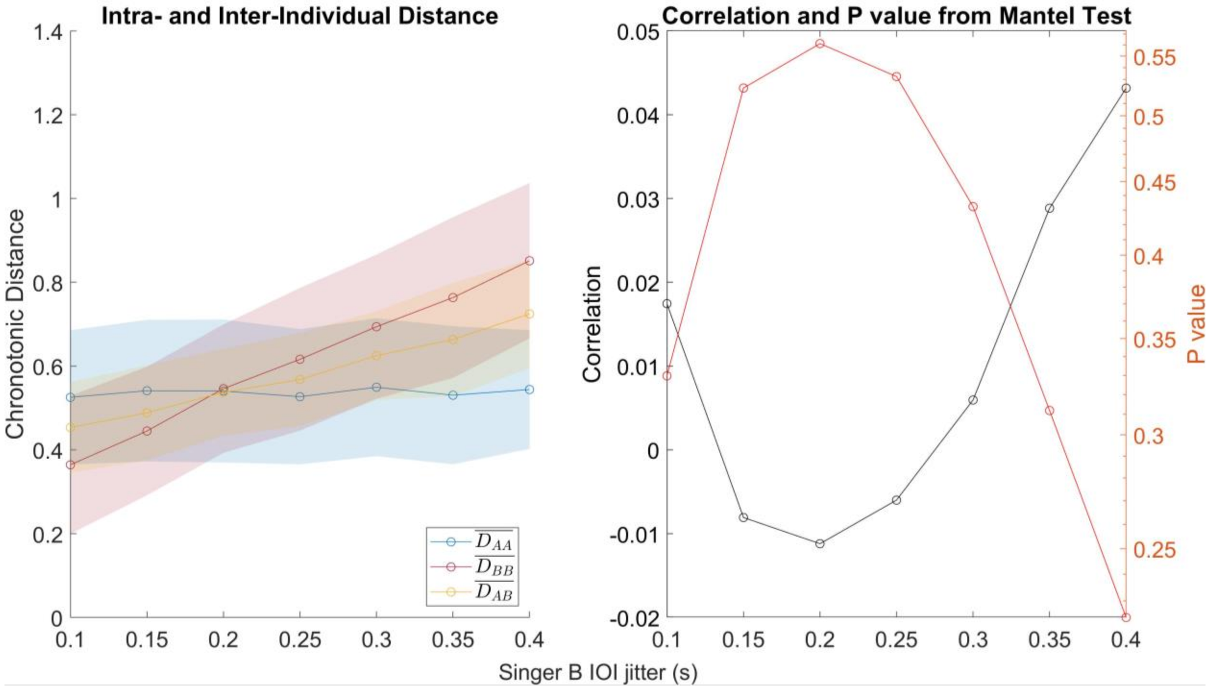


Figure 2.20 **MODEL RUN 1.** *Dependence of scaled chronotonic distances, Mantel test correlation and p-value on IOI jitter of Singer B, with Singer A's IOI jitter held constant at 0.2 s. Plotted lines and shaded areas show the mean and 95% confidence intervals for 100 model runs. Both singers' phrase rhythm sets (N=10) were based on PV 4A. Low correlation and high p-values indicate that the Mantel test would give no evidence to reject the null hypothesis of comparable intra- and inter-individual rhythmic variabilities.*

The fact that discrepancies in IOI jitter do not result in significant correlation is important as it means that such discrepancies should not prevent systematic differences in rhythmic pattern and template phrase duration being detectable. Such resilience to jitter discrepancy was tested directly by modelling the dependence of correlation and *p*-value on IOI jitter discrepancy against a background of systematic difference between the two singers (MODEL RUN 2, Figure 2.21). In this case the introduction of these systematic rhythmic differences did not alter the intra-individual differences but increased the inter-individual distances, leading to larger correlation values and lower *p*-values (<0.001). This indicates that these levels of systematic difference would be easily detectable by the Mantel test for a wide range of jitter discrepancy.

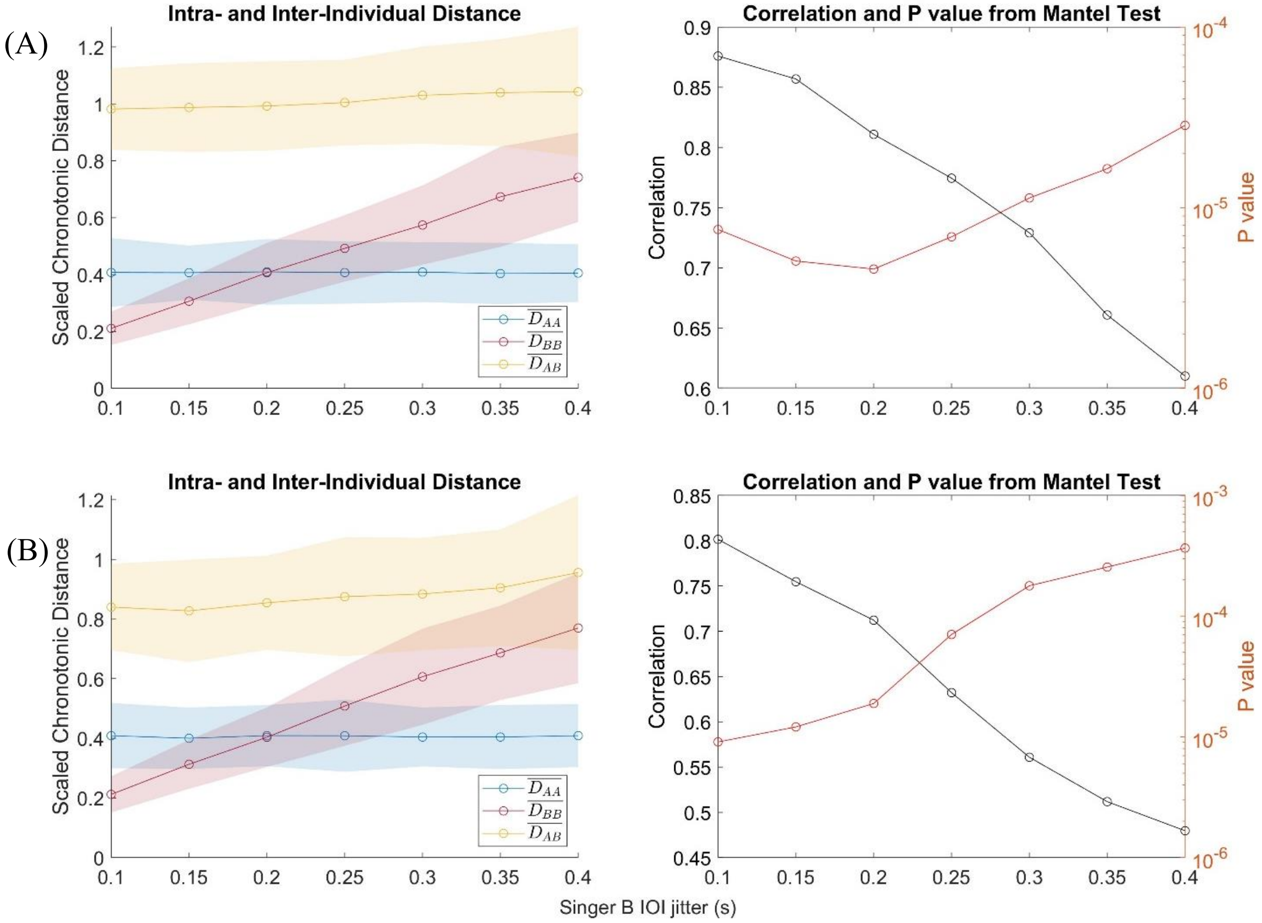


Figure 2.21 MODEL RUN 2. Dependence of scaled chronotonic distances, Mantel test correlation and p-*value on IOI jitter of Singer B, with Singer A's IOI jitter held constant at 0.2 s. Plotted lines and shaded areas show the mean and 95% confidence intervals for 100 model runs. Both singers' phrase rhythm sets (N=10) were based on Phrase Variant 4A. Model runs explored systematic differences in (A) IOI pattern (+1.5s in* IOI_1*) and (B) phrase duration (+10%). In both cases the high correlation and low* p-*values indicate that the Mantel test would return very strong evidence for rejecting the null hypothesis of comparable intra- and inter-individual rhythmic variabilities, in line with the model scenarios (i.e., systematic differences).*

Sensitivity testing showed that inter-individual variability was strongly sensitive to rhythmic pattern differences and to template PD (MODEL RUN 3, Figure 2.22): an elongation in IOI_1 of 0.75 s or change in phrase duration of +/-10% was detected at a significance level of 0.001. Furthermore, intra-individual scaled chronotonic distance was independent of the magnitude of systematic difference, indicating that this quantity can provide a measure of IOI jitter in the presence of systematic difference.

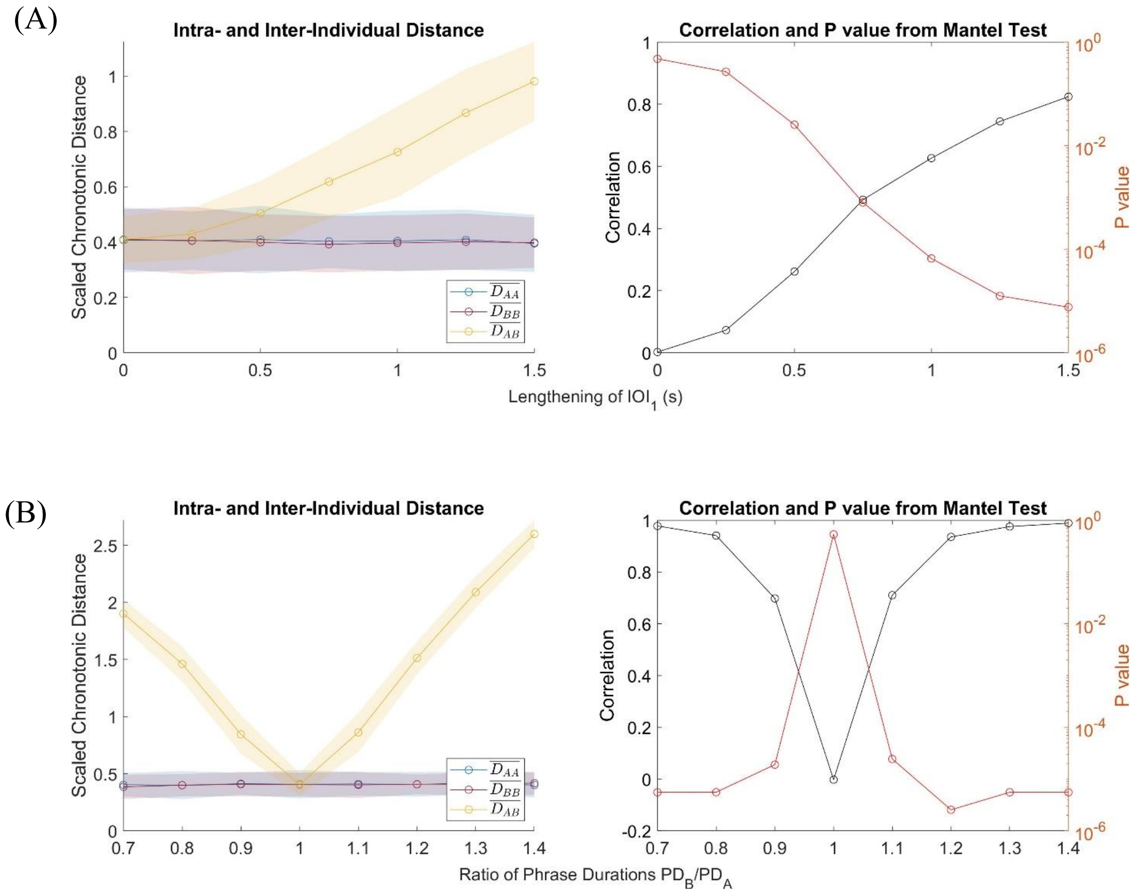


Figure 2.22 MODEL RUN 3. Dependence of scaled chronotonic distances, Mantel Test cor and p-*value on systematic differences between Singers A and B, where both possess the sa of IOI jitter. Plotted lines and shaded areas show the mean and 95% confidence intervals model runs. Both singers' phrase rhythm sets (N=10) were based on Phrase Variant systematic differences are introduced to Singer B: (A) lengthening of* IOI_1 *and (B) vari phrase duration ratio. Relatively minor systematic differences resulted in high correlation* p-*values, indicating that the Mantel Test would return strong evidence for rejecting hypothesis of comparable inter- and intra-individual rhythmic variabilities, in line with th scenarios. Equivalent model runs for Phrase Variant 2A (not shown here) resulted in very levels of correlation and* p-*values.*

Comparisons for PV Mixtures

To test the robustness of the findings for single PVs when extended to PV mixtures, simulations of dependency of pairwise distances on IOI jitter (MODEL RUN 1) and systematic rhythmic differences (MODEL RUN 3) were repeated for PV mixtures (MODEL RUNS 4 and 5). As for the single PV case, intra-individual distances remained approximately linearly dependent on IOI jitter, and discrepancies in IOI jitter did not lead to significant correlation (MODEL RUN 4, see SI §2.7.7). However, sensitivities to systematic differences were lowered when compared with the case of single PVs: only slightly for the low contrast pair, more substantially for the high contrast variants (MODEL RUN 5, see SI §2.7.7).

To explore the sensitivity of the method for PV mixtures, I carried out simulations with variation in PV diversity (MODEL RUNS 6 and 7) rather than IOI jitter (MODEL RUNS 1 and 2). Increasing PV diversity led to increased intra-individual distances, but as with IOI jitter did not lead to significant correlation between distance and singer difference matrices (MODEL RUN 6, Figure 2.23). This had the implication that systematic differences in IOI pattern and template PD could still easily be detected, especially for low contrast PVs (MODEL RUN 7, Figure 2.24, 4A/4Asplit). For high contrast PVs (MODEL RUN 7, Figure 2.24, 2A/2A-) correlations were lower but still significant at the $P < 0.05$ level (differences in template PD) or $P < 0.01$ (IOI_1 lengthening). For both comparison experiments for mixed PVs, confidence levels for the high contrast PVs were found to be about twice the size as those for the low contrast PVs. This suggests that in such cases the method is less sensitive to systematic differences. Further simulations would be required to assess whether this reduction in sensitivity could be overcome by larger sample sizes.

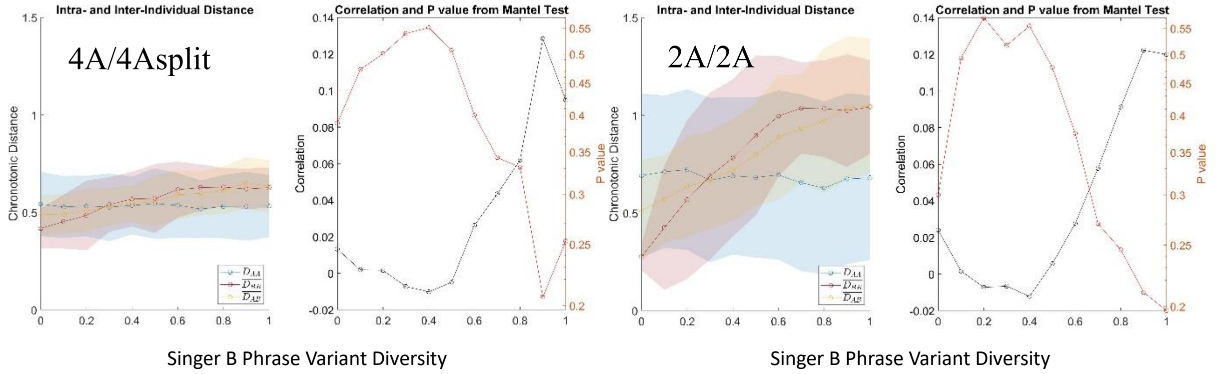


Figure 2.23 MODEL RUN 6. Dependence of scaled chronotonic distance, Mantel test correlation and p-value on Singer B Phrase Variant Diversity (PVD), with Singer A's PVD held at 0.3. Plotted lines and shaded areas show the mean and 95% confidence intervals for 100 model runs. Each singer's theme contained a mix of two phrase variants (4A/4Asplit or 2A/2A-), N=10

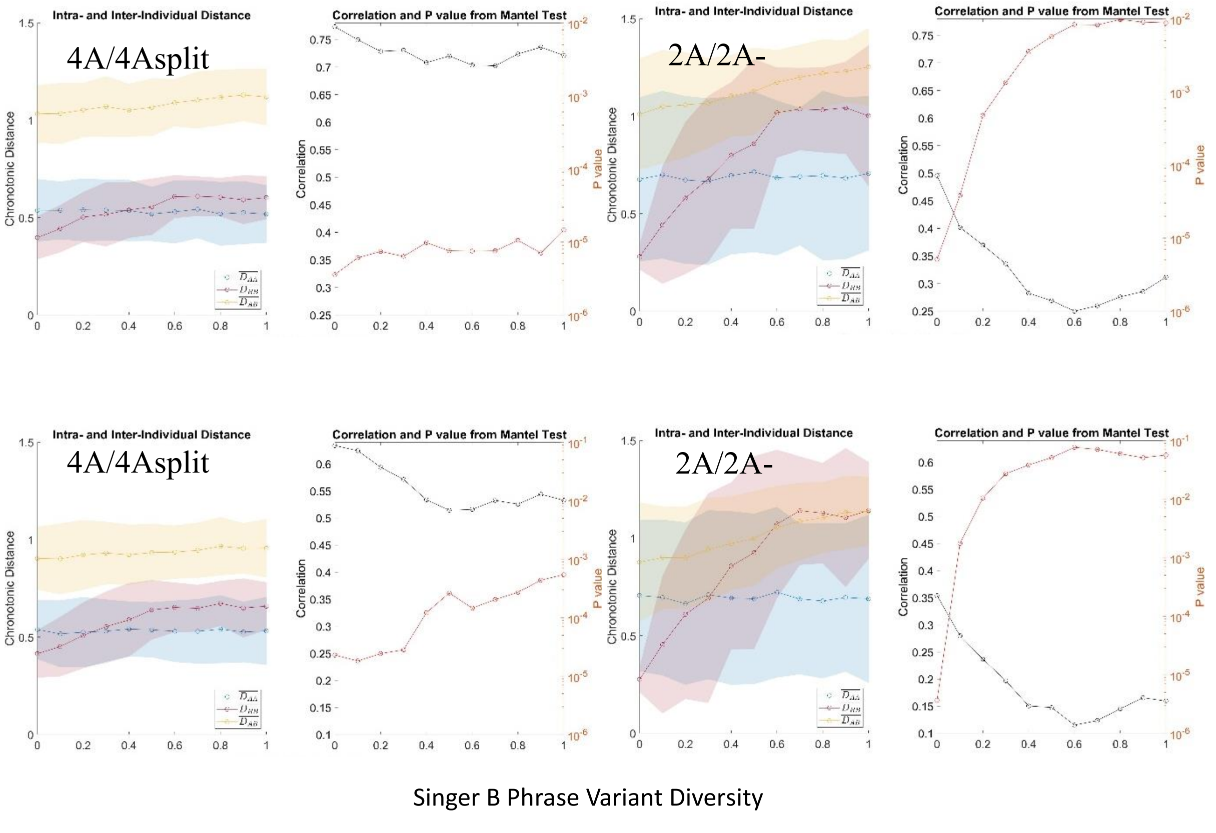


Figure 2.24 MODEL RUN 7. Dependence of scaled chronotonic distance, Mantel test correlation and p-value on Singer B Phrase Variant Diversity (PVD), with Singer A's PVD held at 0.3. Plotted lines and shaded areas show the mean and 95% confidence intervals for 100 model runs. Each singer's theme contained a mix of two phrase variants (4A/4Asplit or 2A/2A-), N=10. Model runs explored systematic differences in (A) IOI pattern (+1.5 s in IOI$_1$) and (B) phrase duration (+10%).

Conclusions

To summarize, Monte Carlo simulations showed a double dissociation between unsystematic variability and systematic rhythmic differences across a wide range of model conditions, implying that the method described here is capable of capturing both. Scaled chronotonic distance offered a similar performance to scaled Euclidean distance in single PV comparisons (SI §2.7.6).

For themes consisting of single PVs, small systematic rhythmic differences were detected against typical levels of IOI jitter, even where these levels were different between individuals (Figure 2.21, Figure 2.22). Based on an examination of intra-individual distances and their confidence limits, discrepancies in IOI jitter appeared to be detectable against typical levels of systematic difference (Figure 2.21), though further statistical testing would be required to validate this conclusion.

In the case of phrase mixtures where both singers were equally diverse, the sensitivity to discrepancies in IOI jitter (i.e., differences in D_{AA} and D_{BB}) was severely impacted for highly contrasting PVs (SI §2.7.7). Similar to the effect of IOI jitter on single PVs, discrepancies in phrase variant diversities (PVD) led to differences in intra-individual but not excess inter-individual variability, with such differences being greater for themes containing more highly contrasting PVs (Figure 2.23). For phrase mixtures, the method was still sensitive to systematic rhythmic differences in the case of typical levels of random variability in PVD, even where such levels were different (Figure 2.24). However, this sensitivity to systematic difference was somewhat lowered with respect to single PVs, especially for differences in phrase duration for themes containing more highly contrasting PVs. Large discrepancies in IOI jitter (assessed via intra-individual distance) appeared to be detectable against typical levels of systematic difference, though as for single PVs further statistical testing would be required to validate this conclusion.

Finally, inspection of the model runs revealed that where the phrase rhythms in one set differed systematically in some way from those in another (and discrepancies in IOI jitter were not too great) we found $\overline{D_{AB}} \geq max(\overline{D_{AA}}, \overline{D_{BB}})$. If the phrase rhythms only differed through IOI jitter, $mean(\overline{D_{AA}}, \overline{D_{BB}}) < \overline{D_{AB}} < max(\overline{D_{AA}}, \overline{D_{BB}})$.

Hence, I conclude that (1) varying levels of intra-individual distances resulting from variation in IOI jitter and PV did not substantially affect excess inter-individual distance (Equation 2.19) and hence the correlation between distance and individual difference matrices is an appropriate gauge of systematic inter-individual differences, (2) varying levels of systematic differences did not substantially affect intra-individual distances, and so discrepancies between individuals reflect differences in random variation in IOI jitter and PVD. In terms of the model parameters, we can write that

intra-individual distance: $\overline{CDS}_{AA} \sim f(\sigma^{A}_{IOI}, TD^{A})$

$\overline{CDS}_{BB} \sim f(\sigma^{B}_{IOI}, TD^{B})$

excess inter-individual distance: $\overline{CDS}_{AB} - mean(\overline{CDS}_{AA}, \overline{CDS}_{BB}) \sim f([\Delta IOI]^{AB}, k^{AB})$

(3) The double dissociation between systematic difference and WRV suggests that use of Welch's *t*-test to determine the significance of differing intra-individual distances will provide information on levels of WRV among different singers.

2.5.4 Results for Song Observations

For shared PV 3A°, correlations and *p*-values returned by the Mantel test provided very strong evidence ($P < 0.0001$) that inter-individual variability D_{AB} for the two singers exceeded intra-individual variability (WRV) (Figure 2.25; centre point compared to mean level of the outer points). Simulation results (§2.5.3) suggest that excess inter-individual variability results from systematic differences in phrase rhythm rather than different levels of WRV, and this is borne out here by the phrase rhythm box plots (Figure 2.11, Figure 2.12). Furthermore, Welch's *t*-test showed that the two singers differ significantly ($\alpha = 0.01$) in their levels of WRV (Figure 2.25; outer points). This confirms impressions given by visualizations (Figure 2.11, Figure 2.12), and the relative values of other measures of intra-individual variability (Table 2.8).

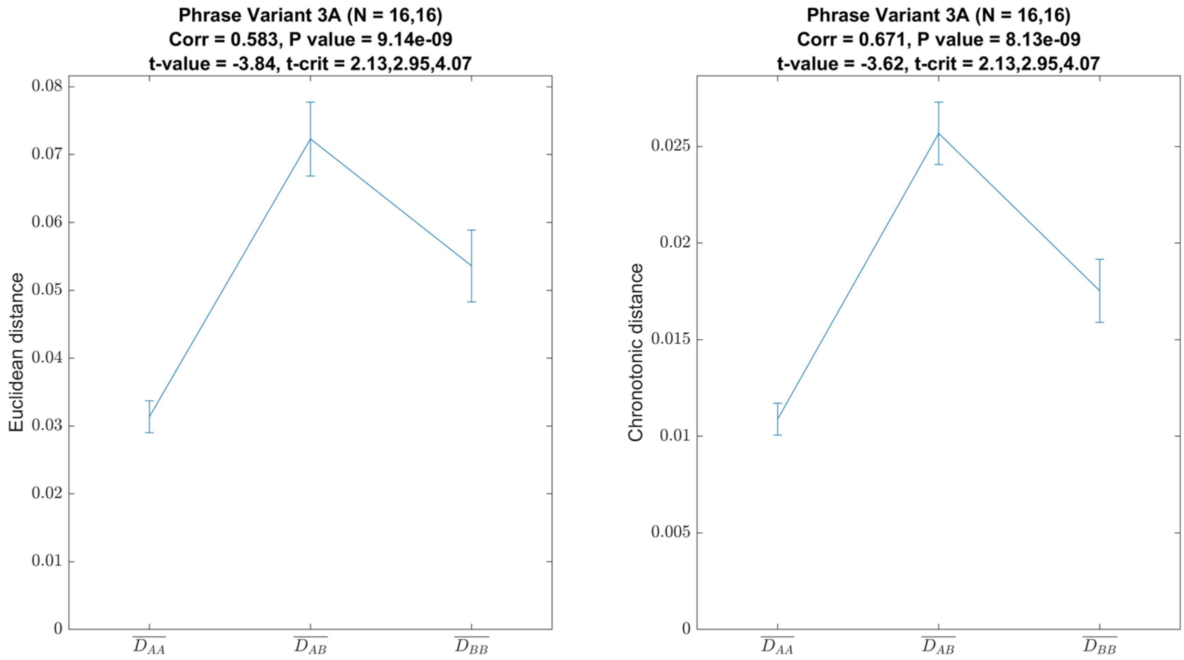
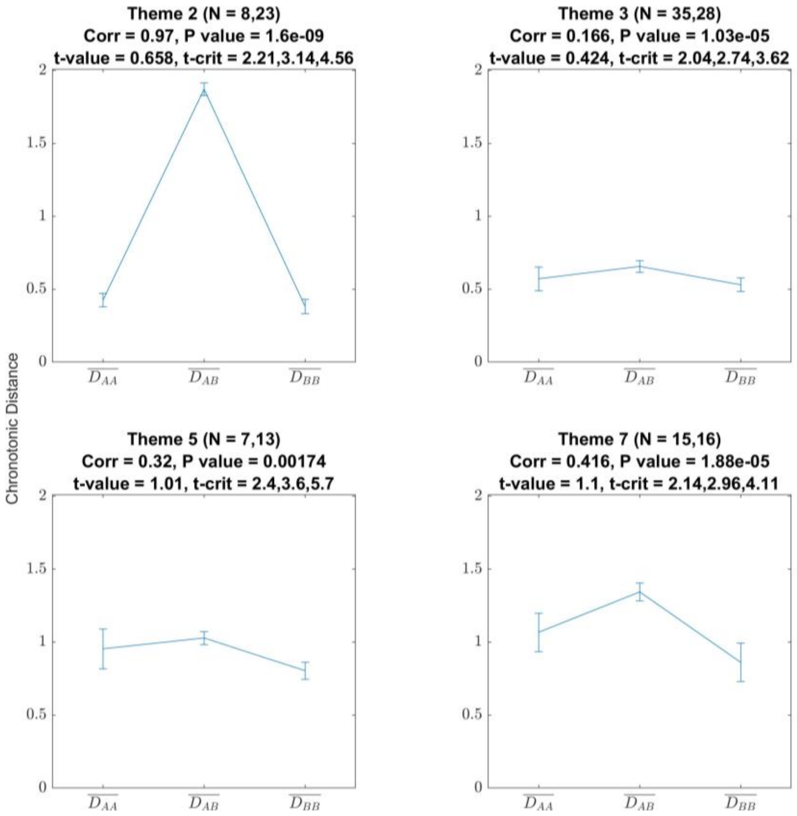


Figure 2.25 Mean pairwise normalized Euclidean and chronotonic distances for PV 3A°. Intra-individual variability is represented by the outer two points (where A stands for Singer 190919 and B for Singer 190924), inter-individual variability by the centre point. Error bars show standard errors calculated using bootstrapping (10,000 replacements). Correlation and p-values are returned from the Mantel test (Pearson coefficients, 100,000 permutations). Critical t-values for Welch's t-test are given for α = 0.05, 0.01 and 0.001. The high level of agreement for Euclidean and chronotonic distance measures was predicted by model simulations (SI 2.7.6).

Applying the same analysis at theme level gave mixed results (Figure 2.26). With regard to inter-individual variabilities, the themes fall into two groups. For Themes 2, 3, 5 and 7 there is strong evidence that inter-individual variabilities are significantly larger than intra-individual variabilities ($P < 0.01$). Simulations (§2.5.3) suggest that such excess inter-individual variability can be interpreted as resulting from systematic differences in phrase duration or rhythmic pattern, although the existence of high contrast PVs makes this interpretation less certain. The effect of different PV repertoires (systematic difference in PV diversity), which was not examined explicitly in the simulations, may also contribute to inter-individual variability. For the remaining Themes 1, 4 and 6, we cannot reject the null hypothesis of equal levels of intra- and inter-individual variability.

Focusing on levels of rhythmic consistency (with respect to both IOI jitter and PV diversity), we see that neither singer stands out as a more precise performer (we cannot reject the null hypothesis of equal levels of intra-individual variability, with $|t'| \leq t'_\alpha$ for $\alpha = 0.05$). Themes 2 and 3 are notably more rhythmically consistent than the others for both individuals (low values of D_{AA}, D_{BB}). Monte Carlo simulations described above (§2.5.1), coupled with the finding that both individuals have similar levels of IOI jitter overall (§2.4.3), could suggest that differences from one theme to another result from varying levels of PV diversity. However, given that there may be marked discrepancies in IOI jitter for specific PVs (as we saw for 3A°), we cannot rule out the possibility that such discrepancies also exist at theme level, contributing to inter-theme differences.

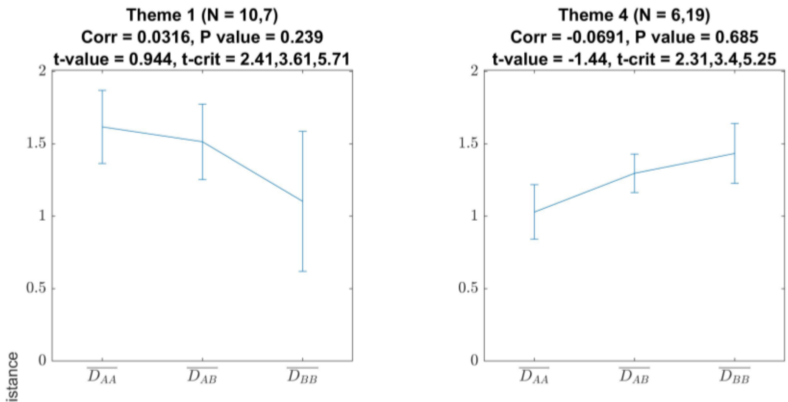

Theme 2 (N = 8,23)
Corr = 0.97, P value = 1.6e-09
t-value = 0.658, t-crit = 2.21,3.14,4.56

Theme 3 (N = 35,28)
Corr = 0.166, P value = 1.03e-05
t-value = 0.424, t-crit = 2.04,2.74,3.62

Theme 5 (N = 7,13)
Corr = 0.32, P value = 0.00174
t-value = 1.01, t-crit = 2.4,3.6,5.7

Theme 7 (N = 15,16)
Corr = 0.416, P value = 1.88e-05
t-value = 1.1, t-crit = 2.14,2.96,4.11

P < 0.01


Theme 1 (N = 10,7)
Corr = 0.0316, P value = 0.239
t-value = 0.944, t-crit = 2.41,3.61,5.71

Theme 4 (N = 6,19)
Corr = -0.0691, P value = 0.685
t-value = -1.44, t-crit = 2.31,3.4,5.25

P > 0.1

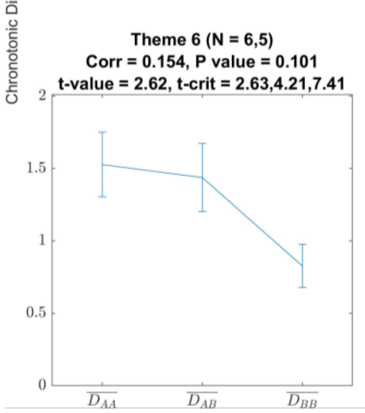

Theme 6 (N = 6,5)
Corr = 0.154, P value = 0.101
t-value = 2.62, t-crit = 2.63,4.21,7.41

Figure 2.26 Mean pairwise scaled chronotonic distances for all themes. Intra-individual variability is represented by the outer two points (A = Singer 190919, B = Singer 190924), inter-individual variability by the centre point. Error bars show standard errors calculated using bootstrapping (10,000 replacements). Correlation and p-*values are returned from the Mantel test (Pearson coefficients, 100,000 permutations). Critical t-values for Welch's* t-*test are given for α = 0.05, 0.01 and 0.001.*

2.6 Discussion

In this chapter I developed and tested a suite of methods drawn from bioacoustics and empirical musicology in order to visualize (§2.3) and quantify (§§2.4–2.5) the variability of repetitive rhythms. Musicological tools were restricted to those that did not assume an underlying isochronous pulse or regular beat. I used them in an exploratory analysis of humpback whale song, demonstrating their scope for examining inter- and intra-individual rhythmic variabilities of single and mixtures of phrase variants (PVs). These methods are applicable to rhythmic phenomena across species and are not limited to the auditory domain. To the useful survey of rhythmic complexity measures for bioacoustics by Ravignani and Norton (2017) I added additional representations (chronotonic diagram, phrase raster plot) and measures (Lempel-Ziv Complexity, anisochrony, chronotonic distance) and introduced statistical tests for comparing inter- and intra-individual variabilities following previous work on sperm whale codas (Rendell and Whitehead, 2003a; Antunes *et al.*, 2011; Gero *et al.*, 2016). In contrast with the recent review of methods for rhythm analysis by Burchardt and Knörnschild (2020), which concentrated on isochronous sequences, my focus has been on the repetitive heterochronous rhythms that are commonly present in human music, poetry and in the vocalizations of other species including the thrush nightingale (Rothenberg *et al.*, 2014; Roeske *et al.*, 2020) and the humpback whale (Handel *et al.*, 2009, 2012; Schneider and Mercado III, 2019). Unlike the study on humpback song rhythm by Schneider and Mercado III (2019), my methods are focused on the rhythmic analysis of song for which additional hierarchical information has been incorporated. In this section I evaluate visualizations (§2.6.1), measurements of intra-individual variability (§2.6.2) and inter-individual variability (§2.6.3) against the aims and criteria summarized above (§2.1.3).

2.6.1 Visualizing the Variability of Repetitive Phrase Rhythms

I begin by evaluating the visualizations used for humpback song rhythms (Table 2.13), starting with the simple IOI timeseries for song coded to theme level (Figure 2.4). This representation was recommended by Schneider and Mercado III (2019), alongside the phase portrait, for its ability to reveal overall patterns in the presence of noise, jitter and rhythmic variations. Compared with Schneider and Mercado III's visualization (2019, Figure 2), I have made inter-theme comparisons easier by adding the mean IOI level. The IOI timeseries qualitatively reveals repeating structures and different levels of intra- and inter-individual variability. Furthermore, it can reveal trends taking place on a variety of timescales, including those belonging to so-called shifting themes (Payne and Payne, 1985), the humpback equivalent of gradual tempo changes in human music. These trends are also visible in phase portraits (Rothenberg *et al.*, 2014), which provide a compact visualization of the level of rhythmic consistency of a particular timeseries, though phase portraits do not allow unambiguous distinctions to be made among the various sources of rhythmic variability and do not offer 1:1 mapping (e.g., [2,3,4,2] would produce the same diagram as [3,4,2,2]).

Among visualizations of humpback song coded to phrase level, the phrase raster plot is a highly useful and flexible tool. It is the only visualization here to show the relative amount of sound and silence in each IOI, which might provide valuable information for further investigation. Although it is better suited to showing the variability in *onset* position rather than variability in IOI, this also means that when plotted in time order it provides a timeseries of durations of the next hierarchical level above the IOI. Hence it may reveal trends in phrase durations and also in the particulars of phrase rhythms (Janney *et al.*, 2016). For humpback song it also displays tendencies in theme composition, e.g., whether particular PVs occur at the starts or ends of themes. Sorting with respect to phrase duration or to the length of the first IOI could also uncover patterns in large amounts of multivariate data (Hyland Bruno and Tchernichovski, 2019). Alongside the phrase raster plot, the phrase rhythm box plot provides a very compact and easy to grasp picture both of the structure of a particular rhythm and statistically well-defined levels of variability in each of its elements. Versions of this plot have been used in recent rhythmic analyses of performances of North Indian and Balkan musics (Goldberg, 2015; Clayton, 2020). The phrase rhythm box plot is limited however to describing the variability in rhythms possessing the same number of onsets, and I therefore suggest that plots of multiple chronotonic functions be used to show the analogous rhythmic structure and variability in PV mixtures, i.e., in families of related rhythms (although such a plot cannot contain the same precise statistical information).

Table 2.13 Evaluation of visual representations of rhythmic structure and variability. Filled circles indicate advantageous features. Empty circles indicate positive responses that are especially context-dependent. E.g., Western Musical Notation is not universally understood, yet for many musicians and musicologists it is the most straightforward way to communicate rhythms.

Representation **(Visualizations)**	**Can be used without phrase coding**	**Easy to interpret**	**Free of musical assumptions**	**Avoids numerical approximation in [IOI]**	**Leads directly to measure of variability**	**1:1 mapping of IOI sequence**	**Compact representation of variability**
Western Musical Notation	●	○				●	
Time Unit Box System	●	●	○			●	
Phase Portrait	●		●	●			●
IOI timeseries	●	●	●	●	●	●	○
Chronotonic Diagram		●	●	●	●	●	●
Phrase Raster Plot		●	●	●		●	
Phrase Rhythm Box Plot		●	●	●	●	●	●

2.6.2 Quantifying Within-Individual Phrase Rhythm Variability

For humpback song I tested a range of quantitative measures of phrase variability (§2.4), including indirect measures which did not require coding to phrase level (Table 2.14). I begin by considering direct assessments of rhythmic variability in repetitions of a single PV, which bears on the question of whether auditory templates (as defined through phrase coding) possess a stable rhythmic aspect or ‘rhythmic template’ (as assumed in my Monte Carlo simulations). I note here that IOIs are not directly represented in coding, and although broad categories of song unit duration were employed I did not use fine-scale durational categories to distinguish among PVs during coding, attempting in this way to avoid assumptions about the level of rhythmic variability in a given PV. The high levels of rhythmic consistency found (mean coefficient of variation in IOI for both singers < 8%) is evidence for stable rhythmic templates.

The most direct way to capture the rhythmic variability in a single PV is the mean standard deviation in IOI across all song unit positions. On the assumption, supported for the humpback song explored in this chapter (SI §2.7.3), that σ_{IOI} is independent of IOI for a given song unit position this measure allows meaningful comparison across different variants. However, in the case of sets of related rhythms containing different numbers of onsets, as for mixtures of PVs in humpback song themes, a different approach is required as there is no non-arbitrary way to align one phrase with another. For human music, sets of related rhythms include cases where a rhythm is slightly altered by the deliberate or accidental omission or introduction of one or more onsets, which may or may not have much impact on the feel of the rhythm. If rhythms can be grouped into families, we could follow the approach of Tanguiane (1993) to establish a “root rhythm” within each family, describing others as elaborations. Alternatively, the rhythm with the most onsets could be taken as the stereotype and in the calculation of means and standard deviations those performed rhythms with fewer onsets would be dealt with by excluding ‘missing’ IOIs from statistical analyses. However, such an approach will only work where notes can be unambiguously assigned to onset positions; in other cases (which will include some human music and certainly humpback song), we may know nothing about intention, whether such variant rhythms should count as belonging to the same or different families, and hence what the putative root rhythm might be. Indeed, inspection of phrase raster plots for the humpback song themes explored here seems to confirm that no one rhythm forms the basis of all the others.

Table 2.14 Evaluation of measures of rhythmic variability. Filled circles indicate advantageous features. Empty circles indicate positive responses that are especially context-dependent.

Measure of Variability	**Robust to phrase coding ambiguity**	**Can be evaluated for mixed rhythms**	**Sensitive to IOI ordering**	**Insensitive to base rhythms**	**Sensitive to IOI jitter**	**Unambiguous response to PV diversity**
Mean standard deviation or coefficient of variation in [IOI] across all song unit positions: $\overline{\sigma_{IOI}}$, $\overline{CV_{IOI}}$			●	●	●	NA
Mean Euclidean Distance ($\overline{\text{ED}}$)			●		●	NA
Mean Scaled Euclidean Distance ($\overline{\text{EDS}}$)			●	●	●	NA
Mean Chronotonic Distance ($\overline{\text{CD}}$)		●	●		●	●
Mean Normalized Chronotonic Distance ($\overline{\text{CDN}}$)		○	●	○	●	●
Mean Scaled Chronotonic Distance ($\overline{\text{CDS}}$)		●	●	○	●	●
IOI timeseries standard deviation (sdIOI)	●	●			●	
Scaled Mean Anisochrony (ANISOS)	●	●			●	
Normalized Pairwise Variability Index (NPVI)	●	●	○		○	
Lempel-Ziv Complexity (LZC)	●	●	●	●	●	●

Pairwise chronotonic distance methods allow us to capture the rhythmic variability in such cases. Here I examined three versions of chronotonic distance (CD, CDN, CDS), initially comparing them against $\overline{\sigma_{IOI}}$ in the case of single PVs (Table 2.9). For the sample humpback data, CDN (measuring the difference between IOI vectors normalized to phrase duration) correlated best with $\overline{\sigma_{IOI}}$, but CDN should not be used with PV mixtures (themes) unless all PVs are roughly comparable in terms of phrase duration and numbers of song units. This is because perceptual or production similarities between rhythms (e.g., shared subphrases) will be lost, because there is little overlap in the normalized chronotonic functions. In general, CDS (scaled chronotonic distance), which preserves such overlap yet is less sensitive than CD to small differences in phrase duration or number of song units, should be preferred. In the case of humpback themes, the use of CDS revealed a wide range of intra-individual variabilities among theme types and singers (§2.5.4). Because CDS is sensitive to both IOI jitter and PV diversity we cannot separate their contributions. Pairwise distance methods also offer a less assumption-ridden approach to selecting a 'root rhythm', as a median phrase rhythm can be selected on the basis of minimum distance to all other phrases in the set (§2.4.1), analogous to work on similarity between rhythmic stereotypes (Toussaint *et al.*, 2011) or humpback song considered as symbol string (Garland *et al.*, 2012).

For human music or other systems where identifying and categorizing phrase rhythms is time-consuming or subject to ambiguity, it may be preferable to avoid this process. I explored the use of indirect alternatives in the specific case of humpback song, the key finding from these simulations and observations being that Lempel-Ziv Complexity (LZC) is a promising candidate for capturing levels of IOI jitter. This is the first time that LZC has been used to assess rhythmical variability in the repetitive rhythms of nonhuman animal song, although Toussaint (2020) employed it to measure rhythmical complexity among rhythmic stereotypes in human music, and it was proposed by Suzuki *et al.* (2006) and explored by Kershenbaum and Garland (2015) as a way of measuring the sequence complexity of humpback song. My modelling of LZC on grid size and theme duration (SI §2.7.2) suggests that the grid size should be set to a level larger than IOI jitter (an approximate value of which must therefore be established by other means), and comparisons restricted to themes of similar duration. The simulations of LZC showed the same general dependence on timeseries duration as simulations of electroencephalogram (EEG) signals (Hu *et al.*, 2006) and the further development of LZC as a measure of jitter might benefit from the normalization scheme proposed by these authors.

The simulations of IOI timeseries standard deviation (sdIOI) and scaled mean anisochrony (ANISOS) showed that these zeroth order measures could only be used to capture the rhythmical variability resulting from IOI jitter in the case of single repetitive rhythms that were not too anisochronous. The correlation between the normalized Pairwise Variability Index (NPVI) and IOI jitter was poorer than for these zeroth order measures, except for isochronous sequences. However, the fact that sdIOI, ANISOS and NPVI all increased monotonically with IOI jitter supports previous suggestions that NPVI has little additional explanatory power over zeroth order measures (Toussaint, 2013; Ravignani and Norton, 2017). The similar dependence of these three measures on IOI jitter also backs up claims that NPVI may be useful in detecting departures from isochrony (Schneider and Mercado III, 2019; Burchardt and Knörnschild, 2020). This is further supported by the shared inter-theme variation of NPVI with sdIOI and ANISOS in my humpback song observations.

When it comes to studying systematic rhythmic differences between individuals or populations, NPVI has been a popular choice in linguistics and empirical musicology (Grabe and Low, 2002; Patel and Daniele, 2003; Raju *et al.*, 2010; London and Jones, 2011; Daniele and Patel, 2013; Hansen *et al.*, 2016; Temperley, 2017). I did not perform significance testing on inter-individual differences in levels of sdIOI, ANISOS or NPVI, but it is certainly possible that these measures could reveal systematic differences in IOI pattern, selection of PVs, and preferences for anisochrony. What Ravignani and Norton (2017) state to be a weakness of NPVI, that it is “not very robust to different speakers,” would here prove to be an advantage.

For the complex mixtures of rhythmic templates seen in many acoustic systems, including humpback song, these theme-level measures would only be useful measures of IOI jitter if they respond unequivocally to PV diversity. In my simulations the responses of sdIOI, ANISO and NPVI were determined by the PVs present, whereas LZC was unaffected by PV diversity. Hence only LZC can be used for the estimation of IOI jitter in repetitive mixed rhythms, and its use could be advantageous over methods such as pairwise chronotonic distance when identifying phrases is time-consuming or ambiguous. Here, LZC as a quantitative measure of variability could be complementary to the more qualitative information provided by IOI timeseries and phase portraits. Even where phrase coding has been carried out, however, the use of LZC remains recommended for mixed repetitive rhythms, as its insensitivity to PV diversity (and sensitivity to IOI jitter) makes it a useful supplement to scaled chronotonic distance (which is sensitive to both).

2.6.3 Comparing Between and Within-Individual Phrase Rhythm Variabilities

Starting from the chronotonic function representation of rhythm I defined a scaled chronotonic distance (CDS) to study rhythmic variability at the level of PV and theme. Like other distance measures, CDS is a single number capturing the variability in a set of multivariate data. Unlike the Euclidean distance previously used to study rhythmic variability in sperm whale coda production (Rendell and Whitehead, 2003a; Antunes *et al.*, 2011; Gero *et al.*, 2016) and other measures applied to recurrent rhythmic patterns in human music performances (Hellmer and Madison, 2015; Clayton, 2020), CDS can be calculated for sets of rhythms consisting of differing numbers of song units, allowing it to be used on themes made up of a range of PVs. Past studies on cyclic rhythms in human music have established that chronotonic distance correlates well with human judgements of perceptual dissimilarity (Guastavino *et al.*, 2009) and can be used to generate phylogenetic trees that agree well with conclusions based on historical musicology (Díaz-Báñez *et al.*, 2004; Toussaint, 2006). We cannot assume that these findings transfer to any assessments of rhythmic dissimilarity made by humpback whales due to perceptual and cognitive differences (e.g., humpbacks may not be able to perceive a beat). However, the fact that the chronotonic representation does not rest on human musical assumptions (such as the existence of a beat or common fast pulse) gives grounds for arguing that chronotonic distances are a good tool for investigating rhythmic variability in such cases.

I used Monte Carlo simulations of repeated phrases to investigate a technique combining chronotonic distance with the Mantel test to detect systematic rhythmic differences and unsystematic rhythmic variability. Although my simulations were for just two singers, the technique is readily scaled up to any number of singers. The main findings were that (1) unsystematic variability stemming from IOI jitter and PV diversity did not prevent the Mantel test being sensitive to systematic differences, although the sensitivity was lowered for sets consisting of mixtures of highly contrasting variants, (2) the existence of systematic difference did not substantially influence levels of intra-individual variability. This double dissociation means that the technique appears to be a promising means of assessing both relative levels of unsystematic variability and the existence of systematic differences among individuals. The simulations also showed that the technique is robust with regards to concerns regarding elevated levels of Type I errors in the Mantel test (SI §2.7.4). Further simulations would be needed to determine the dependence of the sensitivity of the technique on sample size.

Although these simulations were limited to sets of rhythms characteristic of humpback song, I would argue that this technique shows promise for the study of the variability of any repetitive complex (heterochronous) rhythms. It therefore supplements existing methods used for the study of expressive timing and microtiming deviations in music performance (Repp, 1990; Clarke, 1999; Goldberg, 2015; Hellmer and Madison, 2015; Clayton, 2020; Demos *et al.*, 2020). For the humpback song data analysed here (at the level of PV and theme) I found significant systematic rhythmic differences between individuals but no consistent discrepancy in WRV. That rhythmic differences are important in the individuation of human performers is a commonplace among musicians and fans, and has been demonstrated empirically in a number of studies (Prögler, 1995; Van Vugt *et al.*, 2013; Hellmer and Madison, 2015). The role of such differences may be equally important in humpbacks, a species which relies predominantly on sound for communication, potentially allowing individual recognition (Hafner *et al.*, 1979) and/or contributing to the cultural evolution of song forms.

2.6.4 Future Work

Having explored a range of visualizations and quantitative methods for studying rhythmic variability, and developed them for an exploratory analysis of humpback phrase rhythms, we can now move to specific questions of biological and musical interest. For example, in the study of the development or evolution of human and nonhuman animal rhythms in the laboratory (Saar and Mitra, 2008; Glaze and Troyer, 2013; Ravignani, Delgado, *et al.*, 2016), the methods could be used to track rhythmic consistency and aid phylogenetic investigations into song transmission. For humpback whales, the methods could complement ongoing work into the repetitiveness of song structure (Murray *et al.*, 2018; Allen *et al.*, 2019) and the spatiotemporal patterns of its evolution and revolution (Garland *et al.*, 2011, 2017; Allen *et al.*, 2018; Owen *et al.*, 2019; Zandberg *et al.*, 2021). In the study of auditory or melodic templates in nonhuman animal song and human music (Glaze and Troyer, 2006; Konishi, 2010; Kroher and Díaz-Báñez, 2019; Tyack, 2020), the methods described here would allow the rhythmic aspect of such templates to be examined, allowing researchers to ask questions about whether rhythm serves as an independent means of organizing or embedding information in vocalizations. This is often assumed in human music, and there is some evidence that it occurs in humpback whale song (Handel *et al.*, 2009). Finally, the rhythmic variability of the temporal structure of vocalizations has been shown to play an important role in sexual selection across a range of taxa (§2.1.3), but has yet to be investigated in humpback whales. In the next chapter I use a selection of the methods described here, applying them to a much larger database of song phrases, to ask whether rhythmic variability may have evolved as a signal or serve as a cue for individual quality and/or permit discrimination among individuals.

2.7 Supplementary Information

2.7.1 Derivation of the maximum value of mean anisochrony

The mean proportional deviation from isochrony, mean anisochrony (ANISO), for an IOI timeseries of duration DUR and number of song units NSU is given by (§2.4.2)

$$\overline{ANISO} = \frac{1}{NSU}\sum_{i=1}^{NSU}\left|\frac{IOI_i}{\overline{IOI}} - 1\right| = \frac{1}{NSU}\sum_{i=1}^{NSU}\left|\frac{IOI_i}{DUR/NSU} - 1\right|$$

This takes a minimum value (0) for isochronous rhythms, where all IOIs are equal to *DUR/NSU*. Its maximum value will be found as a single interval tends towards occupying the whole phrase duration, and the other intervals tend towards zero, i.e., where all intervals are as far as they can be from the isochronous value. In the limiting case, this can be expressed through taking $IOI_1=DUR$ and all other IOIs=0. Substituting into the equation above gives

$$max(\overline{ANISO}) = \frac{1}{NSU}\left(\left|\frac{IOI_1}{DUR/NSU} - 1\right| + \sum_{i=2}^{NSU}\left|\frac{IOI_i}{DUR/NSU} - 1\right|\right)$$

$$= \frac{1}{NSU}\left(\left|\frac{DUR}{DUR/NSU} - 1\right| + \sum_{i=2}^{NSU}\left|\frac{0}{DUR/NSU} - 1\right|\right)$$

$$= \frac{1}{NSU}\left(|NSU - 1| + \sum_{i=2}^{NSU}|0 - 1|\right)$$

$$= \frac{1}{NSU}(NSU - 1 + NSU - 1)$$

$$= \frac{2NSU - 2}{NSU}$$

$$= 2 - \frac{2}{NSU}$$

2.7.2 Dependence of Lempel-Ziv Complexity on Grid Size and Theme Duration

Simulations using the model described in the main text (§2.4.2) show that Lempel-Ziv Complexity (LZC) depends on the grid size chosen to represent the IOI vector (Figure 2.27). As the grid size increases the value of LZC increases (implying that more structure is found), and the precision of a single value decreases (confidence intervals become wider). The shape of the LZC vs IOI jitter curve also changes, so that there is little increase in complexity with jitter once the IOI jitter is comparable with grid size. This implies that if the LZC is to serve as an effective proxy for IOI jitter, the grid size chosen should exceed the level of IOI jitter.

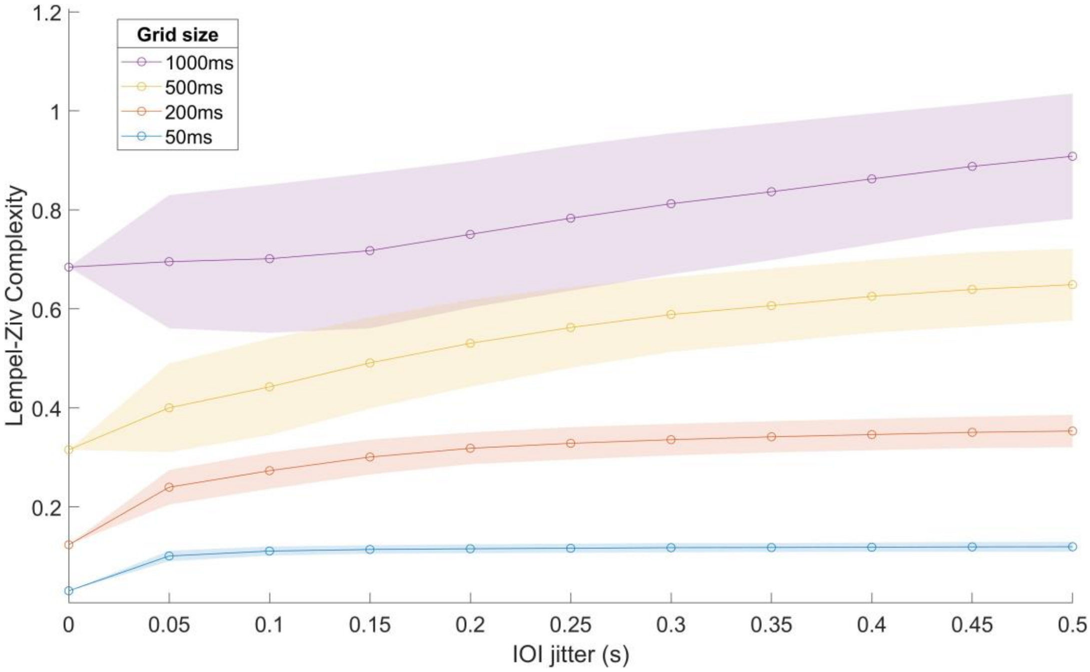


Figure 2.27 Simulated LZC dependence on IOI jitter, for Phrase Variant 4A (10 phrases, theme duration = 174 s). The mean and 2σ confidence intervals are plotted for a range of spacings of the grid used to generate the binary onset string (1000 model runs).

A further possible complication in the interpretation of LZC is its dependence on theme duration. Simulations showed that for short (< c.100 s) themes comprised of a single PV LZC depends strongly on theme duration (Figure 2.28), and that at shorter durations PV differences are greater and confidence limits wider. Observed values of LZC plotted against theme length

(Figure 2.29) confirm that there is both a degree of extra scatter at shorter theme durations and a general decrease with theme duration. As most theme durations in this dataset are under 100 s this somewhat restricts the use of LZC as an index of IOI jitter.

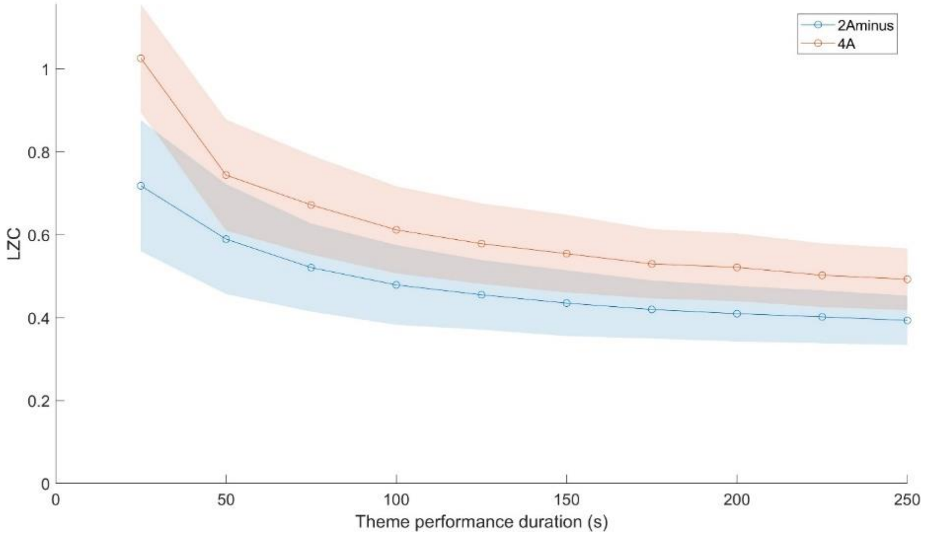


Figure 2.28 Simulations of Lempel-Ziv Complexity as a function of theme duration for repeated phrase rhythm templates based on PV 2A- and 4A. IOI jitter was set to 0.2 s. The model was run 1000 times and the resulting 95% confidence limits plotted as filled polygons. Grid size = 0.5 s.

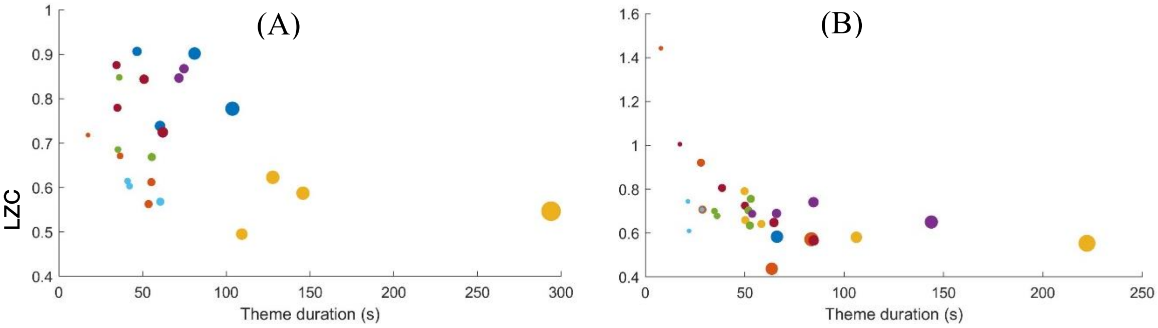


Figure 2.29 Observed LZC values at region level for (A) 190919 and (B) 190924. Each marker represents a single theme, colour-coded by theme and scaled to the number of phrases. Grid size = 0.5 s. As theme durations increase, LZC decreases as predicted by modelling of algorithm convergence.

2.7.3 Monte-Carlo Simulations: Assumptions and Phrase Variants

For Monte-Carlo simulations of indirect measures of intra-individual phrase variability (§2.4.2) and the comparison of inter- with intra-individual phrase variabilities (§2.5.1), PV templates were modelled using observed mean IOI vectors (Table 2.15). I assumed that Within-individual Rhythmic Variability (WRV) was driven by IOI jitter (see SI §2.7.5). For simplicity I made the further assumption, supported by an analysis of the variance in IOI versus mean IOI for single song unit positions (Figure 2.30), that IOI jitter was not dependent on IOI and was thus the same at all song unit positions. Future work could examine the effect of including a dependence on mean IOI.

Table 2.15 PVs used in Monte-Carlo simulations. PV labels and IOI vectors correspond to initial coding of the 190919 and 190924 song recordings.

PV	**NSU**	**Phrase Duration (s)**	**[IOI]**	**ANISO**	**ANISOS**
2A	5	18.2	[4.4 4 3.1 3.5 3.2]	0.123	0.077
2Aminus	3	9.1	[3.5 2.8 2.8]	0.103	0.077
3A	7	18.3	[2.77 2.6 2.17 1.79 2.93 2.39 3.61]	0.163	0.095
4A	6	17.4	[2.5 2.9 2.3 1.5 5.1 3.1]	0.276	0.166
4Asplit	7	17.4	[2.5 2.9 2.3 1.5 2.55 2.55 3.1]	0.135	0.079
4Aminus	5	17.4	[2.5 2.9 2.3 1.5 8.2]	0.543	0.339
4Aiso	6	17.4	[2.9 2.9 2.9 2.9 2.9 2.9]	0	0
5A	5	18.0	[3.8 2.7 5.0 2.0 4.5]	0.278	0.174
7B	3	10.5	[4.1 1.8 4.6]	0.324	0.243
7C	5	17.7	[3.6 2.8 4.8 3.5 3.0]	0.149	0.093

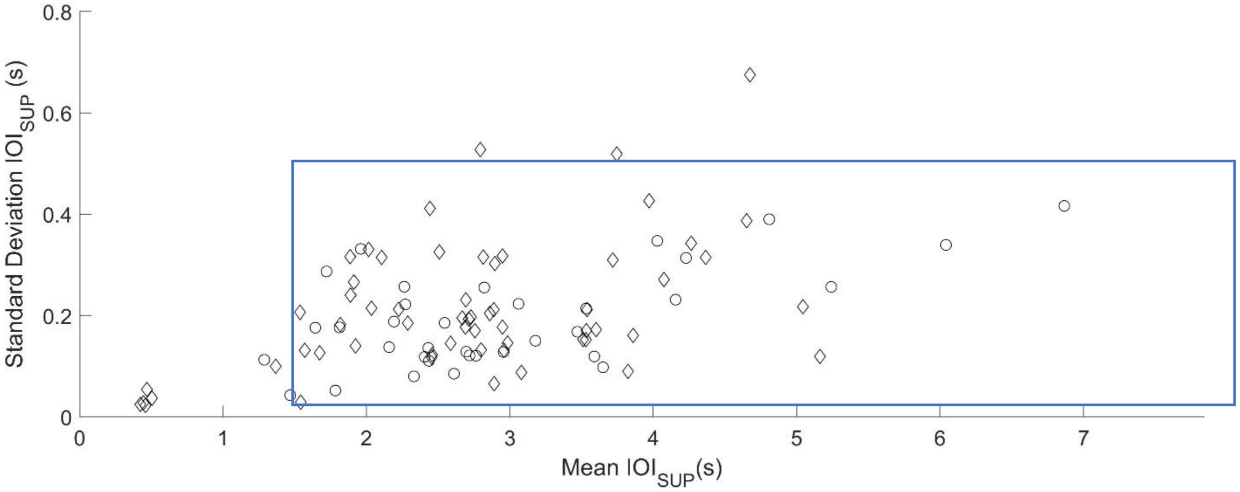


Figure 2.30 Dependence of the standard deviation of IOI for single song unit positions on the value of the mean IOI at that song unit position, for all phrase variants ($N \geq 3$) and both singers (190919 circles, 190924 diamonds). A single statistical outlier (standard deviation IOI$_{SUP}$ > 2 s) has been excluded. The distribution suggests that IOI jitter is approximately independent of mean IOI. The range of IOI values and mean IOI standard deviations used in Monte Carlo simulations is boxed.

2.7.4 Prevalence of Type I Errors in Mantel Test Outcomes

The prevalence of Type I errors was assessed by running simulations for individual singers using the same phrase rhythm template, with no difference in IOI jitter or systematic rhythmic differences. For chronotonic distance measures the proportion of model runs returning a *p*-value below the stipulated significance level was at the level expected from chance alone (Table 2.16). This provides justification for the use of the Mantel test against arguments that it may be subject to high levels of Type I error (Guillot and Rousset, 2013).

Table 2.16 Proportion of model runs for which the Mantel test (10,000 permutations) returned a positive correlation with a p-*value below the stipulated significance level, implying a rejection of the null hypothesis that inter-individual variability was no higher than intra-individual variability. Each model run compared phrase rhythm sets (N=10) from two singers, based on PV 4A (IOI jitter = 0.2 s). The model was run 100,000 times.*

	Observed Type I Error Rate		
Significance Level	CD	CDN	CDS
0.05	0.0501	0.0495	0.0503
0.01	0.0098	0.0100	0.0099
0.001	0.0011	0.0011	0.0011

2.7.5 Distinguishing Among Sources of Unsystematic Rhythmic Variability

Before carrying out the modelling described in this chapter I wanted to validate the assumption that unsystematic variability in phrase rhythms could be captured through IOI jitter. As song unit production is a many-component process, and humpback song is hierarchical, it may not be sufficient simply either to add random noise to onset position following Ravignani and Norton (2017) or to individual IOIs (as might be appropriate for vocalizations with simpler structures such as sperm whale codas). The generative model described here was built to distinguish among sources of timing variability which may arise from different components or at different stages of song unit production. I assumed that variability within each phrase rhythm is independent, i.e., that deviations from the template in one phrase do not affect the next. Three such components were considered in my model (Table 2.17):

(1) ‘Top-down’: independent variation at the phrase level roughly analogous to tempo shifts in human music, giving rise to proportional scaling of the phrase template (‘PD jitter’).

(2) ‘Bottom-up’: independent variation of individual IOIs (‘IOI jitter’). On its own this component results in increasingly large variability in the relative onset positions through the production of a phrase (analogous to a random walk), and hence also a variability in phrase duration.

(3) ‘Self-correction’: a monitoring system which compared the just-produced song unit position against the stored template, and makes changes to the next song unit to be produced to compensate for deviations from the stored template (ΔIOI). The strength of such self-correction was variable; at its strongest it prevented IOI jitter from propagating forwards through the phrase by producing the next song unit at the time it would have occurred had no previous deviation taken place.

Table 2.17 Equations describing the three components of the generative model, in which subscript 'i' denotes the phrase, 'm' the song unit position within each phrase, and '0' the phrase template. In each component i = 1,2,3,…,N_{phrase}. For top-down and bottom-up components m=1,2,3,…,NSU; for self-correction m = 2,3,….,NSU.

Top-down (PD jitter)	$[IOI_m]_i = [IOI_m]_0 \times \left(1 + \frac{\varepsilon_i}{PD}\right)$	$\varepsilon_i \sim N(0, \sigma_{PD})$
Bottom-up (IOI jitter)	$[IOI_m]_i = [IOI_m]_0 + [\varepsilon_m]_i$	$[\varepsilon_m]_i \sim N\left(0, \sigma_{IOI_m}\right)$
Self-correction	$IOI_{1,i} = IOI_{1,0} + \varepsilon_{1,i}$ $rOnset_{1,i} = IOI_{1,i}$ $\Delta IOI_{m,i} = rOnset_{m-1} - rOnset_{m-1,0}$ $[rOnset_m]_i = [rOnset_{m-1}]_i + \Delta IOI_{m,i} + IOI_{m,0} + [\varepsilon_m]_i$ $[IOI_m]_i = [rOnset_m]_i - [rOnset_{m-1}]_i$	$\varepsilon_{1,i} \sim N\left(0, \sigma_{IOI_1}\right)$ $[\varepsilon_m]_i \sim N\left(0, \sigma_{IOI_m}\right)$

The trick to distinguishing among these components relies on the insight that the variance in successive relative onset positions within a phrase evolves differently through the phrase depending on the relative magnitude of these three components. This is because the onset is given by the sum of IOIs up to that onset position, and hence the onset variance at a particular position is given by the sum of variances and covariances in the IOIs (Glaze and Troyer, 2006); coupled with the fact that the covariances differ for the different components. In the case of the bottom-up component the IOIs vary independently from one another and covariance is zero; for the top-down component, all IOIs are proportionately scaled, resulting in positive covariance between IOIs; and finally a self-correction system leads to negative covariance between adjacent IOIs.[10]

If we define a quantity RSD, the ratio of the standard deviation of the relative onset to the standard deviation of the relative onset assuming independent IOIs, we can use this to distinguish among the different components. RSD will be equal to 1 for a fully bottom-up model. Where proportional scaling dominates, the RSD will take values above 1, where self-correction dominates the RSD will take values below 1.

$$RSD_{rOnset j} = \frac{\sigma_{rOnset j}}{\sqrt{\left(\sum_{i=1}^{j}\left(\sigma_{IOI_i}\right)^2\right)}}$$

Plotting the confidence limits for different extreme versions of the model (top-down, bottom-up with and without self-correction) allowed us to assess whether the observed RSD was explained better by PD or IOI jitter, and whether self-correction is present. The observed mean IOI vector, standard deviations (in relative onset and IOI) and sample size for six PVs were used as inputs to the model to calculate the expected RSD as a function of song unit position (Figure 2.31). In these cases the observed pattern of variance was consistent with independent IOI jitter alone, although contributions from the other components cannot be ruled out. A limitation of the use of the RSD is that on its own it cannot allow the relative strengths of each component to be quantified. However, given that the observed variability of RSD is consistent with IOI jitter alone, this component alone has been used in all further modelling.

---

[10] In Glaze and Troyer's three-stage model of zebra finch song the final stage allows for variation in the motor periphery (and/or measurement uncertainties) that that results in a component of IOI variation with no impact on timing of the next onset. Their model builds on an earlier influential model that separated central variation from delays at the periphery (Wing and Kristofferson, 1973). Glaze and Troyer's third stage leads to negative covariance between IOIs within a phrase, similar to that arising from the third component of my model. In fact my model simulation for full self-correction is mathematically equivalent to allowing periphery jitter to be the only source of timing variability.

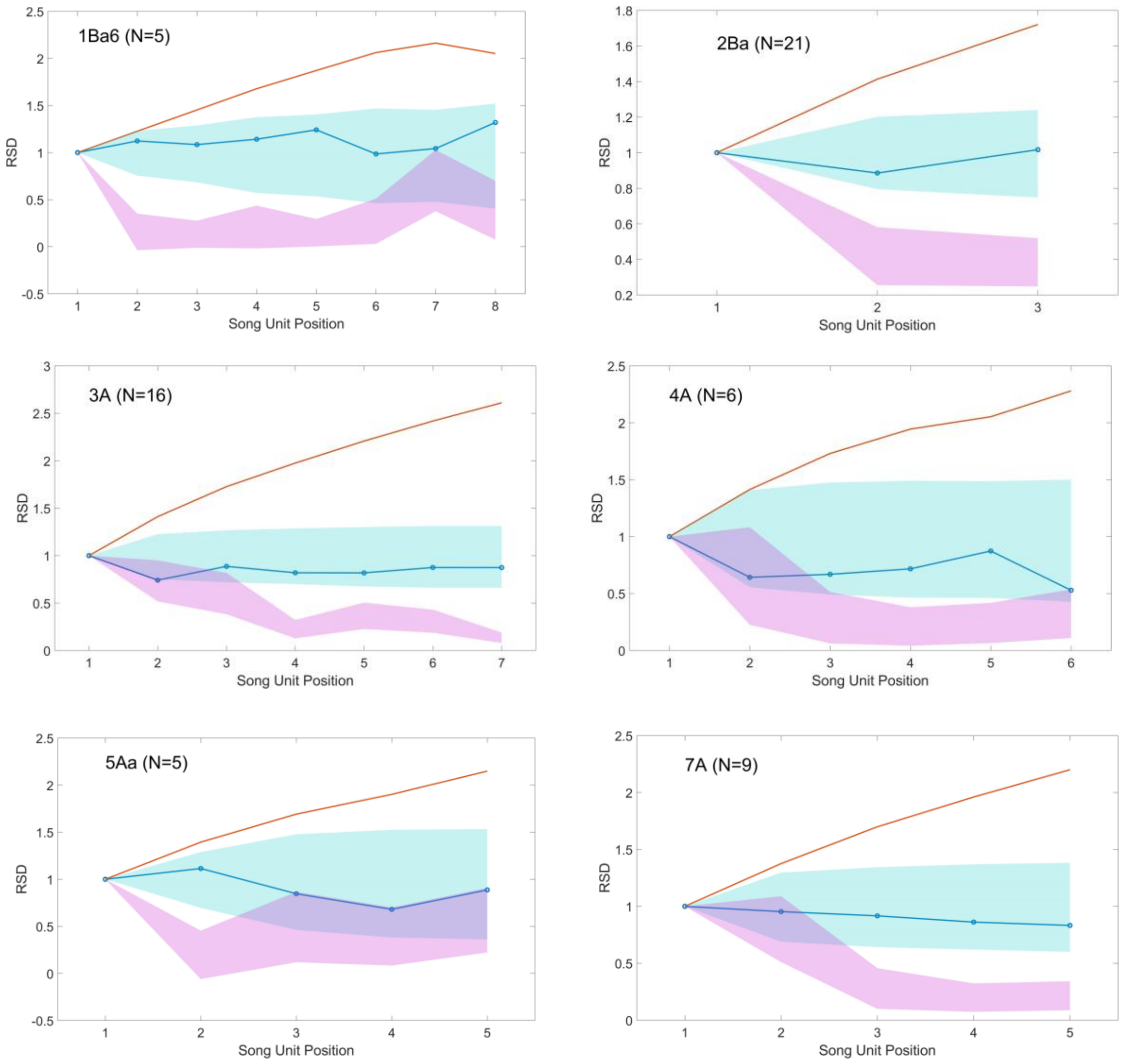


Figure 2.31 Observed variability in onset position could be explained by IOI jitter. Observations of RSD from Singer 190924 (blue line and markers) were consistent with simulated values (NRUN=1000) for which the source of variability was IOI jitter alone (pale blue region, 2σ limits). Further simulations for PD jitter alone (red line) or IOI jitter with full self-correction (purple region, 2σ limits) suggested that PD jitter or self-correction alone could not explain the observed variability.

2.7.6 Comparing Euclidean and Chronotonic Distances

A commonly used measure of distance is the Euclidean distance, which for rhythm vectors $[IOI]^a$ and $[IOI]^b$ is defined as the straight line distance between those vectors plotted as points in *NSU*-dimensional space, that is

$$ed_{ab} = \sqrt{\sum_{i=1}^{NSU} \left(IOI_i^a - IOI_i^b\right)^2}$$

for phrases containing *NSU* song units. Both vectors must have the same number of elements, so the Euclidean distance is most suited to comparing repeated phrases from the same PV. Where the only source of variability within a set of phrases is IOI jitter independent of the magnitude of the IOI, the Euclidean distance should be scaled to provide a measure of this jitter independent of the number of song units in the phrase and of rhythmic patterning:

$$eds_{ab} = ed_{ab}/\sqrt{NSU}$$

To validate the use of chronotonic distance for single PVs, I tested it against Euclidean distance, simulating sets of phrase rhythms for pairs of singers (§2.5.1) and running a Mantel test on correlation between distance and individual matrices. Simulations were run to explore the sensitivity of the Mantel test to discrepancies in IOI jitter between the two singers (MODEL RUN 1, Figure 2.32) and systematic differences in IOI pattern and phrase duration (MODEL RUN 3, Figure 2.33). In each case the calculated correlations and *p*-values returned by the Mantel test were highly similar, though with some evidence that the Euclidean distance is slightly more sensitive to changes in IOI pattern.

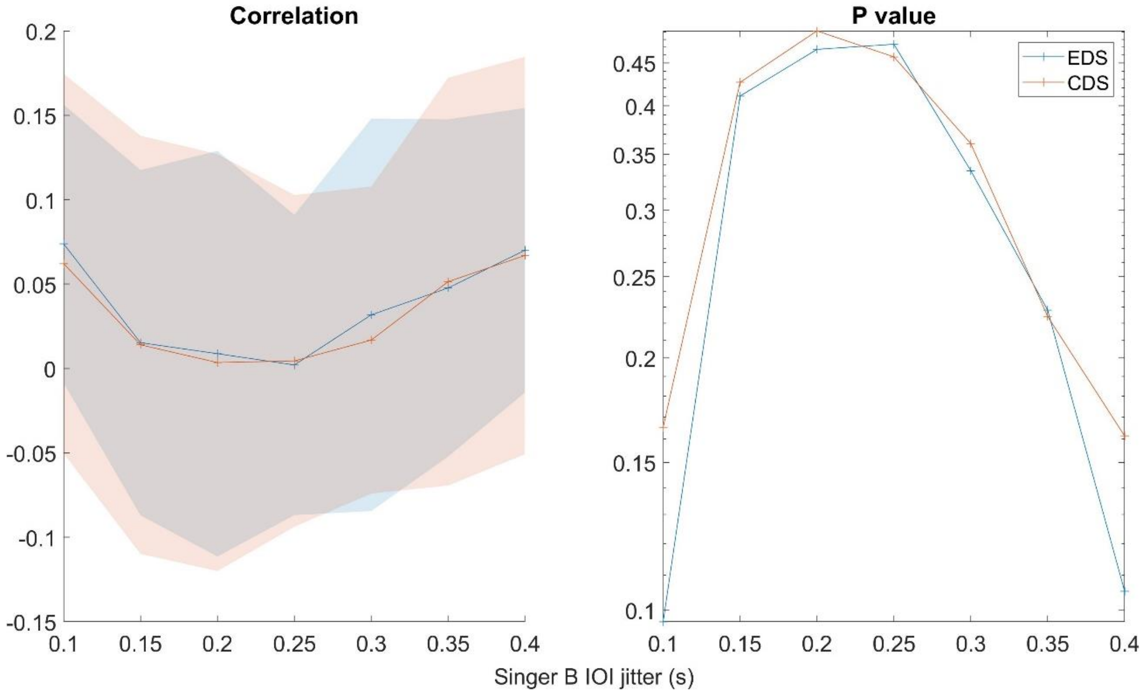


Figure 2.32 MODEL RUN 1. *Dependence of Mantel test correlation and* p-*value on IOI jitter of Singer B for scaled Euclidean Distance (blue) and scaled Chronotonic distance (red). Singer A's IOI jitter was held constant at 0.2 s. Plotted lines and shaded areas show the mean and 95% confidence intervals for 100 model runs. Both singers' phrase rhythm sets (N=10) were based on Phrase Variant 4A. Low correlation and high P values indicate that for both measures the Mantel test would give no evidence to reject the null hypothesis of comparable intra- and inter-individual rhythmic variabilities.*

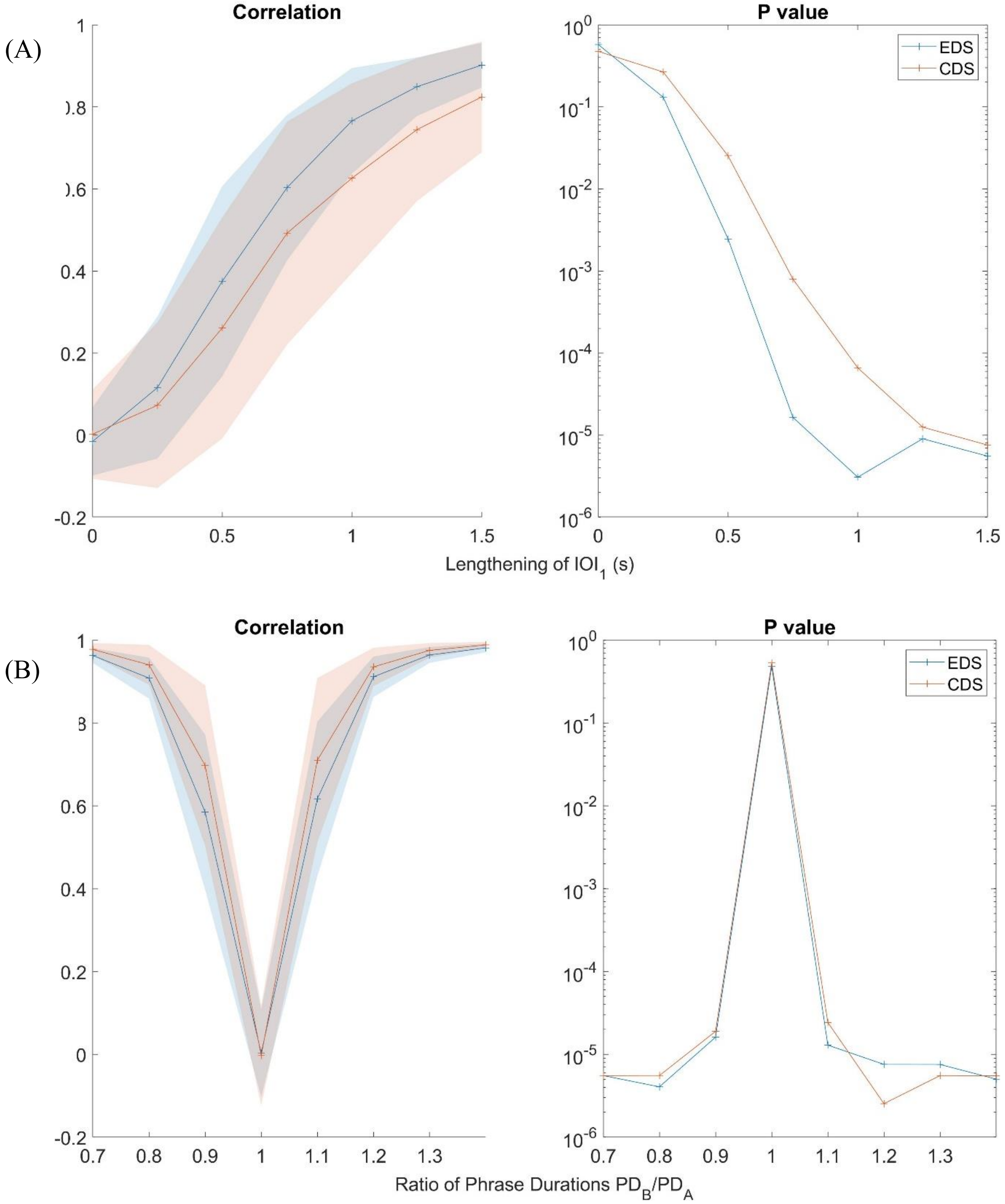


Figure 2.33 MODEL RUN 3. Dependence of Mantel Test correlation and p-*value on systematic differences between Singers A and B for scaled Euclidean Distance (blue) and scaled Chronotonic distance (red). Plotted lines and shaded areas show the mean and 95% confidence intervals for 100 model runs. Both singers' phrase rhythm sets (N=10) were based on Phrase Variant 4A and possess the same level of IOI jitter, but systematic differences are introduced to Singer B: (A) lengthening of* IOI_1 *and (B) variation in phrase duration ratio (lower plot). Relatively minor systematic differences resulted in high correlation and low P values for both measures, indicating that the Mantel Test would return strong evidence for rejecting the null hypothesis of comparable inter- and intra-individual rhythmic variabilities, in line with the model scenarios.*

The of dependence of Euclidean and chronotonic distance measures on IOI jitter was examined using Monte Carlo simulations (§2.4.2) for a range of observed PVs (Figure 2.34). The approximately linear relationship between distance and IOI jitter was least affected by the base rhythm (IOI pattern) for the scaled Euclidean distance, and among the chronotonic distance measures, the scaled chronotonic distance.

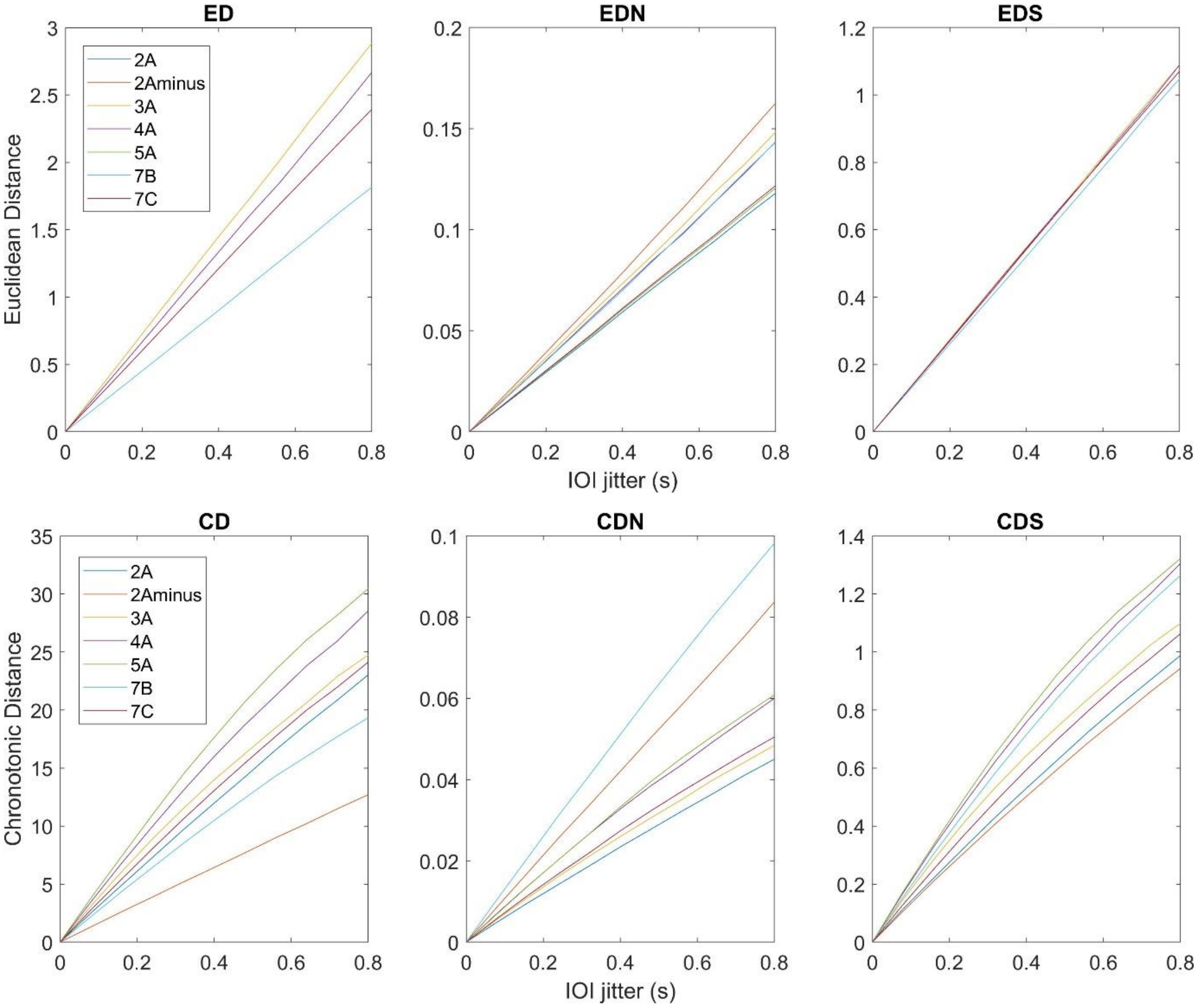


Figure 2.34 Dependence of mean pairwise Euclidean and chronotonic distance measures on IOI jitter, for range of humpback song phrase variants. EDN and CDN are for phrases normalized to phrase duration, EDS and CDS are scaled to compensate for differing numbers of song units. Distances were calculated for sets of 10 phrases and averaged over 1000 model runs.

2.7.7 Simulating the Rhythmic Variability of Phrase Variant Mixtures

To test the robustness of the findings for single PVs when extended to PV mixtures, Monte Carlo simulations of dependency of pairwise distances on IOI jitter (MODEL RUN 1, Figure 2.20) and systematic rhythmic differences (MODEL RUN 3, Figure 2.22) were repeated for PV mixtures (MODEL RUNS 4 and 5, Figure 2.35). Sensitivities were lowered slightly for PV mixtures compared with single PVs. For example, the IOI elongation required to give a significant correlation at $P < 0.001$ is 0.75 s for 4A alone, but 1.0 s in the 4A/4Asplit mixture. Intra-individual distances for the high contrast variants were found to have a slight dependency on the ratio of mean phrase durations, though such effects are small for differences within ± 20%, implying that in general discrepancies in intra-individual distance can still be used as a guide to differences in IOI jitter.

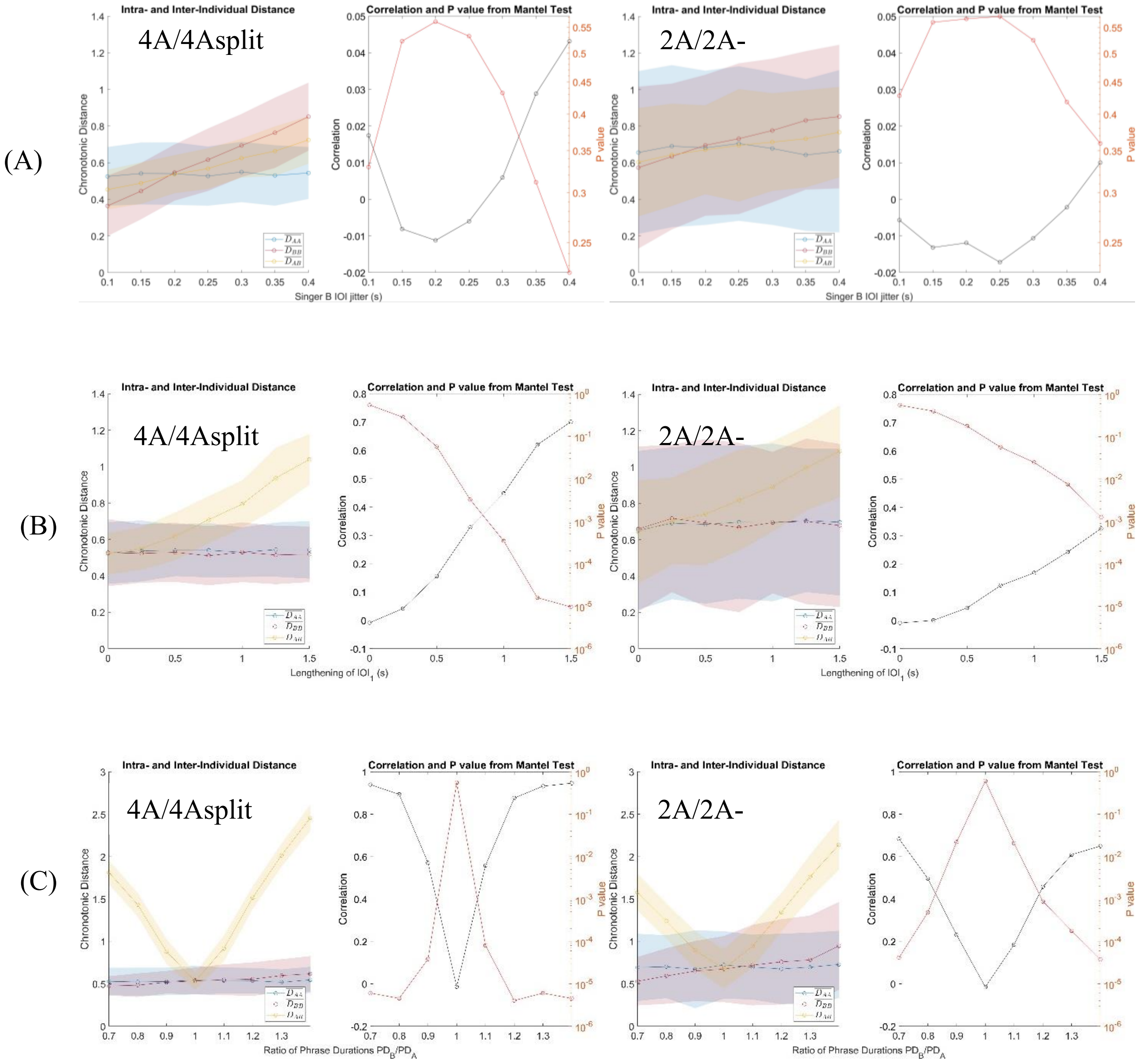


Figure 2.35 Dependence of scaled chronotonic distance on (A) IOI jitter of Singer B (Singer A's IOI jitter is held constant at 0.2 s), (B) IOI_1 elongation, (C) mean Phrase Duration ratio, with correlation and P values returned by the Mantel test. Plotted lines and shaded areas show the mean and 95% confidence intervals for 100 model runs. Each singer's theme contained a mix of two PVs (4A/4Asplit or 2A/2A-), N=10, with phrase variant diversity set to 0.3.

Chapter 3 Humpback Whales Sing Shared Song Phrases with Equal Rhythmic Consistency but Individually Distinctive Rhythms

Abstract. Timing consistency (precision) in music and song has been demonstrated to be pleasing to the listener, as for the musical rhythms of jazz, funk and samba in humans, or to correlate with singer quality, as for singing rates in the great tit (*Parus major*). Moreover, individually distinctive rhythmic patterns may communicate identity, as demonstrated in the tambourine dove (*Turtur tympanistria*). In humpback whales (*Megaptera novaeangliae*), it is well-established that a given population shares a common song repertoire consisting of a stereotyped sequence of themes. In each theme, phrases are repeated with small individual variation, leading to song marked by the steady slow pulse of its phrase duration, long-recognized as a highly consistent measure of song timing. However, the consistency with which humpback singers produce the fixed rhythmic patterns known as 'phrase variants' is unknown. In this study I analysed the rhythmic structure of humpback whale song phrases by measuring the inter-onset intervals between 15,908 song units from 13 hours of song produced by ten singers in waters off Mo'orea, French Polynesia (September-November 2019). Using multiple regression (linear mixed-effects models) and multivariate distance methods[11] I quantified the rhythmic variability within and among individual singers in order (1) to measure within-individual levels of rhythmic consistency and (2) to assess for individually distinctive rhythms, i.e., statistically significant rhythmic differences among individual singers in performance of the same phrase variants. My analysis was unable to distinguish individuals through the level of within-individual rhythmic consistency with which they sing specific phrases. This makes it unlikely that rhythmic consistency can serve to signal quality. However, I did find evidence

---

[11] Linear mixed-effects models allowed the statistical significance of correlations between a 'response variable' (here, rhythmic variability) and multiple 'effects' (here, Singer and phrase variant) to be established. Multivariate distance methods were used to calculate the rhythmic similarity of pairs of phrase rhythms through a comparison of IOI values at each song unit position (§2.4.2).

consistent with the hypothesis that individual whales sing shared song phrases with their own distinctive rhythms. This means that it is possible that rhythms could be used for individual identification, although I did not test whether songs could be assigned, with confidence, to a particular singer. Finally, I defined a measure of 'rhythmic individuality' which quantified how different a singer's rhythms were to population means, and was able to reject the hypothesis that rhythmic individuality varied randomly across phrase variants. This supports the interpretation that individuals who have markedly distinct rhythms in one phrase variant are also distinctive in others.

3.1 Introduction

3.1.1 Repetition in Acoustic Communication

Repetitive rhythms are ubiquitous across the animal kingdom in signals used for acoustic communication and display. In humans, rhythmic repetition tends to distinguish music from language (Fitch, 2006; Huron, 2007; Margulis, 2014) and is one of few musical features to be truly cross-cultural (Savage *et al.*, 2015). Single rhythmic motifs can run through extended compositions (e.g., Ravel's Bolero) and unify entire repertoires (e.g., all sambas). For humans and nonhumans, the precise repetition of rhythmic patterns underlies observed 'statistical universals' such as the presence of categorical rhythms (Roeske *et al.*, 2020; De Gregorio *et al.*, 2021) and restricted sets of inter-onset intervals (IOIs, Table 2.1) (Savage *et al.*, 2015; Ravignani *et al.*, 2019; Schneider and Mercado III, 2019; Helble *et al.*, 2020) including isochrony (Table 2.1) (Watkins *et al.*, 1987; Ravignani and Madison, 2017; Burchardt and Knörnschild, 2020). As a communication strategy, rhythmic repetition is displayed by many vocal species, where sound patterns may be as simple as a single sound type emitted at regular intervals, or as complex as a diverse repertoire of multi-element phrases. Such redundancy in the signal aids its detection and correct identification (Bradbury and Vehrencamp, 2011, p. 305), and may be increased in noisy environments, as shown by chaffinches (*Fringilla coelebs*) that live near waterfalls (Brumm and Slater, 2006) and in the response to high winds of king penguins (*Aptenodytes patagonicus*, Lengagne *et al.*, 1999). Repetition aids memorization and learning in humans (Margulis, 2014) and is assumed to do so in other vocal learners including humpback whales (*Megaptera novaeangliae,* Handel *et al.*, 2009; Schneider and Mercado III, 2019).

Among cetaceans, repetitive sound patterns are found both in calls expressing group identity, such as sperm whale (*Physeter macrocephalus*) codas (Whitehead *et al.*, 1998), and in the mysticete vocalizations, known as song, for which a range of functions has been proposed including long-range contact, sexual advertisement and individual identification (Dudzinski and Gregg, 2018). In humpback whales, the elaborate song is sung only by males (Glockner, 1983), predominantly sexually mature (Herman *et al.*, 2013). This suggests that song may have been produced through sexual selection (Garland *et al.*, 2017), but which of many possible functions it fulfils in the mating system, such as individual mate attraction or the mediation of male-male interactions, remains unclear (Herman, 2017). Mysticete song varies in complexity and diversity, from the singlet, doublet and triplet patterns produced by fin whales (*Balaenoptera physalus*, Širović *et al.*, 2017), through blue whale (*Balaenoptera musculus*) phrases comprising fixed arrangements of a few sound types (Stafford et al., 2011), to the multiphonic compositions of minke whales (*Balaenoptera acutorostrata,* Gedamke et al., 2001) and the large song repertoires of bowhead whales (*Balaena mysticetus,* Stafford et al., 2008). But it is the songs of humpback whales that have received the most attention, standing out on account of their nested hierarchical structure (Payne and McVay, 1971). Sounds judged to be continuous to the human ear ('units') are arranged in 'phrases', phrases are repeated to form 'themes', with a stereotyped population-wide sequence of themes known as the 'song'. Humpback songs are sung with immediate repetition, with the resulting 'song sessions' documented to last for up to 22 hours (Winn and Winn, 1978). Singing usually continues during the periodic surfacing to breathe that occurs every 5–20 minutes (Chu, 1988). In humpback song, then, repetition occurs on many levels. Yet it is the phrase that has been described as "the salient element of repetition" (Cholewiak *et al.*, 2013) for its higher consistency of duration than theme or song (Thompson, 1981; Frumhoff, 1983; Payne *et al.*, 1983; Cerchio, 1993; Cerchio *et al.*, 2001). The repetition of phrases within a theme typically occurs with some variation in unit type and number, leading to the grouping of phrases into types and variants (Cholewiak *et al.*, 2013; Murray *et al.*, 2018). Phrase variants (PVs), that is fixed sequences of unit types, are repeated within and between individuals (Murray *et al.*, 2018), with highly stereotyped duration evidenced by coefficients of variability of only a few percent (Thompson, 1981; Cerchio, 1993; Cerchio *et al.*, 2001). The classification of units, phrases and themes, a form of sequential analysis (Ravignani and Norton, 2017), has enabled the discovery of an unusual song feature: gradual or sudden changes at multiple levels of song hierarchy resulting in the complete replacement of song types within populations (Payne and Payne, 1985). Unit and phrase characteristics change progressively from one year to the next (Payne *et al.*, 1983; Payne and Payne, 1985; Cerchio *et al.*, 2001), with changes in individual singers (Guinee *et al.*, 1983) tracking those in the population. Such gradual cultural evolution is known to be supplemented in the South Pacific by 'cultural revolutions' involving rapid population-wide song replacement (Noad *et al.*, 2000; Garland *et al.*, 2011). The observed spatiotemporal patterns of song use (Garland *et al.*, 2011, 2015; Garland, Gedamke, *et al.*, 2013; Owen *et al.*, 2019; Schulze *et al.*, 2022), backed by modelling studies (Mcloughlin *et al.*, 2018; Zandberg *et al.*, 2021) and detailed mechanistic analysis (Garland *et al.*, 2017) are consistent with the hypothesis that distinct breeding populations learn songs from each other on feeding grounds and migration routes. Such vocal learning in which population-wide conformity is coupled with an apparent novelty bias leads to intra-generational patterns of change resembling human fashions for new musical genres (Acerbi and Bentley, 2014; Mauch *et al.*, 2015; Youngblood, 2019), also described in village indigobirds (*Vidua chalybeate*, Payne, 1985), yellow-rumped caciques (*Cacicus cela vitellinus*, Trainer, 1989) and corn buntings (*Emberiza (Miliaria) calandra*, McGregor, Holland and Shepherd, 1997).

3.1.2 Rhythms of Acoustic Communication

In the study of acoustic communication and display, the analysis of temporal features including rhythm (Ravignani and Norton, 2017) can usefully supplement the sequential analysis that enabled the discoveries discussed above, as demonstrated by work on the temporal patterning of sperm whale codas (Weilgart and Whitehead, 1993) that led to the discovery of an oceanspanning social structure founded on ‘vocal clans’ (Whitehead *et al.*, 1998; Rendell and Whitehead, 2003b). In humans, rhythm has long been regarded as a foundational aspect of musical structure (Sachs, 1953; Cooper and Meyer, 1960; Hasty, 1997; Toussaint, 2020). Operationalized as sequences of IOIs (McAuley, 2010), the rhythms of both music and speech have been the target of extensive empirical research (e.g., Grabe and Low, 2002; Daniele and Patel, 2013; Ravignani, Delgado, *et al.*, 2016; Jacoby and McDermott, 2017; Ravignani and Norton, 2017; Clayton, 2020; Toussaint, 2020), with studies on human capacities for rhythm production, perception and entrainment forming a recent focus for investigators interested in the evolutionary origins of human musicality (Patel, 2006, 2014; Patel *et al.*, 2009; Honing, 2012; Ravignani *et al.*, 2014; Ravignani and Madison, 2017; Kotz *et al.*, 2018). Here, comparative work is essential, and research into rhythm in birds (Patel *et al.*, 2009; Norton and Scharff, 2016; ten Cate and Spierings, 2019; Roeske *et al.*, 2020; Xing *et al.*, 2022), primates (Large and Gray, 2015; Hattori and Tomonaga, 2020; De Gregorio *et al.*, 2021) and pinnipeds (Cook *et al.*, 2013; Ravignani, Fitch, *et al.*, 2016) has revealed intriguing parallels with human music. For example, the existence of categorical rhythms in human music and nonhuman vocalizations (Ravignani, Delgado, *et al.*, 2016; Jacoby and McDermott, 2017; Roeske *et al.*, 2020; De Gregorio *et al.*, 2021) may reflect the operation of shared constraints on vocal learning.

Work on the rhythms of mysticete song has been largely restricted to that in fin whales, whose song was initially considered to consist of an isochronous pulse (Watkins et al., 1987), but is now known to contain patterns of two or three sound types with up to two distinct IOIs (Širović et al., 2017; Helble et al., 2020; Wood and Širović, 2022). Although humpback song researchers have occasionally examined variability in song, theme and phrase durations in conjunction with sequential analysis (Thompson, 1981; Frumhoff, 1983; Payne et al., 1983; Payne and Payne, 1985; Cerchio, 1993; Cerchio et al., 2001), and the independence of song unit durations and inter-unit silences has been demonstrated for a small dataset (Handel et al., 2009), rhythm understood as patterns of IOIs has been the target of very limited research. An initial exploration of rhythm and tempo by Schneider and Mercado III (2019) revealed repetitive isochronous and heterochronous (Table 2.1) rhythms but omitted song coding and hence did not explicitly address the striking regularity of phrase rhythms that are so evident in spectrograms of song (Figure 1.1, Figure 4.2) and have frequently attracted the attention of cetologists (Payne and McVay, 1971; Thompson, 1981; Handel *et al.*, 2009).

3.1.3 Repetition and Variability

Consistency (Precision) in Temporal Features of Song

Although phrases are repeated so consistently by humpbacks that their rhythms have been occasionally described qualitatively as “monotonous” (Payne and McVay, 1971; Thompson, 1981), they of course display the between- and within-individual variability[12] that may be a clue to the functions and production mechanisms of song. Achieving precise repetitions of signals may pose performance challenges for any species (Byers, 2007; Botero *et al.*, 2009; de Kort *et al.*, 2009) and hence high song consistency or precision (low within-individual variability) may serve as an honest signal of a high quality singer (Podos and Nowicki, 2004; Sakata and Vehrencamp, 2012). As Botero and de Kort have put it, “consistency cannot be exaggerated” (2013, p. 283). Field studies across multiple taxa have found the consistency of selected temporal features of vocalizations to be correlated with singer quality, for example in the great tit (*Parus major*, Lambrechts and Dhondt, 1986) and Antarctic leopard seal (*Hydrurga leptonyx*, Rogers, 2017), and reproductive success, as in the blue tit (*Parus caeruleus*, Poesel *et al.*, 2001), chestnut-sided warbler (*Dendroica pensylvanica*, Byers, 2007) and rock hyrax (*Procavia capensis*, Demartsev *et al.*, 2017). However, listeners may not always choose higher consistency, as demonstrated in the field cricket (*Gryllus bimaculatus*, Simmons and Zuk,

---

[12] In Chapter 2 I also use the equivalent terms inter- and intra-individual variability.

1992) and African frog (*Hyperolius Marmoratus*, Jennions *et al.*, 1995). This matters because mate choice plays an important role in the evolution of sexual display (Rosenthal, 2017; Prum, 2018; Ryan, 2018).

Individual Distinctiveness

The measurement of between- and within-individual variability of signal features can also reveal individually distinctive signal characteristics, which may permit individual recognition when maintained consistently over time (Sandoval *et al.*, 2014). This possibility has been established in the case of temporal features of acoustic displays in a wide range of animals, including the ruffed grouse (*Bonasa umbellus*, Garcia *et al.*, 2012), the tambourine dove (*Turtur tympanistria*, Osiejuk *et al.*, 2019), and most relevant here, in selected codas of the sperm whale (Antunes *et al.*, 2011; Gero *et al.*, 2016; Oliveira *et al.*, 2016). For some species, such as the emperor penguin (*Aptenodytes forsteri*), temporal features had to be combined with spectral features to permit individual discrimination (Robisson *et al.*, 1993). Whether or not the possibility of recognition is actually exploited by the species in question must be tested through playback experiments: in the corncrake (*Crex crex*), for example, individually distinctive vocal rhythms did not elicit behaviour indicative of individual recognition (Budka and Osiejuk, 2014). A recent playback study on northern elephant seals (*Mirounga angustirostris*, Mathevon *et al.*, 2017), by contrast, demonstrated for the first time that individual identity can be encoded in temporal features of mammalian vocalizations. In the case of humpback whales, acoustic identification of individuals may seem unlikely to be reliable where individual song is subject to continual evolution within the constraints of group conformity (Herman, 2017, p. 1805). However, although rhythmic variability has not been investigated, recent assessments of the compositional variability of repetitive patterns of song units have found high levels of between-individual variability for some parts of singers' repertoires (Murray *et al.*, 2018; Lamoni *et al.*, 2023), suggesting that individual differences may advertise singer quality or carry information on individual identity.

Rhythmic Consistency and Individual Distinctiveness in Human Performance

In studies of microtiming deviations in the performance of human rhythms, it has become common to distinguish between within-individual variability, referred to as random or unsystematic variability, and the recurrent or systematic rhythmic differences that distinguish individuals or styles (such as swung quavers in jazz) (Kvifte, 2007; Davies *et al.*, 2013; Hellmer and Madison, 2015). Human audiences, measured in some musical genres under laboratory conditions, have exhibited preferences for low within-individual variability (high rhythmic consistency or precision) (Davies *et al.*, 2013; Frühauf *et al.*, 2013; Datseris *et al.*, 2019), although other factors can be more important (Senn *et al.*, 2018). In addition, systematic rhythmic differences among individual performers ('individual distinctiveness') have been found in contexts ranging from the performance of scales, where "pianistic fingerprints" allowed recognition (Van Vugt *et al.*, 2013), to the degree of swing in jazz performance (Corcoran and Frieler, 2021). In general, musicians performing what are taken to be the same rhythms differ both systematically and unsystematically in their performances: they produce individually distinctive patterns of IOIs with different levels of consistency or precision (Hellmer and Madison, 2015). This is likely to be true for nonhuman animal repetitions of shared acoustic patterns, but standard methods comparing between- and within-individual variability do not conceptually separate rhythmic consistency and individual distinctiveness. Between-individual variability can arise from both individual distinctiveness (systematic differences) *and* different levels of rhythmic consistency (SI §3.6.9).

In this chapter I combine standard humpback song coding methods with the analysis of IOIs to examine variability in the rhythms of shared PVs, i.e., those PVs sung by multiple individual humpback whales, investigating both rhythmic consistency and individual distinctiveness. I draw from and build on methods developed in Chapter 2. First, I measure the levels of within-individual rhythmic variability (WRV) for all shared PVs in each individual singer's repertoire. This assesses an individual singer's rhythmic consistency, or precision, for specific PVs. To determine whether WRV is driven by differences among singers, or differences among PVs, I ask whether individual singers significantly differ from the grand mean across all singers (H1), and whether specific PVs significantly differ from the grand mean across all PVs (H2). If individual singers are significantly more or less precise (lower or higher WRV) than the grand mean, for a range of PVs, then this could be a sign of a quality index based on generalized timing ability. Conversely, if specific PVs are significantly more or less precise than others, for a range of singers, then this may be associated with specific challenges to precise execution posed by those variants. Regarding systematic rhythmic differences, I then test the hypotheses that singers sing shared PVs with individually distinctive rhythms (H3), and whether singers tend to be consistently far away from population mean rhythms for all PVs in their repertoires (H4). If shared PVs are individually rhythmically distinctive, then they may play a role in individual recognition, although I do not test here whether songs may be confidently assigned to a particular singer. Finally, if singers consistently deviate from population mean rhythms, this may be evidence for a novelty bias operating in the transmission of song.

3.2 Methods

3.2.1 Humpback Song Recording and Coding

Songs were recorded from September-November 2019 off the east coast of Mo'orea, French Polynesia. Two OceanInstruments SoundTrap ST300STD recorders were tethered on a single cable to the ocean floor at depths of c. 27 m and c. 28 m in ~30 m of water, approximately 800 m from the shore and 150 m outside the reef, at 17°33' S, 149°46' W. The lower device recorded continuously for 17 days, while the upper recorder was on a duty cycle (30 mins/2 hours), resulting in over 800 hours of stored audio.

On the basis of signal to noise ratio, I selected ten continuous song sessions (Table 3.1) for analysis, excluding sessions recorded less than 12 hours apart to reduce the possibility that recordings came from the same singer. Although this cannot guarantee that no singer was sampled more than once, satellite-tagging of humpback whales on diverse other wintering grounds indicates that individuals move tens of kilometres per day (Guzman *et al.*, 2013; Andriolo *et al.*, 2014; Dulau *et al.*, 2017), including those engaged on "slow, area-restricted movements" (Guzman and Felix, 2017). For example, the median daily distance travelled by 12 humpbacks tagged at Réunion August–October 2013 was 65 km, with a range of 23–102 km (Dulau *et al.*, 2017, data from Table 1). Thus it seems highly unlikely that the same singer would stay within range of my recorders for more than 12 hours, a conclusion bolstered by the observed individual residence times for Mo'orea of three days or less (Gales *et al.*, 2011, p. 15). Following previous studies (Garland *et al.*, 2011, 2012, 2017; Garland, Noad, *et al.*, 2013), I manually analysed spectrograms in Raven 1.6 Pro (K. Lisa Yang Centre for Conservation Bioacoustics at the Cornell Lab of Ornithology, 2021) to code the song sessions down to unit level. Based on acoustic features such as frequency contour, timbre and duration, I classified each sound to a unit type, delineating recurring patterns of units as phrases. I assigned phrases to themes on the basis of shared unit types, labelling themes 1, 2, 3, etc. in the order encountered in analysis. Within themes, I classified phrases into phrase type and phrase variant (PV) on the basis of structural similarities, with each type (nA, nB, etc.) containing a number of variants (nAa, nAb, nA-, nA’, etc.) to capture fine-scale variability. I followed the rule (Cholewiak *et al.*, 2013; Murray *et al.*, 2018) that all phrases in a variant should have identical unit type sequences. For example, phrases in Theme 7 consisted of ‘moans’ alternating with higher frequency ‘cries’ and ‘whistles’. I classed all phrases containing three moans as Phrase Type 7A (including variants PV 7A°, 7A’, 7A”), and all phrases containing two moans as Phrase Type 7B (including variants PV 7B°, 7B’, 7B+). I provide further details of the classification scheme in Appendix B. Here I focus on rhythmic variability within and among PVs.

Table 3.1 Coded song comprised ten sessions recorded at least 12 hours apart. The number of song cycles was established by examining the theme sequence and counting a new cycle each time any theme repeated. For example, theme sequence ***3****412756****3****12****1****75634****1*** *would contain three cycles.*

Recording Date (yyyy.mm.dd)	Singer ID	Recorder	Coded Duration	# Song Cycles	Theme							
					1	2	3	4	5	6	7	8
2019.09.19	190919	lower	2h18	20	●	●	●	●	●	●	●	
2019.09.21	190921	lower	1h14	11	●	●	●	●	●	●	●	●
2019.09.22	190922	lower	3h00	27	●	●	●	●	●	●	●	●
2019.09.24	190924	lower	2h10	24	●	●	●	●	●	●	●	
2019.09.25	190925	lower	2h30	65	●	●			●		●	●
2019.10.04	191004	upper	0h30	3	●	●	●		●	●	●	
2019.10.06	191006	upper	0h24	1				●				●
2019.10.08	191008	upper	0h30	1						●		●
2019.11.02	191102	upper	0h10	2		●	●				●	
2019.11.15	191115	upper	0h10	1	●		●		●	●	●	

3.2.2 Extracting Song Unit Timing and Constructing Inter-Onset Interval Vectors

I wrote custom MATLAB version 9.10.0 (R2021a) (The MathWorks Inc., 2021) code for the semi-automated extraction of unit onsets from the song recordings, applying a double-threshold detection algorithm (Pace *et al.*, 2010) to the signal amplitude-squared time series to identify candidate units (§2.2.2). I manually inspected spectrograms to check for and remove false positive detections, and flagged phrases containing false negatives (units missed by the detector operating with on/off thresholds of 3.01/1.76 dB) to exclude them from analysis (SI §3.6.1). Following previous studies, I also eliminated transitional phrases, i.e., phrases occurring at theme boundaries which contained unit types from previous and subsequent themes (Payne *et al.*, 1983), because these are known to be atypically variable (Garland *et al.*, 2017), and phrases including 'surface ratchets' (Winn and Winn, 1978) because these were often attenuated and hence harder to categorize (Garland *et al.*, 2017). The elimination of these phrases (c.2.5 % of the total) made no difference to the main findings. Finally, I calculated an inter-onset interval (IOI) vector for all phrases (Figure 3.1).

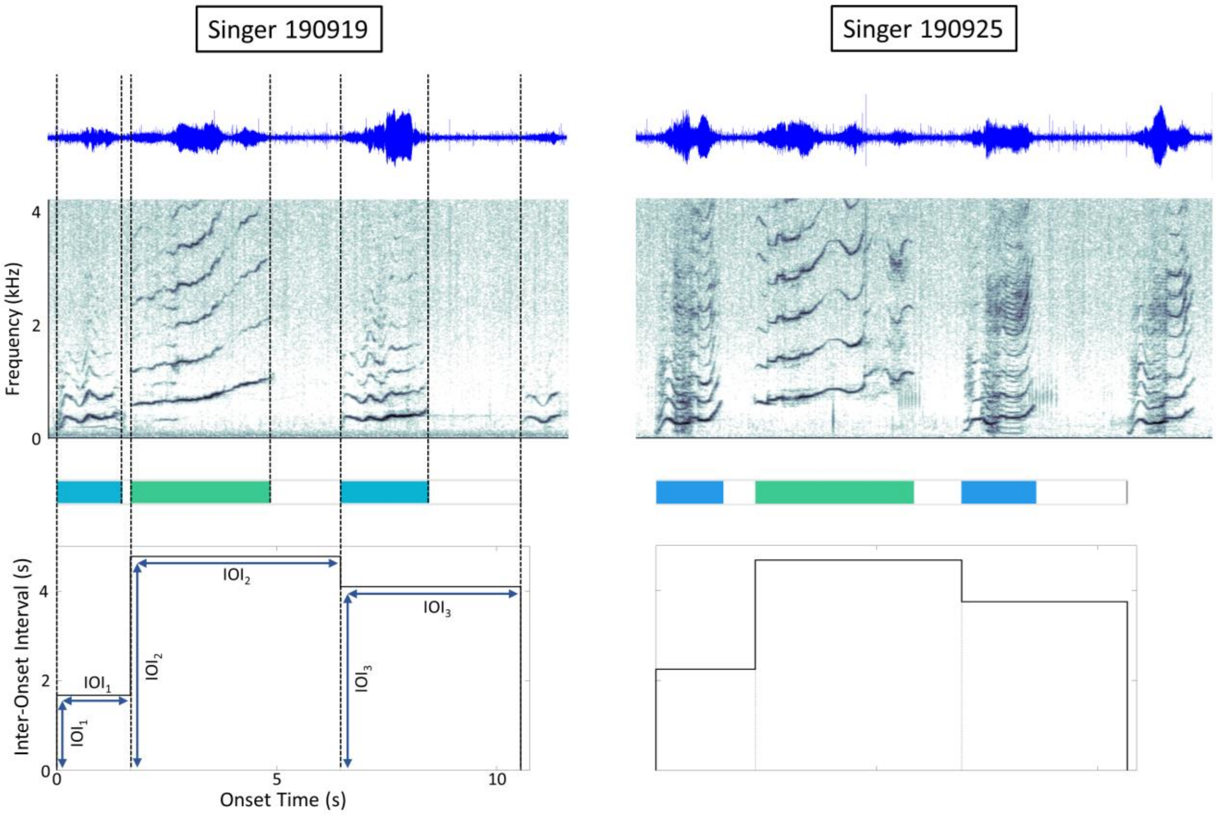


Figure 3.1 Phrases of the same shared variant (7B°) from different singers vary in waveform and spectral detail but have similar rhythms. The song units shown in the waveforms and spectrograms include the first unit of the following phrase. The rhythm of each phrase was characterized with an IOI vector $[IOI_1, IOI_2, IOI_3]$*, which measures the durations between unit onsets and can be readily compared by plotting IOI against onset time in 'chronotonic diagrams' (Toussaint, 2006) (lower panels). The 'phrase raster plots' (centre panels) reveal how each IOI divides into sound and silence, with units colour-coded by peak frequency. Dashed vertical lines show the start and end of each unit determined by the threshold detector.*

3.2.3 Measuring Rhythmic Variability

For each individual singer's set of phrases belonging to a specific PV (an 'observation'), I calculated the within-individual rhythmic variability (WRV) as the mean across all song unit positions of the sample standard deviation (σ_{IOI}^{i}) of the IOI at each song unit position i (Equation 3.1, in which NSU_{α} is the number of song units in PV α).

Equation 3.1

$$WRV(A,\alpha) = \overline{\sigma_{IOI}}(A,\alpha) = \frac{1}{NSU_{\alpha}} \sum_{i=1}^{NSU_{\alpha}} \sigma_{IOI}^{i}\,(A,\alpha)$$

WRV, which varies from one singer to another, measures the rhythmic consistency (precision) with which an individual singer repeats a PV. But singers may sing a PV with the same rhythmic consistency and yet differ on account of systematic rhythmic differences (Figure 3.2) resulting from individually distinctive rhythms. In order to compare among different singers' renditions of a shared PV, I distinguished between tests (H1, H2) that directly compare levels of WRV (§3.2.4), and those (H3, H4) that are sensitive to systematic rhythmic differences (§3.2.5).

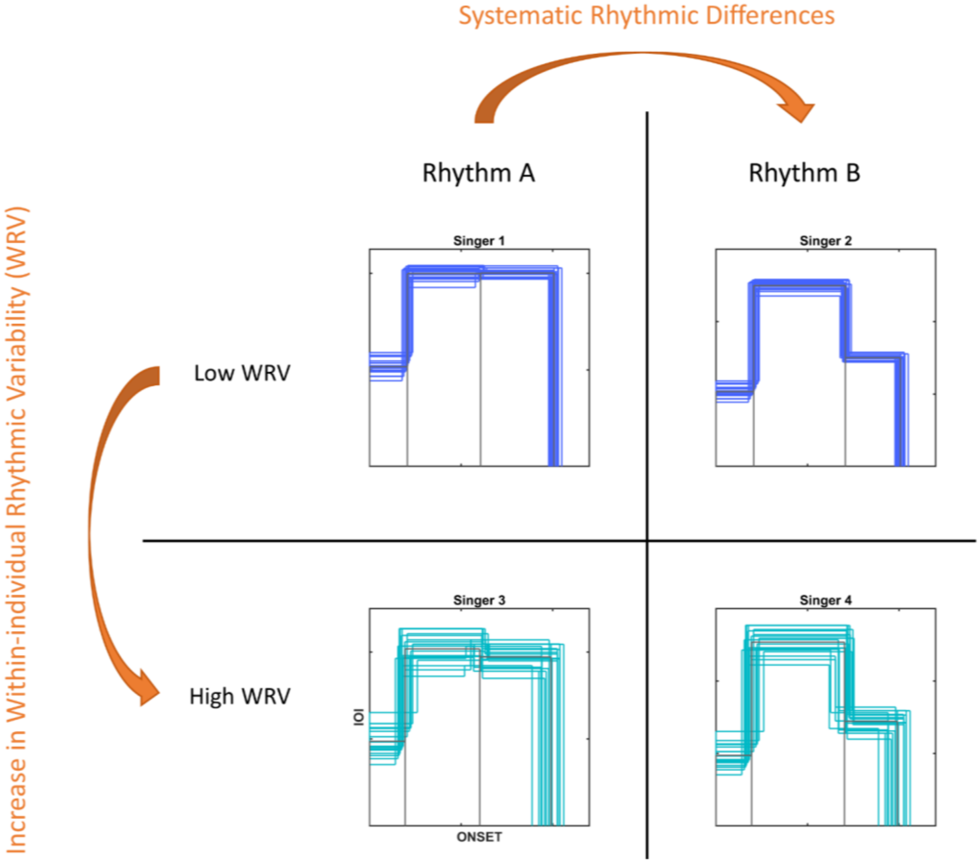


Figure 3.2 Different levels of within-individual rhythmic variability (WRV) are distinguished from individually distinctive rhythms A and B. In these simulated phrase rhythms, Gaussian 'IOI jitter' was applied to each song unit position for 20 repetitions of a phrase template similar to PV 7B°. Singers 1 and 2 share the same level of WRV ($\sigma_{IOI} = 0.1\ s$*), whereas Singers 3 and 4 share a higher level of WRV (*$\sigma_{IOI} = 0.3\ s$*). The phrase rhythms of Singers 1 and 3, based on Rhythm A (* $IOI = [2\ 4\ 4]$*), are systematically different to those produced by Singers 2 and 4, based on Rhythm B (*$IOI = [2\ 5\ 3]$*). I used linear mixed-effects modelling (§3.2.4) to assess between-individual differences in WRV. This method is insensitive to systematic rhythmic differences. I tested for such individually distinctive rhythms using multivariate pairwise distance methods (§3.2.5).*

3.2.4 How does Within-individual Rhythmic Variability depend on Singer and Phrase Variant?

As a measure of rhythmic precision, WRV may be determined both by properties of the singer (such as a generalized timing ability) and by performance challenges offered by the specific PV. To distinguish these contributions, I tested two hypotheses. Firstly (H1) that individual singers consistently vary in WRV across PVs within their repertoires (that is, some singers are rhythmically looser for all their PVs, and other singers are tighter), and secondly (H2) that specific PVs consistently vary in WRV across all singers singing that PV (that is, for some PVs individual singers have loose rhythm, and for other PVs they are tighter). For H1, the null was that individual singers do not differ significantly from the unweighted grand mean (of repertoire means). For H2, the null was that specific PVs do not differ significantly from the unweighted grand mean (of population means).

The observed song data can be thought of as the outcome of a factorial experimental design in which multiple singers sang a range of PVs, with an unbalanced dataset (not all singers sang all PVs). Under these circumstances the statistical significance of differences to the grand mean were assessed with linear mixed-effects modelling. Mixed-effects modelling allowed sources of variation (among singers and among PVs) to be taken into account simultaneously; furthermore, missing observations would not substantially reduce sample sizes (Brown, 2021). I tested H1 in MATLAB version 9.10.0 (R2021a) (The MathWorks Inc., 2021) using the function **fitlme** with the singer set as the fixed effect μ and PV as a random effect u (*Equation 3.2*). This assessed H1 for *this* set of singers, and assumed that the PVs represented in their song sessions were drawn at random from a larger population-wide repertoire (Brown, 2021). The converse procedure was followed for H2 (Equation 3.3). All fits were weighted with bootstrapped standard errors (SI §3.6.6).

Equation 3.2

$$WRV(A,\alpha) \sim \mu_o(H1) + \mu_{singer(A)} + u_{PV(\alpha)} + e(A,\alpha)$$

Equation 3.3

$$WRV(A,\alpha) \sim \mu_o(H2) + \mu_{PV(\alpha)} + u_{singer(A)} + e(A,\alpha)$$

To obtain regression coefficients μ_{singer} (μ_{PV}) that represented the difference between that singer (PV) and the grand mean (the intercept μ_o), dummy variables for the categorical effects were created using effects coding (Aiken and West, 1991, pp. 127–128; The MathWorks Inc., 2023): each level of the categorical variable was coded with ones or zeros, except for the final category, which was coded with -1's. For k levels of the fixed effect, the model fitted k-1 coefficients (i.e., for each level except for the final level) and calculated associated p-values against the null hypothesis that the coefficient is zero. I obtained the coefficient and p-value for this final level by repeating the fit with the level order reversed. For hypothesis H1, p-values for individual singers indicated whether those individuals differed significantly in WRV to the grand mean, i.e., were relatively high or low in WRV for all PVs in their repertoires. For hypothesis H2, p-values for PVs indicated whether those PVs differed significantly in WRV to the grand mean, i.e., were consistently higher or lower in WRV for all singers singing those PVs.

My data contained recordings of 136 PVs, of which 56 could be said to be shared in the sense that they were sung at least once by a minimum of two singers. To avoid underestimating within-individual variability at small sample sizes, for each singer I selected only those PVs sung at least five times in its song session. This resulted in 282 singer-PV observations dispersed across 51 PVs and 10 singers (range: 5–168 phrases, see SI §3.6.2). The outcome of the modelling was robust to the value of this initial criterion (SI §3.6.5). To test hypotheses H1 and H2, I used different subsets of these observations to make optimal use of the available song data (SI §3.6.2). In testing for differences among singers (H1), I used the criterion that each singer's repertoire should contain at least three PVs. To test for differences among PVs (H2), the criterion was reversed such that each PV be sung by at least three singers. In addition, both hypotheses were tested on a common subset (DS_{restr}) selected using the more restrictive requirements that each singer had a repertoire of at least three PVs, and that each of these PVs was shared by at least three of those singers. The outcome of the restrictions was to exclude several PVs which were not widely shared and to exclude two singers whose observations were drawn from a small number of PVs (Table 3.2).

Table 3.2 'Relaxed' (DS_{H1}, DS_{H2}) and 'restricted' (DS_{restr}) datasets drawn from full song data (282 observations containing at least five phrases). Relaxed datasets were constructed using the criterion that the relevant fixed effect contain at least three groups. The restricted dataset was constructed applying the same criterion to both singers and PVs.

Dataset	**Min. Singers per PV**	**Min. PVs per Singer**	**$N_{observation}$**	**N_{phrase}**	**N_{singer}**	**N_{PV}**	**PVs per Singer mean (std)**	**Singers per PV mean (std)**
DS_{H1}	1	3	96	1949	7	48	15.83 (4.40)	2.72 (1.37)
DS_{H2}	3	1	49	1183	8	13	14.47 (4.82)	3.94 (0.94)
DS_{restr}	3	3	42	1128	6	11	7.00 (2.00)	3.82 (0.60)

3.2.5 Rhythmic Individuality: Do Singers have their own Distinctive Phrase Rhythms?

To investigate whether singers performed shared PVs with individually distinctive rhythms, I measured rhythmic differences between all pairs of phrase rhythms. For a given PV, I compared the rhythmic differences between phrases sung by different singers with those sung by the same singer. According to my rhythmic individuality hypothesis (H3), the between-individual rhythmic differences were expected to be larger than the within-individual differences. The null hypothesis was that between- and within-individual rhythmic differences would be statistically indistinguishable. Rejecting the null for H3 for a given PV would be consistent with the hypothesis of rhythmic individuality for that PV only. To restrict comparisons to PVs which were well-balanced in terms of contributions from a range of singers I applied the same criteria used above to select subset DS_{H2}. To account for multiple tests, significance was Bonferroni corrected to 0.05/13.

To measure the rhythmic differences between two phrase rhythms, I calculated the Euclidean distance between IOI vectors, i.e., the length of the shortest line between them in multidimensional space, divided by the square root of the number of song unit positions to allow for meaningful comparison among PVs (§2.7.6). For each PV, I constructed a distance matrix $\overline{\overline{D}}$ for all pairwise distances d_{ij} for the full set of N phrase performances.

Equation 3.4

$$\overline{\overline{D}} = \begin{pmatrix} d_{11} & \cdots & d_{1j} & \cdots & d_{1N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{i1} & \cdots & d_{ij} & \cdots & d_{iN} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{N1} & \cdots & d_{Nj} & \cdots & d_{NN} \end{pmatrix}$$

A parallel ‘singer difference’ matrix $\overline{\overline{F}}$ was constructed containing zeros to identify phrases from the same singer, and ones for phrases from different singers. I then calculated the Pearson correlation between $\overline{\overline{D}}$ and $\overline{\overline{F}}$, and followed Mantel’s permutation method (§2.5.2, Mantel, 1967) to estimate the probability (*p*-value) that this correlation could have occurred through chance, using 10,000 permutations. Intuitively, the more individually distinctive each singer’s rhythm, the greater the difference between between-individual and within-individual distances, and hence the larger the correlation coefficient between matrices. However, between-individual distances also rise as the amount of within-individual variability increases, and large discrepancies in WRV could potentially lead to significant between-singer differences. In initial explorations of song data (§2.5), I showed using Monte Carlo simulations that for typical PVs and observed levels of WRV, statistically significant correlations were unlikely to arise from between-singer discrepancies in WRV alone. On the other hand, significant correlations did result from systematic rhythmic differences in simulated IOI vectors, allowing observed correlations to be interpreted as arising from systematic rhythmic differences. However, further simulations based on the wider range of singer WRVs encountered in the full dataset have showed that small (< 0.15) but statistically significant ($P < 0.05$) correlations can arise in the absence of systematic rhythmic difference (SI §3.6.9), and thus that the p-value alone from the Mantel test does not allow us to distinguish between systematic rhythmic difference and WRV discrepancy. To overcome this difficulty, I ran Monte Carlo simulations (100 model runs) based on the observed population mean rhythm and individual levels of WRV for each singer-PV observation, in order to estimate confidence limits for the correlation value for a scenario in which there were no systematic rhythmic differences between individuals. Where correlation values retrieved from the observed phrase data lay above these confidence limits, I took them to indicate significant systematic rhythmic differences.

When the Mantel tests allowed me to reject the null hypothesis for H3, I investigated a further hypothesis (H4), testing whether between-singer systematic rhythmic differences could be generalized across PVs, e.g., whether singers tending to stand out as rhythmically distinct for one PV would do so for others. The null was that singers whose mean phrase rhythms are significantly above or below the population mean for one PV are no more likely to differ significantly from the population mean for another PV than expected by chance. As with the testing of H1 (Equation 3.2), linear mixed-effects modelling was used with singer as fixed effect and PV as random effect, but with 'rhythmic individuality' as the response variable. My measure of rhythmic individuality $RI(A, \alpha)$ was based on the difference between a singer's mean phrase rhythm, or rhythm template $[\overline{IOI}](A, \alpha)$ and the population mean phrase rhythm. This difference, representing the distinctiveness of Singer A's rhythm for PV α, was ratioed to the population mean WRV to account for variations in WRV among PVs (Equation 3.5, in which $N_{singer}(\alpha)$ stands for the number of singers singing PV α).

Equation 3.5

$$RI(A,\alpha) = rms\left([\overline{IOI}](A,\alpha) - \frac{\sum_{A=1}^{N_{singer}(\alpha)} [\overline{IOI}](A,\alpha)}{N_{singer}(\alpha)} \right) \Big/ \left(\frac{\sum_{A=1}^{N_{singer}(\alpha)} WRV\,(A,\alpha)}{N_{singer}(\alpha)} \right)$$

RI thus acts as an individual-level measure of rhythmic distinctiveness, complementing the population-level measure given by the correlation between pairwise rhythm distance and singer difference matrices. Indeed, my simulations have shown that for typical song data this correlation is approximately linearly dependent on RI (SI §3.6.9).

3.3 Results

3.3.1 Overview of Song Structure and Phrase Variant Timing

Song coding revealed a population-wide song made up of eight themes, with transition probability analysis suggesting wide variation in sequence order and/or frequent omission of themes, as is typical with most humpback songs (Figure 3.3). Seven themes were sung by at least 50% of the singers (SI §3.6.2), the exception (Theme 8) containing low and long pulsed song units known as ‘surface ratchets’ (Winn and Winn, 1978). Theme 3 stood out as containing an unusually high number of phrases per theme repetition (7.3, compared with a mean of 3.0 across all other themes). Themes 1, 2 and 7, whose phrases shared structural features, were particularly variable in ordering, with 2-1-7 occurring almost as frequently as 1-2-7 (30 vs 40 occurrences). Moreover, these themes also demonstrated occasional reversals such as 1-2-1 and 2-1-2. Although most singers produced a majority of the themes, the songs of Singers 191006 and 191008 were limited to material from a single theme.

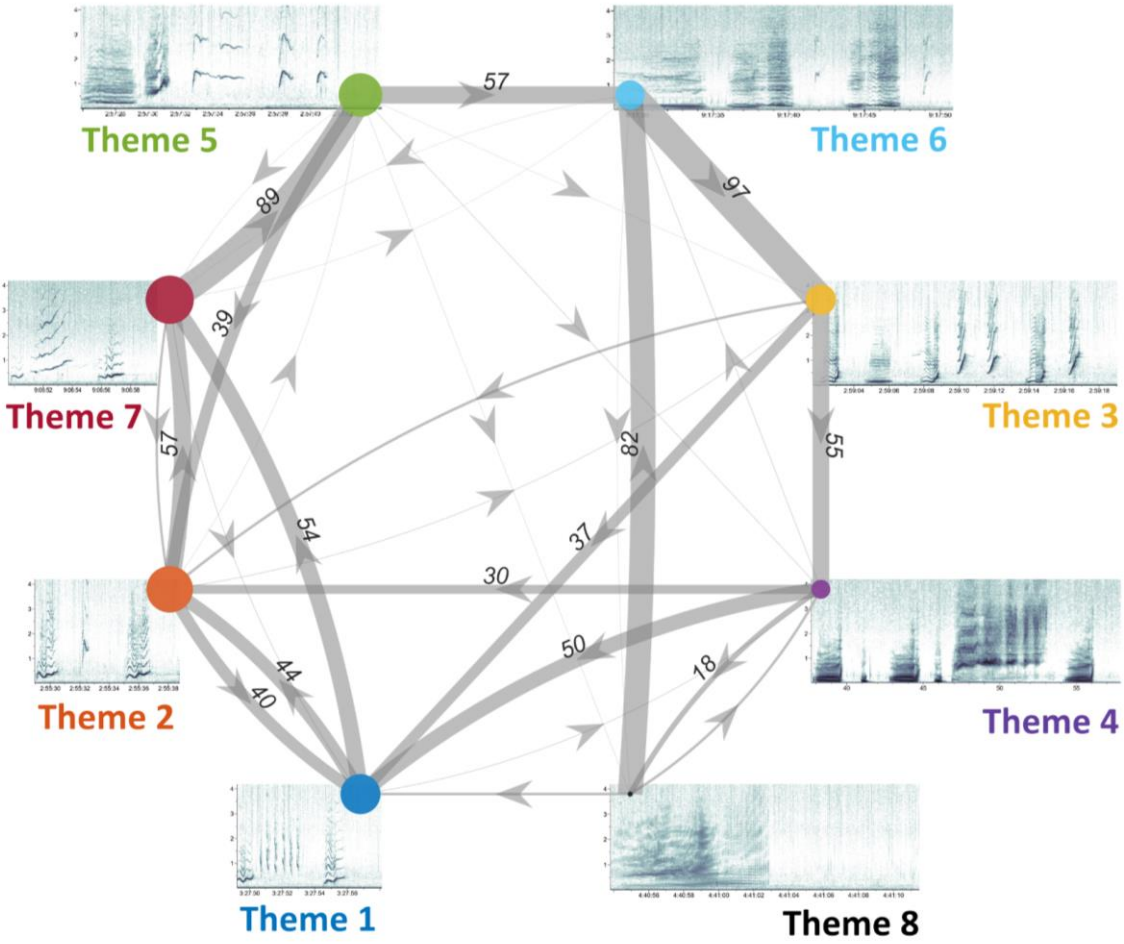


Figure 3.3 Population song structure (10 singers) revealing multiple paths through the eight themes. In this theme transition network, the proportion of transitions is indicated by edge thickness (labelled %), with infrequent transitions (< 10%) unlabelled. Node marker size is scaled by number of theme performances in the full dataset. Spectrograms (frequency range 0–4 kHz, phrase durations c.10–25 s) display characteristic PVs for each theme. The numbering of themes is arbitrary, representing the order in which they were encountered during analysis.

Each theme comprised the repetition of a number of phrase variants (PVs), with some themes dominated by one or two variants and others much more diverse (SI §3.6.2). In my analysis, I aggregated the phrase repetitions of specific PVs for each singer: these sets of phrases displayed consistent rhythms (Figure 3.4), with population mean WRV of 0.30 (0.10) seconds across all singer-PV observations in the restricted dataset (coefficient of variation 9.7 (2.6) %). To compare my data against previous studies, I calculated the sample standard deviation of the phrase durations for each observation as a measure of within-individual phrase duration variability. Across all observations the population mean of this measure was 0.54 (0.30) seconds (coefficient of variation of 3.3 (1.5) %), consistent with literature values (Frumhoff, 1983; Payne *et al.*, 1983; Cerchio, 1993; Cerchio *et al.*, 2001). Short-term trends in IOI and/or phrase duration, that in some cases continued through intervening phrases of other variants, can also be seen, for example, by the gradual shortening and lengthening of the phrases for Singer 1909019. In the longer continuous recordings (2–3 hours), there were also occasional short periods of song marked by strikingly low and high variabilities. This demonstrated the value of using such lengthy recordings in the analysis of variability to generate informative samples.

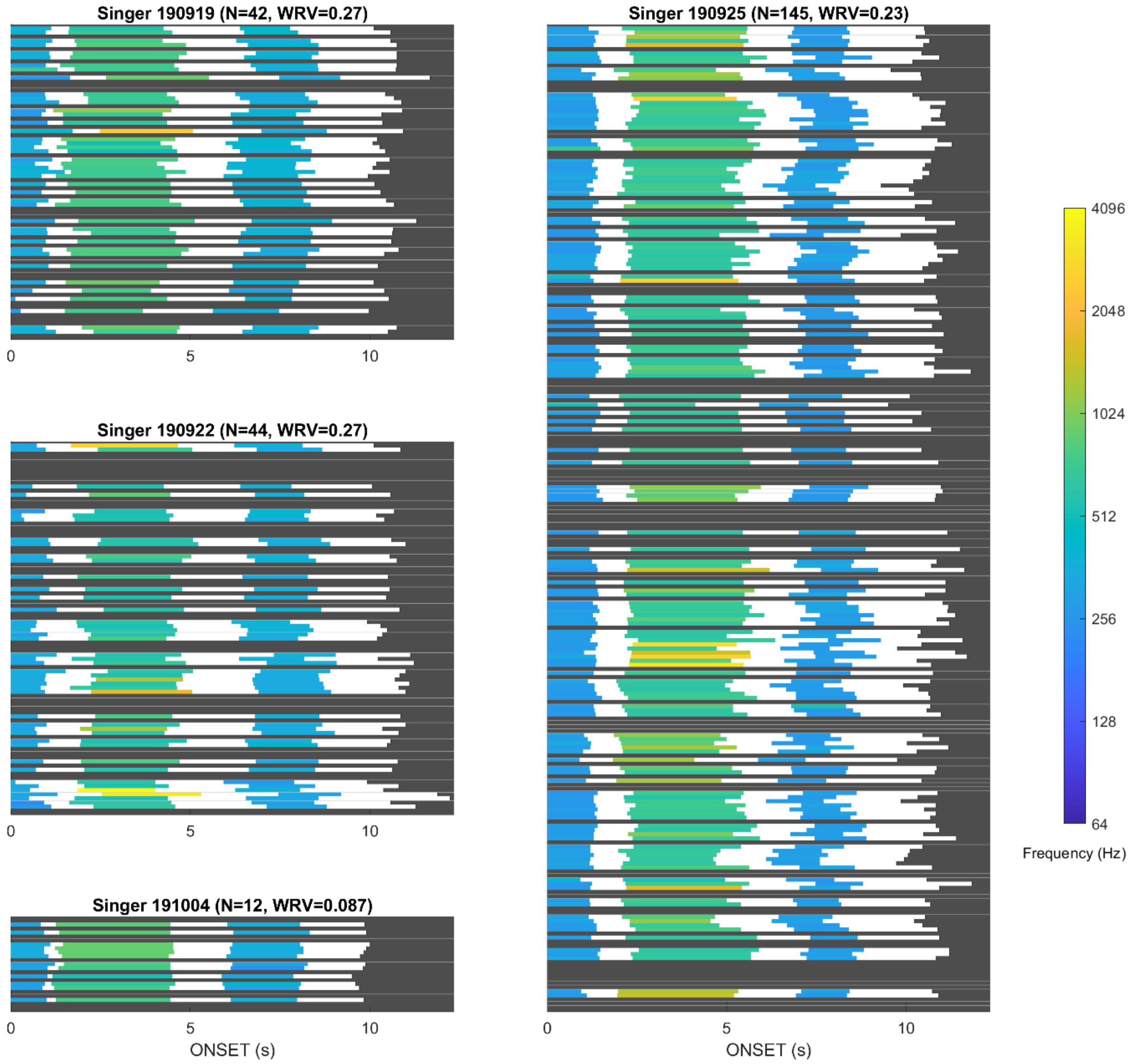


Figure 3.4 Phrase rhythms vary within and between singers, for the same PV. In each raster plot, a singer's renditions of PV 7B° are plotted horizontally, with the time series running top to bottom like a musical score. Each song unit is shown as a coloured rectangle, coded by peak frequency and followed by white inter-unit silences. Dark grey strips represent alternative PVs in Theme 7, removed for clarity (for full time series, see SI 3.6.2). Thin grey lines indicate where material belonging to other themes has been excised. Values of WRV measure the within-individual rhythmic variability for each observation.

3.3.2 Phrase Rhythm Variability Within and Between Individuals

The phrase rhythms for specific PVs varied both within and among individual singers. In the following sections I present the full results of hypothesis testing on both within-individual (§3.3.3) and between-individual (§3.3.4) rhythmic variability but here I show the details of phrase rhythms for a typical subset of observations (Figure 3.5). For these observations, selected for illustration on the basis of comparable sample sizes and to ensure overlap between singers and PVs, we see that WRV appears to vary more consistently from one PV to another than from one singer to another, an impression confirmed by linear mixed-effects modelling (§3.3.3). Systematic rhythmic differences among singers are also apparent, some quite pronounced, such as the variation in the relative size of the third and fourth IOIs in PV 7C°. Such systematic differences were found to be statistically significant for most PVs (§3.3.4), and furthermore, individual singers tended to be more or less distinctive with respect to the population across all of the PVs in their repertoire. I note here that the systematic differences in rhythm from one PV to another were clearly evident and were not tested for in my analysis.

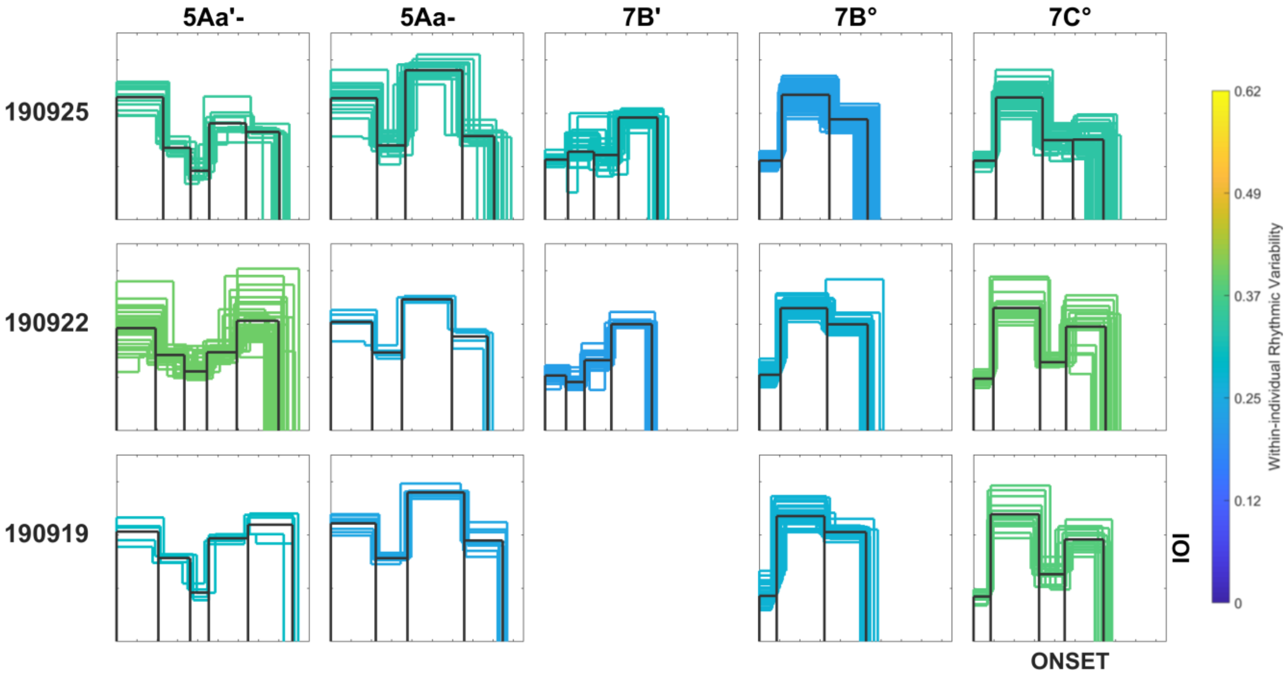


Figure 3.5 The phrase rhythms of humpback whale singers are individually distinctive but do not display consistent precision across PVs. In these typical observations, repetitions of specific PVs by three singers can be compared by reading down a column. Each chronotonic diagram (§2.3) presents individual (colour-coded by WRV) and mean (dark grey) phrase rhythms with one rectangle per song unit. IOI is plotted against onset time on a common axis (maxONSET=19s, maxIOI=7s). Singer 190919, PV 7B' is excluded due to the small sample size ($N_{phrase} = 3$). Levels of WRV suggest that individuals are not consistently rhythmically precise across this subset of PVs. For example, Singer 190922 is more variable than 190925 for 5Aa'-, 7B° and 7C°, but the reverse is true for 5Aa- and 7B'. On the other hand, WRV appears to vary consistently from one PV to another: e.g., 7B° is consistently lower in WRV than 7C° for all singers shown. Specific sites of high between-individual rhythmic variability are evident, for example, in the relative IOI values of the third and fourth song units of PV 7C°.

3.3.3 Differences in Within-individual Rhythmic Variability

I found that singers performed the different PVs in their song repertoire with a wide range of WRVs (Figure 3.6). The similarity of repertoire means (taken across all PVs in a singer's repertoire) and variation in population means (taken across all singers of a particular PV) suggested that WRV varies more among PVs than it does among singers. This was confirmed by linear mixed-effects modelling of the relaxed dataset (Figure 3.7). This showed that on the one hand there was little evidence for rejecting the null hypothesis of equal WRV values for each singer: only Singer 191004 was significantly different to the grand mean. Moreover, this difference depended on one influential statistical outlier (Singer 191004, PV 7B°). I concluded, then, that the variation among singers was largely not consistent across PVs. On the other hand, the null hypothesis of equal WRVs for each PV could be rejected because 4/13 PVs differed significantly from the grand mean (becoming 5/13 when the statistical outlier was excluded). This lends support to the hypothesis that variation among PVs was generally consistent across all singers. These results were unchanged when modelling the restricted dataset (SI §3.6.3).

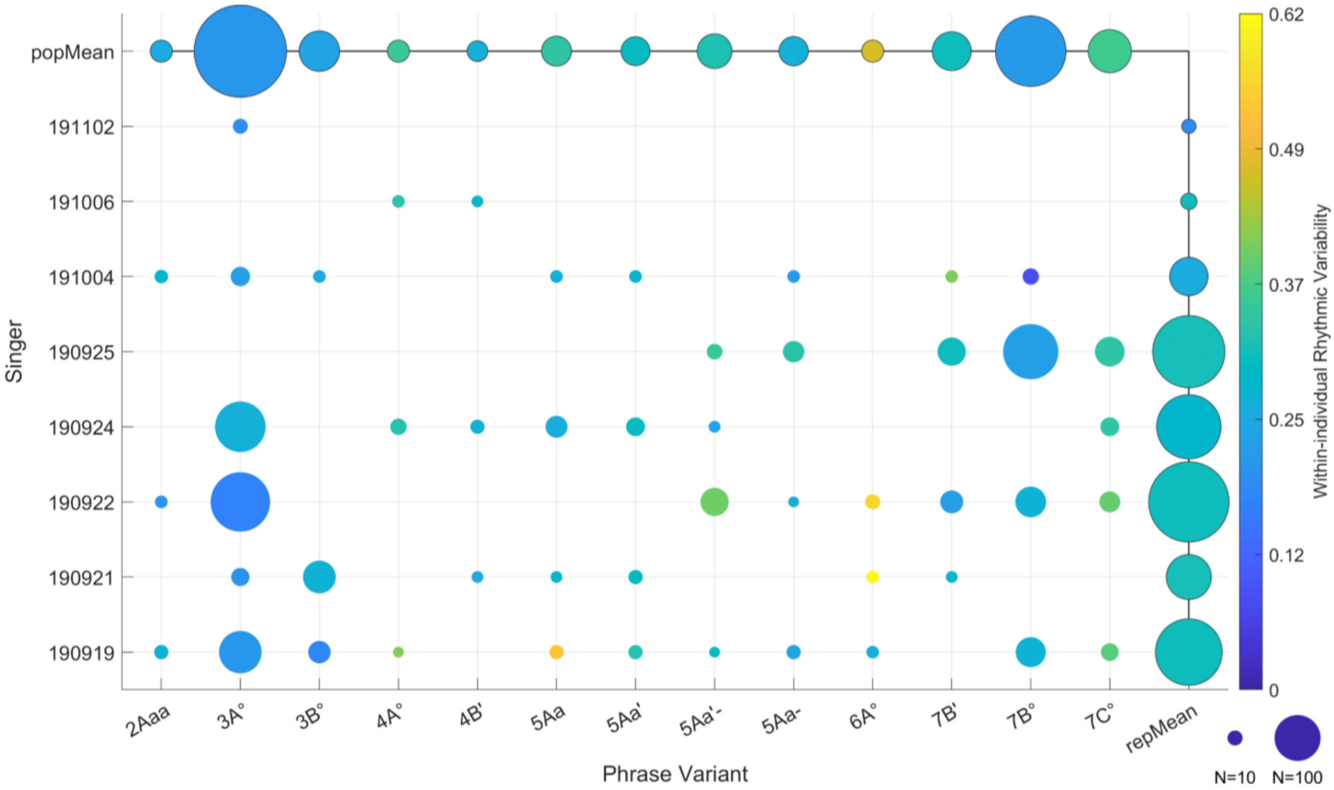


Figure 3.6 Phrases vary in WRV among singers and their shared PVs. Modelling confirmed that differences among PVs tend to be common to all singers (e.g., PV 7C° has a higher WRV than 7B°, 6A° is higher than 3A°). However, singers are not consistently more or less variable across their repertoire of PVs (e.g., Singer 190922 has lower than average WRV values for PVs 2Aaa and 3A°, but higher than average WRV values for PVs 5Aa'-, 6A° and 7C°). Observations here (N_{obs}=49) comprise dataset DS_{H2}, i.e., all PVs sung by at least three singers ($N_{phrase} \geq 5$). Marker areas are scaled by N_{phrase} (range 5–420 including means).

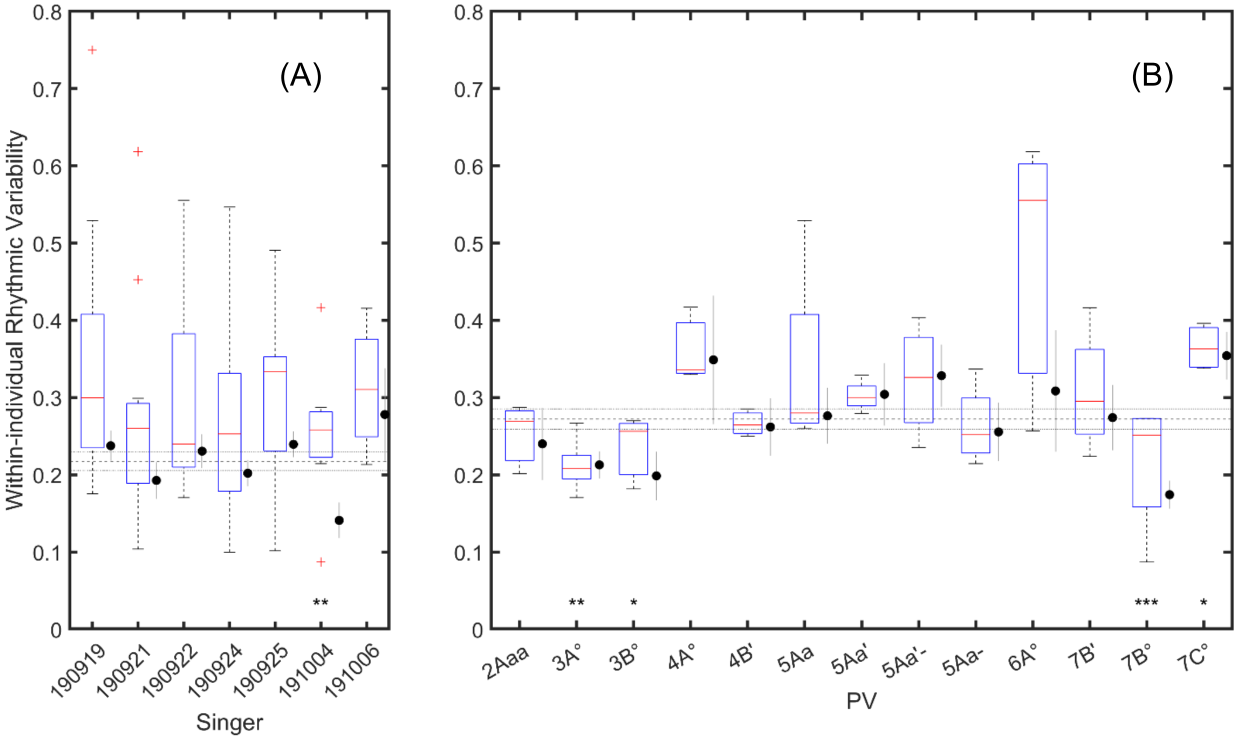


Figure 3.7 Singers do not vary consistently in WRV across PVs (A); but PVs do vary consistently in WRV across singers (B). The box plots summarize the observations in the relaxed datasets by (A) Singer (96 observations, 1949 phrases) and (B) PV (49 observations, 1183 phrases), with the horizontal dashed lines showing the intercept (unweighted grand mean) and associated error returned by modelling (A) Singer and (B) PV as fixed effect, with PV and Singer respectively as random effect (A: WRV ~ 1 + Singer + (1|PV), B: WRV ~ 1 + PV + (1|Singer)). Offset markers show the predicted values returned by modelling; with fixed effect levels coded for significant differences to the unweighted grand mean (*** $P < 0.001$, ** $0.001 \le P < 0.01$, * $0.01 \le P < 0.05$). For both (A) and (B) the fitted variance in the random effect was negligible; omitting the random effect made no difference to the fitted coefficients and p-values for the fixed effect. To allow the detail of (A) to be visible, I set the y-axis limits so that one statistical outlier (Singer 190925, PV 5Ab--) was excluded from the plot (WRV = 1.84 s). It was included in the model fits.

Having found that singer variability in WRV was inconsistent across PVs, and that PV variability is likely a population-wide trait, I explored whether WRV could be predicted by (1) the degree to which a PV was repeated immediately (mean number of reiterations), (2) how far PV rhythm was from isochrony (mean scaled phrase anisochrony, §2.4.2), and (3) the mean song unit peak frequency (calculated in Raven 1.6). In the first case, it might be expected that rhythmic consistency (precision) would be higher for phrases which were repeated in longer series, analogous to a 'blocked' practice schedule with less "contextual interference" (Brady, 2004). This reflects a common assumption made by instructors of human music students, though there is evidence that it does not apply to learners at all stages (Stambaugh, 2011). In the case of anisochrony, I expected to find that PVs with rhythms further from isochrony would have higher levels of WRV, under the assumption that a regularly repeating IOI is less demanding to sing than patterns with irregular intervals. Finally, it has been proposed that phrases containing higher frequency units are more variable, given their possible use in short-range courting (Murray *et al.*, 2018). Hence, I predicted that phrases characterized by a higher mean song unit frequency would be more variable. The dependence of WRV on these features was modelled using a mixed-effects model for the relaxed dataset DS_{H2} (in which each PV is sung by at least three singers), with each feature as a fixed effect and Singer as random effect (Table 3.3).

Table 3.3 Specification of models for post-hoc exploration into possible explanations for the dependence of WRV on PV.

Model	**Specification of Fixed and Random Effects**
1	WRV ~ 1 + PV + (1|Singer)
2	WRV ~ 1 + meanReiterations + (1|Singer)
3	WRV ~ 1 + meanScaledAnisochrony + (1|Singer)
4	WRV ~ 1 + meanPeakUnitFrequency + (1|Singer)
5	WRV ~ 1 + meanScaledAnisochrony + meanPeakUnitFrequnecy + (1|Singer)

Testing of Model 2 showed that the null hypothesis could not be rejected, i.e., that there was no evidence that WRV depended on the mean number of reiterations. On the other hand, mean anisochrony and mean peak frequency were statistically significant when modelled individually (Model 3, Model 4) and hence they were both included in Model 5 (Table 3.4), which was tested with and without the statistical outlier Singer 191004 PV 7B°.

Table 3.4 The dependence of WRV on selected PV characteristics was explored using linear mixed-effects modelling. Model 1 corresponds to the fitting of H2 without accounting for these characteristics. The gradient of WRV plotted against each standardized PV characteristic is reported as coefficients $\beta_1, \beta_2, \beta_3$. The p-*values of significant effects are coded* $*** P < 0.001$, $** 0.001 \leq P < 0.01$, $* 0.01 \leq P < 0.05$. *Akaike Information Criterion (AIC) values justified a model combining anisochrony and peak frequency.*

		Mean Reiterations		**Mean Scaled Anisochrony**		**Mean Song Unit Peak Frequency**		
Model	Singer 191004 PV 7B°	β_1	*p*-value	β_2	*p*-value	β_3	*p*-value	AIC
1	●	–	–	–	–	–	–	-81.9
2	●	-0.0109	0.173	–	–	–	–	-70.7
3	●	–	–	-0.0175	*	–	–	-74.6
4	●	–	–	–	–	0.0488	***	-84.6
5	●	–	–	-0.0138	*	0.0453	***	-87.3
5		–	–	0.0232	***	0.0313	***	-118.5

For the full dataset, AIC values suggest that mean scaled anisochrony and mean unit frequency together (Model 5) better explain the variability in the WRV values than PV alone (Model 1), giving some support to the suggestions that WRV depends on anisochrony and song unit frequency. However, the direction of the effect on anisochrony was the opposite sign to that expected, the result of a single observation (Singer 191004, PV 7B°), anomalously low in WRV. Excluding it resulted in a change of sign in β_2 (Figure 3.8)

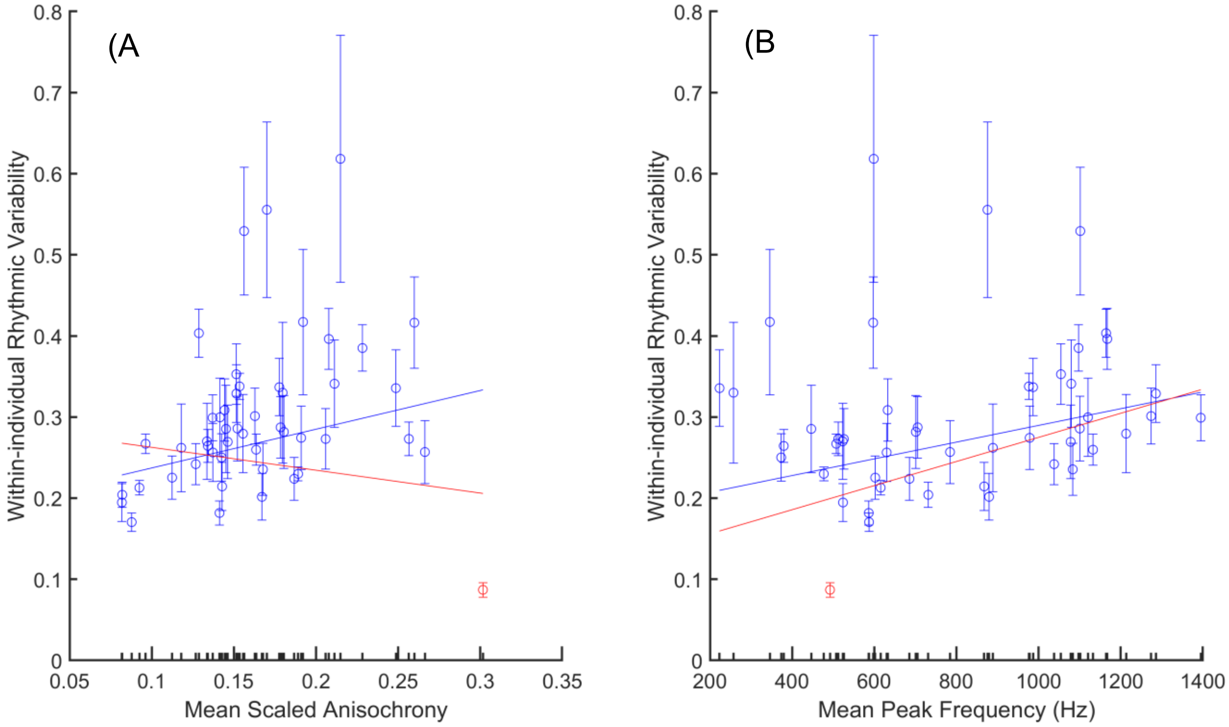


Figure 3.8 WRV varies systematically with PV features: (A) mean scaled anisochrony (§2.4.2), (B) mean song unit peak frequency (measured in Raven Pro 1.6). Markers are plotted for each singer's performance of a shared PV. Partial dependence plots are shown for the two variables fitted in the Models 5: in red with all data points included, and blue with the statistical outlier (Singer 191004 PV 7B°, red point) excluded. Fits were weighted using the bootstrapped errors indicated (SI 3.6.6).

3.3.4 Differences in Rhythmic Individuality

For all but one of the shared PVs tested, I found evidence for systematic rhythmic differences among singers. Mantel tests showed that pairwise distances between phrase rhythms were correlated with singer difference, at a significance level (Bonferroni correction applied) of 0.001 (9/13 PVs), 0.01 (2/13), or 0.05 (1/13) (Figure 3.9). This is overwhelming evidence for rejecting the null hypothesis of equal between- and within-individual variability, and accepting that the between-individual differences in rhythms apparent in chronotonic diagrams (Figure 3.4) exceed within-individual variability. Moreover, for all shared PVs for which between-individual variability significantly exceeded within-individual variability, the observed correlations were significantly higher than simulated values that assumed no systematic rhythmic difference (Figure 3.9). This shows that the correlation is highly unlikely to be the result of discrepancies in WRV, lending further support to H3, i.e., that singers sing shared PVs with individually distinctive rhythms.

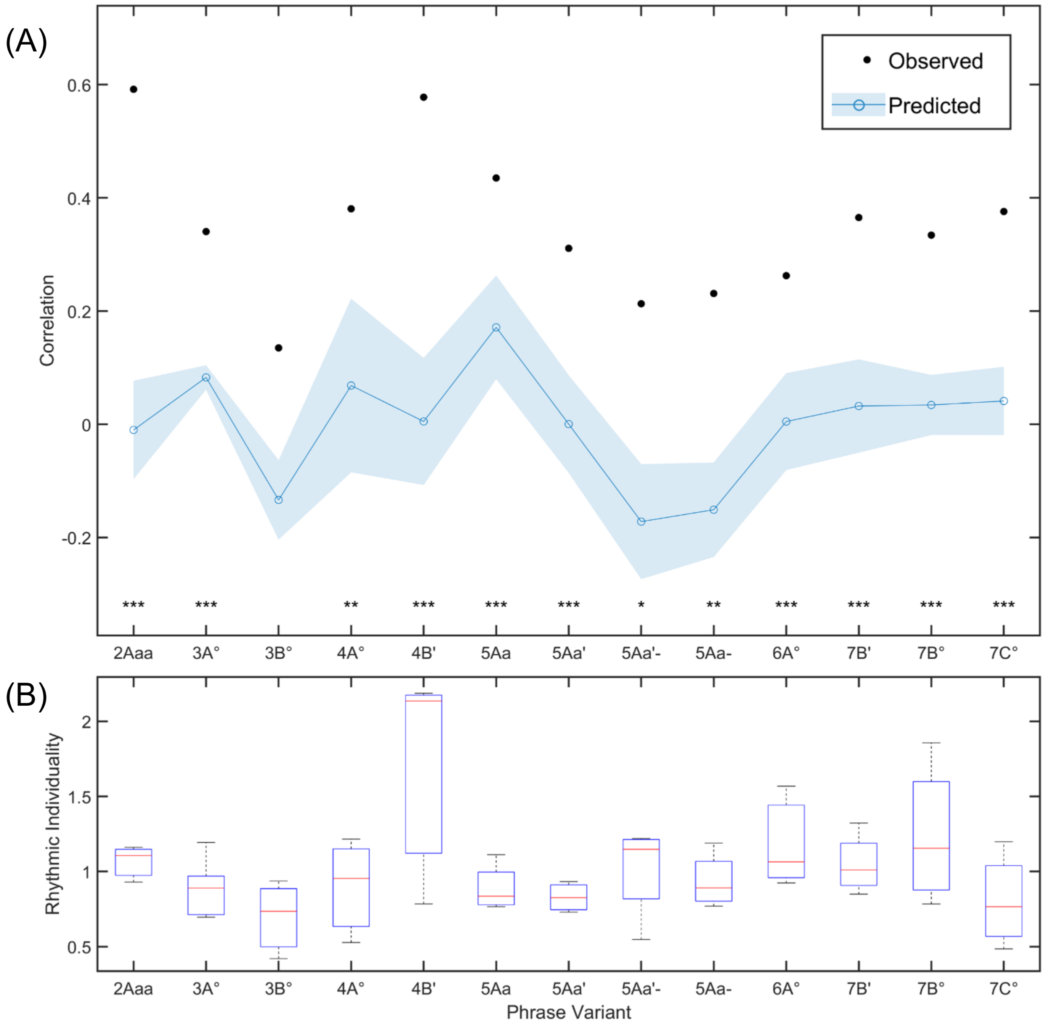


*Figure 3.9 Humpback whales sing shared PVs with individually distinctive rhythms. (A) The Mantel test applied to observed phrase rhythms indicated that between-individual rhythmic variability was higher than within-individual variability (**** $P < 0.001$, **** $0.001 \leq P < 0.01$, *** $0.01 \leq P < 0.05$). *Observed correlations (filled black circles) are compared with predictions from simulations based on actual numbers of singers, sample sizes, and levels of WRV, in which individual mean rhythms for each singer were replaced with a single population mean rhythm. The simulated correlations (blue open circles, shaded area showing mean and 95% confidence intervals) show what is predicted in the absence of individually distinctive rhythms. The higher levels of observed compared with simulated correlations thus indicate individually distinctive rhythms. (B) An independent measure of Rhythmic Individuality, calculated for all Singer-PV observations, closely tracked the Mantel test correlations, giving further support to the hypothesis (H3) of singer distinctiveness. The box-plot for each PV shows the distribution across the population.*

Finally, linear mixed-effects modelling of Rhythmic Individuality has shown that I can reject the null hypothesis that singers vary randomly in their degree of difference to population means, with 4/6 singers significantly different to the grand mean (Figure 3.10). These results support the proposal (H4) that singers vary consistently across PVs in how far their mean phrase rhythms (rhythm templates) are from the population mean rhythm (SI §3.6.8). For example, if a singer is close to the population mean rhythm for one PV, it is likely that this is true for other PVs in that singer's repertoire.

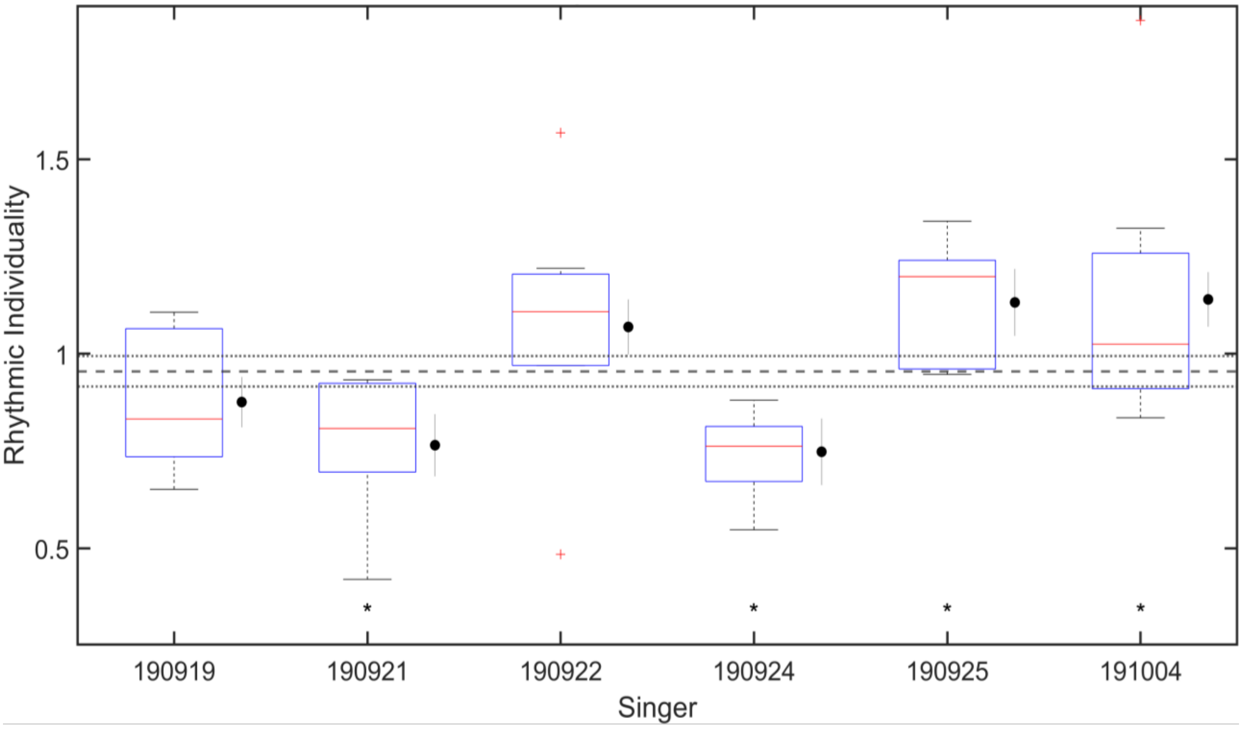


Figure 3.10 Rhythmic Individuality (RI) of singers varies consistently across PVs: singers with the highest values have phrase rhythms which are furthest from the population mean rhythms. The box plots summarize the observations (N=42) in the restricted dataset with the horizontal dashed line showing the intercept (unweighted grand mean) returned by modelling Singer as fixed effect and PV as random effect (RI ~ 1 + Singer + (1|PV)). *Offset markers show the predicted values returned by modelling; fixed effect levels are coded for significant differences to the unweighted grand mean (**** $P < 0.001$, ** $0.001 \leq P < 0.01$, * $0.01 \leq P < 0.05$).

3.4 Discussion

This analysis of humpback whale phrase rhythms failed to reject the null hypothesis (against H1) that the identity of a singer does not influence the level of Within-individual Rhythmic Variability (WRV) in phrase repetitions consistently across phrase variants (PVs). Rather, the observed differences in WRV from one PV to another were consistent *across singers* (H2), and may be partially explained by the effects of phrase anisochrony and song unit frequency. Although singer identity seems not to be a significant factor determining WRV across PVs, the results of Mantel tests on between- and within-individual variability were consistent with the hypothesis (H3) that individual singers sang shared PVs with their own distinctive rhythms. Furthermore, mixed-effects modelling supported the hypothesis (H4) that individual singers maintained their levels of rhythmic individuality (measured as the difference between individual rhythm templates and population mean rhythms) across PVs. The overall picture is one in which individual singers sing shared PVs with their own individually distinctive rhythms, but with equal levels of rhythmic precision.

3.4.1 Song Structure (Themes)

Besides the rhythmic variability that is the focus of this study, I also found evidence for considerable between- and within-individual variability in theme ordering (Figure 3.3), including both the omission of themes, and the existence of reversals. These departures from the simplified picture of humpback song defined by theme sequence occur in both the South and North Pacific (Frumhoff, 1983; Payne *et al.*, 1983; Helweg *et al.*, 1992, 1998; Cholewiak *et al.*, 2013; Garland *et al.*, 2017), though reversals in which a theme is immediately returned to are less commonly reported. Reversals in my data (theme sequences 1-2-1 and 2-1-2) took place between themes containing similar phrase types, consisting of two or three modulated moans with intervening high frequency units. These structures, also shared with Theme 7, could imply radiation from a common ancestral theme, and indeed published song data from French

Polynesia 2016–2018 shows a theme (therein named Theme 5) that appears to combine song unit types found in my Themes 1 and 2 (Schulze *et al.*, 2022). This evolution in phrase type structure, likely gradual, illustrates the judgement calls that need to be made when assigning theme/phrase type labels to data collected in a single season: Themes 1, 2, 7 could have been grouped as different phrase types within a single theme. However, this high-level organization would not affect my analysis of rhythmic variability at the fine-scale level of PV.

3.4.2 Phrase Rhythms

Humpback whales are rhythmically precise singers. Their level of timing variability for both phrase durations and IOIs (WRV) is lower than that for equivalent measures in other species, and bettered only by expert human musical performance (Table 3.5). In the case of humpback song, rhythmic consistency (precision) for repetitions of the same PV was high even when interrupted by other PVs and themes (Figure 3.4) and decreased only slightly for observations made over longer periods of time (SI §3.6.7). This offers indirect evidence in support of a vocal learning model in which rhythm (IOI pattern), above and beyond unit durations and sequence order, is an essential aspect of auditory templates. It supports the received view that human coding of humpback song in terms of phrase types and variants picks out a level of organization salient to the humpbacks themselves (Cholewiak *et al.*, 2013), and further implies that the delineation and classification of phrases might be aided by taking rhythmic structure into account (Handel *et al.*, 2009). This may well be true too for the songs of other mysticetes, and is an argument for recording IOIs alongside the commonly measured unit durations. A related issue is whether rhythm serves as a means of organizing song units independent of song unit sequence ('rhythmic autonomy'). In many forms of human music, different rhythms applied to the same sequences of pitch-intervals (and vice versa), are considered to be different melodies,[13] although recurrent minor variations in rhythm may rather be characterized as different performance styles of the same melody (Seashore, 1938; Dodson, 2020). Similarly, the combinatorial properties of the elements of human language are typically analysed separately from their rhythms, but their independence has been questioned for birdsong (Mol *et al.*, 2017; Hyland Bruno and Tchernichovski, 2019). In humpback whale song, the independence of duration and inter-unit silences for the same unit type used in different phrase types (Handel *et al.*, 2009) offers limited evidence for rhythmic autonomy. The between-individual systematic rhythmic differences (same unit sequence, different rhythms) found in the current study could also be taken as additional evidence for rhythmic autonomy, but I would characterize these as performance variations analogous to the ways that different human singers perform the same melodies (whether jazz standards, pop ballads or operatic arias). Further work is also required to examine whether such rhythmic differences are driven by minor between-individual variability in unit types, e.g., differences in unit durations or frequencies. Whether looking at the same unit type between PVs, or between singers, the first steps would be to see whether differences in unit features (duration, frequency, etc.) and sequence ordering are correlated with rhythmic differences. If not, this would support the hypothesis that, as in much human music, rhythm serves as an independent level of organization of song.

---

[13] For example, Barlow and Morgenstern's 'A Dictionary of Musical Themes' (1949) contains eight melodies, each with a different rhythm, whose first four pitches when appropriately transposed match those heard at the start of Beethoven's Fifth Symphony.

Table 3.5 Timing variability observed for the durations of a multi-element phrase (termed 'coda', 'song', 'call', 'measure') and, where available, for elements of that phrase. Figures are given for the mean (standard deviation) coefficient of variation (CV) across multiple individuals and vocalization types, derived from data reported in the source.

Species	**Vocalization type**	**phrase duration CV (%)**	**inter-onset interval CV (%)**	**Source**
Humpback whale	song phrase	3.3 (1.5)	9.7 (2.6)	this study
Bowhead whale	song	30.5 (11.7)	†	(Johnson *et al.*, 2015)
Fin whale	20 Hz pulse	N/A	18.5 (5.4)	(Watkins *et al.*, 1987)
Sperm whale	coda	21.2 (11.0)	c.5–10*	(Weilgart and Whitehead, 1993)
Chaffinch	multisyllabic song	6.8 (1.2)	†	(Riebel and Slater, 2003)
Emperor penguin	multisyllabic mutual display call	3.2 (1.9)	10.5 (4.2)	(Robisson *et al.*, 1993)
Human	bar/beat lengths in Aksak rhythms (recording session)	4.6 (0.8)	5.4 (2.7)	(Clayton, 2016)
Human	bar/beat lengths in Aksak rhythms (CD recording)	1.7 (0.8)	4.2 (1.1)	(Goldberg, 2015)
Human	Two-tap patterns in tapping experiment	c.2–4*	c.4–13*	(Semjen and Ivry, 2001)

[*- mean and standard deviation in the CV could be calculated but would require measurements to be made from a figure; † - data unavailable in paper]

3.4.3 Within-individual Rhythmic Variability

Might the rhythmic precision of phrase rhythms communicate information about a singer's singing abilities? I hypothesized (H1) that a singer's WRV would be consistent across shared variants, implying a generalized timing ability. If true, then WRV could serve as an honest signal of this ability, which might in turn reflect underlying singer attributes such as age or cognitive health. However, my analysis did not support H1, but was consistent with WRV determined by PV rather than by singer identity (Figure 3.7). This could be because there is no generalized timing ability, or because phrases are not rhythmically challenging (i.e., not testing of such a generalized timing ability). In either case it seems that WRV may rather be determined by properties of the phrase, perhaps including anisochrony and song unit frequency, which pose cognitive or other challenges that cannot be overcome in a general way by singers. Post-hoc exploration suggested that PVs were less rhythmically precise if their song units were higher in frequency, but the results were ambiguous regarding anisochrony, with the direction of the dependence determined by a single statistical outlier (Singer 191004 PV 7B°).

If we consider the audience perspective then there are further alternative explanations for the main findings. The modelling of listener discrimination in the presence of noise has shown that rhythms closer to isochrony may be easier to judge against internal templates (van den Broek and Todd, 2003), but as we have seen humpback phrase rhythms are in general far from isochrony (Figure 3.4, Figure 3.8). Another factor that could make WRV hard to gauge is the constantly-evolving nature of phrase variants and types (Payne and Guinee, 1983; Payne *et al.*, 1983; Payne and Payne, 1985; Cerchio *et al.*, 2001). Both these considerations imply that rhythmic variability may be too difficult to judge, which removes the mechanism offering evolutionary rewards to competition in this domain, and consequently we should not be too surprised that singers do not vary consistently across PVs. Whatever the explanation, the lack of dependence on singer makes it unlikely that WRV could be used as a reliable guide to timing ability or play the role in sexual selection found for consistency in assorted temporal features in multiple taxa (Lambrechts and Dhondt, 1986; Simmons and Zuk, 1992; Jennions *et al.*, 1995; Poesel *et al.*, 2001; Byers, 2007; Demartsev *et al.*, 2017; Rogers, 2017). Cross-species comparisons here may be of rather limited value as none of these studies targeted a generalized timing ability, but it could be more relevant that none of the vocalizations studied belonged to a song culture in which rhythms were constantly changing. In the case of humpback whales, whose songs do fit into such a culture, my findings from the study of WRV indirectly support the view that shared PVs may serve more to convey group membership than individual quality (Murray *et al.*, 2018). But this is not the end of the story, because in addition to differing levels in WRV I also examined whether humpback whales sing shared PVs with systematically different rhythms.

3.4.4 Individually Distinctive Rhythms

Here, I found evidence consistent with the existence of individually distinctive rhythms (systematic rhythmic differences) for all but one shared PVs (§3.3.4). Thus, although WRV cannot communicate individuality, the possibility exists that individually distinctive rhythms might facilitate individual recognition and the role of shared PVs would not be limited to signalling group membership. The subtlety of this interpretation of between-individual variability requires the conceptual separation of individually distinctive rhythms from WRV, and I recommend that bioacoustics follows empirical musicology in adopting it more widely. Although I have not tested whether the systematic differences observed would allow reliable individual-level recognition (this is left for a future application of discriminant function analysis or similar), if found to do so this would contrast with the historical finding that phrase duration was an unreliable predictor of singer identity (Thompson, 1981). Given that multivariate phrase rhythms likely contain more information than simple phrase duration, this would be unsurprising. Early work showing that a combination of song unit duration and spectral properties allowed reliable discrimination among individual humpback singers (Hafner *et al.*, 1979). However, this study compared songs from different years and hence may have been confounded by song evolution; a later study of unit properties concluded that individual differences “did not provide useful discrimination” among individuals (Macknight *et al.*, 2001). Nonetheless, if individually distinctive rhythms were to play a role in the communication of individual identity among humpback whales, then this would align them with other marine mammals (Antunes *et al.*, 2011; Gero *et al.*, 2016; Oliveira *et al.*, 2016; Mathevon *et al.*, 2017) and human musicians, where systematic between-individual differences with a magnitude sometimes exceeding within-individual rhythmic variability (Hellmer and Madison, 2015), and in some cases enabling recognition (Van Vugt *et al.*, 2013), exist across a range of genres and contexts (Seashore, 1938; Prögler, 1995; Frane, 2017; Dodson, 2020; Corcoran and Frieler, 2021).

We might usefully compare humpback whale phrase rhythms with the stereotyped rhythms of the broadband click sequences (codas) produced by sperm whales. Like humpback phrases, sperm whale codas are comprised of a few (typically 3–12) sounds separated by silences (Weilgart and Whitehead, 1993). Unlike humpback phrases, codas are timbrally uniform, shorter in duration, and stable from one year to the next (Rendell and Whitehead, 2005). They are known to function as markers of social identity (Rendell and Whitehead, 2003b; Hersh *et al.*, 2022), with timing differences within certain coda types also permitting individuals to be distinguished (Antunes *et al.*, 2011; Gero *et al.*, 2016; Oliveira *et al.*, 2016). It is true that the communicative functions played by sperm whale codas, produced predominantly by females (Marcoux *et al.*, 2006; Frantzis and Alexiadou, 2008), are most likely different to that played by humpback whale phrases (sung by males only), especially when we consider that sperm whales have more highly-structured social systems than humpback whales (Gero and Rendell, 2015), with lifelong relationships among related females (Gero *et al.*, 2015). However, evidence is coming to light of multi-year pair bonds in some populations of humpback whales (Ramp *et al.*, 2010; Weinrich, 2011; Wray *et al.*, 2021), perhaps facilitating co-ordinated foraging techniques (Wray *et al.*, 2021). For communication within the enduring female-female dyads observed in western North Atlantic feeding grounds (Ramp *et al.*, 2010; Weinrich, 2011), song can clearly play no role, but it has been proposed that the social calls (produced by both females and males) that are stable on decadal time scales yet acoustically variable may enable individual recognition (Fournet *et al.*, 2018). Nonetheless, the existence of groups containing males whose associations last at least seven days (Ramp *et al.*, 2010; Weinrich, 2011), and the singing of males whilst escorting mother-calf pairs (Darling *et al.*, 2006; Herman *et al.*, 2013) makes it at least possible that the rhythmical signatures in song (alongside variations in pitch, contour and timbre) could play a role in enabling individual recognition of male humpbacks. Playback experiments and/or longitudinal studies would be required for confirmation, but in any case, individually distinctive rhythms might still offer a useful means for human researchers to track individual whales.

Going beyond individually distinctive rhythms within specific PVs, I have also shown for my song data that the singers' rhythmic distance from population norms is consistent across PVs (Figure 3.10). What could be behind this? It is worth reiterating that previous research into shared PVs and stereotyped theme sequences, evolving over time within a population, strongly supports the interpretation of humpback song as an example of a changing vocal culture spread by social learning and competing learning biases for conformity and novelty (Payne *et al.*, 1983; Payne and Payne, 1985; Garland *et al.*, 2011; Zandberg *et al.*, 2021). The systematic rhythmic differences observed in this study could be taken as individual differences either in the 'what' or the 'how' of song production, depending on whether rhythm is conceptualized as belonging to the identity of a PV (Handel *et al.*, 2009) or to its manner of singing, but here we can be agnostic: all we have to go on are the (audio recordings of) songs themselves. We can still ask what might underlie the observed cross-PV rhythmic differences, and I divide the possible explanations into those arising from population mixing, and those describing changes made to phrase rhythms during learning, memorization and production. The first is perhaps the simplest: that the singers in my sample include a majority group of individuals who have been more closely associated than the other ‘outsiders’. Given vocal learning and a conformity bias, this would result in the outsiders’ phrase rhythms being more individually distinctive (i.e., standing out against the overall population mean rhythms). It could also be expected that those singers observed closer together in time would be more likely to be similar to the population means (lower levels of Rhythmic Individuality). My observations of Rhythmic Individuality were made of five singers observed within seven days, and a sixth observed nine days after the last of the five. We might expect, therefore, that the rhythms of the first five singers would be more similar, and there is some evidence to support this. Of these five singers, two possess values of Rhythmic Individuality significantly lower than the grand mean, implying that their phrase rhythms are more similar than average to the population mean rhythms. The sixth singer is significantly above the grand mean, as would be expected from an outsider. However, one singer in the first group is also significantly higher than the grand mean, so these results are somewhat inconclusive regarding this explanation (population mixing).

The second kind of explanation covers changes introduced during transmission arising from (1) learning biases (novelty, prestige) affecting which rhythms are adopted (such as in the case of a singer hearing two rhythmically different versions of a PV from two different singers), and (2) any tendency to introduce rhythmic changes to what is heard, whether such changes are accidental slips in memorization or production, deliberate improvisation or innovation, or the result of general vocal capacities or habits (e.g., an increased inter-unit silence at the end of a phrase, or a tendency towards longer/shorter average IOIs[14]). For example, a singer who tended towards a relatively high novelty bias would tend to copy unusual rhythms and hence stand out from population norms (possess higher levels of Rhythmic Individuality across a range of PVs). The same would be observed for a singer tending to alter existing song rhythms (whether deliberately or accidentally). Although I cannot use my observations to distinguish between learning biases and tendencies, the fact that the dataset contains singers significantly above and below the grand mean offers some support to the hypothesis that such general biases or tendencies both exist in and vary among humpback singers. Agent-based modelling studies of humpback song evolution have shown that, at the level of theme sequence, changes termed ‘production errors’ are required over and above spatial mixing to give rise to realistic patterns of song evolution but that a uniform novelty bias tended to dominate model outcomes, leading to “unrealistically variable song sequences” (Mcloughlin *et al.*, 2018, p. 13). To make further progress here would seem to require multi-level (theme, phrase) modelling of humpback song together with a more detailed account of bias (Mcloughlin *et al.*, 2018). This, coupled with observations tracking the associations and songs of individual humpbacks could shed light on the role of learning biases in mate choice. For example, observations of males approaching other singers have been interpreted as evidence that males are “prospecting for females” (Smith *et al.*, 2008), and thus that song mediates male-female rather than male-male interactions. If a singer is found to be escorting a female, and cannot be displaced, then a good strategy could be to copy the song of this successful male (Garland and McGregor, 2020). Such a prestige bias has been seen in songbirds with evolving song cultures (Payne, 1985; Williams and Lachlan, 2022), and if found in humpbacks would be an exciting illustration of convergent (biological) evolution (Garland *et al.*, 2021). Here it seems to me essential that the intricacies

---

[14] See §2.5 for simulations showing that Mantel’s procedure is sensitive to small changes in phrase duration (which would result from differences in average IOI).

of song production and development are examined at the individual level, for only this, when combined with the measurement of individual reproductive success, will allow a better understanding of the functions and origin of the humpback whale's elaborate, rhythmically precise, and individually distinctive song.

3.5 Conclusions and Future Work

In this study, I analysed the phrase rhythms of humpback song data collected from multiple singers over a two month period during the 2019 breeding season in Mo'orea, French Polynesia. To assess whether rhythmic precision might be a signal for individual quality or a generalized timing ability, I used linear mixed-effects modelling to examine the dependence of Within-individual Rhythmic Variability (WRV) on singer and PV. This modelling failed to reject the null hypothesis that the level of WRV did not depend on singer, showing that it is unlikely that rhythmic precision signals individual quality. However, further modelling produced evidence that WRV was dependent on the PV, and in a post-hoc analysis I found that this dependence could be explained partially by the presence of high frequency song units in the phrase. Regarding systematic rhythmic differences, I showed using Mantel tests and Monte Carlo simulations that for 12/13 PVs there was strong evidence that singers possessed individually distinctive rhythms, which makes it possible that they could be used for individual recognition. Finally, I defined a measure of Rhythmic Individuality that measured the distinctiveness of an individual singer's rhythms with respect to the population mean rhythm. With linear mixed-effects modelling, I showed that Rhythmic Individuality varied consistently across PVs, which is indirect evidence for a novelty bias or tendency to introduce rhythmic changes.

These findings were derived from a single season and geographical location. Although they represent a good snapshot of this time period and population, further work is required to extend the coverage to multiple seasons and populations, to confirm whether the findings are generally true of humpback whale song. Future studies of rhythmic consistency could repeat the analysis at the level of 'motifs' (Lamoni *et al.*, 2023), which may assist in showing whether phrase types, phrase variants or motifs are more important in the encoding of auditory templates. They could also address the question of how rhythmic consistency varies in time. In my data short stretches of low and high variability (Figure 3.4) suggests that even if WRV cannot signal quality, it might function as an indicator of motivation or arousal, as has been suggested of chaffinch song (Riebel and Slater, 2003). In the longer recordings, the data could be divided into chunks in order to examine whether there are trends in rhythmic consistency through the song session, which might be another sign of male quality (Lambrechts and Dhondt, 1986; Poesel *et al.*, 2001).

Turning to the individually distinctive rhythms found in my data, one next step would be to use discriminant analysis or similar to discover whether these permit individual singers to be confidently identified through such rhythmic differences. It would also be interesting to compare the potential of rhythmic and spectral information for this purpose, as for some species both have been shown to be required (Robisson *et al.*, 1993). If rhythmic information does permit individual recognition, then it would also be important to show that the rhythms are stable enough over days or weeks for them to be practically useful in the typical social groupings of humpback whales. To establish this would require multiple recordings of the same singer across the season. Finally, all measurements of individually distinctive rhythms in my study were based on *absolute* IOIs. Results not shown here suggest that the scaling of IOIs to phrase duration, analogous to removing the effects of tempo changes in human music, made no difference to the main findings of this study. Further investigation could provide valuable data for the debate over whether the rhythmic perception of nonhuman animals is more sensitive to absolute or relative timings (ten Cate *et al.*, 2016; ten Cate and Spierings, 2019).

3.6 Supplementary Information

3.6.1 Song Unit Detection

Coded song sessions were screened to exclude phrases identified as (1) transitional, i.e., those occurring on a border between two themes and containing a mixture of song units from both themes (Payne *et al.*, 1983), (2) occurring during probable surfacings, such as phrases containing ‘surface ratchets’ (Winn and Winn, 1978), or (3) containing low SNR units missed by the automated threshold detector (Table 3.6). Song units in the remaining phrases had a mean SNR of 17.0 (std 8.5).

Table 3.6 Coded song sessions varied in duration (10–180 minutes), but only those including widely-shared PVs were used in the datasets for hypothesis testing (in bold). Approximately 20% of the coded phrases were removed from the analysis: those identified as (1) transitional, (2) taking place during probable surfacing events, or (3) containing low SNR units that were missed by the automated threshold detector (MNU). A small proportion (< 0.4%) of units had start or end times adjusted manually in Raven (MASU).

Singer	Duration	#Phrase0	#Phrase1	#SU0	#SU1	#MNU	#MASU	mean SNR across song units used in analysis
190919	**2h18**	**526**	**335**	**2824**	**1878**	**174**	**2**	**15.6**
190921	**1h14**	**257**	**212**	**1638**	**1323**	**42**	**18**	**15.6**
190922	**3h00**	**669**	**437**	**3589**	**2401**	**189**	**8**	**11.6**
190924	**2h10**	**511**	**479**	**3025**	**2837**	**11**	**29**	**22.7**
190925	**2h30**	**713**	**637**	**2968**	**2543**	**46**	**2**	**17.9**
191004	**0h30**	**120**	**103**	**664**	**560**	**15**	**0**	**18.0**
191006	**0h24**	**73**	**40**	**484**	**295**	**22**	**2**	**15.5**
191008	0h30	41	39	292	280	0	0	18.1
191102	**0h10**	**35**	**20**	**200**	**116**	**13**	**0**	**11.9**
191115	0h10	37	26	224	156	11	0	12.6
TOTALS	12h56	2982	2328	15908	12389	523	61	17.0

- #Phrase0 = total number of phrases coded
- #Phrase1 = #Phrase0 after excluding phrases coded as transitional, surfacing, or containing song units missed by the detector (‘Missing Next Unit’).
- #SU0 = total number of song units coded
- #SU1 = total number of song units contained in #Phrase1
- #MNU = number of song units coded as Missing Next Unit (i.e., case where a song unit could be heard aurally but fell below the threshold of the automated detector).
- #MASU = number of song units coded as ‘Manually Adjusted Song Unit’ (e.g., cases where fish sounds overlapped with humpback song units).

3.6.2 Song Structure

The song sessions belonging to the analyzed singers included eight themes, of which Theme 8 was the only theme not sung by a majority of the singers (Table 3.7). Most phrases in this theme contained only a single surface ratchet. Phrases analysed as part of the relaxed or restricted datasets were drawn from 48 PVs and eight singers (Table 3.8).

Table 3.7 Frequency of theme repetitions by Singer, after exclusion of phrases containing song units missed by the threshold detector. With the exception of Theme 8, all themes were sung by at least 50% of the singers. In the continuous dataset only the song session from Singer 190925 stands out for being deficient in themes. In the non-continuous dataset the song sessions from Singers 191006 and 191008 are predominantly made up from single themes and would conventionally be classed as 'aberrant' (Frumhoff, 1983).

Theme	190919	190921	190922	190924	190925	191004	191006	191008	191102	191115	TOTAL themes	TOTAL phrases	MEAN phrases/ theme
1	16	7	11	11	32	4				1	82	203	2.48
2	13	1	8	24	65	4			2		117	292	2.50
3	17	12	27	19		4			1	1	81	592	7.31
4	8	11	8	14			2				43	182	4.23
5	16	6	18	19	49	4				1	113	332	2.94
6	16	11	23	19		3		2		1	75	154	2.05
7	16	7	25	19	58	4			2	1	132	539	4.08
8		5	3		1			1			10	34	3.40
Totals	102	60	123	125	205	23	2	3	5	5	653	2328	3.57

Table 3.8 Phrases used in model fits (observations in bold), by Singer and PV, ordered by population total. For each Singer, table entries are the number of phrase repetitions aggregated over the whole song session. The PVs and Singers included in the relaxed (DS_{H1}, DS_{H2}) and restricted (DS_{restr}) datasets are indicated. Singer 191008's song session was highly aberrant (predominantly made up of a single unshared PV).

Phrase Variant	**190919**	**190921**	**190922**	**190924**	**190925**	**191004**	**191006**	**191008**	**191102**	**191115**	**Pop. Total**	**DS_{H1}**	**DS_{H2}**	**DS_{restr}**
1B0	**5**		**5**	2	2	1					*15*	●		
1Ba1	2		**7**			1					*10*	●		
1Ba2	1	**7**			3						*11*	●		
1Ba4	1	1		4	**8**	**7**				1	*22*	●		
1Ba5	4	2		**12**	**13**	4					*35*	●		
1Ba6	1			**29**	**13**	1					*44*	●		
1Ba7		1		**8**	**14**						*23*	●		
1Ba8				4	**8**						*12*	●		
2Aa	**19**	2	1	2	1						*25*	●		
2Aaa	**9**		**7**			**8**					*24*	●	●	●
2Ba	1		2	**66**	**31**				1		*101*	●		
2Baa				**5**	**126**				1		*132*	●		
3A°	**85**	**15**	**168**	**121**		**17**			**10**	4	*420*	●	●	●
3A-	3		**6**	**10**							*19*	●		
3A+			3	**5**							*8*	●		
3AA				**7**							*7*	●		
3AAA	5		**13**	1					1		*20*	●		
3B°	**23**	**50**				**7**					*80*	●	●	●
3B---		**11**									*11*	●		
4A°	**5**		2	**12**			**7**				*26*	●	●	
4A'	2	1	2	**8**			**10**				*23*	●		
4A''	1	2	**14**				**12**				*29*	●		
4A-		1	1	**9**			1				*12*	●		
4B°		4		**13**			1				*18*	●		
4B'	1	**6**		**9**			**6**				*22*	●	●	
4B''		**13**					1				*14*	●		
4B-4	2	**5**									*7*	●		
5Aa	**9**	**6**		**22**		**7**				1	*45*	●	●	●
5Aa'	**9**	**9**		**16**		**7**					*41*	●	●	●
5Aa'-	**5**		**37**	**6**	**11**	2				2	*63*	●	●	●
5Aa-	**9**		**5**	4	**21**	**7**					*46*	●	●	●
5Aa-+b					**7**						*7*	●		
5Ab'-					**20**						*20*	●		
5Ab-	2				**79**						*81*	●		
5Ab--					**5**						*5*	●		
5Ab-+b					**7**						*7*	●		
6A°	**7**	**7**	**10**	3		3					*30*	●	●	●
6A--		1	**9**								*10*	●		
6A+				**7**							*7*	●		
6B°	**10**	3	**29**			3					*45*	●		
6B'	**6**	4				1					*11*	●		
7A°	**11**	3		**44**							*58*	●		
7A'	3	3		**7**					2		*15*	●		
7B°	**42**	2	**44**	2	**149**	**12**			1	1	*253*	●	●	●
7B'	3	**6**	**24**		**37**	**7**			3	4	*84*	●	●	●
7C°	**14**		**20**	**16**	**41**	2					*93*	●	●	●
7C-					**6**						*6*	●		
7CC	1	**5**			2	2					*10*	●		
Singer total	*301*	*170*	*409*	*454*	*604*	*99*	*38*	*0*	*19*	*13*	*2101*			
DS_{H1}	●	●	●	●	●	●	●							
DS_{H2}	●	●	●	●	●	●			●					
DS_{restr}	●	●	●	●	●	●								

Transition probability analysis of individual song sessions revealed complex variations in the production and ordering of shared PVs, bringing the Theme and PV summaries (Table 3.7, Table 3.8) to life. For example, Singer 190924 (Figure 3.11) produced a wide variety of PVs, but with all song themes following in a strict order with the exception of reversals occurring between Themes 1 and 2 and occasional omissions of Theme 4.

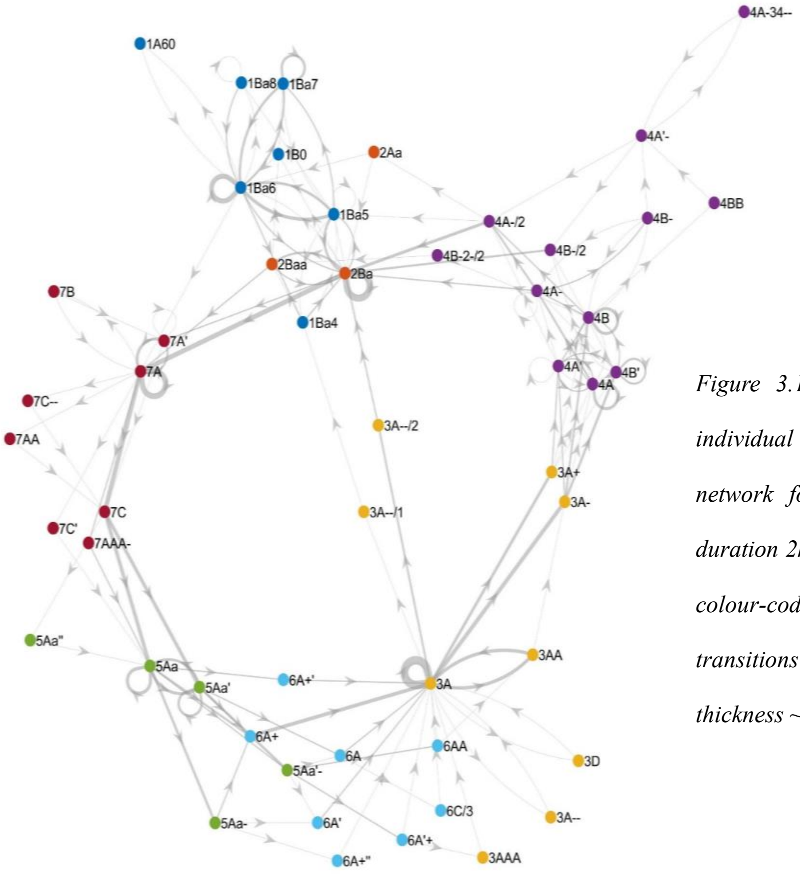


Figure 3.11 PVs were used freely during individual themes. In this phrase transition network for Singer 190924 (song session duration 2h10), nodes represent PVs (N=52), colour-coded by theme, and edges represent transitions (N=501) between PVs. Edge thickness ~ log*(number of transitions).*

The analysis of phrase rhythms for single themes, using phrase raster plots, was capable of revealing multiple subtle trends in unit onsets and phrase durations, and relationships between different PVs (Figure 3.12). For example, near the start of the song session for Singer 190925, we see a progressive lengthening and some splitting of the second song unit, but without this splitting affecting the next inter-unit interval (boxed in red). The phrase duration remains roughly constant, implying negative covariance between IOI2 and IOI3 (7B°), or IOI2+IOI3 and IOI4 (7B’). A little later, we see a progressive shortening of the second IOI, coupled with a shortening of the phrase duration (boxed in purple). These trends may be analogous to the deliberate manipulations of tempi and rhythm by human musicians; similar parallels have been observed in birdsong (Rothenberg *et al.*, 2014; Taylor, 2017b). Raster plots also revealed unusual PV sequences used by individual singers, such as the habit shared by Singers 190922 and 191004 of ending a theme with a longer PV.

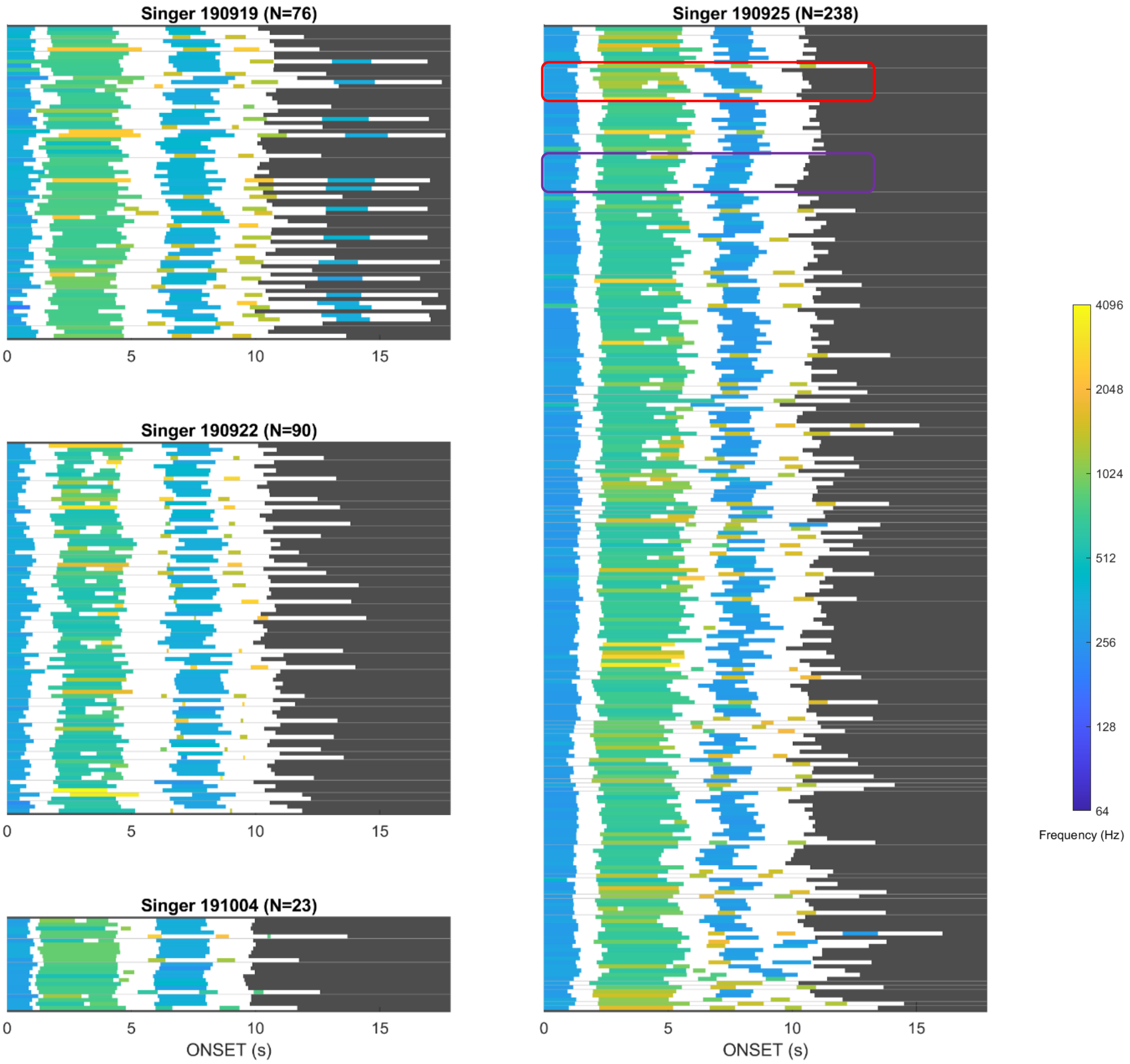


Figure 3.12 Phrase rhythms vary within and between singers. Phrases may occur as part of music-like progressions (red and purple boxes). In each panel, a singer's performances of phrases from Theme 7 are plotted horizontally, with the time series running top to bottom like a musical score. Each song unit is represented as a coloured rectangle, coded by peak frequency and followed by white inter-unit silences.

3.6.3 WRV Fits – Observations and Fits for Restricted Dataset

In addition to the relaxed datasets for which results are provided in the main text, a restricted dataset was selected by more stringent criteria (§0) and used to test both hypotheses H1 and H2. The use of a restricted dataset did not change the main findings of the linear mixed-effects modelling (§3.3.3). Singers did not differ consistently from the unweighted grand mean, but PVs did vary consistently across singers (Figure 3.13). As with the results for the relaxed dataset, the exclusion of the statistical outlier (Singer 191004, PV 7B°) from the dataset reinforced these conclusions: for H1, *no* singers were significantly different to the grand mean; for H2, 5/11 were significantly different to the grand mean.

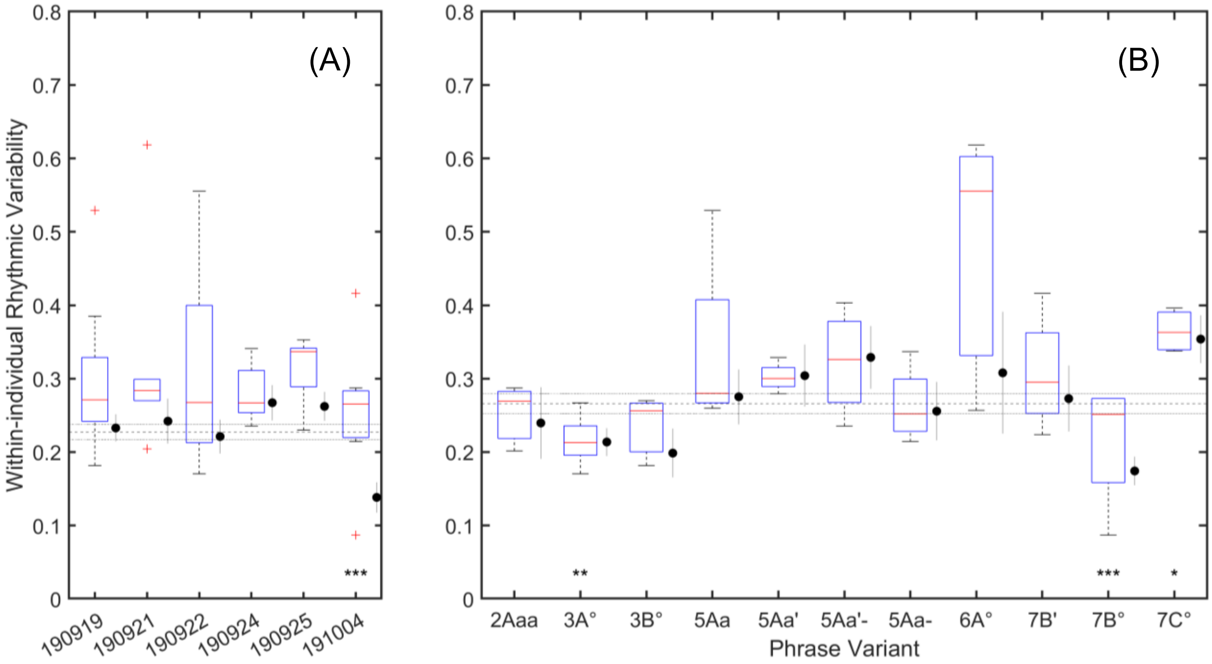


*Figure 3.13 There is little evidence that WRV varies consistently among singers, but support for the hypothesis that it varies consistently among PVs. The box plots summarize the observations in the restricted dataset (N_{obs}=42) by (A) Singer and (B) PV, with the horizontal dotted line showing the intercept (unweighted grand mean) returned by modelling (A) Singer and (B) PV as a fixed effect, with PV and Singer respectively as random effect. Offset markers show the predicted values returned by modelling; with fixed effect levels coded for significant differences to the unweighted grand mean (*** $P < 0.001$, ** $0.001 \leq P < 0.01$, * $0.01 \leq P < 0.05$).*

3.6.4 WRV Fits – Residuals

Residuals to linear mixed-effects modelling of the relaxed datasets for hypotheses H1 (Figure 3.14) and H2 (Figure 3.15) showed some deviations from normality for extreme fitted values, but these deviations fell within typical levels. There was no tendency to heteroscedasticity.

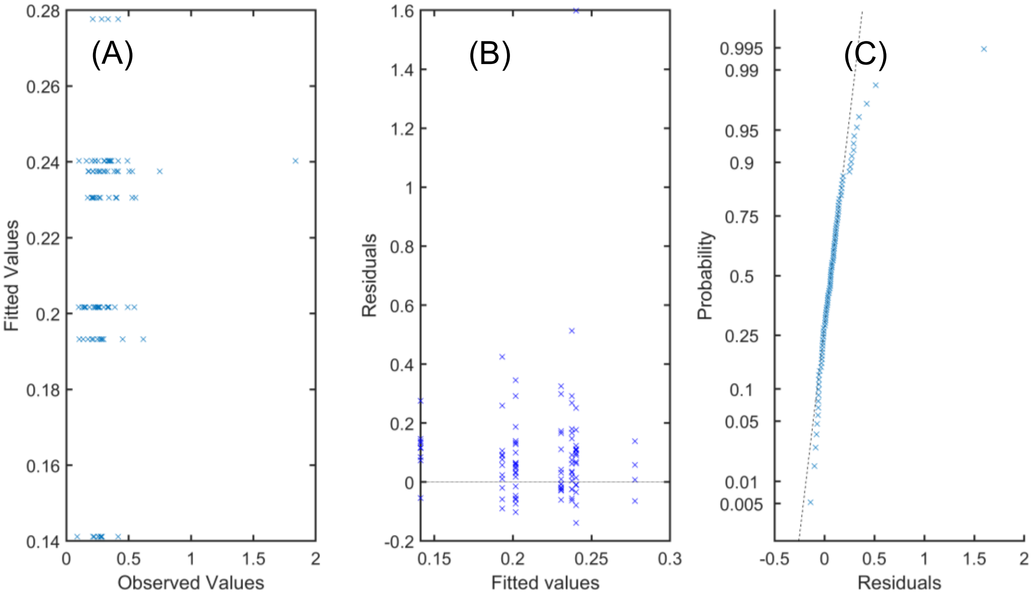


Figure 3.14 (A) Fitted versus observed values, (B) Residuals versus fitted values, and (C) Normal probability plot of residuals for linear mixed-effects modelling of hypothesis H1 (dataset DS$_{H1}$).

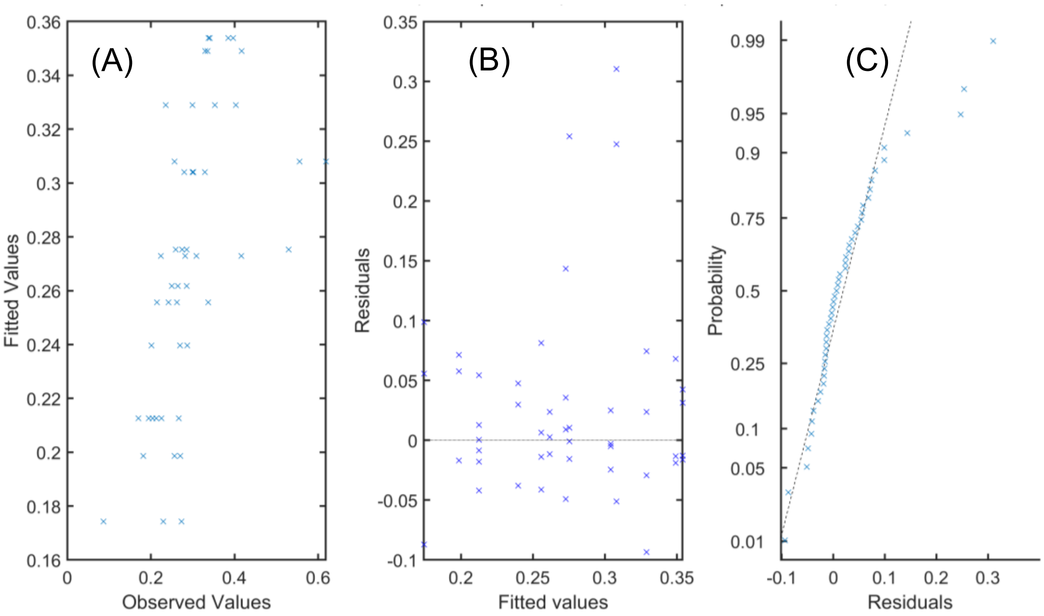


Figure 3.15 (A) Fitted versus observed values, (B) Residuals versus fitted values, and (C) Normal probability plot of residuals for linear mixed-effects modelling of hypothesis H2 (dataset DS$_{H2}$).

3.6.5 WRV Fits – Sensitivity of Modelling to Observation Size

To test the sensitivity of modelling on the initial criterion for creating an observation (i.e., the minimum number of phrase repetitions for any given singer and PV, $N_{phrasemin}$), this number was varied above and below the value ($N_{phrasemin} = 5$) used for the results reported in the main text (§3.3.3), resulting in limited and expanded datasets respectively (generated using the "relaxed" criteria, §0). Linear mixed-effects modelling for both H1 and H2 led to similar outcomes for all datasets (Table 3.9), implying that my conclusions are robust to the value of this initial criterion.

Table 3.9 The effect of minimum observation size on the significance of hypothesis testing for H1 and H2 obtained using linear mixed-effects modelling. R^2 values are ordinary (adjusted). Significance levels for fitted coefficients coded as *** $P < 0.001$, ** $0.001 \le P < 0.01$, * $0.01 \le P < 0.05$.

	Do singers differ significantly from the grand mean? (H1) (N_{singer} = number of levels of fixed effect with coefficients significantly different to zero)								**Do PVs differ significantly from the grand mean? (H2)** (N_{PV} = number of levels of fixed effect with coefficients significantly different to zero)							
$N_{phrasemin}$	N_{obs}	N_{singer}	N_{singer} ***	N_{singer} **	N_{singer} *	Intercept	Random effect sd	R^2	N_{obs}	N_{PV}	N_{PV} ***	N_{PV} **	N_{PV} *	Intercept	Random effect sd	R^2
3	137	9	2	0	0	0.22	0	0.24 (0.19)	76	19	4	1	2	0.26	0	0.51 (0.36)
5	96	7	0	1	0	0.22	0	0.16 (0.10)	49	13	1	1	2	0.27	0	0.51 (0.35)
7	77	7	0	1	0	0.22	0	0.17 (0.10)	36	10	1	0	2	0.26	0	0.47 (0.29)

3.6.6 WRV Fits – Weighted Linear Mixed-Effects Modelling

My measure of within-individual rhythmic variability (WRV) in a singer's repetitions of PV α is the mean sample standard deviation across all song unit positions. Given that the standard error in a sample standard deviation of a normal distribution depends approximately linearly on the size of the standard deviation itself (Equation 3.6, Harding *et al.*, 2014), heteroscedasticity might be expected.

Equation 3.6

$$SE_{\sigma} \approx \frac{\sigma}{\sqrt{2(N-1)}}$$

In these circumstances one recommended approach is to weight each observation with its variance (https://online.stat.psu.edu/stat501/lesson/13/13.1), i.e., the square of the standard error. To estimate the variance we can use the standard result (Equation 3.7), taking into account the number of song unit positions (NSU_{α}). Providing that each standard deviation is independent then standard errors of WRV would be given by

Equation 3.7

$$SE_{WRV} \approx \frac{WRV}{\sqrt{2NSU_{\alpha}\left(N_{phrase}-1\right)}}$$

However, given that IOIs may be neither independent (§2.7.5) nor normally distributed, I estimated standard errors using bootstrapping (Sokal and Rohlf, 1995), with 5000 resamplings. The bootstrapped errors lay close to but tended to exceed those errors estimated using assumptions of normality and independence (Equation 3.7, Figure 3.16), and were used to weight the linear mixed-effects model ($weight = (bootstrapped\ error)^{-2}$) for all results reported in the main text (§3.3.3).

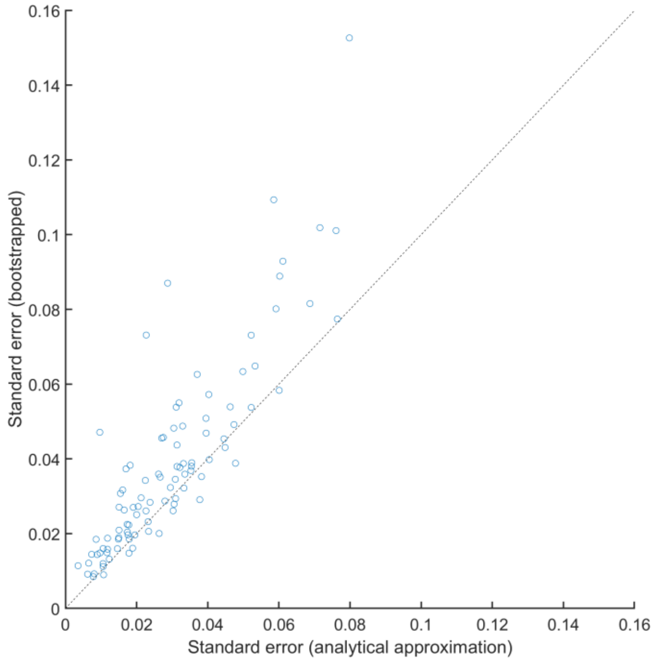


Figure 3.16 Bootstrapped standard errors compared with errors calculated using analytical approximation (Equation 3.7), for all observations ($N_{phrase} \geq 5$). The figure does not show two outliers with large errors, which also fall close to the plotted y=x line).

I also tested the use of observation size alone as a weighting factor (observation weight = N_{phrase}). In fact, heteroscedasticity was not observed even in unweighted fits (Figure 3.17), and I found that the main findings from the modelling were not dependent on the weighting regime (Table 3.10).

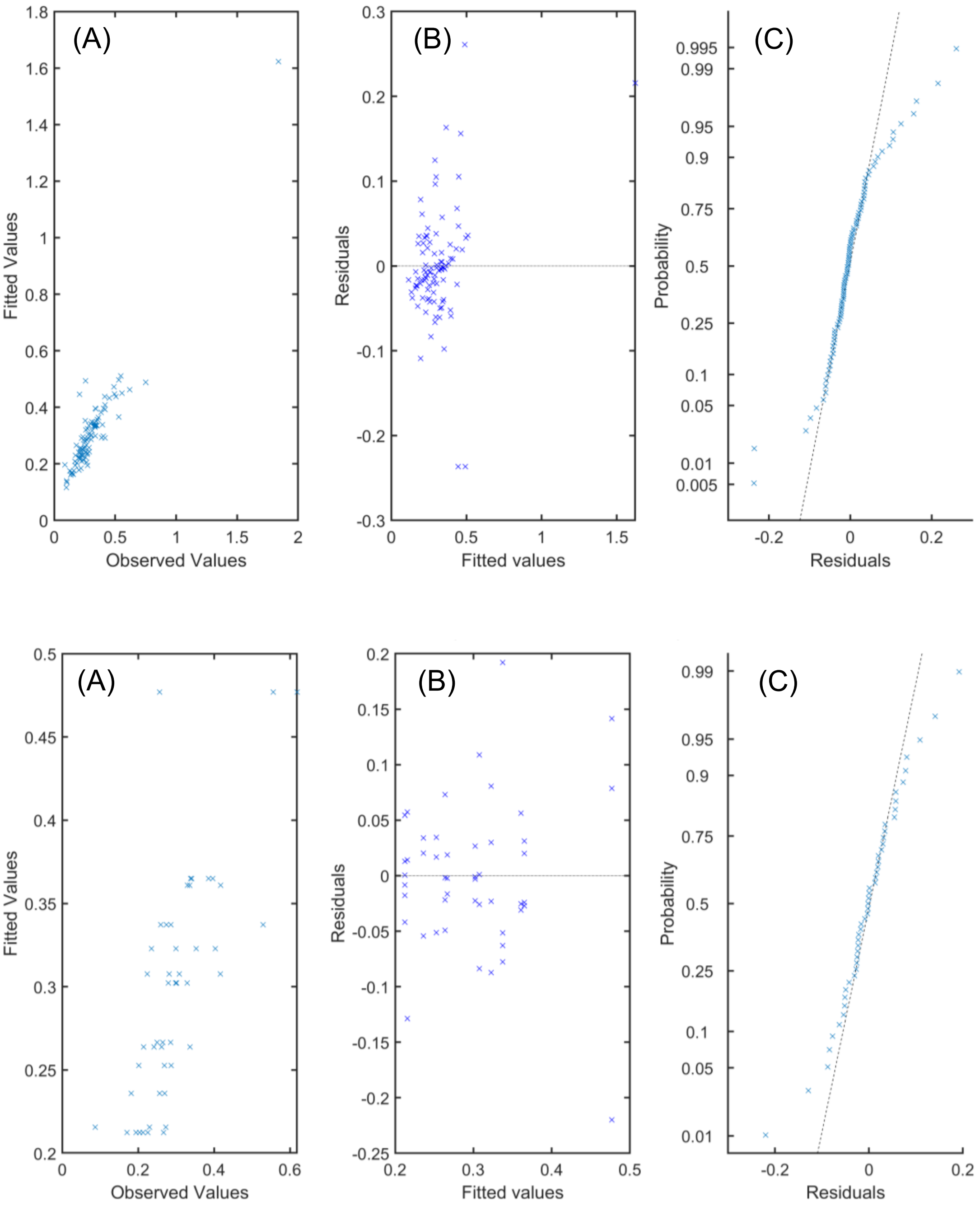


Figure 3.17 (A) Fitted versus observed values, (B) Residuals versus fitted values, and (C) Normal probability plot of residuals for unweighted linear mixed-effects modelling of hypotheses H1 (upper panels, dataset DS_{H1}) and H2 (lower panels, dataset DS_{H2}).

Table 3.10 The effect of weighting on the significance of hypothesis testing for H1 and H2 obtained using linear mixed-effects modelling (relaxed datasets), with three different weighting regimes. (A) Unweighted modelling. (B) $Observation\,weight = N_{phrase}$ *(C)* $Observation\,weight = (bootstrapped\,error)^{-2}$. *For each weighting regime, I have reported the number of levels of the fixed effect significantly different to the grand mean, the value of the intercept, and the standard deviation of the random effect.* R^2 *values are ordinary (adjusted). Significance levels for fitted coefficients coded as* *** $P < 0.001$, ** $0.001 \leq P < 0.01$, * $0.01 \leq P < 0.05$.

	Do singers differ significantly from the grand mean? (H1) (NN_{singer} = number of levels of fixed effect with coefficients significantly different to zero)							**Do PVs differ significantly from the grand mean? (H2)** (N_{PV} = number of levels of fixed effect with coefficients significantly different to zero)						
Weighting	N_{singer} ***	N_{singer} **	N_{singer} *	Intercept	Random effect sd	R^2	AIC	N_{PV} ***	N_{PV} **	N_{PV} *	Intercept	Random effect sd	R^2	AIC
A	0	0	0	0.32	0.22	0.85 (0.84)	-53.6	1	1	1	0.30	0	0.52 (0.36)	-97.2
B	0	0	0	0.29	0.071	0.24 (0.19)	-71.2	3	0	1	0.31	0	0.61 (0.48)	-99.6
C	0	1	0	0.22	0	0.16 (0.10)	-142	1	1	2	0.27	0	0.51 (0.35)	-81.9

3.6.7 WRV Fits – Signal to Noise Ratio and Aggregation Strategy

Song unit SNR varied (SI §3.6.1, Table 3.6), so I checked that WRV was not dependent on onset detection uncertainty. Given the main findings (no consistent differences among singers, consistent differences among PVs), I grouped observations by PV and examined whether WRV depended on mean SNR (calculated across all phrases in each observation). The absence of any general trend in WRV with SNR across different PVs (Figure 3.18A), is consistent with the threshold detector working well to the lowest levels of SNR used in this study. Similarly, I examined whether WRV depended on the number of phrases in an observation, again grouping observations by PV (Figure 3.18B), and again there was no apparent general trend in WRV with observation size. Finally, I checked that the WRV was not biased by the temporal spread of the phrases contained within each observation. If song evolution happens on short timescales, then observations whose phrases were collected closer in time would likely have a lower WRV than one whose phrases were more widely dispersed. As a measure of temporal spread I calculated the mean pairwise Euclidean distance of the phrase start times. Once again, no general pattern of WRV was seen across PVs (Figure 3.18C).

This visual assessment was checked for H1 and H2 using linear mixed-effects modelling on the relaxed datasets, including SNR, observation size and temporal spread as fixed effects alongside Singer and PV respectively (Table 3.11). Including these parameters did not change the main findings. For H1, the *p*-values for the new fixed effects fitted singly showed that in no case could the null hypothesis, i.e., that the relevant coefficient was zero, be rejected (Table 3.12). For H2, the fixed effects of observation size and temporal spread were found to be statistically significant ($P < 0.01$), but small in size (Table 3.13). In addition to providing confidence that the WRVs derived from observations collected from different singers are not biased by SNR, observation size or song session duration, these results suggest that any evolution of the phrase rhythms within the timescale of these song sessions (up to three hours) is minimal.

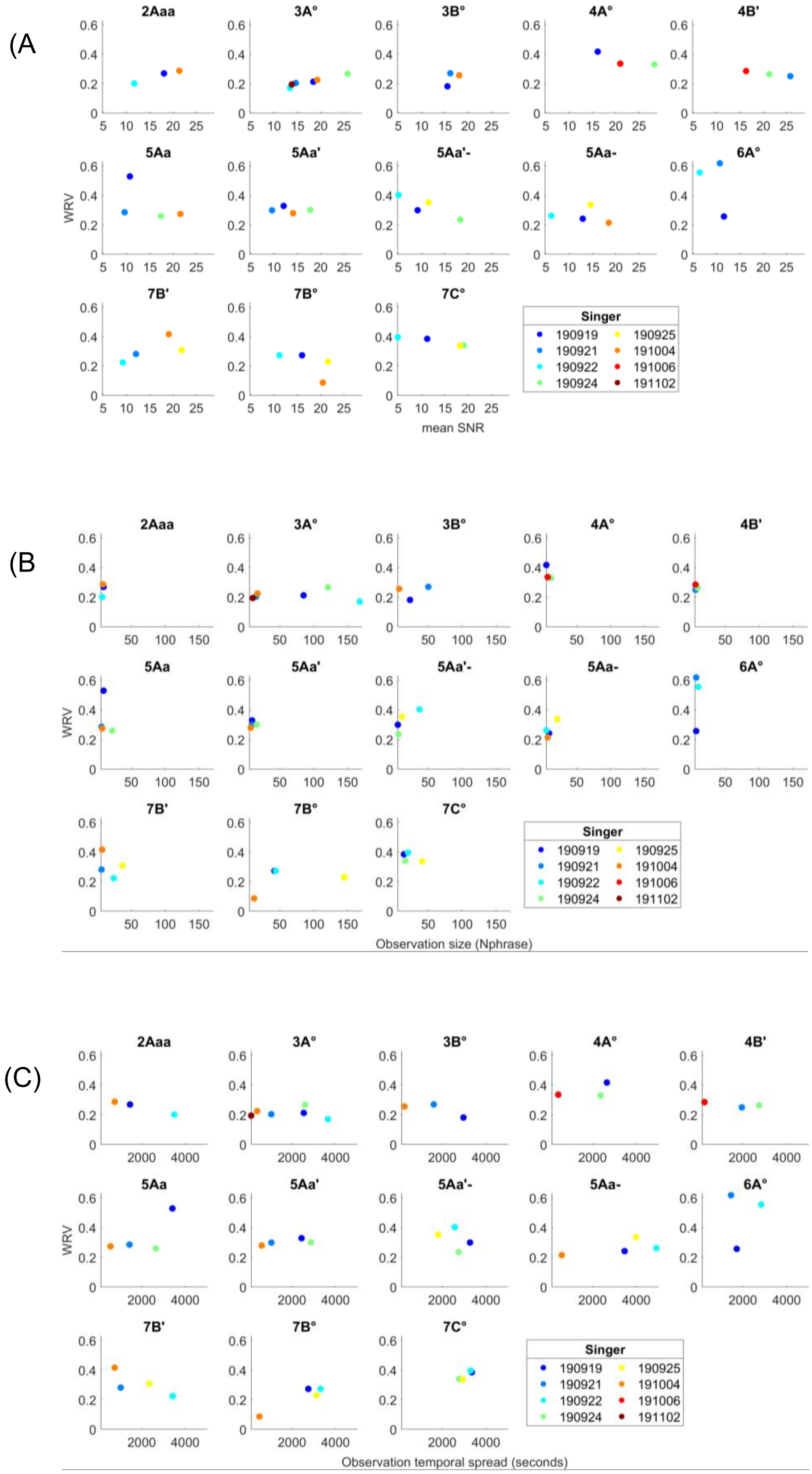


Figure 3.18 For a given PV, WRV does not appear to vary in a systematic way against (A) mean SNR, (B), Observation size, (C) Observation temporal spread. When tested with linear mixed-effects modelling of H2, the dependence of WRV on observation size and observation temporal spread was found to be statistically significant, but small (Table 3.13). Including these parameters in the modelling of H2 did not alter the interpretation of the results discussed in the main text (§3.3.3).

Table 3.11 Model specifications for testing WRV dependence on mean SNR levels, observation size and temporal spread.

Model	**Specification of Fixed and Random Effects**	**Dataset**
1	WRV ~ 1 + Singer + (1|PV)	DS_{H1}
2	WRV ~ 1 + Singer + meanSNR + (1|PV)	DS_{H1}
3	WRV ~ 1 + Singer + N_{phrase} + (1|PV)	DS_{H1}
4	WRV ~ 1 + Singer + temporalSpread + (1|PV)	DS_{H1}
5	WRV ~ 1 + PV + (1|Singer)	DS_{H2}
6	WRV ~ 1 + PV + meanSNR + (1|Singer)	DS_{H2}
7	WRV ~ 1 + PV + N_{phrase} + (1|Singer)	DS_{H2}
8	WRV ~ 1 + PV + temporalSpread + (1|Singer)	DS_{H2}
9	WRV ~ 1 + PV + N_{phrase} + temporalSpread + (1|Singer)	DS_{H2}

*Table 3.12 Linear mixed effects modelling of H1 for DS_{H1}, including standardized mean SNR, observation size, and temporal spread (Models 2–4) as fixed effects added individually to Model 1 (Singer as fixed effect and PV as random effect), showed no support for accepting that these parameters were responsible for the observed variation in WRV. Significance levels for fitted coefficients coded as *** $P < 0.001$, ** $0.001 \leq P < 0.01$, * $0.01 \leq P < 0.05$.*

	Mean SNR		**Observation Size (N_{phrase})**		**Temporal Spread**		**Singer fixed effect** (N_{singer} = number of levels with coefficients significantly different to zero)			
Model	β_1	*p*-value	β_2	*p*-value	β_3	*p*-value	N_{singer} ***	N_{singer} **	N_{singer} *	AIC
1	–	–	–	–	–	–	0	1	0	-142.1
2	4.60×10^{-4}	0.966	–	–	–	–	0	1	0	-140.1
3	–	–	-6.07×10^{-4}	0.894	–	–	0	1	0	-140.1
4	–	–	–	–	0.0270	0.101	0	0	0	-142.8

Table 3.13 Linear mixed effects modelling of H2 for DS_{H2}*, including standardized mean SNR, observation size, and temporal spread (Models 6–8) as fixed effects added individually to Model 5 (PV as fixed effect and Singer as random effect), showed that although observation size and temporal spread were statistically significant, their effects were small and did not alter the overall main finding. Including both observation size and temporal spread (Model 9) did not improve the model fit, most likely due to the weak correlation of these characteristics (*$R^2 = 0.0917$*,* p-*value =0.0344). Significance levels for fitted coefficients coded as* $*** P < 0.001$, $** 0.001 \leq P < 0.01$, $* 0.01 \leq P < 0.05$.

Model	**Mean SNR**		**Observation Size (N_{phrase})**		**Temporal Spread**		**PV fixed effect** (N_{PV} = number of levels with coefficients significantly different to zero)			**AIC**
	β_1	*p*-value	β_2	*p*-value	β_3	*p*-value	N_{PV} ***	N_{PV} **	N_{PV} *	
5	–	–	–	–	–	–	1	1	2	-81.9
6	0.0102	0.363	–	–	–	–	1	2	1	-80.7
7	–	–	0.0167	**	–	–	2	1	1	-89.7
8	–	–	–	–	0.0244	**	2	1	1	-89.3
9	–	–	0.00999	0.241	0.0125	0.323	2	0	2	-88.7

3.6.8 Individual and Population Mean Phrase Rhythms

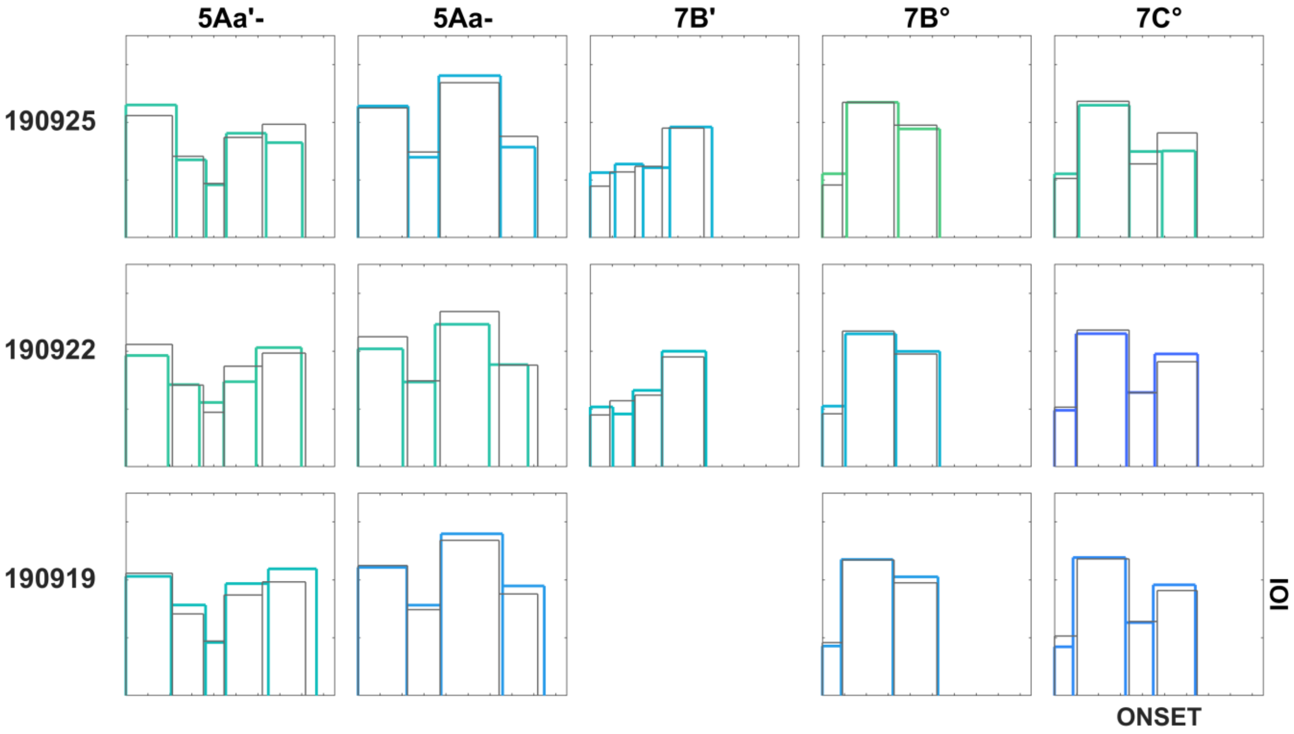


Figure 3.19 Individual mean phrase rhythms (colour-coded by Rhythmic Individuality) differ from population mean phrase rhythms (medium grey). In this subset of typical observations, the mean phrase rhythms produced by three singers for specific PVs can be compared by reading down a column. In each chronotonic diagram IOI is plotted against onset time on a common axis (maxONSET=19s, maxIOI=7s). The observation for Singer 190919, PV 7B' is excluded due to the small sample size ($N_{phrase} = 3$)*. Systematic rhythmic differences among individual singers may be seen by comparing the shapes of the colour-coded mean phrase rhythms in the same column, with warmer colours indicating that individual mean rhythms lie further from the population mean rhythms (higher Rhythmic Individuality). Results from linear mixed-effects modelling (§3.3.4) offered support for the hypothesis (H4) that levels of Rhythmic Individuality displayed by individual singers are consistent across PVs in their repertoire. In this subset, Singer 190925 is significantly higher (p-value < 0.05) than the grand mean, as suggested by the preponderance of greens over blues.*

3.6.9 Simulations of Rhythmic Individuality

I used Monte Carlo simulations (§2.5.1) to evaluate whether observed correlations between rhythm distance and singer difference matrices can be used to gauge systematic rhythmic difference, measuring pairwise distances with the scaled Euclidean Distance (§2.7.6). To encompass the maximum WRV discrepancies observed in the full dataset, I simulated phrases from three singers across an expanded WRV range. Under these conditions I found that at extreme ratios of WRV the correlations were significant (p-value < 0.05) by the Mantel test, even for phrase repetitions based on the same rhythmic template, showing that the p-value alone cannot be used as criterion for systematic difference (Table 3.14). However, when there was no systematic rhythmic difference, correlations remained low (< 0.15) whatever the WRV discrepancy. Further, the correlation was found to be approximately linearly dependent on systematic difference measured either with the input change to the IOI vector, or with Rhythmic Individuality (§3.2.5) (Figure 3.20AB). So, it is justifiable to use correlation as a population-level gauge of systematic difference. Finally, the steep dependence of Rhythmic Individuality on the input change to the IOI vector (Figure 3.20C) confirms that Rhythmic Individuality, which can be derived from observations, is a useful measure of an individual singer's distance from the population norm.

Table 3.14 Systematic Rhythmic Differences must be diagnosed through values of the correlation, not solely from the p-value *returned by the Mantel test. Monte Carlo simulations of phrase repetitions (N=10), for a population of three singers, showed that high levels of WRV discrepancies in the absence of systematic difference can lead to* p-values *< 0.05, although as these were associated with low correlation values (< 0.15) higher correlation values are diagnostic of systematic difference. IOI vector from 4A°. Two WRV ranges were tested (A: 0.05–0.3 s; B: 0.25–0.65 s). The phrase rhythms of one 'rhythmically distinctive' singer were shifted systematically by adding the vector k*[1 -1 1 -1 1 -1] (A: k = 0–0.3; B: k = 0–0.9, providing a range of systematic rhythmic difference leading to values of Rhythmic Individuality c.0.2–1 for the distinctive singer). For each scenario, 100 model runs were performed, with 1000 permutations of matrices for the Mantel test.*

Mantel test results for simulated phrase rhythms		**Systematic Rhythmic Difference**	
		None	Moderate (Rhythmic Individuality ~ 1 for distinctive singer)
WRV discrepancy	None	p-value > 0.3 correlation < 0.01	p-value ~ 1×10^{-6} correlation > 0.45
	Maximum observed	$1\times10^{-6} < p$-value < 0.05 correlation < 0.15	p-value ~ 1×10^{-6} correlation > 0.35

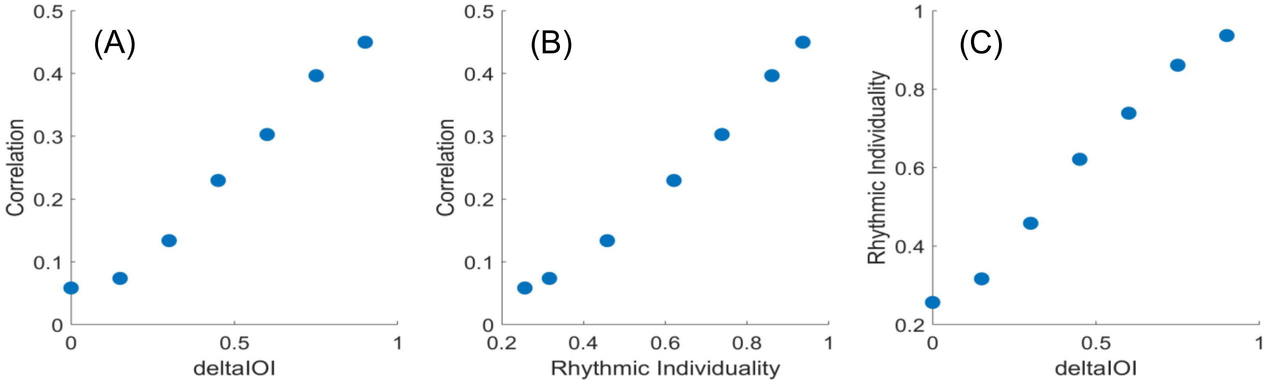


Figure 3.20 The correlation returned by the Mantel test for simulated phrase rhythms was approximately linearly dependent on the level of rhythmic distinctiveness, assessed as (A) input to the simulation as the amount of systematic difference applied to the rhythmic template of one singer ('deltaIOI'), or (B) the calculated Rhythmic Individuality from the simulated phrase rhythms. (C) The observed dependence of Rhythmic Individuality on deltaIOI for this singer confirms that it is a useful probe of an individual's level of difference to the population norm. Mantel test results (mean values from 100 runs of the model) are from simulated phrase performances of three singers, each with a different value of WRV (0.25, 0.45, 0.65 s). Each singer's performance was of 10 phrases of PV 4A°, with varying levels of systematic rhythmic difference (deltaIOI) applied to one singer only.

Chapter 4 Towards A Multispecies Heterophony

Abstract. In this chapter I provide a commentary on my portfolio, discussing my (re)compositional processes alongside aesthetic and ethical goals. I also aim to show how these processes and goals are integrated with the overall theme of this thesis: the representation and misrepresentation of repetition and variability in the rhythms of humpback song. All of my musical works were inspired and informed by the sounds and structures of humpback whale song, some were created in collaboration with other human musicians. I begin with a brief introduction to the circumstances surrounding the composition and performance of the portfolio works. I then describe my engagement with humpback song through the processes of de-compositional listening, transcription and recomposition, during which I created and manipulated multiple representations of humpback song. From there I critique the process of objectification, drawing on insights from Critical Theory and postcolonial environmental ethics into the logics of domination and extractivism that continue to structure many of the interactions between humans and other-than-human worlds. This has allowed me to deepen my earlier analysis of 'whale music' in which I argued that certain works could be read as manifesting human exceptionalism, and that others avoided it by embodying a 'critical anthropomorphism.' With this critique, I will challenge the presumption that no ethical questions arise from the musical use of other-than-human sounds. I also develop a form of 'aesthetic rationality' through a set of strategies for those composers who are concerned to avoid binary anthropocentrism in their compositional processes and works. In my own case aiming at critical anthropomorphism and following these strategies has led to what I interpret as an emergent 'multispecies heterophony', an overlapping non-hierarchical sounding of difference. This temporal form, on the one hand, echoes the asynchronous chorus of the collective singing of humpback whales, and, on the other, provides a promising musical model for a much-needed ecological thinking.

4.1 Introduction

In an era characterized by mounting global crises it is easy to focus on existential threats to humanity, and ignore implications for the lives of other animals. It is also easy to focus on the practicalities of the massive changes to human behaviour that are undoubtedly required for future multispecies flourishing, forgetting that those being asked to change need to care about doing so. Whereas relentless bad news delivered in "information dump" mode (Morton, 2018) seems ill-suited at generating such motivation, the arts may be better-placed to nurture more ecologically-attuned ways of thinking and feeling, and to promote what philosopher Donna Haraway (2016) terms a response-ability in audiences. The non-representational arts, in particular, may offer a less didactic route towards re-imagining the relationships between human members of industrialized societies and other species, thereby challenging the human exceptionalism that underwrites the current era, commonly referred to as the Anthropocene.[15]

Contrasting with the ceaseless flow of information presented in words and images, music appears an appropriate channel for opening humans up to a more embodied appreciation of their interdependence with other animals, to "multispecies worldings" (van Dooren *et al.*, 2016). Non-verbal music cannot dump information, but can open ears to difference through suggestion or evocation. Its ability to promote transformation is well-evidenced by the 'whale music' (Ritts, 2017) and the field recordings of humpback whale song that inspired it, that in the 1970s became the soundtrack to Greenpeace's successful campaign for a moratorium on commercial whaling (Burnett, 2012). However, musicians and audiences alike have often oversimplified the differences and similarities between human musics and cetacean vocalizations (Feldman, 2021; South, 2022; Susam-Saraeva, forthcoming). At one extreme, the

---

[15] Many scholars (e.g., Crist, 2013; Haraway, 2016; Moore, 2017) have argued that the term 'Anthropocene' in fact manifests a human exceptionalism, reinforcing the hyperseparation of human and other-than-human. It also unfairly groups all humans together as equally responsible for the damage that has been unequally inflicted on the planetary ecosystems. Alternative suggestions include the 'Capitalocene'.

exaggeration of difference leads to the exoticization of whale sounds, with whales themselves regarded as mysterious underwater aliens (‘anthropodenial’). At the other, difference is eliminated through an imposed identity in which whale sounds are assimilated into conventional human musical contexts, the whales viewed as though they were fully human musicians (‘naïve anthropomorphism’). Both extremes appear to share a “binary anthropocentrism” (Martinelli, 2009, p. 19) that fails to value or “understand [other] animals on their own terms” (de Waal, 1999, p. 273). I have previously analysed examples of whale music that avoided this binary thinking (South, 2022), adopting biologist Gordon Burghardt’s ‘critical anthropomorphism’ (1991) to describe their treatment of difference and similarity. Where difference was taken seriously and combined with the temporal entwining of human musicking with humpback whale song field recordings, I found that such works achieved an aesthetics of “difference without distance.”

Whether or not manifesting tendencies to anthropodenial or naïve anthropomorphism, the commodification of whale sounds has sometimes been used to satisfy human interests (Ritts, 2017), generating “new aural tapestries to get high with” (Feldman, 2021). Although the human motivations behind the 1970s production of whale music are likely consistent with my own desire to nurture ecological modes of thinking and feeling, there are uncomfortable parallels with industrial capitalism’s rendering of ‘nature’ into ‘resource’ that rests on human exceptionalism. Just as it has been a short step from commercial whaling to commercial whale-watching as profit-making enterprises (Reeves, 2015), it might be considered a short step from extracting oil to extracting sound. In each case, the interests of *Homo sapiens* comes first, with the whale managed or dominated by humans; its oil, attractive behaviours, and vocalizations regarded as resources open for appropriation. I extend the comparison. Historical musicologist Rachel Mundy (2018) has drawn attention to analogies between the collection of birdsong by biologists and naturalists, and the collection of Indigenous musics by early comparative musicologists. The ‘subordinated’ (Plumwood, 2002) groups are different, but Mundy links anthropocentrism with ethnocentrism as part of a network of self-reinforcing hierarchical binaries, a position shared with ecologically oriented feminist, critical and postcolonial thought (e.g., Plumwood, 1993, 2003; Agamben, 2004; Huggan, 2004). Moreover, the extractivist logic of colonialism has been identified operating in multiple domains: “Not just substances and labour, but forms of creative practice and knowledge, too, are extracted, exchanged, and monetized” (Clark, 2021, p. 7). Putting this together, I arrive at the conclusion that the creation of whale music not only risks manifesting binary anthropocentrism, but may also end up mirroring the colonial and extractivist structures common to modernity’s anthropocentric treatment of nature and ethnocentric treatment of Indigenous musics.

The Indigenous sound studies scholar Dylan Robinson has criticized ethnocentric approaches to Indigenous musics for their “hungry listening [that] prioritizes the capture and certainty of information over the affective feel, timbre, touch, and texture of sound” (2020, p. 38). In Chapters 2 and 3 I presented a scientific mode of engagement with humpback whale song that might be described in the same way. In this chapter I move to a more reflective mode of writing in order to critique my own practical re-engagement with these other aspects of humpback sounds, analysing my own compositional processes to show how I attempted to maintain a ‘fidelity’ to humpback song, pursuing a critical anthropomorphism that “embraces the full complexities of sameness and difference” (Taylor, 2017b, p. 36). With the term fidelity I intend to convey the dual connotations of (1) the accuracy with which my works represent features of humpback song,[16] and (2) the loyalty out of which I show my respect for its singers. This second point brings us to the explicitly ethical aspect of my work. As ‘interspecies mediator’ between whale singers and human listeners, and against the backdrop of continued existential threats to cetaceans (§1.3.1), I acknowledge a responsibility to attend “to relations of power

---

[16] This is something held in common with the scientific research I report in Chapters 2 and 3.

and asymmetry” (Susam-Saraeva, forthcoming). Regarding both meanings of fidelity, I have been interested in the avoidance of ‘misrepresentation’. The misrepresentation of whale song and its singers may not cause (direct or indirect) harm, but I accept that the disrespectful treatment of song can amount to an ethical wrong, following current work in animal ethics on agency, dignity and the intrinsic value of other animals (Humphreys, 2016; Wichert and Nussbaum, 2019; Abbate, 2020; Nussbaum, 2022).

In this chapter my focus is on the representation of temporal and structural aspects of humpback whale song for the purposes of composition and performance. After a brief account of the circumstances of composition of my portfolio works (§4.2), I turn to my initial encounter with humpback whale song as apprentice bioacoustician, learning to categorize the sounds of a specific singer (§4.3). The process of ear-training was dependent on the spectrograms that are used universally among bioacousticians and share formal properties with Western musical notation. I question whether such representations can avoid anthropocentrism, and further, whether the scientific objectivity sometimes claimed for them is either attainable or desirable. Ethical questions arise regarding fidelity to structure. For example, what balance should be struck between repetition and variability, and does explicitly granting a certain creative autonomy to the human performers of my works implicitly overrule any possessed by the original humpback singers? The representation of song via transcription into musical notation (§4.4) also brought up anthropomorphic assumptions of musical rhythm and metre that I aimed to identify and avoid in my own work. My portfolio works combine transcriptions with field recordings edited in Digital Audio Workstations, thereby ‘objectifying’ humpback song, and potentially risking misrepresentation through decontextualization and consequent domestication or exoticization (§4.5). I also question whether excessive compositional manipulation exploits a power imbalance and fails to respect the agency and intrinsic value of other animals.[17] However, I have no desire to engage in the kind of "ideological gate-keeping" (Bellman, 2011, p. 427) that would condemn all use of other-than-human sounds. No matter how composers work, I accept that in general the incorporation of other-than-human sounds in human music is likely to increase human appreciation of them and the music-like capacities of their producers. Hence, I turn to the reception and interpretation of whale music by human audiences, and close this chapter with a reading of my own works (§4.6), focusing on their heterophonic form. This form, that well describes the asynchronous sounding together of similar-sounding variants found in the collective singing of humpback whales, is one I chose deliberately in some of my works, and emerged unintentionally in others. I propose that a nonhierarchical 'multispecies heterophony' allows the exercise of an aesthetics of "difference without distance," (South, 2022) perhaps raising the possibility of an ethics of "proximity without indifference" (South, 2022). This is to say, an ethics in which I acknowledge my living co-presence in shared habitats with other sentient agents as "dense ecologies of selves" (Kohn, 2013, p. 193), but where this awareness does not lead to the deadness of familiarization: I must still care, I must listen not just to, but with, each Other.[18]

4.2 Portfolio Works

I composed my portfolio works (Portfolio Contents, Table 4.1) whilst carrying out the scientific research described in Chapters 2 and 3, with the earliest pieces coming from a period of obsessive listening to a single humpback whale song recording made in September 2015 (Singer 150918, included in Portfolio). Whilst learning how to categorize the humpback sounds

---

[17] In this chapter I consider the ethical implications of the intrinsic value of other animals. In some cultures there are additional considerations arising from relationships between human and other animals, which could serve as further constraints on the free use of other-than-human sounds. For example, (Andean, 2014) reports that ethnomusicologist Steven Feld was "reprimanded for a casual reference to the call of a particular bird, unaware that the local culture believes this call to be the voices of ancestors, and was therefore to be treated with deference" (p. 177). I am not aware of any such cultural constraints on the use of humpback whale song.

[18] I capitalize 'Other' to emphasize that all other sentient agents possess a radical alterity that exceeds what I can grasp, and to signal my support for arguments that a Levinasian ethics can be extended beyond the human (Calarco, 2008; Atterton, 2011, 2018).

for scientific purposes (§4.3.1), I was also beginning to transcribe them for my instruments. The difficulties involved in reproducing the sounds of the humpback gave me the title for '(im)possible gestures', composed for the Scottish Clarinet Quartet. With this piece, I aimed at creating an experience for the audience in which they were surrounded by an "asynchronous chorus" (Herman, 2017) of clarinet-singers, who gradually transition from unsynchronized mimetic material into more conventional uses of their instruments, and more conventional musical interaction. It was important to me that my score followed the humpback song structure and empirical rules of variation I had observed.

Table 4.1 Works contained in portfolio (see Portfolio Contents*) and discussed in this chapter (§§4.3.2, 4.3.4, 4.4.3, 4.5.2, 4.6), listed in chronological order of composition.*

Work	**Instrumentation**	**Composer**	**Performers**
'(im)possible gestures' (2019)	Four Clarinets optionally doubling Bass Clarinet	Alex South	Scottish Clarinet Quartet
'The Path of the Unseen Whale' (2019)	Bass Clarinet, Live Electronics and Fixed Media	Alex South	Alex South, Humpback Singer 150918
'Underwater Rain (Farehau Humpback, 26.10.2019)' (2020)	Electroacoustic (Field recordings, Bass Clarinet, Piano)	Alex South	Alex South, Humpback Singer 191026
'Entanglement' (2020)	Electroacoustic (Viola, Clarinet, Bass Clarinet)	Alex South with Katherine Wren	Alex South, Katherine Wren
'Whale, Bow, Echo' (2021)	Violin, Cello, Fixed Media, Projections	Alex South with Sequoia	Sequoia, Humpback Singer 190924
'Submergence' (2022)	Clarinet, Fixed Media	Jan Foote, with realization of solo clarinet by Alex South	Alex South

Around this time, I was reading Philip Hoare's *Leviathan or, The Whale* (2008). In homage to this wonderful book, and also to refer to the effects of anthropogenic noise pollution on cetaceans (e.g., Tyack, 2009), I used Hoare's characterization of an Indigenous view of flukeprints as the title for my next piece, 'The Path of the Unseen Whale'. Written for solo performance, it combines electronic processing of live bass clarinet alongside fixed media tracks constructed from the gaps between humpback sounds on the 2015 recording, re-using some of the transcriptions made for '(im)possible gestures'. As in that piece, there is a transition from humpback-like imitation to a human melody (built from the pitches I had heard in the field recording), the imitated humpback song gradually transformed into wailing alarm sirens.

By the time of composition of 'Underwater Rain', live performances had been halted due to COVID-19. I was asked to contribute an electroacoustic piece to a pack of mindfulness resources for returning university students. I had recently been struck by the expressive qualities I heard in a humpback song possessing "shifting themes" (Payne and Payne, 1985) containing progressive changes in pitch and duration (Singer 191026, included in Portfolio, Figure 4.1, Figure 4.2). Aided by visual representations and quantitative measurements of rhythm (Chapter 2), I was becoming more aware of the subtle variations of timing that are hard to hear at the typical tempo of humpback song, and I explored these in 'Underwater Rain', setting the flexible whale song against a more rigid human bassline (§4.4.3).

With the end of the first lockdown it was possible once again for musicians to rehearse together. I spent time in the studio with violist Katherine Wren, improvising on the transcriptions I had made for 'Underwater Rain'. I was interested in performing the same material at different tempos in an effort to experience a range of temporal dynamics: I wondered what they might be like for the humpback singer, whose perceptual present might be assumed to be longer than ours (Thompson, 1981). At slow tempos the song transcription elicited dirge-like responses; at faster tempos our rhythms became less measured, more embodied, almost dance-like. I made

‘Entanglement’ from our recorded improvisations, when we decided to submit a piece for an album marking the annual ‘Remembrance Day for Lost Species’. The lamenting quality of our slower improvisations fused with my awareness of the risks to cetaceans posed by entanglement in fishing gear (e.g., Moore, 2021), and this in turn influenced the way I layered up and mixed the recordings.

Collaborative processes were also at the heart of ‘Whale, Bow, Echo’, composed for violin and cello duo Sequoia. Initially, I had planned to explore rhythms driven by constraints of embodiment and bowing technique, but during the first workshops the players’ pleasure in responding mimetically to a field recording drove us in a new direction. Structurally the first part of the piece resembles the first part of ‘(im)possible gestures’, with the two players performing transcriptions of humpback song according to a codified set of possible variations. The freedom of these live variations is juxtaposed with more mechanical repetitions created in a Digital Audio Workstation, which steadily grow, somewhat like the echoes in ‘The Path of the Unseen Whale’. They are supplemented with additional resonating layers which emphasize the harmonic consonance found in the original whale song. Finally, emerging out of this mounting sonic opacity the original field recording is heard. It is a tribute to the mimetic skills of the Sequoia players that is hard to hear when the humpback starts to sing.

The last work in my portfolio is more loosely based on humpback whale song. I had provided composer Jan Foote with a schematized humpback song structure and recorded bass clarinet multiphonics for him. He wrote ‘Submergence’, a piece for four bass clarinets, in which each humpback sound is represented by unique combinations of multiphonics. These combinations continually morph as the piece progresses, evolving much more rapidly than happens in a typical performance of a humpback whale song, but perhaps representative of what might take place across a whole season. Still in lockdown conditions, I made and mixed a studio recording, playing all of the parts. We had discussed a version in which this recording would form a backing for an additional live part, and when the opportunity for live performance arrived I found myself working mainly by ear, improvising tremolos, colour trills and multiphonics to ruffle and reanimate the recorded surface. My notated part and a recording of this version for live performance is included in my portfolio.

4.3 The Time of Representation

My portfolio of compositions was shaped by multiple listenings to the sounds of whales and humans, including commercial whale song recordings (Payne, 1970, 1977; Rothenberg, 2015) and dozens of existing works of whale music, i.e., human music influenced by or incorporating field recordings of whale song (Appendix A). Above all, for a period of several weeks towards the beginning of my PhD, I repetitively listened to an 11-minute humpback song recording (Singer 150918, included in Portfolio), made in September 2015 off the coast of Mo'orea, French Polynesia, as part of learning how to code humpback song for the purposes of quantitative bioacoustical analysis (§2.2.1). This pre-compositional "listening for objectivity" (Mundy, 2018) was visually accompanied and aided by the rendering of sound into spectrograms. Here I focus on this initial training in listening and coding, discuss how the resulting symbolic representations of humpback song ('song codings') were used for compositional purposes, and critique the use of spectrograms, questioning their assumed objectivity and assessing their tendency to objectification.

4.3.1 Categorical Listening and Song Coding

The transcription of the vocalizations of humpback song into hierarchically structured symbolic representations (song codings) has played a central role in the growth of knowledge of humpback whale behaviour, cognition and culture. Since the start of scientific interest in humpback whale song, researchers have described the population-wide 'songs' as built of a characteristic sequence of 'themes', each consisting of repeated 'phrases' and 'subphrases'

(Payne *et al.*, 1983).[19] Phrases and subphrases consist of distinctive patterns of a few ‘units’, i.e., those individual sounds that are “continuous to our ears” (Payne and McVay, 1971). To my ear the complex pitch and timbral profiles of humpback units led to my hearing them less as melodic fragments or notes, and more akin to the sonic gestures theorized in electroacoustic music, the “spectral shapes and shape-sequences created by the energy of physical and vocal articulation” (Smalley, 1986). The transition from one theme to another often involves an aurally striking change in the units employed, in addition to changes in their arrangement. As we have seen (Chapters 2, 3), in humpback whale song the repetition of phrases within a theme includes different kinds of variability, and furthermore the theme sequence is by no means fixed. What I emphasize here is that there is also considerable variability (e.g., in pitch, timbre, duration, subunit structure) in the performance of units that scientists group together as one unit type.

In learning to code, I had first to learn to listen, to acquire a set of perceptual categories, establishing experiential acoustic prototypes (Rosch, 1973). Each unit type was given a name (e.g., ‘cry’, ‘moan’), where possible taken from a pre-existing dictionary derived from the analysis of hundreds of songs from South Pacific humpbacks (Garland *et al.*, 2017), combined with continual reflection and consultation on where to draw the boundaries, an intersubjective re-negotiation of the borders between one unit type and another as more units were encountered. In my initial encounter with the French Polynesian humpback song, increasing familiarity with the recording led me towards hearing increasing degrees of differentiation among units (e.g., moans might be modulated, ascending, or descending). This ‘ear training’

---

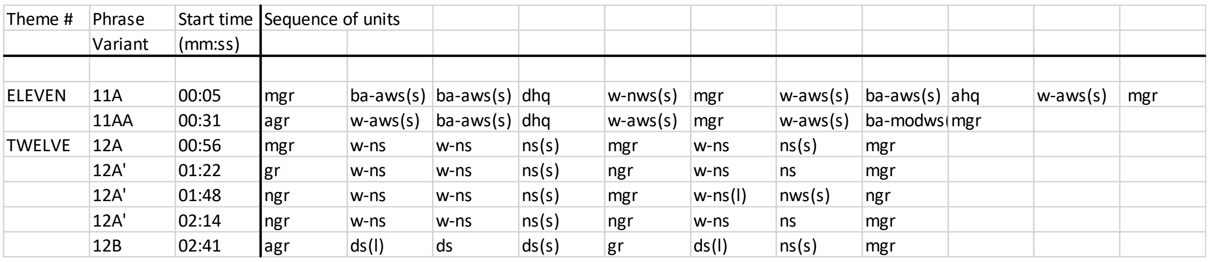
[19] This terminology may appear anthropomorphic. It seems to me that the use by ethologists and zoömusicologists does not display a naïve anthropomorphism: there are features that these vocalizations share with much human music termed song, for example, repetition on multiple hierarchical levels (§1.2.2). However, it is likely that the precise terminology of themes and phrases owes much to the Western musical training of Katy Payne, a scientist who has played an important and arguably under-recognized role in the interpretation of humpback whale song (Mundy, forthcoming).

might be summarized as an iterative process of listening, describing, listening again, and refining descriptions in order to decide what counted as ‘the same’, where nothing was ever really identical, and where my descriptions relied on my own human perceptual and cognitive faculties. Evidently this reliance entails a certain degree of anthropocentrism: I could never be certain that what I or another human counted as the same would correspond to what was counted the same by a humpback whale. But I would argue that this is a form of “default anthropocentrism” (Martinelli, 2009, p. 19) that is unavoidable and unproblematic, a consequence of species-specific *Umwelten*.[20] As Martinelli points out, it does not necessarily lead us to a “binary anthropocentrism,” which divides humans from all other animals and “puts the observer, and the group the observer belongs to, in a superior position in relation to the group observed.” (2009, p. 19).

Unit categorization proceeded in step with the delineation of units into phrases and themes. Theme boundaries were generally straightforward to identify, as they involved a switch from one set of unit types to another. Within themes, though, it was not always easy to identify what should be counted as a phrase. Some themes included near-exact repetition of patterns of units, but in others phrases could only be identified by seeking resemblances across multiple theme performances. Phrase variants could be related to each other through processes including the addition, deletion or modification of units, and the addition or deletion of entire subphrases. The process of delineation involved numerous provisional decisions, returned to in the light of further listening to multiple singers (Chapters 2 and 3). The song coding for each field recording was stored in a spreadsheet (Figure 4.1).

---

[20] Plumwood refers to this as “human epistemic locatedness” (2002, p. 132) and argues that it does not entail that humans need consider only their own interests.



Theme #	Phrase Variant	Start time (mm:ss)	Sequence of units										
ELEVEN	11A	00:05	mgr	ba-aws(s)	ba-aws(s)	dhq	w-nws(s)	mgr	w-aws(s)	ba-aws(s)	ahq	w-aws(s)	mgr
	11AA	00:31	agr	w-aws(s)	ba-aws(s)	dhq	w-aws(s)	mgr	w-aws(s)	ba-modws(	mgr		
TWELVE	12A	00:56	mgr	w-ns	w-ns	ns(s)	mgr	w-ns	ns(s)	mgr			
	12A'	01:22	gr	w-ns	w-ns	ns(s)	ngr	w-ns	ns	mgr			
	12A'	01:48	ngr	w-ns	w-ns	ns(s)	mgr	w-ns(l)	nws(s)	ngr			
	12A'	02:14	ngr	w-ns	w-ns	ns(s)	ngr	w-ns	ns	mgr			
	12B	02:41	agr	ds(l)	ds	ds(s)	gr	ds(l)	ns(s)	mgr			

Figure 4.1 Song coding for three minutes of humpback song from Singer 191026 (recording included in Portfolio), consisting of sixty units delineated into seven phrases belonging to two themes. Song units are given in individual cells with one phrase per row, with unit names abbreviated. E.g., the phrase beginning at 1m48 contains a mixture of modulated and n-shaped groans (mgr, ngr), short n-shaped shrieks and whistles (ns(s),nws(s), and a compound unit consisting of a whoop joined to an n-shaped shriek (w-ns). The corresponding spectrogram is given in Figure 4.2.

4.3.2 Re-composing With Song Codings: Phrase Variability

This initial ear-training left me with an analytical way of hearing and listening to humpback whale song that was related to aspects of human musicking, e.g., when listening back to a recorded improvisation that might serve as the basis for a composition. In both cases the language used to communicate my experience focuses on gestural qualities generated through timbre, register, pitch contour, dynamic level, and duration, rather than on concepts of absolute pitch, precise intervals or rhythms. In perceiving and recognizing sonic gestures (units) as belonging to categories, and hearing phrases as variants of one another, I call this close listening 'de-compositional' in order to acknowledge that humpback song appears to share compositional (structural) similarities with some forms of human music (Payne, 1995, 2000; Rothenberg, 2008b; Martinelli, 2009). I intend also to convey some of the positive connotations of decomposition, which leads to products that may be put together anew, re-composed.

The different 're-compositions' in my portfolio were based on unit categorizations and song codings resulting from the first stage of bioacoustical analysis. A major concern was how to incorporate the variability found in humpback whale song: my experience of coding many hours of songs had led me to concur with the view that "much of the structure in these songs varies highly from individual to individual – even between successive songs sung by the same individual" (Payne, 1995, p. 148). At the degree of granularity used in my codings, repetition at the level of unit and phrase was rarely exact. Besides my concern of 'fidelity' to this variability for the sake of zoömusicological accuracy, it seemed to me this brought up an ethical issue, in that my treatment of variability would project my view of humpback singers. Do they display the monotonous and mechanical repetition of 'the same', or do they manifest the agency and autonomy attributed to human musicians?

In '(im)possible gestures', 'The Path of the Unseen Whale' and 'Whale, Bow, Echo', I aimed to recapture this variability without predetermining its order or extent, and provided performers with generalized phrase templates and empirical rules of variability. By comparison, in 'Underwater Rain' and 'Entanglement' I retained the precise sequence of phrase variants from the analysis of a three-minute stretch of song (Singer 191026, included in Portfolio) (Figure 4.1), in order to preserve the sense of purposeful progression I heard in the pitch sequence of the lower units against the gradually descending upper units.[21] For 'Submergence', I provided composer Jan Foote with an illustrative schema containing two themes, each containing three repetitions of a standardized and very abstract phrase template ('AABBC' and 'DDEE'). Jan's piece for four bass clarinets was built on this structure, with each unit (e.g., 'A') represented by a particular combination of multiphonics. He specified variability through subtle variations in part-writing; further variability emerged in the process of my recording the score as some of the multiphonics were unstable and challenging to repeat with precision. The resulting minor fluctuations in pitch, pitch contour, and timbre, as non-deliberate, therefore offer a different kind of variability to that prescribed in '(im)possible gestures' and 'Whale, Bow, Echo'.

---

[21] Composer Emily Doolittle has taken similarly contrasting approaches. She transcribed individual birdsongs in works inspired by the blackbird (Doolittle, 1999), bobolink (Doolittle, 2018b, 2018a) and hermit thrush (Doolittle, 2009), but in works based on nightingale song (Doolittle, 2022) and Scottish garden bird songs (Doolittle, 2020a) she supplied general instructions to performers, allowing them to determine the appropriate form of variability.

4.3.3 Spectrograms and objectivity

I now turn to spectrograms, the visual representations of song used to assist in its coding (Appendix B). They mediate between sound recordings and the song codings used as the basis for much humpback song research. Spectrograms have played a central role in bioacoustics since the 1950s, when they rapidly replaced musical notation (Marler, 2004; Mundy, 2018) and transformed the scientific study of birdsong.[22] Spectrograms appeared to offer an objective representation of sound that could avoid the anthropomorphism of conventional musical notation, in which frequency, timing and timbre are compelled to conform to human musicality and culture (§4.4). Moreover, spectrograms seemed to bypass the default anthropocentrism of human hearing. They can represent sounds below or above its range, and portray more detail than humans are typically able to hear towards the extremes of this range.[23] Where change occurs on timescales faster than can be followed (e.g., in the songs of birds or insects), the spectrogram enables "sound microscopy" (Szőke *et al.*, 1969). Conversely, where patterns occur on timescales that are difficult to hold in short-term memory, as for humpback whale song, spectrograms of long stretches of song (e.g., Figure 4.2) allow viewers to "immediately grasp rhythm, shape, and form" (Rothenberg, 2008b, p. 140), and contributed to the initial identification of the repetitive structure of humpback vocalizations *as song* (Rothenberg, 2008b).

---

[22] The use of musical notation to aid scientific research currently appears to be restricted to studies in which bioacousticians have collaborated with zoömusicologists (Doolittle and Brumm, 2012; Janney *et al.*, 2016). Taylor (2017b) points out that the sounds of her study species, the pied butcherbird, are well-suited to musical notation. By contrast, the sounds of humpback whales are somewhat resistant to accurate transcription (§4.4.2).

[23] This is not so important for the sounds of humpback whales, which almost all fall within the range of a grand piano, i.e., within typical human capabilities, but is essential for the study of the vocalizations of many other species, from mice to elephants and including the infrasonic song of blue whales (Payne, 1977).

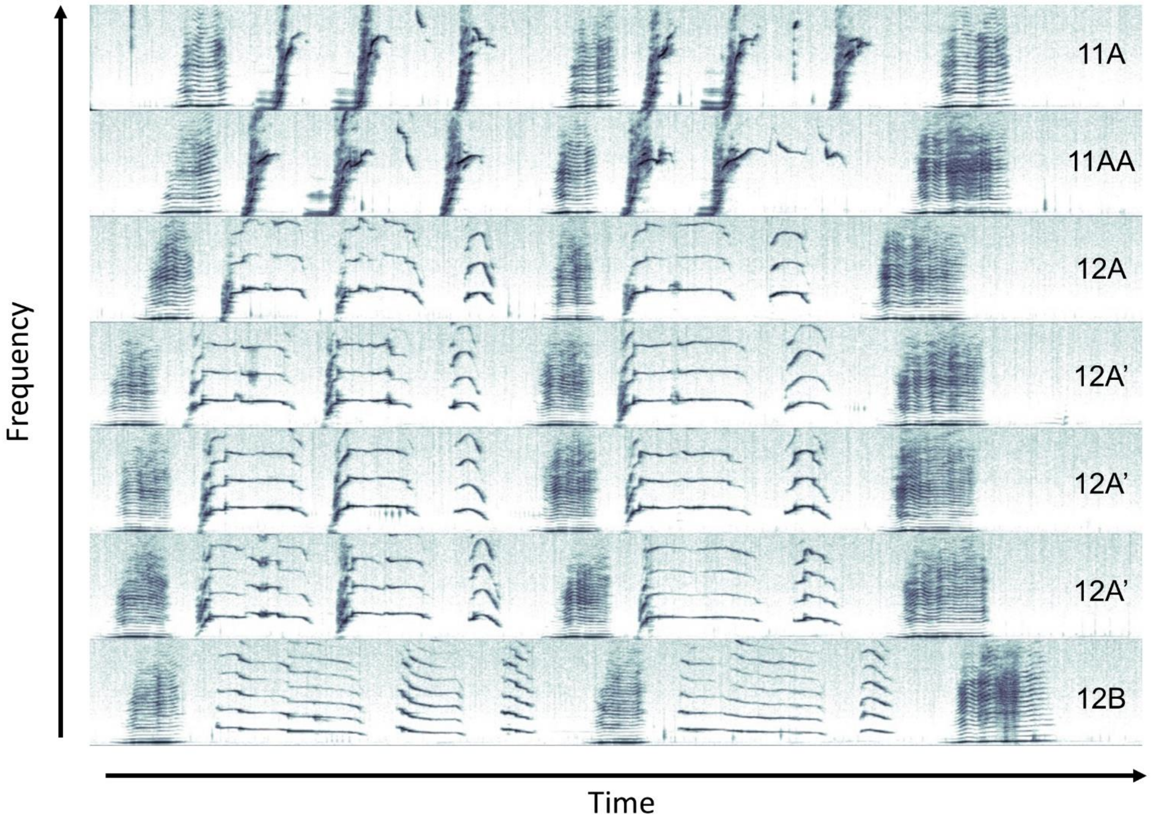


Figure 4.2 Multi-panel spectrogram showing three minutes of humpback song from Singer 191026 (recording included in Portfolio), coded into seven phrases consisting of 8–11 units. Each phrase is plotted on a single panel with common axes (time: 0–26 seconds, frequency 0–4 kHz linear scaling) and is labelled with its variant name. Phrases were sung in the order top to bottom.

Concerns with anthropocentrism aside, the fact that spectrograms are static representations rather than dynamic processes provides pragmatic advantages over human hearing. Whether assessing short or long-range structure, spectrograms readily permit the rapid visual comparison of individual sounds and sound patterns from different parts of a recording. In short, spectrograms supplement aural recognition and memory, and assist analysis: the categorization of units, the delineation of phrases, themes and songs, assessing the extent of variability at all hierarchical levels, and visualizing temporal relationships such as the repetitiveness of phrase rhythms. Finally, the production of spectrograms using computer software (such as Raven 1.6) allows the rapid quantification of multiple acoustic measures, removing the need for human judgment (e.g., of the balance of harmonics within a sound) and thus expanding their claims to objectivity (and hence, for science, in their superiority over the human ear.) For these reasons, for some bioacousticians spectrograms may not only supplement but even come to supplant the human ear. As eminent ornithologist Donald Kroodsma has put it: “I have well-trained eyes, and it is with my eyes that I hear.” (quoted in Mundy, 2018, p. 140). The use of spectrograms does raise the question of the approach to be taken where ear and eye disagree. Zoömusicologist Hollis Taylor, who has perfect pitch, reports that in her transcriptions into musical notation, made initially for the purposes of analysis, she defers to her ear to settle uncertainty (2017b, p. 76). Musicologist Rachel Mundy has argued that the scientific use of spectrograms still involves aural judgments and interpretation, which are “invisible in the shadow of the spectrograph’s apparent objectivity” (2018, p. 130). She provides ornithologist William Thorpe’s analysis of chaffinch (*Fringilla coelebs*) song (Thorpe, 1954) as example: his subdivision of the more or less continuous song into three distinct phrases seems to owe little to its spectrographic representation and much to his own auditory perception.

I argue that we need to keep eyes *and* ears open. Although spectrograms assist with coding, there is also a danger that they can mislead. For example, they provide no spatial or other contextual information that would allow the viewer to isolate the represented sounds of the focal species from those originating from other sources. Furthermore, the freezing of dynamic sonic process into static visual object means making decisions over how to visually represent temporal features, invoking cross-modal metaphors (such as high/low for frequency) and unavoidably indexing general and individual features of human embodiment (e.g., visual acuity, colour perception). Moreover, choices made in rendering the audible (and inaudible) visible may be task-dependent, whether suppressing ‘noise’ via display settings (brightness, contrast) or performing the unavoidable trade-off between temporal and spectral resolution. In short, spectrograms are representations *for* a human perceiver with a job to do.

Decisions may be grounded on anthropocentric assumptions or disciplinary convention. One example is the choice of how to map frequency. With a linear scaling, a given absolute difference in frequency (e.g., 1 kHz) has the same magnitude (approx. 3.5 mm in Figure 4.2) for all frequencies, whereas a logarithmic scaling preserves the magnitude of relative frequency (e.g., 2:1, corresponding to the musical interval of an octave) across all frequencies. A linear scaling is the standard in bioacoustics, leading to spectrographic images in which the 'just noticeable difference' Δf in the frequency domain is the same in absolute terms at low and high frequencies. However, auditory frequency discrimination in humans is roughly constant in *relative* terms (Weber fraction, $\Delta f/f$) from 0.5–4 kHz (Sek and Moore, 1995), meaning that for human perceivers the linear scaling of the spectrogram given above (Figure 4.2) will tend to exaggerate differences at high frequencies (visually-discriminable frequencies may not be aurally discriminable) and vice versa (at low frequencies, aurally discriminable frequencies may not be visually-discriminable) (Figure 4.3).[24]

The constancy of relative frequency discrimination in humans may underly the important role of relative pitch in the perception of melody, which seems in turn to explain the approximately logarithmic frequency scaling employed in staff notation. Measurements of auditory frequency discrimination for dozens of species of birds and terrestrial mammals (Fay, 1988, pp. 269–275, 451–457), and bottlenose dolphins (Nachtigall *et al.*, 2000, p. 354) all indicate approximately constant Weber fractions across wide frequency ranges.[25] No such measurements exist for

---

[24] Outside of this range, corresponding approximately to the pitch range C5–C8, the Weber fraction expressing the just noticeable difference increases for humans, especially at higher frequencies (Sek and Moore, 1995). This worsens the problem of misrepresentation by a linear scale for frequencies > 4 kHz, and slightly reduces it at frequencies < 500 Hz.

[25] This raises the question of why it is that other-than-human animals may have a "lack of facility with relative pitch" (Patel, 2008, p. 396), here I can only note that the current empirical evidence is mixed (Hoeschele, 2017).

humpback whales, but it seems reasonable to assume that their auditory frequency discrimination functions as for dolphins, humans and other mammals. The linear scaling used by bioacousticians, perhaps originating in an attempt to avoid anthropocentrism, could well be leading to the misrepresentation of humpback sound, for example when categorizations are based on or influenced by the visual features of spectrograms.

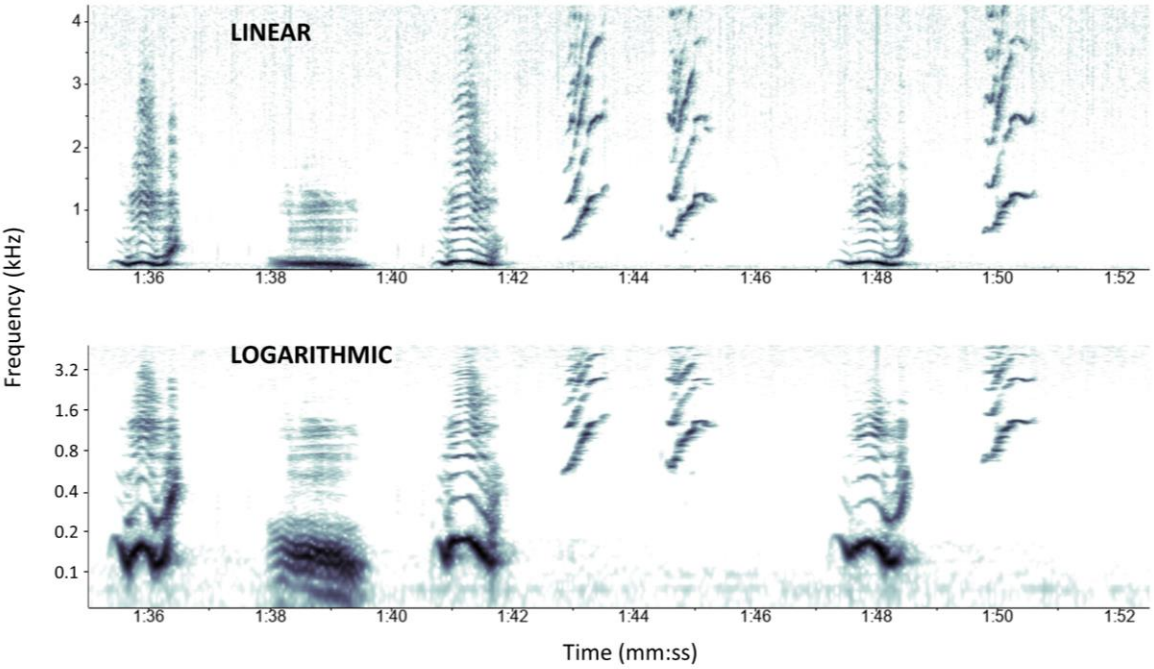


Figure 4.3 Spectrograms of humpback song PV 3A° for the pitch range C2–C8 (Singer 190919, included in Portfolio). With the logarithmic frequency scaling (lower panel), vertical distance is proportional to the size of the interval, meaning that the representation better matches human pitch perception and it is easier to see what can be heard. For example, the third and fourth song units both span a similar interval (approximately an octave). On the linear frequency scaling (upper panel), the melodic (frequency) contour in the lower frequency song units is hard to discern.

In the case of the song unit categorization I acquired through Ellen Garland (Garland *et al.*, 2011, 2012, 2017; Garland, Noad, *et al.*, 2013), her emphasis on staying with the sound rather than the spectrogram, and the use of frequency boundaries based on constant frequency ratios (250 Hz, 500 Hz, 1 kHz) to mark major distinctions among categorizations (groans, moans, cries, whistles), goes a long way towards alleviating concerns that these categorizations do not account for humpback hearing. However, the recent use of neural network-based image recognition methods to categorize sounds from spectrograms with linear frequency scaling (e.g., Zhong *et al.*, 2020; Allen *et al.*, 2021) may be problematic if the spectrograms are interpreted as being meaningful to the sound producer, as is any naïve interpretation of spectrograms as being a more accurate representation of sounds than that provided by the human ear. An attempt to eliminate the human auditory system from the picture, so to speak, may thus also eliminate the auditory system of the species concerned. In the time domain, this attempt can also be found in the choice of extreme temporal magnification to provide an image of sound "in reality" (Szőke, 1987, quoted in Loch, 2018, p. 149). In both cases, we risk forgetting that the sounds we study have been made to be heard, are not just pressure waves in a compressible medium. In conclusion: although spectrograms present an image, a visual 'object', it is scarcely an 'objective' picture of sound as a pressure wave, let alone of sound as it arises from the interaction of pressure wave and perceptual system. Visible differences may not be salient to the producer or receiver of the sound, and conversely, salient audible differences may not be visible.

Full objectivity may be unattainable. Is it desirable? In the case of the ethological study of the vocalizations made by animals other than humans, it certainly appears reasonable to assert that we ought not to import anthropocentric or anthropomorphic assumptions of how such vocalizations would be heard by the species under study. Again, we must not forget that such vocalizations are made to be heard. Here, we might consider a parallel use of a device similar to the spectrograph in the field of ethnomusicology. For a period in the 1950s and 1960s, the melograph appeared to offer the chance of an automated transcription of songs collected during fieldwork, relieving the ethnomusicologist of the arduous task of transcribing musics with unfamiliar tuning systems and rhythmic patterns (Mundy, 2018).[26] However, enthusiasm waned in the light of Nazir Jairazbhoy's critique that highlighted the gap between what would be "heard" by the melograph and by "an experienced listener of a particular musical idiom" (1977, p. 269). Jairazbhoy further argued that the chief benefit of an automated transcription would be to reveal some of the limitations of and assumptions in our listening: "to throw light on what we do *not* 'hear,' what we change in the process of 'hearing,' or what we take for granted." (1977, p. 270).

When it comes to listening to the sounds made by other animals, bridging the gap identified by Jairazbhoy becomes impossible, as the subjectivity of the experienced listener of another species is arguably unknowable (e.g., Nagel, 1974), and is certainly inaccessible to ethnomusicology's participant-observation methodology (Sorce Keller, 2012). This, for the bioacoustician, is an argument for remaining with the (careful) use of the spectrogram, bearing in mind the lessons learnt from Thorpe's chaffinch analysis (what I hear influences what I see) and Jairazbhoy's critique of the melograph (what I hear depends on me, what I see can tell me about what I do not hear). However, as a performer-composer working with whale song, I must also bear in mind that my music is written to be listened to by humans, not whales, and its aesthetic success or failure rests on human perception.

4.3.4 Spectrograms and subjectivity

So far I have considered spectrograms in their use as visual aids in the production of the song codings that I used for bioacoustical and compositional purposes. Next I turn to their striking similarities with musical scores. To any musician who uses notation in which pitch is plotted against time, the spectrogram may readily be viewed as a kind of score, its "tentative

---

[26] Automated transcription tools are widespread today, and may perhaps be configured to account for perceptual capacities. But it's hard to see how they can possibly account for every individual's learning history and cultural background, which shapes *how* they hear and listen to sound and music.

hieroglyphics" (Rothenberg, 2008b, p. 138) appearing to us as the script of a gestural musical language.[27] Here, the usual function of a spectrogram, of *describing* sounds, begins to take on that belonging to the musical score, of *prescribing* sound-making actions. This applies both to its framing, i.e., the pitch-time axes, and its foreground: the visual tracings of sounds are reminiscent of symbols used to describe musical tones and prescribe performance techniques. Furthermore, the repetition of the units and phrase patterns in the spectrogram of humpback whale song suggests a music-like ordering. The interpretation will depend on musical experience. Practitioners or historians of early music may see the resemblance with neumes, precursors to Western staff notation (Rothenberg, 2008b, p. 139) (Figure 4.4). Players or composers of contemporary Western classical music may be more struck by similarities with the symbols used to represent timbre or extended techniques. Improvisers accustomed to graphic scores are already accustomed to using images to facilitate the creation of music.

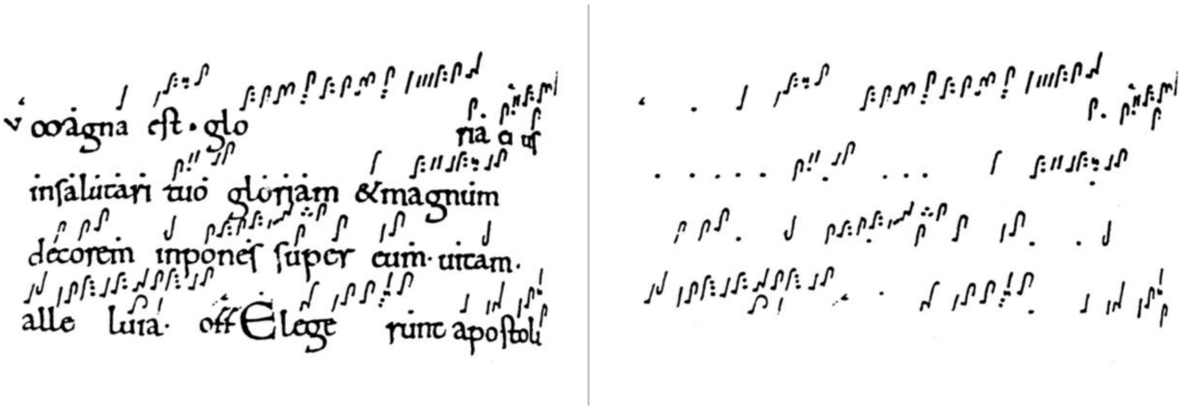


Figure 4.4 Anglo-Saxon Offertory Elegerunt, from Levy (1990) Figure 3. Erasing the words (right) brings out the visual resemblance between neumes and the spectrographic traces of whale song units.

This ambiguity of the spectrogram, situated between record of and prompt to sound, is well illustrated by Rothenberg and Deal's (2015) whale song notation, in which individual units are replaced by coloured graphic elements based on a composite spectrogram for each unit type (Figure 4.5). This reduction in variability through a similar process of categorization to that

---

[27] This is perhaps especially true of the early hand-traced spectrograms (Thorpe, 1954; Payne and McVay, 1971).

used in song coding leads to a perspicuous visualization of repetition and hierarchical structure. It is arguably a form of analysis that has its roots both in music theory and bioacoustics. The similarity between spectrogram and musical score is highlighted through the over-plotting of staves and the choice of a logarithmic frequency scale for the spectrogram.[28] For some viewers this may immediately solicit inner or external rehearsal of the represented sounds, for others the stave will only act to connote a more general musicality. A revised version of their notation, "to be played by all readers" (Rothenberg, 2023, p. 290), recently featured in an article about humpback song in *National Geographic* (May 2021). In the revision, the coloured glyphs were superimposed on Western musical notation to create a hybrid score.

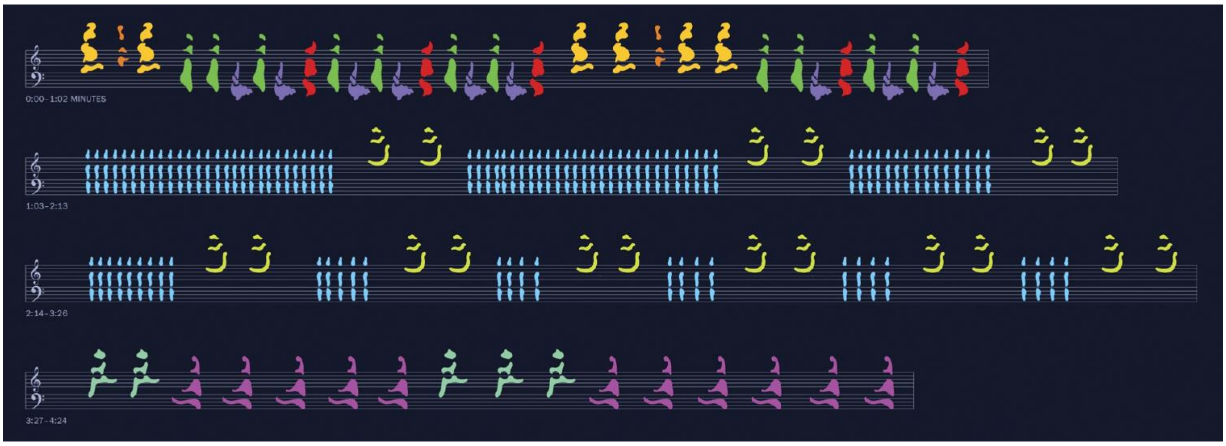


Figure 4.5 "Morphological" notation for humpback song based on spectrograms, with song unit types distinguished through shape and colour. Detail from Rothenberg and Deal (2015) Figure 5.

Given the suggestive forms of the spectrographic traces of humpback song, it is not surprising to find their influence on musical scores in the form of graphic notation (e.g., Souster, 1972; Powell, 1993; Sanders, 2021). I provided spectrograms in my score to 'Whale, Bow, Echo' (Figure 4.6) partly for their function as "tentative hieroglyphics", i.e., as visual prompts to sound-making, partly to remind the original performers of their role during the piece's development.

---

[28] Fully accurate representation of the lines of the stave would require two different spacings, as they subtend intervals consisting of either three or four semitones.

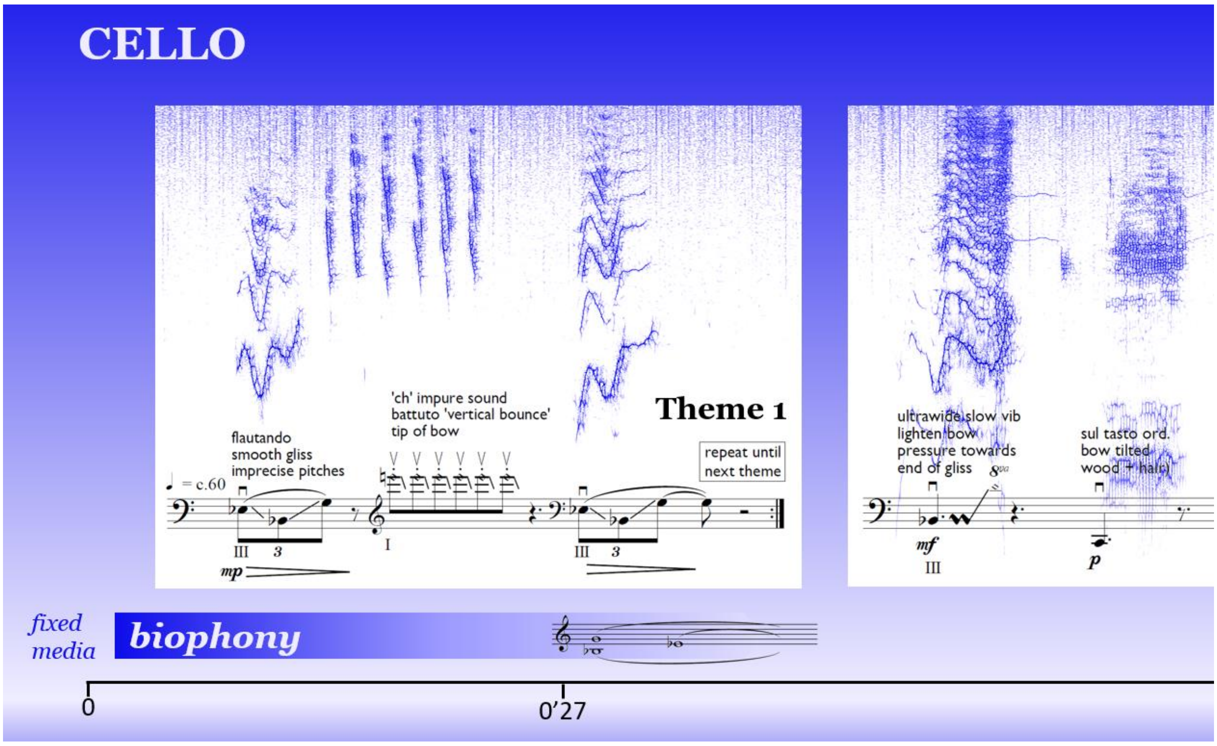


Figure 4.6 Detail from score to 'Whale, Bow, Echo', showing a transcription approximately aligned with the spectrogram (plotted on a logarithmic scale). The score includes three time scalings: (1) the spectrogram, (2) the transcription (approximately proportional), (3) the fixed media transcription with linear time line to assist players with co-ordination with the fixed media track.

In early workshops, the translations of the humpback sounds to sonic gestures on violin and cello were achieved by the players responding mimetically while listening to whale song, guided by the spectrograms. Initially I had included over-plotted staves, seeing the spectrogram as half-way to a transcription of pitch and rhythm, but removed them when it became clear that the players wanted a more conventional notation of the sonic gestures. The specific transcriptions, techniques and notations were arrived at through our joint discussion. As a comparison of the transcription and the spectrogram reveals, we settled on a relatively low level of granularity in notating similar-sounding song units. For example in Theme 1 the first and last units are notated in the same way despite minor pitch differences (Figure 4.6). As in the case of the scoring of '(im)possible gestures', the accompanying instructions make it clear that variability should be reinstated in performance. In both works the aim was to avoid a mechanical repetition in the live parts. What the playing scores do not show is the existence in the fixed media track of looped 'echoes' of the players' parts, created from separately-recorded versions of each phrase variant. In the final piece, the echoes are brought in following the first live playing of each phrase variant, gradually becoming more present in the overall mix. Although subject to increasing reverb, these electronic echoes *are* fixed in rhythm and pitch, and present a contrast with the freedom granted to the performers. I hoped in this way to emphasize the importance of organic, sentient expressivity. I also exploited this contrast between organic flexibility and a mechanical rhythmic fixity in 'The Path of the Unseen Whale', where again electronic echoes (this time created live through interactive electronics, and subjected to filters to gradually transform timbre) are presented against the variability of a live performer.

4.4 Transcribing Humpback Song

4.4.1 Anthropomorphism and Anthropocentrism in Musical Transcription

I have already mentioned that until the 1950s musical notation was in widespread use in the scientific study of birdsong. Although now rare in the scientific literature, the practice of musical transcription has remained in common use among zoömusicologists (Szőke *et al.*, 1969; Mâche, 1992; Rothenberg, 2005; Doolittle and Brumm, 2012; Taylor, 2017b; Loch, 2018), where it may signify the zoömusicological goal of taking the vocalizations of other animals seriously as "music-like" (Doolittle and Gingras, 2015), or as music "other than with the quotation marks" (Mâche, 1992). It is also, of course, widely used by composers in works informed by other-than-human animal sounds, for whom an accurate representation of a particular song performance may be of paramount aesthetic importance (Taylor, 2017a; Liebermann, 2020). Here we should distinguish between the making of a transcription and its subsequent treatment during composition, as the question of fidelity can be rather different in each case.[29] As Messiaen commented,

> "I've used bird songs in two different ways: either by trying to outline the most exact musical portrait possible, or by treating the bird song as malleable material. [. . .] Personally, I'm very proud of the exactitude of my work [. . .] I assure you that everything is real; but, obviously, I'm the one who hears, and involuntarily I inject my reproductions of the songs with something of my manner and method of listening" (1967, quoted in Taylor 2014).

Messiaen's awareness of his own contribution to listening has been validated in a comparison between his own and a more neutral transcription of songs of the pied butcherbird (Taylor, 2014). Taylor, working from the same recordings as Messiaen, used spectrograms to aid the measurement of frequency and duration and concluded that "what Messiaen heard, saw, thought and felt, was filtered by and through his personal musical vocabulary… [he] does not wait until the moment of composition to transform the birdsong he notates" (2014, 98). She shows in detail how Messiaen stretched pitch intervals, omitted portamenti, and considerably simplified rhythms through quantization. Although this may be an extreme case, Taylor's findings are in line with prior Messiaen scholarship and support other comparative work on a wide range of historical transcriptions of birdsong (Szőke *et al.*, 1969; Loch, 2018; Mundy, 2018; Doolittle, 2020b, 2020c). The point I wish to make is that the processes both of transcription and composition are carried out in a human musical language and system of notation that have not evolved and cannot be fully repurposed to deal with the rhythms,

---

[29] Taylor comments that her "transcriptions sit first and foremost inside birdsong analysis, not inside composition." (2017b, p. 61). She also refers to Seeger's distinction between ethnographic transcriptions intended as 'blueprints' and 'reports'. "[My notations] are often called upon to be both, but not at the same time. The 'report' serves as the basis of my analysis. My violin performances of pied butcherbird song are based on the 'blueprint,' where I allow myself the (re)composerly freedom to adjust the notation to my and the instrument's needs and sensibilities." (2017b, p. 76).

intervallic relationships, and timbres of the songs of birds, and thus their products are inevitably shaped to some degree by anthropomorphism. Such a default anthropomorphism, that I suggest many transcribers are aware of, may be conceptually distinguished from the shaping of transcriptions through deeply embedded beliefs. For example, Doolittle (2020b, 2020c) has showed that historical transcriptions and descriptions of hermit thrush song tracked the changing cultural associations of that particular species in the North American imagination, and assumptions about birdsong in general. The latter include the problematic equivalence drawn between birdsong and so-called ‘primitive’ musics also explored by Mundy (2018), revealing assumptions about the superiority of Western Art Music that manage at the same time to be both ethnocentric and anthropocentric.

4.4.2 Whale Song Transcriptions and Whale Music

The examples given so far of musical transcription have all been taken from birdsong, and if it is perhaps not surprising that the use of birdsong in human music is far more widespread than the use of cetacean vocalizations,[30] it is noteworthy that with a single exception (Rothenberg, 2023), the many (zoö)musicological discussions of whale song have avoided transcriptions (Payne, 1995, 2000; Nollman, 1999, 2008; Rothenberg, 2008b, 2008c; Martinelli, 2009; Sorce Keller, 2012; Rothenberg and Deal, 2015; Grover Friedlander, 2020; Cook, 2022; South, 2022). This reticence may be for reasons related to disciplinary or thematic context, or because whale vocalizations were first made public via spectrograms (Payne and McVay, 1971), but I suspect that a contributing factor, at least in the case of humpback song, is the challenge posed by the

---

[30] Over a hundred musical scores and albums or tracks inspired by the sounds of whales and dolphins have appeared since the 1970s (Appendix A), but this is a tiny fraction of those inspired by birdsong (for an overview, see Doolittle, 2008). The disparity has recently been further augmented by the 172 new recorded works of the ‘The Birdsong Project’ (Poster and Reagan, no date).

prevalence of smoothly-varying pitch contours (which I refer to as ‘portamento’) and a range of different ‘noisy’ sounds.

Turning from description to prescription, several musical compositions include notated sections of humpback song (Hovhaness, 1970; Souster, 1972; McLean, 1977; Powell, 1993; Gonneville, 1998; Bright, 2007; Doolittle, 2007b; Sanders, 2020). Responses to the technical challenges of reproducing whale sounds via human musical notation have included the choice of instruments such as the trombone or ondes Martenot, with their capacity for portamento (Figure 4.7), extended techniques (Souster, 1972; McLean, 1977; Powell, 1993; Doolittle, 2007b; Sanders, 2020), graphical notation (Souster, 1972; Powell, 1993; Sanders, 2021) and electronic processing (Souster, 1972). There is occasionally implicit or explicit recognition of the difficulties. For example, Souster specifies that “In Aria (for viola only) the notation is as exact a representation as possible of the song of the hump-backed whale and should be adhered to *as closely as possible*, using a recording of the actual song as an aid during practice” (1972, emphasis added). Bright requests his singers to “[i]mitate whale approximately but not exactly – a symbiotic interaction[,] a sound somewhere between whales and voices” (2007), and Doolittle writes “[m]irror whale song as closely as possible (including ‘out of tune’ notes). Of course it won’t be possible to mimic it perfectly: relish the little clashes between your version and the whale’s!” (2007b, p. 4). These scores bear witness to a conscious recognition of the anthropomorphic constraints of musical notation, and I turn next to consider those that I have been concerned with in my own compositions: rhythm and metre, repetition and variability. The question I address is whether an unavoidable anthropomorphism stemming from the cultural limitations of musical notation necessarily leads to the expression of a naïve anthropomorphism, i.e., a projection of a broader human musicality into humpback song or its singers, in the results.

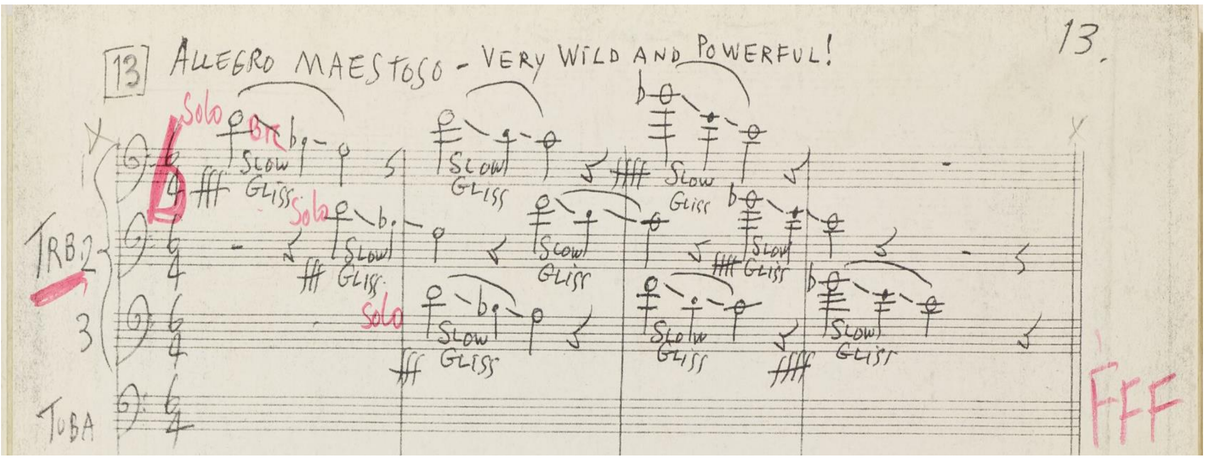


Figure 4.7 Imitative writing for trombones by Alan Hovhaness, in 'And God Created Great Whales' (1970).

4.4.3 Anthropomorphism in Rhythm and Metre

Constant pulse, constant metre: '(im)possible gestures'

I begin with my first pieces inspired by the 2015 song used for the purposes of ear-training, composed before I had become focused on the temporal aspects of song (Singer 150918, recording included in Portfolio). In '(im)possible gestures', I was not striving for a precise reproduction of the rhythms: I used 4/4 metre for all four notated themes, notated inter-unit silences so that units predominantly begin on the start of a crotchet (quarter-note) beat, often at the start of a bar (Figure 4.8), and articulated the longer units with the beat.[31,32] There is a clear presumption that the four performers would each be feeling the beat (to the flexible tempo stipulated in the score), though there was no collective time-keeping required. This ordinary notational device points to a foundational element of human musicking, the ability to entrain to isochronous beats, whether external or self-generated (Clayton *et al.*, 2005). Although

---

[31] This approach to temporal notation has been used by Rothenberg (2023, pp. 290–291).

[32] My goals in this piece were primarily elsewhere: exploring instrumental techniques capable of evoking the song unit timbres and timbral contrasts I was hearing, generating an immersive quality for the audience, and contrasting the kinds of variability found in humpback heterophony (where individual singers do not appear to coordinate their songs), and in human collective song (where interaction and coordination is a consistent cross-cultural norm).

audiomotor entrainment has been demonstrated by some other animals (Wilson and Cook, 2016), no attempts have been made to assess it in whales or dolphins. It is clear that my adherence to a beat introduced a naïve anthropomorphism into the rhythms of the work, at least at the level of the score and for the performer. However, it is arguable that most audiences will not hear it in the work, particularly as all four performers are playing different musical lines at different and flexible tempos.

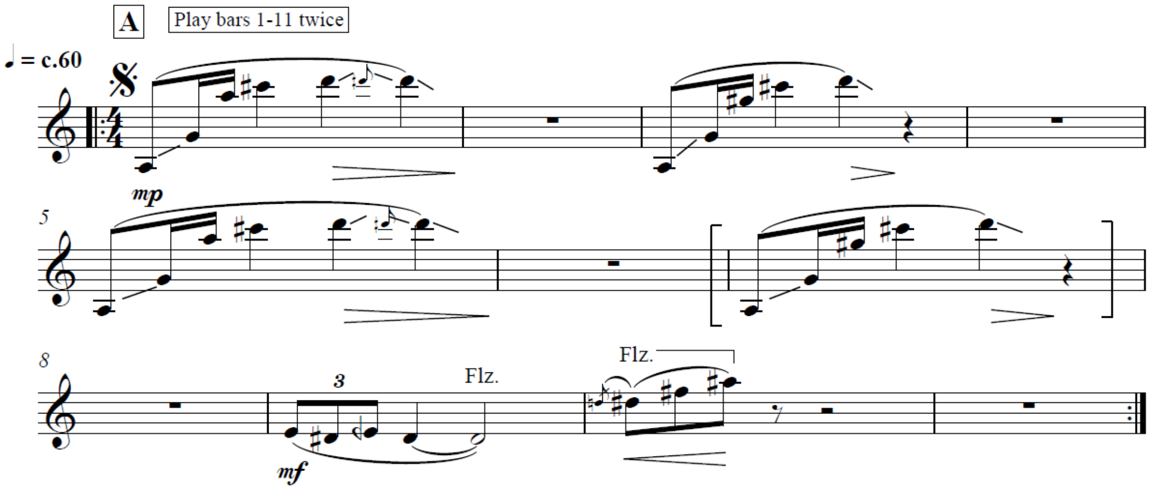


Figure 4.8 Section of score from '(im)possible gestures' (2019). The square brackets indicate optional material. The metronome mark results in performed phrases that are approximately three times longer than the time taken by the humpback whale.

Multiple beat lengths, varying metres, common fast pulse

By the time of composition of 'Underwater Rain' and 'Entanglement', both based on the song whose coding (Figure 4.1) and spectrogram (Figure 4.2) I have already discussed (Singer 191026, recording included in Portfolio), my scientific interest had settled on representing and quantifying the variability of humpback phrase rhythms. I had developed software to measure inter-onset intervals (IOIs) between song units, and produce visualizations of the resulting timings (Chapter 2). In addition, I was investigating the question of whether any of the 'statistical universals' (Savage *et al.*, 2015) of the rhythms of human music could be found in the humpback songs I had been listening to in my exploratory bioacoustical analysis. The existence of something like metre, previously noted in humpback song by others (Handel *et al.*, 2009), appeared clear: phrases and subphrases are of quite consistent length, even where they vary in precise song unit content, with song units occurring in roughly the same positions (Figure 4.2). There seemed to be clear parallels, that is, between the level of the measure in human music, sometimes conceptualized as a ‘common slow pulse’ (Kvifte, 2007) and a phrase or subphrase in humpback song.

In my initial mimetic investigation of a particular sequence of phrases with very low timing variability (Figure 4.9A), I had found a way of playing their rhythms which employed a counting method frequently used by performers of musics with irregular metres, in which metres may change from bar to bar but are usually subtended with a common fast pulse (Table 2.1). I used beats of two different lengths, subdivisible by two or three (Figure 4.9B), and irregular metres within the phrase to articulate the subphrase structure (Figure 4.9C). This development allowed a more accurate representation of rhythms than with the single beat length used for ‘(im)possible gestures’ (Figure 4.8): a musicological parallel could be drawn with composer Béla Bartók’s revising of transcriptions of Rumanian folk music (Figure 4.10) following his encounter with the principles of Bulgarian rhythm in the pages of Vasil Stoin’s ‘*Grundriss der Metrik und Rhvthmik der bulgarischen Volksmusik*’ (Rice, 2000, p. 201).

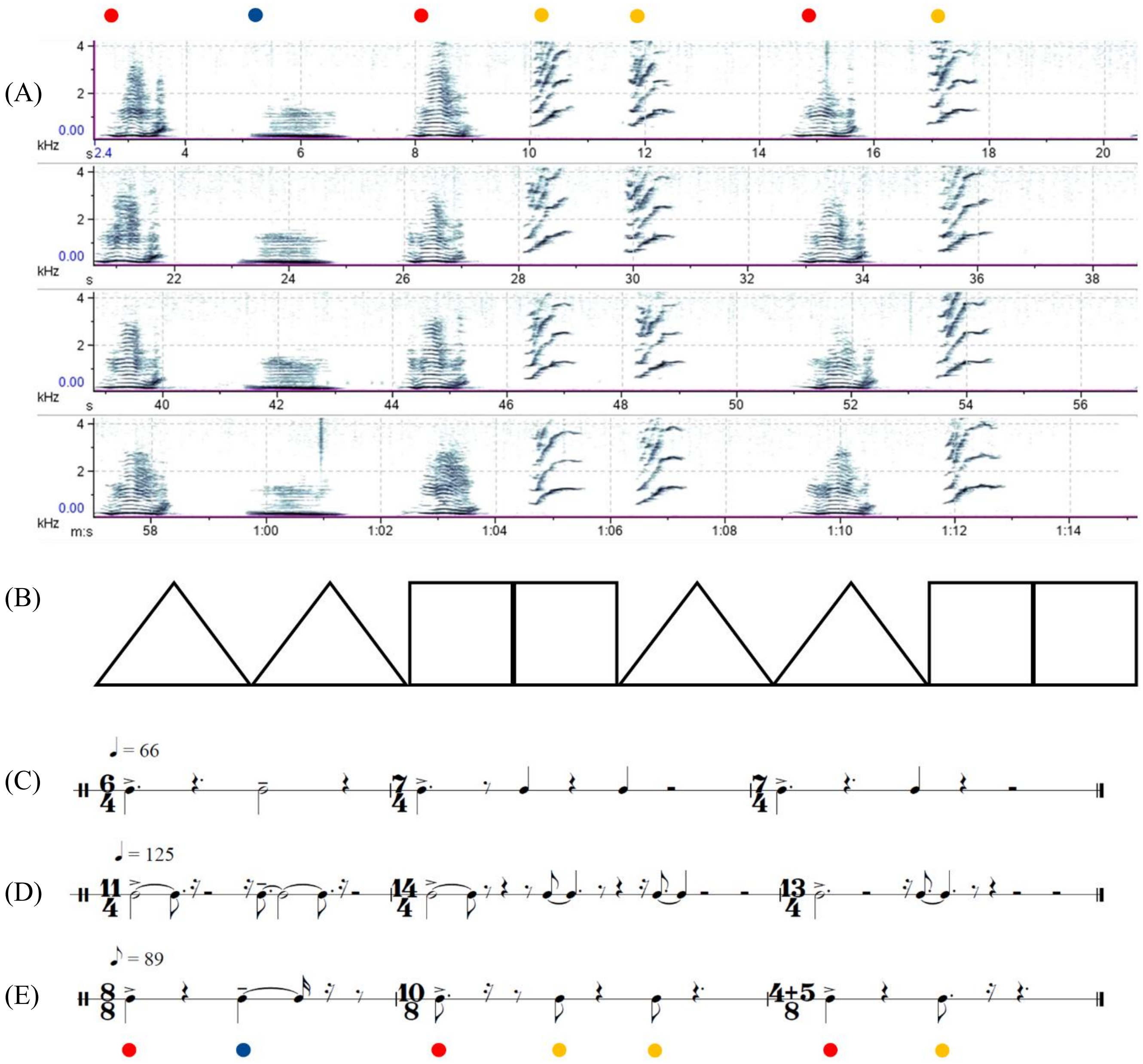
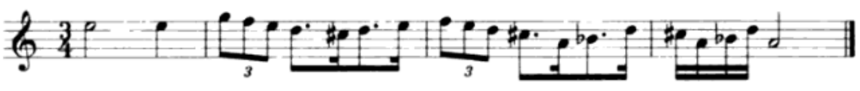


Figure 4.9 Representations of humpback whale song from Singer 190919 (recording included in Portfolio). (A) Spectrogram. This shows 1m13 of continuous song comprising the performance of four phrases (Phrase Variant 3A°, see Appendix B), each containing seven song units, with one phrase per line. The coloured dots indicate the broad song unit categories (red = ‘moan’, blue = ‘croak’, yellow = ‘ascending cry’). The high precision (WRV = 0.064 s) of the repeated phrase rhythm is confirmed by the near-vertical alignment of the song unit onsets. (B) An approximate felt phrase rhythm with two beat lengths in ratio 3:2, shown with triangles and squares. This rhythm can be used for each line of the spectrogram. (C) Transcription of the phrase rhythm shown in (B). Bar lines indicate subphrases beginning with the same kind of song unit. (D) A more accurate transcription, created by calculating a “mean phrase rhythm” from the four phrases, and then finding a whole number ratio (11:14:13) that fit the length of the subphrases and was appropriate for human performance. (E) Transcription performed by applying Norton and Scharff’s (2016) algorithm to the mean phrase rhythm, with onsets matched to the quaver (eighth-note) common fast pulse.

Example 14.7. Bartók, Romanian Folk Dances from Hungary, for Piano, no. 4 (1915), mm. 3–6.

Example 14.8. Bartók, *Rumanian Folk Music*, Vol. 1, no. 175, mm. 5–8.

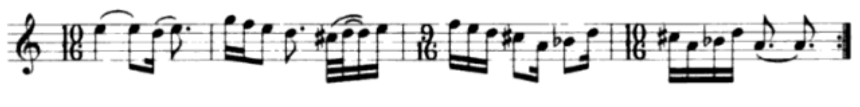



Figure 4.10 Bartók's transcriptions, separated by at least twelve years, of the same Rumanian folk dance recording (Rice, 2000, p.201). The earlier transcription (Ex. 14.7) employed a regular metre, one beat length and repetitive rhythms. The revised transcription (Ex. 14.8) used irregular metre and rhythms based on two beat lengths.

My initial ‘felt’ rhythm (Figure 4.9B,C), with a ratio of 6:7:7 for the three subphrases, corresponded approximately to the humpback performance of the four phrases, but I wanted to get closer and avoid any anthropomorphic projections arising from my embodied response to the whale song. To do this, I measured the inter-onset intervals and created a mean phrase rhythm for the four phrases by averaging the intervals at each song unit position. Working at the level of the subphrase, I used a trial and error method to find whole-number ratios that were closer to those calculated from the mean rhythm. This gave me ratios of 11:14:13, but onsets that did not fall onto the metrical common fast pulse of a crotchet. The notation (Figure 4.9D) was no more complex than many contemporary scores, and would allow an accurate reproduction in performance, albeit one that would initially at least lack the sense of ease that was achievable with onsets coinciding with beats. To return this sense of ease, and yet maintain rhythmic fidelity, I employed an algorithm developed to analyse the timing of zebra finch song (Norton and Scharff, 2016). The algorithm finds a best fit common fast pulse by projecting a series of fixed grids across the onsets and minimizing the difference between the grid and the onsets. As a composer, the value of this algorithm was that the calculated common fast pulse made possible a rhythmic transcription that was accurate, positioned onsets at the beginnings of beats, and could be carried out rapidly for any length of song (Figure 4.9E). Furthermore, it did not involve any assumptions regarding metre or beat length.

The rhythms of 'Underwater Rain' and 'Entanglement'

For the 191026 song used in 'Underwater Rain' and 'Entanglement', I applied Norton and Scharff's (2016) algorithm to mean phrase rhythms for each theme. In 'Underwater Rain', an electroacoustic piece commissioned as part of a pack of mindfulness resources for university students, my goal was to provide a piece that would provoke and reward a close and sustained listening. I wanted the listener to be primarily focused on the whale song, and structured my composition around the fact that I heard the song in two contrasting parts: a low bass line accompanied by a high and expressive melody (Figure 4.11). I used a three-minute section of

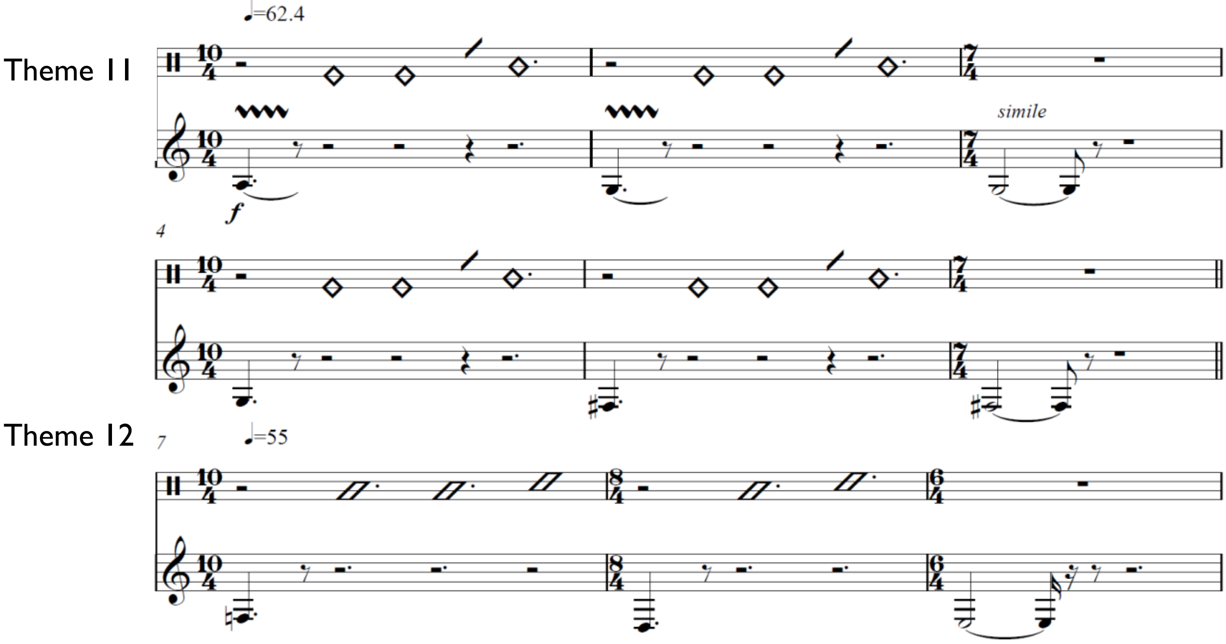



Figure 4.11 Transcription of the first three phrases from the recording of Singer 191026 (Figures 4.1, 4.2, recording included in Portfolio), employing onset extraction (§2.2.2), and Norton and Scharff's (2016) algorithm applied to mean phrase rhythms. This transcription was used in 'Underwater Rain' and 'Entanglement'. The pitches are transposed for B-flat clarinet.

the 191026 song, slowed down by a factor of two and played twice, adding minimal electronic processing to introduce additional variation without taking the focus away from the humpback. The second time around (c.8m00–14m40), the low song units are supplemented by bass clarinet notes (Figure 4.11, lower stave). I recorded the bass clarinet to a click-track, giving me a humanly-performed layer whose onsets shifted ahead and behind those of the whale song due to the difference between the whale's performed rhythms and the mean phrase rhythms used for the transcription. I wanted to show that the expressivity that I perceived in the melodic content of the high song units (Figure 4.11, upper stave) was also present in a rhythmic flexibility that was hard to hear due to the slow pacing of the song. I hoped to present a reversal of a human exceptionalist view that privileges humans as artistic. In this piece, various delay effects applied both to bass clarinet and whale song provide a third layer of temporal organization, one of fixed short-term repetition. Field recordings of waves and rain made in French Polynesia contributed quasi-random rhythms to the temporal complexity.

For 'Entanglement', violist Katherine Wren and I recorded a number of improvisations based on the 191026 transcription (Figure 4.11), at a range of different tempos, and to varying degrees of freedom. During some takes we both played material on both staves, in others we took one stave each. In this piece, the transcription was used only as a starting point, and during the improvising stage we individually and freely elaborated on the transcription, especially when interpreting the graphic symbols on the upper stave. Some strategies for interaction were predetermined (e.g., the clarinet solo over viola drone that opens and closes the piece), most emerged during performance. In the final piece, I selected, cropped and layered up different takes, including some minor manipulation of playback rates, to create up to seven different simultaneous tempos. I was aiming to achieve a sense of extremely free heterophony, hoping for the listener to hear that there are overlapping versions of closely-related material. The final result has only selective similarities with the original humpback song, but nonetheless the approximate rhythms, and pitch progression of the lower stave, are present throughout.

My transcription of humpback song rhythms for ‘Whale, Bow, Echo’ (Figure 4.6) was also achieved using Norton and Scharff’s (2016) algorithm. With ‘Underwater Rain’, ‘Entanglement’ and ‘Whale, Bow, Echo’, I therefore moved away from an anthropomorphic projection of entrainment to a single beat length. Nevertheless, the use of a common fast pulse to define onsets, although facilitating reasonably accurate reproduction of humpback song rhythms, remained an anthropomorphic projection into the score and performance.[33] As with ‘(im)possible gestures’, however, it seems likely to me that the common fast pulse will not be heard by the audience. The rhythms, although repetitive at the phrase level, predominantly comprise irregular patterns of onsets with intervals of a few seconds (considerably longer in the case of the bass clarinet line in ‘Underwater Rain’).

4.5 The Time of Re-Composition: The Objectification of Humpback Sounds

In this chapter so far, I have described how I have employed a range of representations of humpback sounds, enabling me to reorganize and sonically transform them in my compositions. In this section I argue that the processes generating these representations should be termed ‘objectification’, to acknowledge that they and the manipulation they permit share in the logic of domination underlying past and present human exploitation both of other humans and the other-than-human world (§4.5.1). I develop Theodor W. Adorno’s notion of ‘aesthetic rationality’ through Val Plumwood’s analysis of binary dualisms in order to outline strategies that may assist composers who wish to avoid the continuation of these logics, deepening my

---

[33] In ‘Underwater Rain’ and ‘Entanglement’ the transcription can be performed by using two beat lengths related in the ratio 2:3, such as is commonly found in human musics. Empirical studies have shown that entrainment is achievable in music whose performance includes two or three beat lengths, even at fast tempos and in the absence of a common fast pulse (Polak, 2010; Polak *et al.*, 2016). For ‘Whale, Bow, Echo’ the rhythmic transcription is more complex, involving up to four beat lengths which are not related by small integer ratios. I decided not to simplify it for ease of performance.

earlier recommendations to avoid anthropodenial and naïve anthropomorphism (§4.5.2). I then turn to my portfolio works, analysing the kinds of manipulation used, and exploring the effects of the ensuing decontextualization and recontextualization, discussing tendencies to exoticization and domestication and assessing the works against the principles of a respectful aesthetic rationality (§4.5.3).

4.5.1 Does Representing Song Objectify Singers?

In this thesis I have assumed that we should take seriously other-than-human animal agency, subjectivity, dignity and intrinsic value, accepting arguments made in animal studies (e.g., Haraway, 2008; Calarco, 2015; Despret, 2016), animal ethics (e.g., White, 2007; Humphreys, 2016; Sebo, 2017; Wichert and Nussbaum, 2019; Abbate, 2020; Nussbaum, 2022), environmental ethics (e.g., Plumwood, 2002) and cognitive ethology (e.g., Allen and Bekoff, 1997; Burghardt, 1997; Griffin, 1998; de Waal, 2016). I maintain that a strong case can be made that the singers of the songs on my field recordings were not just causally responsible for the sounds captured, but that they also possess, perhaps collectively, some rights of authorship or ownership over their sounds (even if such rights do not amount to the rights we give humans, and accepting that 'rights' may not be the most useful way to describe the ethical claim). Further, I accept Abbate's (2020) argument that other-than-human animals can be wronged, even when they are not harmed (either directly, or indirectly). Her examples include the mistreatment of animal corpses,[34] the mocking of animals through dressing them in human clothes, or misrepresenting their natures through inaccurate anthropomorphism. Abbate distinguishes between subject welfare and experiential welfare, and argues that an "[a]ction X is a dignitary wrong if the object of X has inherent value, yet is viewed or treated as a tool,

---

[34] Roger Payne describes a case in which the corpse of a porpoise had been mistreated. "It had been mutilated. Someone had hacked off its flukes for a souvenir. Two others had carved their initials deeply into its side and someone had stuck a cigar butt in its blowhole" (2013). According to Payne, the experience of finding of the corpse led him to his life of activism and biological research into cetaceans.

commodity, or consumable *or* if the object of X is not treated in accordance with the subject's specific nature." (2020, p. 775). For Abbate, the concept of a dignitary wrong can also be applied in the cases where the object of X *represents* a creature. Given the plausibility that recordings of humpback whale song represent the singer, or for general audiences stand as representations of the whole class of humpback singers, disrespectful treatment of such recordings, according to Abbate, would therefore be classed as "representational dignitary wrongs" (2020, p. 783).

The arguments put forward by Abbate from the discipline of animal ethics are complemented by arguments I propose here. I will use the term 'objectification' to describe the conversion of humpback song into manipulable representations. I take the concept from Critical Theory, in which it is synonymous with 'reification' and associated with what social theorist Max Weber (1864–1920) had called 'instrumental rationality', which we use to reason about the most efficient means to achieve a given end. For philosopher and musicologist Theodor W. Adorno (1903–1969), influenced by philosopher György Luckács (1885–1971), "[r]eification is a mode of perception which is prevalent wherever instrumental rationality and capitalism predominate. Instrumental rationality assumes that the world is made up of things we can manipulate" (O'Connor, 2013, pp. 201–202). Objectification describes the way in which commodity thinking comes to dominate those that live in the "over-administered world" (Adorno, 2002, p. 53) of modern Western society, transforming their very experience of self, world, and others:

> "the relationships of human beings, including the relationship of individuals to themselves, have themselves been bewitched by the objectification of mind. Individuals shrink to the nodal points of conventional reactions and the modes of operation objectively expected of them. Animism had endowed things with souls; industrialism makes souls into things" (Horkheimer and Adorno, 2002, p. 21).

With my usage of objectification I intend to connote both the literal object-making that representational processes enact through the conversion of the flux of sound into song unit names, audio clips and so on, and the 'objectifying' consciousness of the composer who henceforth takes such representations as freely available material for composition, without any acknowledgment that the sounds originally carried a function, and intrinsic value, for their singers and listeners. It should be clear that I intend objectification to carry the same negative charge as it carries in critical thought.[35] I claim that if we wish to take seriously the agency of other animals, the practice of objectification has ethical implications for the collection and use of field recordings of their sounds. Here, a starting point is offered by the way that postcolonial thought has begun to address Western practices of collecting recordings of Indigenous musics, as carried out by ethnomusicologists, individual musicians, and recording companies. Although the 'subordinated' (Plumwood, 2002) groups are not equivalent, the power structures share in the same logic, as I argue below (§4.5.2). In the context of Indigenous musics, past and ongoing exploitation and disempowerment have sharpened ethical questions of intellectual property and cultural appropriation (Feld, 1996; Young, 2008; Coleman *et al.*, 2009; Bellman, 2011; Andean, 2014; Naylor, 2014; D'Agostino, 2020; Robinson, 2020; White, no date). This is a live and complex debate, but one might reasonably argue that the objectification and use of Indigenous musics by Western composers, their transformation from "oral performance and cultural participation… into material commodity and circulable representation" (Feld, 1996, p. 13), has often reinforced inequity, created caricatures, and disrespected the agency of the producers. Although the details of the debate in the case of other-than-human animal sounds

---

[35] Ironically, given the recent history of exploitation of cetaceans, objectification is particularly straightforward for humpback whale vocalizations, where units are vocalized in staccato fashion, each gesture separated from others by silence. In electroacoustic composition especially this results in a clear danger of allowing instrumental rationality to predominate over other forms of response.

are different, and these differences matter, it seems a clear case of human exceptionalism to assume that what is problematic for humans is unproblematic for other animals.

It is notable that Adorno himself identified the operation of objectification throughout artistic activity and aesthetic experience, most notoriously within the ‘culture industry’ (Horkheimer and Adorno, 2002). Here we might consider the way in which the opening track on Payne’s *Songs of the Humpback Whale* (1970) has been sampled and come to circulate in human culture on dozens of pop songs (Appendix A, Rothenberg, 2008b, Feldman, 2021). However, more relevant here is Adorno’s description of the development of musical languages through notation: “It only became possible for music to develop through graphic mediation, reification, and availability – musical writing is the organon of musical control over nature” (Adorno, 2006, p. 53). That he regarded reification (enabled by instrumental rationality) as at once necessary *and* problematic is clear: “Rationalization, the condition for all autonomous art, is at once its enemy. Notation always also regulates, inhibits, and suppresses whatever it notates and develops” (2006, p. 53). But it is not only notation that involves rationalization, which he recognized as “the source of *all* of art’s means and methods of production” (Adorno, 2002, p.54). Paddison, in his discussion of Adorno’s aesthetics of music, connects this critique with the use of electronics in composition and performance:

> “It could be argued, for example, that the development of musical instrument technology and of performance technique has been characterized by progressive control over all aspects of the means of sound production, towards the ends of the purest possible sound throughout the whole range of the instrument (the element of control is seen at its extreme in the case of the new technology as applied to music).” (Paddison, 1993, p. 139)

Adorno's account of rationality in music did not stop with his identification of an instrumental rationality that operates during composition and performance. Indeed, he claimed that "art denounces the particular essence of a *ratio* that pursues means rather than ends" (Adorno, 2002, p. 330), and proposed a counter-acting "aesthetic rationality" (Adorno, 2002). According to Paddison, Adorno's aesthetic rationality "corresponds to some extent to Weber's value rationality" (1993, p. 139), i.e., "a belief in the intrinsic, absolute value of a particular form of behaviour for its own sake, irrespective of ends or consequences" (Paddison, 1993, p. 136). Moreover, aesthetic rationality has the somewhat paradoxical goal of the "*purposiveness without a purpose"* (Paddison, 1993, p. 140) that Kant had proposed belonged to artworks in his 'Critique of Judgment'. This enables Adorno to argue that aesthetic rationality is inherently opposed to, and even "irrational according to the criteria of" (2002, p. 330), instrumental rationality, and thus that artistic practices and artworks can offer an implicit critique of everyday instrumental rationality: "aesthetic rationality wants to make good on the damage done by nature-dominating rationality" (Adorno, 2002, p. 289).

To unpack this more fully would require an unjustifiably lengthy digression into Adorno's treatment of mimesis,[36] but I see here an opportunity to develop his aesthetic rationality into a form that treats its materials – in this case the objectified elements of humpback whale song – with the respect due to anything that either possesses intrinsic value or represents something that does. This aesthetic rationality, although it operates on these objectified elements as a

---

[36] Mimesis, as the artistic representation of humans, nature or natural processes, has been a key concept in aesthetics since Plato (*The Republic*) and Aristotle (*Poetics; Politics*). Mimesis has been described as "the most elusive notion in all of Adorno's philosophy" (O'Connor, 2013, p. 149), and as "a foundational concept never defined nor argued but always alluded to, by name, as though it had preexisted all the texts" (Jameson, 1990, p. 64). In *Dialectic of Enlightenment* (Horkheimer and Adorno, 2002) it is seen as a pre-rational means through which early humans attempted to gain power over nature (e.g., through shamanic imitation of animal behaviour); in *Aesthetic Theory* (Adorno, 2002) it is used to capture the way in which artists and their artworks identify with nature, and audiences interact with artworks; in *Towards a Theory of Musical Reproduction* (Adorno, 2006) it describes the relation between performer and musical work. I leave for the future a full exploration of how Adorno's notion of mimesis as "the nonconceptual affinity of the subjectively produced with its nonposited other" (Adorno, 2002, p. 54) might strengthen my proposal for an aesthetics of "difference without distance" (South, 2022).

means to an end (the completed work), values their integrity and allows them to determine the shape of what emerges from the creative processes. Such a rationality would value humpback singers and the structures of their songs *for their own sake*, and is just what is lacking in a blind application of instrumental rationality. It provides a way in which composers who work with the sounds of other animals can avoid both the logic of domination and the representational dignitary wrongs described by Abbate (2020).

To specify this aesthetic rationality in a more detailed way, we must take a step beyond Adorno's focus on human culture and on his somewhat abstract and distancing conception of "external nature" (Horkheimer and Adorno, 2002). Although Adorno was critical of the mistreatment of other-than-human animals (e.g., by behaviourist researchers and vivisectionists) and indeed foresaw an eradication of wild animals, he had a restricted view of animal cognition and sentience, describing animal existence as oscillating between desire and satiation, made dreamlike through the lack of human capacities such as concepts, language and complex emotion (Horkheimer and Adorno, 2002, p. 205). Adorno may have "recognized the close linkage between anthropocentric ontologies and ethics and the domination of animals" (Calarco, 2015, p. 34), but his critique of instrumental rationality and objectification comes from a time before any widespread recognition in Western philosophy of the interdependence of humans and the other-than-human. To take us further, I turn to the work of Val Plumwood (1939-2008), a founding voice in feminist philosophy and environmental ethics.

4.5.2 Resituating Humans and Other Animals

According to Plumwood, facing up to contemporary ecological crises requires us to acknowledge and value the interdependence of human and other-than-human worlds. She identified two tasks "that arise from the rationalist hyperseparation of human identity from nature… the tasks of (re)situating humans in ecological terms and non-humans in ethical terms" (Plumwood, 2002, pp. 8–9). These projects are interconnected, she argued, because hyperseparation prevents me from both empathizing with other species, and realizing that the fates of humans and the other-than-human are entwined (worldings are always multispecies worldings). In pursuing these tasks, Plumwood laid bare the logical structure of Western dualist thinking (hegemonic centrism), in a way that has clear parallels with Critical Theory, summarizing dualism as "a relation of separation and domination inscribed and naturalised in culture and characterised by radical exclusion, distancing and opposition between orders construed as systematically higher and lower" (Plumwood, 1993, p. 447). Applying and expanding her initial analysis of the exclusionary logic of sexism (1993), to anthropocentrism and colonialism (Plumwood, 2002, 2003),[37] she found five components in the ideological scaffolding underpinning oppression, including the objectification/instrumental rationality discussed above:

(1) The hyperseparation between the dominator and the dominated that is present in both colonialism and in anthropocentrism. Humans are

> "emphatically separated from nature and animals. From an anthropocentric standpoint, nature is a hyper-separate lower order, lacking any real continuity with the human. This approach stresses heavily those features that make humans different from nature and animals, rather than those we share with them" (Plumwood, 2003, p. 54).

An example of this distancing is the exclusion of other-than-human animal sounds from the domain of music, whether through stipulation, such as when music is defined as "humanly

---

[37] With respect to the development of what I have here called "ecological thinking" (Morton, 2018) and what Plumwood refers to as "environmental culture," she implicitly rejects a universalist environmentalism, explicitly arguing against "the kind of reverse ecological analysis, often originating in reductionist population biology, that reads the reductionism it adopts towards non-human species back into the human context and discounts the vital role of cultural difference" (2002, p. 3).

organized sound" (Blacking, 1973), or through invalid philosophical argument, as in Kivy's (1990, pp. 24–25) denial that birdsong can possess syntax (see Taylor, 2017b, for a rebuttal).

(2) The operation of "denial" or "backgrounding" that comes in to play once the dominated has been radically distanced from the centre, since "to acknowledge dependence on an Other who is seen as unworthy would threaten the One's sense of superiority and apartness" (Plumwood, 2002, p. 104). Thus human interdependence with the other-than-human is forgotten: e.g., "crucial biospheric and other services provided by nature, and the limits they might impose upon human projects, are not considered in accounting or decision-making" (Plumwood, 2003, p. 57).[38] Philosophers and musicologists may regard the structures of birdsong as irrelevant to an understanding of human music (e.g., Kivy, 1990), or at least of minor importance, even where they play a central role in a composer's music (Taylor, 2014, p. 67).

(3) The relational definition of the other (assimilation). Other-than-human animals are characterized in relation, typically negation, of humans.

> "We consider nonhuman animals inferior because they lack, we think, human capacities for abstract thought. But we do not consider as superior those positive capacities that many animals have and which we lack, such as remarkable navigational capacities." (Plumwood, 2003, p. 58).[39]

The sounds of other animals are regarded as defective versions of (often idealized forms of) human music or language, e.g., in failing to "conform to the harmonic rules of human music"

---

[38] Recent discussions of the potential mitigation of climate change brought about by recovering populations of cetaceans have included references to "whale-based carbon sink technology" (Chami *et al.*, 2019), with whales described as "marine ecosystem engineers" (Roman *et al.*, 2014). Although successful conservation efforts would benefit cetacean individuals and species, this way of putting things arguably views whales as tools for the benefit of *Homo sapiens*. Furthermore, so long as whales are not regarded as possessing intrinsic value, such instrumental arguments for whale conservation always have the potential to cut the other way (for example, if cetacean methane emissions turned out to be problematic).

[39] In the Glossary I discuss the use of the term 'nonhuman', which here almost seems a performative contradiction.

(Araya-Salas, 2012), or lacking the recursion said to be unique to human language (Hauser *et al.*, 2002).

(4) The combination of instrumental rationality and objectification in which the dominated are viewed as means to the dominators' ends. Other-than-human entities are denied agency, purpose and ethical considerability, their interests prioritized below those of humans. Instrumental attitudes "reduce nature to raw materials for human projects" (Plumwood, 2003, p. 59). This viewing of the other as resource is also a central element of the postcolonial critique of "extractivism," which has been recognized as a "fundamental expression of global capitalism… based on socio-ecologically destructive processes of subjugation, depletion, and non-reciprocal relations, occurring at all levels of practices." (Chagnon *et al.*, 2022). Although the concept has its origins in the analysis of colonial extraction of natural materials, it has also been extended to ways in which colonizing nations have acted towards the cultural creations and Indigenous knowledge of colonized peoples (Robinson, 2020; Vasquez and Cojtí, 2020; Clark, 2021). It is especially here that my crucial question arises: should the use of other-than-human sounds by human musicians be subject to ethical considerations?

(5) The homogenization (stereotyping) that is blind to diversity in the other. This relation is premised on power imbalance: the dominator replaces sensitivity to difference with power or force (Plumwood, 2002, p. 103). In Western representations of Ba-Benzélé Pygmy music "A single untexted vocalization or falsetto yodel" caricatures "a vast repertory of musical forms and performance styles" (Feld, 1996, p. 26–27). To paraphrase Ronald Reagan, "You've heard one humpback whale, you've heard them all" (original quotation in Plumwood, 2003, p. 55). There are parallels here with the way in which *Songs of the Humpback Whale* (Payne, 1970) has continued to be sampled by so many musicians despite the existence of many other field recordings illustrating the diversity of song (e.g., Payne, 1977; Rothenberg, 2015). Another manifestation of stereotyping is the way in which one recording is taken to be equivalent to

another: fungibility substitutes for individuality. This is a plausible explanation for the existence of multiple recordings of Alan Hovhaness's 'And God Created Great Whales' (1970) featuring different whale song recordings (South, 2022), although a more charitable interpretation would be that Hovhaness was attempting to raise awareness of the diversity of song, but time or circumstances did not permit the composition of new work.

I have taken the time to spell out this analysis because it provides a structured way to interrogate my own compositional relationship to humpback song. I now turn from critique to a set of positive and practical strategies for change, applying Plumwood's (2002) insights into binary anthropocentrism to composition with other-than-human sounds, and building on my earlier proposals on how to avoid anthropodenial and naïve anthropomorphism (South, 2022). I hope that these strategies may be of practical use to composers.

(1) To overcome hyperseparation and avoid either the distancing or exoticization of whale song, we can emphasize the continuity between the music of humans and the sounds of other animals, and between human musicians and other-than-human singers. This is arguably one effect of Roger Payne's release of *Songs of the Humpback Whale* (1970), is part of the implicit praxis of many composers, and has been a central plank in the theoretical project of zoömusicology (South, 2022). As a general principle, I believe that recognizing this continuity through the incorporation of other-than-human sounds into human musics is likely to enhance human appreciation of other-than-human singers and incline us towards less exploitative relationships with them (Rothenberg and Deal, 2015; Grover Friedlander, 2020). We are more likely to empathize with those with whom we feel we have something in common. Nevertheless, in order that this appreciation is not based in a fragile naïve anthropomorphism, we must also attend to difference (South, 2022).

(2) Recognizing continuity is also a step towards bringing other-than-human animal sounds out from the background, in recognizing that they, while not essential to all human musics, nonetheless play important and valued roles in the daily lives of most human beings. Recent scientific studies have shown that the appreciation of birdsong has benefits for human mental well-being (e.g., Ferraro *et al.*, 2020), and even Kivy (1990, p. 71–72) is happy to accept that the sounds of birds are beautiful and provide pleasure. Here, the composer can attempt to undo backgrounding, through the foregrounding of the sounds of other animals, whether in transcribed or recorded form, in their works.

(3) To counter relational definitions of the other-than-human, we require an "affirmation of non-human difference" (Plumwood, 2002, p. 112), as Nussbaum (2022) also argues in her 'capabilities' account of animal ethics. In my context this means that rather than describing other-than-human animal song negatively with respect to human music or language, I remain open to properties or capacities that I do not find in these human uses of sound. The affirmation of difference by the composer might include the refusal to manipulate tunings and durations in order to make nonhuman sounds fit into conventional harmonic or rhythmic contexts, i.e., a rejection of assimilation and domestication. It might also extend to spatial and temporal timescales of performance. In my '(im)possible gestures' I aimed at recreating the effect of being surrounded by multiple asynchronous humpback singers, and in the future I hope to produce one or more durational works lasting for the timescale of typical humpback song sessions.

(4) The avoidance of over-manipulation goes hand in hand with a rejection of objectification and the recognition of intrinsic value in the other-than-human. Plumwood urges that we nurture "some degree of human humility and sensitivity to nature's own creativity and agency… the cultivation or recovery of ways of seeing beings in nature in mind-inhabited ways as other centres of needs and striving" (2002, p. 113). I suggest that composers might allow their work

with other-than-human sounds to be guided by a ‘preservation’ model developed for the respectful use of traditional music in electroacoustic composition, which “[gives] priority to the recognisability of the traditional material and its cultural content” (D’Agostino, 2020, p. 107). This model is consistent with the aesthetic rationality described above (§4.5.1). D’Agostino argued against the exclusive use of transformative or destructive processing, the use of short audio clips, or focusing entirely on intrinsic sonic qualities, but does not shut down other possibilities for creative elaboration: “the aim is to create an artwork that will take into account contemporary languages and trends” (2020, p. 113) rather than simply present field recordings. I have found it useful to distinguish between *sonic selection,* i.e., the initial choice of material; *sonic reorganization*, e.g., altering the temporal order of sounds whether through notational means or the editing of field recordings; and *sonic transformation,* e.g., timbral changes when transcribing for human instruments, or in electroacoustic works: removing noise, adding reverb or delay, altering playback rates. For the composer all are essential to some degree. I accept that there is always going to be a tension between a sensitivity to the intrinsic value in the field recordings and the necessary objectification of other-than-human sounds for the purposes of composition (transcription, the use of audio clips) and a certain level of instrumentalism. The other-than-human sounds *are* being manipulated for human compositional goals, taking the work beyond straight transcriptions or wildlife recordings, even if such manipulation is as minimal as that found in soundscape compositions such as Ferrari’s ‘*Presque rien No.1*’ (1970). The tension may not always be resolved in the sonic content of the work: here additional contextualization may be provided by written texts, as in my ‘The Path of the Unseen Whale’, or alternative artistic modalities (e.g., in multimedia works).

(5) To restore diversity and individuality to the view of other-than-human song, I urge that composers pay closer attention to variability at all levels: within and between species, populations and individuals. Bioacousticians interested in the function of song are interested in

the relative levels of within- and between-individual variability: in my own work I investigated how this plays out with respect to the phrase rhythms of humpback song (Chapter 3). For composers there are multiple aspects of other-than-human sounds where pure diversity could be explored and presented to audiences.

An excellent example where inter-species diversity has been showcased is Blackford and Krause's 'The Great Animal Orchestra' (2014), intended as "a seamless integration of natural and orchestral soundworlds" (Pound, 2014). This skilfully crafted and melodic 30-minute work incorporates a large number of Krause's field recordings from across the globe, and has been described as "atmospheric, dramatic and populist in its approach" (Whitmore, 2014). The variety of sounds on display is indeed impressive, but how does the piece fare when judged against my other principles? Each of the five movements begins with a sample played from a keyboard: in the first the soundscape and gibbon song do little more than set the scene but in the third, fourth and fifth there are direct affective or melodic links with the ensuing orchestral writing. For example, the melody of the musician wren that opens the fifth movement is taken up by the piccolo (matching timbre) and then transformed by the clarinet. In the second movement, the looped recordings of frogs serve as a kind of other-than-human percussion section, setting a tempo and rhythm for the orchestral winds. Here it seems that "each different frog [recording] is assigned to a different [controller] key, meaning that they can be played just like any other instrument in the orchestra, at whatever tempo the conductor chooses" (Pound, 2014). What should we make of it? Blackford and Krause have arguably overcome the distance between human and other-than-human sounds, but I was left uneasy about their integration. It feels rather carefully managed, the animals deprived of autonomy and their sounds domesticated to fit in with the conventional orchestration. I could imagine the animals placed around the orchestra as though in a circus, waiting for the conductor to grant them their moment in the limelight. Although some animals perform for a reasonable duration, others are severely constrained, especially the unfortunate humpback whale, whose contribution is limited to two short song units towards the end of the first movement, fitting neatly into the harmonic context. Admittedly the same hierarchy is present among the human musicians, but it is hard to think of a better metaphor for domination and assimilation. There is also the appearance of a naïve anthropomorphism: that the animals would want or agree to sing in such circumstances. However, there are also sections when the relationship seems more equal, such as when the beaver vocalizations pass into the bassoon in the third movement, or when the common potoo dialogues with the trombone: here I felt a sense of interaction, of the humans being transformed by the other-than-human rather than the reverse. As the accompanying website and interviews with the composer suggest, the intention was surely to challenge anthropocentrism, and this does seem to have been transmitted to the audience: a reviewer concluded that "the message was clear: this is a world that man is doing its best to obliterate; if we don't recognise or celebrate it, it will die" (Evans, 2014).

4.5.3 Musical Manipulation Leads to Decontextualization and Recontextualization

This analysis of Blackford and Krause's work indicates the relevance not only of the degree of manipulation, but also of the musical context into which the other-than-human sounds are placed. In Andean's discussion of ethical issues involved in electroacoustic composition, he rather downplays the importance of contextualization prioritized by soundscape composers, claiming that this "ignores the primordial decontextualizing act: the severing of sound from source" (2014, p. 178).[40] However, I contend that we must still consider later stages of decontextualization, even once this primordial act is acknowledged. Moreover, I assert that every decontextualization is simultaneously a recontextualization: sounds severed from their sources are not left dangling in the middle of nowhere. In the case of other-than-human sounds,

---

[40] In the contrasting paradigm of acousmatic composition, where listeners ideally cannot identify the source of the sound, this primal decontextualization is taken still further.

especially where sonic transformation and reorganization are conservative, de/recontextualization can bring about the perception of exoticization or domestication and hence is of primary importance in shaping responses to and interpretations of a work. Even unedited field recordings, when packaged for aesthetic appreciation and situated within a commercial product in a marketplace, radically recontextualize the sounds with respect to their original function.

Additional questions of the effects of de/recontextualization arise through editing processes, whether or not these involve the introduction of human sonic elements. Selected sounds of a focal species may be displaced to varying degrees from the sonic elements of their social, cultural, and environmental contexts. These include but are not limited to the sounds of other members of the same species, other songs from their own or a population-level repertoire, the sounds of other species, or non-agential processes. Let us compare Payne's *Songs of the Humpback Whale* (1970) with Rothenberg's *New Songs of the Humpback Whale* (2015). In both cases the editing is described in sleeve notes. The older recording is notably noisier, partially no doubt as a result of more primitive recording technology, but also due to different decisions made by Payne and Rothenberg as producers. Payne rejected requests from his record company's engineers to filter out wave sounds (Feldman, 2021), and left in the sounds of boats and dynamite explosions in the first track. Payne's aim, it seems, was not just to present the whale song, but to show himself in the act of listening to it: "As you listen to this recording, I wish only that it could convey to you the pleasant circumstances under which we made it" (Payne, 1970). By comparison, Rothenberg appears to wish us to focus fully on the sounds of the whales. With the exception of one track, noise reduction software has been used to minimize environmental sounds. He describes his approach as one where he "massages the sound to reveal the beauty and detail that lies within" (Rothenberg, 2015). We certainly hear more acoustic structure, more timbral detail, in the result. Coming to it from the older disc, with its splashes, crackles and echoes, Rothenberg's is an astonishing recording. To me, the whales sound closer to hand, more dangerous, more exotic, more radically Other. Perhaps this points to the effects of decontextualization. In removing the whales from their splashy surroundings, Rothenberg has arguably recontextualized them into the silent vacuum of outer space. If to this 21st century listener the original 'Songs' can sometimes sound like a kind of watery Pink Floyd, the new disc could be an alternative soundtrack to Ridley Scott's 'Alien'.

Moving now to more interventionist processes of sonic selection and reorganization performed with human compositional aims, I note that these can alter the music-like context of what is represented. For example, the musical effect on human listeners of a humpback phrase within a theme, where it is heard as an element of repetition with variation, will be fundamentally altered if a single phrase is mechanically looped. Finally, as we saw with 'The Great Animal Orchestra' (Blackford and Krause, 2014), and in other examples I have discussed elsewhere (South, 2022), the positioning of other-than-human sounds in human harmonic and rhythmic contexts strongly influences how audiences hear and interpret them and consequently also the human relationships to other-than-human worlds. Audiences are, in part, led to listen and hear how composers listen and hear, which is one reason why composers bear an ethical responsibility in their use of other-than-human sounds.

I now turn to my portfolio, examining how manipulation in my own compositional processes has led to de/recontextualization and evaluating my works with respect to hyperseparation, backgrounding, relational definition, objectification and stereotyping (§4.5.2). Here I was not aiming to achieve the separation of sound and source considered vital by many composers working in the acousmatic tradition (Emmerson, 1986, p. 28). Rather, as with soundscape compositions, I wanted listeners to be able to attribute sounds of humpback whales to their sources, although this attribution might be gradual or delayed. In 'Whale, Bow, Echo', for example, where the humpback whale sounds eventually emerge out of violin and cello imitations, I hoped that for a short while the listener be uncertain as to their source, before coming to realize that these sounds are not being produced by the human performers.

Sonic selection and reorganization

In the manipulation of structure, I was quite conservative. With the exception of 'The Path of the Unseen Whale', I did not alter the composition of a phrase in terms of its rhythm and pattern of song units. In 'Underwater Rain', an unedited three-minute stretch of song was used to dictate the phrase and theme structure of the accompanying bass clarinet line. This fidelity to the sequence of phrases was also followed in the field recording which appears at the end of 'Whale, Bow, Echo', though in this case three different themes were layered over each other to give the impression of multiple whales singing. By contrast, in '(im)possible gestures' and the live parts of 'Whale, Bow, Echo', phrase variant ordering was determined by the performers and I allowed themes to follow each other in a different sequence to that presented by the field recording. Here we might talk about decontextualization, because sections were extracted from whole humpback songs, or song recordings. Of course this is true. Yet I had to select *something*, and in doing so aimed at being faithful both to the song's multilevel structure and to the complex syntax that seems to operate at the level of the phrase variant (what might be described as the rules of variation). Only in 'The Path of the Unseen Whale' did I drill down further, to the subphrase, where I selected from those transcribed for '(im)possible gestures' on the basis of their suitability for performance on the bass clarinet. In 'Underwater Rain' and 'Whale, Bow, Echo', the humpback recordings were recontextualized against backdrops of humanly produced sounds (live and pre-recorded) and, in 'Underwater Rain', field recordings of wave and rain sounds made in Mo'orea. In an attempt to overcome hyperseparation without eliminating difference, my strategy for both these pieces was that human and humpback sounds would be heard both together and apart.

Sonic transformation

In the transcribing of the song units used in ‘(im)possible gestures’ and ‘The Path of the Unseen Whale’, I transposed some up or down an octave either to bring them into the range of clarinet/bass clarinet, or to facilitate a smoother portamento. Similar procedures were followed during the collaborative work with the players of Sequoia, in arriving at the transcriptions of ‘Whale, Bow, Echo’. In ‘(im)possible gestures’, the duration of the sounds was stretched by a factor of three and their relative durations vastly simplified. In ‘Underwater Rain’, I reduced the playback rate of the field recording by a factor of two, lowering its pitch by an octave. This was in order to emphasize for the human ear what I perceived as expressiveness in the long drawn-out high song units, though the consequent shift of the low song units into the low bass clarinet register was also welcome for the way in which this created the psychoacoustic effect of ‘beating’ between the sounds of bass clarinet and whale. In ‘Underwater Rain’, the manipulation of sound was not restricted to the humpback recording: in its opening section (0m24–1m44) I also took a section of the bass clarinet line down an octave. Both whale song and human produced sounds were further subjected to complex delay effects.

In all pieces involving recorded whale song, I added varying degrees of reverb, partially to compensate for their lo-fi quality, but also in an attempt to simulate an immersive underwater environment.[41] This recontextualization could be interpreted as a response to Andean’s “primordial decontextualizing act” (2014, p. 178) by repositioning the listener. Other processing was carried out to remove non-humpback sounds. My source recordings, made near the reef surrounding Mo’orea, were characterized by varying levels of crackling sounds originating from vast numbers of snapping shrimp (*Alpheidae*), a diverse group comprising 600+ species (Anker *et al.*, 2006). These broadband high frequency sounds “dominate temperate and tropical coastal soundscapes worldwide” (Lillis and Mooney, 2022). A snapping

---

[41] It was hard to escape the influence of *Songs of The Humpback Whale* (Payne, 1970).

shrimp has a specialized claw, “a powerful, multifunctional tool used for defence and aggression,” that when snapped shut generates a short and extremely loud sound that may be heard up to 1 km distant (Anker *et al.*, 2006). It is thought that the crackling may be used by humpback whales as a navigation aid around coastlines during migration (Lillis and Mooney, 2022). In my focus on humpback whale song, I initially treated snapping shrimp sounds as noise. I both attempted to remove the sounds, and treated them as a sonic resource whenever I wanted to simulate an oceanic environment, e.g., in ‘Whale, Bow, Echo’ they can be heard in the introduction leading to the first entries on cello and violin, in ‘Underwater Rain’ in the section immediately before the solo humpback is first heard (1m54). In ‘The Path of the Unseen Whale’, where the fixed media tracks comprise humpback song field recordings with the humpback sounds largely removed, the sounds of the snapping shrimp are more prominent, but were edited without any regard for concern for their integrity. Did this manipulation for my own creative purposes amount to disrespect to the snapping shrimp? I’m not sure: the recorded sounds come from a large number of individual organisms. Furthermore, given their rudimentary neural architecture, it may be justifiable to attribute them a lower degree of agency on Calarco’s ‘continuum’ (2020, p. 8). However, as sentient creatures the snapping shrimp still possess intrinsic value. To recognize this, I now acknowledge their presence on the recording in my programme notes for these pieces (Appendix C).

4.6 Hearing Heterophony

In the preceding sections, I have focused on the activities of the composer, asking whether the objectification and manipulation of other-than-human sounds share in the logics of domination and extractivism. Following Adorno, I accept that a degree of objectification and instrumental rationality is inevitable in any musical language. I have argued that manipulation leads to de/recontextualization, which in turn risks the exoticization or domestication that are symptomatic of anthropodenial or naïve anthropomorphism. I have suggested, however, that these risks may be reduced through an aesthetic rationality that recognizes the intrinsic value of the sound-maker, and compositional strategies designed to overcome the elements of anthropocentrism identified by Plumwood: hyperseparation, backgrounding, assimilation, objectification and homogenization.

I now turn to the completed work, and its reception by an audience. Are the attitudes of the composer, and the procedures they follow, heard in its performance? Where traces of domination are heard in a work, as for example White (no date) has argued of Stockhausen's treatment of samples of non-Western musics in his 'Telemusik', we might ask how they are interpreted. White interprets this work as one that involves "unification through annihilation", the treatment of non-Western musics as revealing Stockhausen's "opinion that they are 'the exotic' and thus essentially identical." In response, I argue that we must not essentialize the audience: "each listener will bring his or her own knowledge to their understanding of the meaning of the material" (Naylor, 2014, p. 110). A multiplicity of interpretations are always available and indeed inevitable. I interpreted the domination/assimilation that can be heard in 'A Great Animal Orchestra' as problematic, but another listener might conclude that its mirroring of extractivist logics could be read as critique. This is one direction taken by Adorno, as when he writes that "Art is modern art through mimesis of the hardened and alienated; only thereby, and not by the refusal of a mute reality, does art become eloquent; this is why art no longer tolerates the innocuous" (2002, p. 21). Alternatively, the musical domination might be celebrated as some form of fantasy, as if the circus I imagined were something to be aimed for, or viewed simply as realism: this world just *is* one in which humans dominate other animals. All this is to accept that the artist must take responsibility for the work, but that ultimately the responsibility for the interpretation of the work is shared with the listener:

> "The work of art is entirely transactional – a cultural negotiation, with artist and audience as the primary agents. As an independent unit, the art work is

> beyond the reach of ethics; as a locus of cultural communication, exchange and interaction, ethics are fully implicated in the very heart of the art work" (Andean, 2014, p. 178).

The existence of multiple interpretations is not in itself problematic; this multiplicity is exactly what gives rise to the sorts of conversations that can challenge and advance the critical self-knowledge of one's own presuppositions. Nonetheless, the ever-present possibility of misinterpretation in nonverbal music – even field recordings can misrepresent their sources – does imply that what a composer writes about their work matters.

Taking this to be true, or at least, plausible, I close this chapter by providing a unitary interpretation of one aspect I take to be central to each of my portfolio works, the polyphonic form in which their musical and music-like lines are interwoven. It has been important to me from the earliest pieces in this portfolio that I be faithful to the structure of the song of the humpback whale, and yet I have been challenged from the beginning by the apparent lack of interaction in the collective singing of humpback whales, commonly understood to be "singing the same song" (Payne and Guinee, 1983). Unlike most groups of human musicians who come together to perform, there seems to be no attempt to temporally coordinate phrases or themes, and limited evidence that the performance of one singer directly affects the performance of another (§1.2.4). This has challenged me because in my practice as a performer and improviser, I have always valued sensitivity and responsiveness. How should I deal with this conflict in my composition? In some instances I have thematized it: in '(im)possible gestures', the absence of interaction between players at the start of the piece transforms into explicit listening and unscripted responsiveness. In 'Underwater Rain', human and whale rhythms proceed along parallel yet never fully synchronized paths. 'Whale, Bow, Echo' enfolds the lack of coordination into its opposite, as the players must attend closely to each other in order avoid falling into synchronicity. The human interactions in the duo improvisations at the heart of

‘Entanglement’ are electronically layered into a heterogeneous multiplicity of unsynchronized audio tracks. In other cases, sensitivity has been channelled towards absence, as in ‘The Path of the Unseen Whale’, in which the final melodic solo of the bass clarinet plays out in the gaps expressing the flukeprints of the unseen whale. Only in ‘Submergence’ does one line really follow another closely in time, and here the responsiveness of the live part is directed towards the animation and occasional subverting of the pre-recorded track rather than its mimicry.

I propose that despite their differences, these multiple forms all share in an emergent “multispecies heterophony” that enacts an overlapping nonhierarchical sounding of difference and thereby answers precisely to my earlier call for an aesthetics of “difference without distance” (South, 2022). Heterophony is a form that is present in many musics worldwide, including Hebridean psalm-singing and Polynesian polyphonic *hīmene tārava*, in which individual voices inscribe their own criss-crossing musical paths on a common ground. Such music is frequently less about aesthetic experience and more about “community made audible and visible… where group participation, sonic power, and individuality within the group are all more important than a homogeneous sound” (Diettrich *et al.*, 2011, p. 16). In my own music-making, I have experienced this “framework of abundance, of tones and subjectivities” (Coleman, 2021, p. 282) most profoundly when playing free jazz in large improvising ensembles. In his postcolonial reformulation, Coleman theorized heterophony as

> “an opaque, decentralized musical texture... [that] opens up new epistemological terrain in the context of experimental improvised music by affording multiple simultaneous subjectivities… interpolating the listener into a dynamic and constantly shifting sonic mesh” (pp. 294–295).

Among my portfolio works, Coleman’s description perhaps applies most precisely to the live performance of ‘(im)possible gestures’ in which the Scottish Clarinet Quartet recording was used as a basis for a group improvisation by the jazz quartet in residence at the Phillip Island Whale Festival (Victoria, Australia, July 2019). Here, I assert that this form, which echoes the “asynchronous chorus” (Herman, 2017) of the collective singing of humpback whales, provides a promising nonhierarchical model for how human members of industrialized societies might better listen to and acknowledge their living co-presence in shared habitats with other sentient agents as “dense ecologies of selves” (Kohn, 2013, p. 193).

4.7 Future Work

As was perhaps inevitable in a project of this nature, the number of musical re-compositions that I was able to complete make up a only small proportion of the ideas that I would have liked to pursue. With my increasing awareness of and interest in the idiosyncrasies of individual humpback singers noted by researchers (Payne, 1995, p. 148), confirmed with evidence at the level of phrase structure (Murray *et al.*, 2018; Lamoni *et al.*, 2023) and now with respect to phrase rhythms (Chapter 3), I hope to write new works that thematize these individual patterns in a more precise way than I achieved in ‘(im)possible gestures’ and ‘Whale, Bow, Echo’. These works might, by drawing on research into how song changes over time (e.g., Guinee *et al.*, 1983; Payne *et al.*, 1983; Payne and Payne, 1985; Cerchio *et al.*, 2001; Garland *et al.*, 2011), portray how individual variations and innovations are taken up and drawn into the evolution of the population-level song. They could be realized in durational form, with changes occurring over several hours or even days. Live performances are what I live for, but COVID-19 showed us new possibilities for online musical collaborations, and the internet is in some ways an ideal platform for this kind of work; I envisage a version of it with multiple online contributors continually updating their version of the song in an international heterophony. From my application of Plumwood’s analysis of anthropocentrism I have realized that it is also important for musical intermediaries to highlight the sounds of those other cetaceans whose voices are perhaps further from human musics. There is a wide diversity of song and non-song vocalizations that are rarely heard outside bioacoustics laboratories, and which could be given human musical settings in order to broaden awareness of the rich soundscapes that lie below the waves.

In addition to these musical possibilities, I intend to continue work on some of the more philosophical and musicological topics that have been opened up by this chapter. One such is that of the representation of time itself. Spectrograms, graphs, musical notations all present time in a linear, spatialized form that philosophers (including Adorno and Henri Bergson) and thinkers on music (e.g., Christopher Hasty, Victor Zuckerkandl, Olivier Messiaen) have all recognized as being inadequate to the task of representing the experiential time of music, whether from the point of view of performer or audience. New avenues (musical, philosophical, even scientific) might be opened up by an analysis of other-than-human rhythms that recognized the thickness of experiential time: its enfolding of the past and anticipations of the future. This direction takes us towards a path already mentioned in this chapter (§4.5.1), that laid out by Adorno in his analysis of an aesthetic rationality that operates against the prevailing instrumental rationality. Adorno held that this aesthetic rationality is bound up with mimesis, as a human capacity enabling our embodied and nonconceptual affinity with nature and art, and I hope to push these ideas further. New artistic experiences, insights into cetacean subjectivity, and empathic connection might be gained by a musical mimetics of the sort brilliantly explored by artists such as Hanna Tuulikki. Perhaps they can be made into participatory artistic practices that will help nurture a grounded human sensitivity and response-ability to the other sentient agents who throng our more-than-human planet.

Glossary

Term	**Definition**
Agency	The ability to act and make choices usually ascribed to humans and often denied to other animals. I take the approach that agency exists on a "continuum from (a) the minimal sense of animals having some effect on history and the world, to (b) more complex individual acts of choice or decision-making, to (c) the maximal sense of animals engaging in fully conscious, intentional agency in collective and organized forms" (Calarco, 2020, p. 8), and assume that humpback whales exhibit agency at least in sense (b). This is consistent with Nussbaum's capabilities approach to animal ethics, which insists on the intrinsic value of "species-specific forms of agency" (2022, p. 117).
Anthropocentrism, Binary	Although humans cannot escape a certain human-centred view of other living and non-living entities, this "default anthropocentrism" (Martinelli, 2009, p. 19) does not necessarily entail either that humans need consider only their own interests (Plumwood, 2002, p.132), or that they should separate themselves off as superior from all other organisms, a position Martinelli terms "binary anthropocentrism" (2009, p. 19). I take binary anthropocentrism to be synonymous with "human exceptionalism".
Anthropodenial	"[T]he *a priori* rejection of shared characteristics between humans and [other] animals when in fact they may exist" (de Waal, 1999).

Term	**Definition**
Anthropomorphism, Critical	I use this term, introduced by ethologist Gordon Burghardt (1991), to describe a stance from which we use all that we know about an animal's sensory world and cognitive abilities, its natural history and phylogeny, in order "to understand [other] animals on their own terms" (de Waal, 1999, p. 273). I have argued that adopting this stance is necessary for the avoidance of human exceptionalism (South, 2022). It allows us to steer between anthropodenial and naïve forms of anthropomorphism.
Anthropomorphism, Naïve	The uncritical and likely inaccurate projection of human characteristics (physical, psychological, socio-cultural) directly onto other animals, without consideration of crucial differences and the limits of the data involved.
Bioacoustics	"A branch of science concerned with the production of sound by and its effects on living organisms" (Merriam-Webster, no date).
Biomusicology	*"Biomusicology* is the biological study of musicality in all its forms. Human *musicality* refers to the set of capacities and proclivities that allows our species to generate and enjoy music in all of its diverse forms" (Fitch, 2018). Biomusicologists tend to be interested in comparative studies of music-like behaviours in other animals insofar as they shed light on the evolution of human musicality.
Coding	In the bioacoustics of cetacean sounds, the representing of sequences of sounds in symbolic form.

Term	**Definition**
Critical Theory	In a narrow sense, Critical Theory applies to the thinkers of the Frankfurt School, including Adorno, Horkheimer, Marcuse and Habermas. In a broader sense, a theory is referred to as ‘critical’ when it seeks liberation for the oppressed. In both senses, “a critical theory provides the descriptive and normative bases for social inquiry aimed at decreasing domination and increasing freedom in all their forms” (Bohman *et al.*, 2021). It has typically involved a combination of empirical research and philosophical argumentation, and in recent decades disciplines such as critical animal studies and ecocriticism have expanded the scope of concern from human beings to encompass other-than-human entities.
Culture	It is possible to define culture as “those group-typical behaviour patterns shared by members of a community that rely on socially learned and transmitted information” (Laland and Hoppitt, 2003, p. 151). This definition is broad enough that it can be, and has been, applied to the behaviours of other-than-human animals. In this sense, the cultural behaviours of cetaceans include songs, migratory routes, and foraging techniques (Whitehead and Rendell, 2015; South *et al.*, 2022).
Ethnomusicology	“Ethnomusicology is the study of why, and how, human beings are musical” (Rice, 2014, p. 1). In this context, musical refers to the capacities possessed by humans that underlie and support their musicking (see below).

Term	**Definition**
Ethology	The scientific study of the behaviour of animals other than human animals, usually in their natural setting.
Heterophony	In musicology, this term has been used to describe the temporal form of human musics, such as Gaelic psalm-singing, in which “multiple individuals simultaneously produce minor variations of the same material without temporal synchronization” (South, 2022, p. 56). Coleman has radicalized the term to reference the “multiple simultaneous subjectivities” involved in free improvisation (2021, p. 295). Its applicability to the collective singing of humpback whales has also been recognized (Sorce Keller, 2012; South, 2022).
Human Exceptionalism	“[T]he idea that humankind is radically different and apart from the rest of nature and from other animals... providing ideological background conditions for inferiorizing animals” (Plumwood, 2007). I take human exceptionalism to be synonymous with ‘binary anthropocentrism’ (see above).
More-than-human	In Chapter 4 I use this term, originally employed by philosopher David Abram (1997), to signal an awareness that human activities are situated within and codependent with a vast array of other living organisms, many of which are agential (in more than a minimal sense) and sentient.
Musicking	The taking part “*in any capacity, in a musical performance*” (Small, 1998, p. 9), including performing, listening, rehearsing, practising and composing.

Term	**Definition**
Other-than-human	Used here to describe animals other than humans, without the implication of lesser value that is often carried by ‘nonhuman’ (Calarco, 2020, p. 100). The description of animals, their subjectivity, agency, sociality, capacities and behaviours as nonhuman may imply the lack of some essential, usually privileged, human quality. In posthumanism and anthropology other-than-human may also extend to non-living entities (Lien and Pálsson, 2021).
Portamento	In Western music, term sometimes used interchangeably with ‘glissando’ to describe the practice of gliding smoothly from one pitch to another. Here I have also applied it to the melodic contours of humpback whale song, which commonly move across intervals of several tones.
Quality	In the study of sexual selection, the ‘quality’ of an individual is the degree to which that individual possesses the kind of traits likely to enhance reproductive success. The quality of a potential mate is often judged through its reproductive advertisement display (see ‘Sexual Selection’).
Rhythm	The pattern of timing in a short sequence of sounds.

Term	**Definition**
Sexual Selection	Sexual selection is "the evolutionary process that arises from competition among members of one sex (usually male) for access to the limiting sex (usually female)" (Bradbury and Vehrencamp, 2011, p. 470). Intrasexual selection may involve direct competition between "courters" (Rosenthal, 2017), e.g., through physical contest, or mediation via a reproductive advertisement display (such as song) that permits the competing sex (typically male) to establish a dominance hierarchy without potentially costly conflict. Intersexual selection involves choice by members of the limiting sex, who may assess the displays of the courters to judge their relative quality (see above). In most species, both intra- and intersexual processes play a role in the selection of a mate (Bradbury and Vehrencamp, 2011, p. 471), and are thought to lead to the evolution of the advertisement display.
Song	In bioacoustics, the definition of 'song' varies by taxon but for birds and cetaceans usually refers to repeated patterns of sounds used as sexual displays and often acquired through learning, rather than the sometimes simpler 'calls' that occur in other behavioural contexts.

Term	**Definition**
Stereotype	In bioacoustics, a stereotyped sequence of sounds is one that is repeated with little variability, whether by an individual or within a population. In Chapter 2 I use ‘rhythmic stereotype’ to refer to the characteristic phrase rhythm of a phrase variant. In Chapter 4 I also use the term ‘stereotyping’ to describe how members of a dominant class deny or disparage individuality or diversity in the subordinated (Plumwood, 2002, pp. 102–103).
Subjectivity	The experiences of other-than-human animals are sometimes reduced to those resulting from sentience, “the capacity of an individual organism to sense and feel things” (Calarco, 2020, p. 121), especially pleasure and pain. But it is plausible that subjectivity, like agency (see above), admits of degrees. For example, subjectivity may additionally involve the possession of conscious preferences, beliefs, memories and expectations that make up what it is to be the *subject-of-a-life*, in philosopher Tom Regan’s formulation (Calarco, 2015, p. 16). In humans and possibly a number of other animals including the bottlenose dolphin (Reiss and Marino, 2001), it further involves reflective self-consciousness, i.e., being aware of oneself *as* oneself (Smith, 2020).
Transcription	The representation of sounds in musical notation.

Term	**Definition**
Umwelt	Literally translated as ‘surrounding world’, or ‘environment’, biologist Jakob von Uexküll’s concept is better rendered “the world of or for some subject” (O’Neil, 2010, p. 35). The *Umwelt* is the lived world of perception and action that is specific to each sentient organism, where objects and other organisms show up in relation to that organism’s goals and needs. The ocean is a very different place for a human swimmer and a humpback whale.
Whale Music	I follow Ritts (2017) in using ‘whale music’ to describe human music inspired by and/or including field recordings of whale song.
Worlding	I use this term to capture the idea that the world of any organism is actively created through its interactions and becomings-with other organisms and its environment. “The notion of worlding insists on the co-constitution, the material-semiotic interplay, that shapes what is” (van Dooren *et al.*, 2016, p. 12). In its material aspect, it is related to biology’s notion of niche construction (Odling-Smee *et al.*, 1996), in its semiotic aspect it connects with von Uexküll’s *Umwelt* (see above) and processes of sense-making through which things come to have meaning for, or matter to, sentient organisms (De Jaegher and Di Paolo, 2007).
Zoömusicology	“[T]he study of music-like aspects of sound communication among non-human animals” (Doolittle and Gingras, 2015).

References

Abbate, C.E. (2020) ‘Valuing animals as they are —Whether they feel it or not’, *European Journal of Philosophy*, 28(3), pp. 770–788. Available at: https://doi.org/10.1111/ejop.12521.

Abram, D. (1997) *The spell of the sensuous: Perception and language in a more-than-human world*. New York : Vintage Books.

Abramson, J.Z., Hernández-Lloreda, M.V., García, L., Colmenares, F., Aboitiz, F. and Call, J. (2018) ‘Imitation of novel conspecific and human speech sounds in the killer whale (Orcinus orca)’, *Proceedings of the Royal Society B: Biological Sciences*, 285(1871), p. 20172171. Available at: https://doi.org/10.1098/rspb.2017.2171.

Acerbi, A. and Bentley, R.A. (2014) ‘Biases in cultural transmission shape the turnover of popular traits’, *Evolution and Human Behavior*, 35(3), pp. 228–236. Available at: https://doi.org/10.1016/j.evolhumbehav.2014.02.003.

Adam, O., Cazau, D., Gandilhon, N., Fabre, B., Laitman, J.T. and Reidenberg, J.S. (2013) ‘New acoustic model for humpback whale sound production’, *Applied Acoustics*, 74(10), pp. 1182–1190. Available at: https://doi.org/10.1016/j.apacoust.2013.04.007.

Adorno, T.W. (2002) *Aesthetic Theory*. Edited by G. Adorno and R. Tiedemann. Translated by R. Hullot-Kentor. London ; New York: Continuum.

Adorno, T.W. (2006) *Towards a Theory of Musical Reproduction: Notes, a Draft and Two Schemata*. Edited by H. Lonitz. Translated by W. Hobin. Cambridge UK ; Malden MA: Polity Press.

Agamben, G. (2004) *The Open: Man and Animal*. Translated by K. Attell. Stanford, California: Stanford University Press (Meridian, crossing aesthetics).

Aiken, L.S. and West, S.G. (1991) *Multiple Regression: Testing and Interpreting Interactions*. London: SAGE Publications.

Allen, A.N., Harvey, M., Harrell, L., Jansen, A., Merkens, K.P., Wall, C.C., Cattiau, J. and Oleson, E.M. (2021) ‘A convolutional neural network for automated detection of humpback whale song in a diverse, long-term passive acoustic dataset', *Frontiers in Marine Science*, 8, p. 607321. Available at: https://doi.org/10.3389/fmars.2021.607321.

Allen, C. and Bekoff, M. (1997) *Species of Mind: The Philosophy and Biology of Cognitive Ethology*. MIT Press.

Allen, J.A., Garland, E.C., Dunlop, R.A. and Noad, M.J. (2018) 'Cultural revolutions reduce complexity in the songs of humpback whales', *Proceedings of the Royal Society B: Biological Sciences*, 285(1891), p. 20182088. Available at: https://doi.org/10.1098/rspb.2018.2088.

Allen, J.A., Garland, E.C., Dunlop, R.A. and Noad, M.J. (2019) 'Network analysis reveals underlying syntactic features in a vocally learnt mammalian display, humpback whale song', *Proceedings of the Royal Society B: Biological Sciences*, 286(1917), p. 20192014. Available at: https://doi.org/10.1098/rspb.2019.2014.

Allen, J.A., Garland, E.C., Garrigue, C., Dunlop, R.A. and Noad, M.J. (2022) 'Song complexity is maintained during inter-population cultural transmission of humpback whale songs', *Scientific Reports*, 12(1), p. 8999. Available at: https://doi.org/10.1038/s41598-022-12784-3.

Altherr, S. and Hodgins, N. (2018) *Small Cetaceans, big problems - A global review of the impacts of hunting on small whales, dolphins and porpoises*. Animal Welfare Institute, Pro Wildlife, Whale and Dolphin Conservation, p. 67. Available at: https://awionline.org/sites/default/files/press_release/files/AWI-ML-Small-Cetaceans-Report.pdf.

Andean, J. (2014) 'Towards an ethics of creative sound', *Organised Sound*, 19(2), pp. 173–181. Available at: https://doi.org/10.1017/S1355771814000119.

Andriolo, A., Zerbini, A.N., MoreiraA, S., Pizzorno, J.L., Danilewicz, D., Maia, Y.G., Mamede, N., De Castro, F.R. and Clapham, P. (2014) 'What do humpback whales Megaptera novaeangliae (Cetartiodactyla: Balaenopteridae) pairs do after tagging?', *Zoologia*, 31(2), pp. 105–113. Available at: https://doi.org/10.1590/S1984-46702014000200001.

Anker, A., Ahyong, S.T., Noel, P.Y. and Palmer, A.R. (2006) ‘Morphological phylogeny of Alpheid shrimps: Parallel preadaptation and the origin of a key morphological innovation, the snapping claw’, *Evolution*, 60(12), pp. 2507–2528. Available at: https://doi.org/10.1111/j.0014-3820.2006.tb01886.x.

Antunes, R., Schulz, T., Gero, S., Whitehead, H., Gordon, J. and Rendell, L. (2011) ‘Individually distinctive acoustic features in sperm whale codas’, *Animal Behaviour*, 81(4), pp. 723–730. Available at: https://doi.org/10.1016/j.anbehav.2010.12.019.

Araya-Salas, M. (2012) ‘Is birdsong music?’, *Significance*, 9(6), pp. 4–7. Available at: https://doi.org/10.1111/j.1740-9713.2012.00613.x.

Atterton, P. (2011) ‘Levinas and our moral responsibility toward other animals’, *Inquiry*, 54(6), pp. 633–649. Available at: https://doi.org/10.1080/0020174X.2011.628186.

Atterton, P. (2018) ‘Levinas’s Humanism and Anthropocentrism’, in M.L. Morgan (ed.) *The Oxford Handbook of Levinas*. Oxford University Press, pp. 709–730. Available at: https://doi.org/10.1093/oxfordhb/9780190455934.013.10.

Au, W.W.L., Pack, A.A., Lammers, M.O., Herman, L.M., Deakos, M.H. and Andrews, K. (2006) ‘Acoustic properties of humpback whale songs’, *The Journal of the Acoustical Society of America*, 120(2), pp. 1103–1110. Available at: https://doi.org/10.1121/1.2211547.

Bailey, N.J., Cremel, T. and South, A. (2014) ‘Using Acoustic Modelling to Design and Print a Microtonal Clarinet’, in *Proceedings of the 9th Conference on Interdisciplinary Musicology. CIM14*, Berlin.

Bailey, N.J., South, A., Evans, W. and Hair, G. (2017) ‘A microtonal wind controller: Building on Yamaha’s technology to facilitate the performance of music based on the “19-EDO” scale’, *Scottish Music Review*, 4.

Barlow, H. and Morgenstern, S. (1949) *A Dictionary of Musical Themes*. London: Williams and Norgate.

Bateson, P. and Laland, K.N. (2013) ‘Tinbergen’s four questions: An appreciation and an update’, *Trends in Ecology & Evolution*, 28(12), pp. 712–718. Available at: https://doi.org/10.1016/j.tree.2013.09.013.

Bellman, J.D. (2011) ‘Musical voyages and their baggage: Orientalism in music and critical musicology’, *The Musical Quarterly*, 94(3), pp. 417–438. Available at: https://doi.org/10.1093/musqtl/gdr014.

Beltran, J.F., Liu, X., Mohanchandra, N. and Toussaint, G.T. (2014) ‘Measuring musical rhythm similarity: Statistical features versus transformation methods’, *International Journal of Pattern Recognition and Artificial Intelligence*, 29(02), p. 1550009. Available at: https://doi.org/10.1142/S0218001415500093.

Bispham, J. (2006) ‘Rhythm in music: What is it? Who has it? And Why?’, *Music Perception*, 24(2), pp. 125–134. Available at: https://doi.org/10.1525/mp.2006.24.2.125.

Blackford, R. and Krause, B. (2014) *The Great Animal Orchestra: Symphony for Orchestra and Wild Soundscapes* [Musical Score]. Monmouth UK: Nimbus Music Publishing.

Blacking, J. (1973) *How Musical is Man?* Seattle: University of Washington Press.

Bohman, J., Flynn, J. and Celikates, R. (2021) ‘Critical Theory’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Spring 2021. Metaphysics Research Lab, Stanford University. Available at: https://plato.stanford.edu/archives/spr2021/entries/critical-theory/ (Accessed: 30 July 2023).

Botero, C.A. and de Kort, S.R. (2013) ‘Learned signals and consistency of delivery: a case against receiver manipulation in animal communication’, in U.E. Stegmann (ed.) *Animal Communication Theory: Information and Influence*. Cambridge: Cambridge University Press, pp. 281–296.

Botero, C.A., Rossman, R.J., Caro, L.M., Stenzler, L.M., Lovette, I.J., De Kort, S.R. and Vehrencamp, S.L. (2009) ‘Syllable type consistency is related to age, social status, and reproductive success in the tropical mockingbird’, *Animal Behaviour*, 77(3), pp. 701–706. Available at: https://doi.org/10.1016/j.anbehav.2008.11.020.

Bradbury, J.W. and Vehrencamp, S.L. (2011) *Principles of Animal Communication*. 2nd ed. Sunderland, MA: Sinauer Associates.

Brady, F. (2004) ‘Contextual interference: A meta-analytic Study’, *Perceptual and Motor Skills*, 99, pp. 116–126. Available at: https://doi.org/10.2466/pms.99.1.116-126.

Bright, C. (2007) *The Last Whale* [Musical Score]. Australian Music Centre. Available at: https://www.australianmusiccentre.com.au/workversion/bright-colin-last-whale/16528.

van den Broek, E. and Todd, P.M. (2003) ‘Piep Piep Piep – Ich Hab’ Dich Lieb: Rhythm as an indicator of mate quality’, in W. Banzhaf, J. Ziegler, T. Christaller, P. Dittrich, and J.T. Kim (eds) *Advances in Artificial Life*. Berlin, Heidelberg: Springer (Lecture Notes in Computer Science), pp. 425–433. Available at: https://doi.org/10.1007/978-3-540-39432-7_45.

van den Broek, E.M.F. and Todd, P.M. (2009) ‘Evolution of rhythm as an indicator of mate quality’, *Musicae Scientiae*, 13(2_suppl), pp. 369–386. Available at: https://doi.org/10.1177/1029864909013002161.

Brown, V.A. (2021) ‘An introduction to linear mixed-effects modeling in R’, *Advances in Methods and Practices in Psychological Science*, 4(1), p. 2515245920960351. Available at: https://doi.org/10.1177/2515245920960351.

Brumm, H. and Slater, P.J.B. (2006) ‘Ambient noise, motor fatigue, and serial redundancy in chaffinch song’, *Behavioral Ecology and Sociobiology*, 60(4), pp. 475–481. Available at: https://doi.org/10.1007/s00265-006-0188-y.

Budka, M. and Osiejuk, T.S. (2014) ‘Individually specific call feature Is not used to neighbour-stranger discrimination: The corncrake case’, *PLOS ONE*, 9(8), p. e104031. Available at: https://doi.org/10.1371/journal.pone.0104031.

Burchardt, L.S. and Knörnschild, M. (2020) ‘Comparison of methods for rhythm analysis of complex animals’ acoustic signals’, *PLOS Computational Biology*, 16(4), p. e1007755. Available at: https://doi.org/10.1371/journal.pcbi.1007755.

Burghardt, G.M. (1991) ‘Cognitive Ethology and Critical Anthropomorphism: A Snake with Two Heads and Hognose Snakes That Play Dead’, in C.A. Ristau (ed.) *Cognitive Ethology: The Minds of Other Animals*. Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 53–90.

Burghardt, G.M. (1997) ‘Amending Tinbergen: A Fifth Aim for Ethology’, in R.W. Mitchell, N.S. Thompson, and H.L. Miles (eds) *Anthropomorphism, Anecdotes, and Animals*. Albany, N.Y.: State University of New York Press, pp. 254–276.

Burnett, D.G. (2012) *The Sounding of the Whale: Science and Cetaceans in the Twentieth Century*. Chicago: The University of Chicago Press.

Bush, K. (1978) *‘Moving’, The Kick Inside* [Vinyl]. EMI Records.

Bush, K. (no date a) *Moving | Kate Bush, Fish People: The Official Site of Kate Bush*. Available at: http://www.katebush.com/discography/moving (Accessed: 11 September 2021).

Bush, K. (no date b) *The Kick Inside | Kate Bush, Fish People: The Official Site of Kate Bush*. Available at: http://www.katebush.com/discography/kick-inside (Accessed: 27 October 2021).

Byers, B.E. (2007) ‘Extrapair paternity in chestnut-sided warblers is correlated with consistent vocal performance’, *Behavioral Ecology*, 18(1), pp. 130–136. Available at: https://doi.org/10.1093/beheco/arl058.

Calarco, M. (2008) *Zoographies: The Question of the Animal from Heidegger to Derrida*. New York: Columbia University Press.

Calarco, M. (2015) *Thinking Through Animals: Identity, Difference, Indistinction*. Stanford, California: Stanford University Press.

Calarco, M.R. (2020) *Animal Studies: The Key Concepts*. London & New York: Routledge.

Catchpole, C.K. and Slater, P.J.B. (2008) *Bird Song: Biological Themes and Variations*. Second Edition. Cambridge England ; New York: Cambridge University Press.

ten Cate, C., Slabbekoorn, H. and Ballintijn, M.R. (2002) ‘Birdsong and male—male competition: Causes and consequences of vocal variability in the collared dove (Streptopelia decaocto)’, in *Advances in the Study of Behavior*. Academic Press, pp. 31–75. Available at: https://doi.org/10.1016/S0065-3454(02)80005-5.

ten Cate, C. and Spierings, M. (2019) ‘Rules, rhythm and grouping: auditory pattern perception by birds’, *Animal Behaviour*, 151, pp. 249–257. Available at: https://doi.org/10.1016/j.anbehav.2018.11.010.

ten Cate, C., Spierings, M., Hubert, J. and Honing, H. (2016) ‘Can birds perceive rhythmic patterns? A review and experiments on a songbird and a parrot species’, *Frontiers in Psychology*, 7(MAY). Available at: https://doi.org/10.3389/fpsyg.2016.00730.

Cazau, D., Adam, O., Laitman, J.T. and Reidenberg, J.S. (2013) ‘Understanding the intentional acoustic behavior of humpback whales: A production-based approach’, *The Journal of the Acoustical Society of America*, 134(3), pp. 2268–2273. Available at: https://doi.org/10.1121/1.4816403.

Cerchio, S. (1993) *Geographic Variation and Cultural Evolution in Songs of Humpback Whales (Megaptera novaeangliae) in the Eastern North Pacific*. Masters. San Jose State University. Available at: http://scholarworks.sjsu.edu/etd_theses/614.

Cerchio, S., Jacobsen, J.K. and Norris, T.F. (2001) ‘Temporal and geographical variation in songs of humpback whales, Megaptera novaeangliae: synchronous change in Hawaiian and Mexican breeding assemblages’, *Animal Behaviour*, 62(2), pp. 313–329. Available at: https://doi.org/10.1006/anbe.2001.1747.

Chagnon, C.W., Durante, F., Gills, B.K., Hagolani-Albov, S.E., Hokkanen, S., Kangasluoma, S.M.J., Konttinen, H., Kröger, M., LaFleur, W., Ollinaho, O. and Vuola, M.P.S. (2022) ‘From extractivism to global extractivism: the evolution of an organizing concept’, *The Journal of Peasant Studies*, 49(4), pp. 760–792. Available at: https://doi.org/10.1080/03066150.2022.2069015.

Chami, R., Cosimano, T.F., Fullenkamp, C. and Oztosun, S. (2019) ‘Nature’s solution to climate change: A strategy to protect whales can limit greenhouse gases and global warming', *Finance & Development*, 0056(004). Available at: https://doi.org/10.5089/9781498316880.022.A011.

Cheung, V.K.M., Harrison, P.M.C., Meyer, L., Pearce, M.T., Haynes, J.-D. and Koelsch, S. (2019) 'Uncertainty and surprise jointly predict musical pleasure and amygdala, hippocampus, and auditory cortex activity', *Current Biology*, 29(23), pp. 4084-4092.e4. Available at: https://doi.org/10.1016/j.cub.2019.09.067.

Cholewiak, D.M., Cerchio, S., Jacobsen, J.K., Urbán-R, J. and Clark, C.W. (2018) 'Songbird dynamics under the sea: Acoustic interactions between humpback whales suggest song mediates male interactions', *Open Science*, 5(2), p. 171298. Available at: https://doi.org/10.1098/rsos.171298.

Cholewiak, D.M., Sousa-Lima, R.S. and Cerchio, S. (2013) 'Humpback whale song hierarchical structure: Historical context and discussion of current classification issues', *Marine Mammal Science*, 29(3), pp. E312–E332. Available at: https://doi.org/10.1111/mms.12005.

Chu, K.C. (1988) 'Dive times and ventilation patterns of singing humpback whales (Megaptera novaeangliae)', *Canadian Journal of Zoology*, 66(6), pp. 1322–1327. Available at: https://doi.org/10.1139/z88-194.

Clark, C.W., Ellison, W.T., Southall, B.L., Hatch, L.T., Van Parijs, S.M., Frankel, A.S. and Ponirakis, D. (2009) 'Acoustic masking in marine ecosystems: intuitions, analysis, and implication', *Marine Ecology Progress Series*, 395, pp. 201–222. Available at: https://doi.org/10.3354/meps08402.

Clark, C.W. and Gagnon, G.J. (2022) 'Baleen Whale Acoustic Ethology', in C.W. Clark and E.C. Garland (eds) *Ethology and Behavioral Ecology of Mysticetes*. Cham: Springer International Publishing (Ethology and Behavioral Ecology of Marine Mammals), pp. 11–43. Available at: https://doi.org/10.1007/978-3-030-98449-6_11.

Clark, C.W. and Garland, E.C. (eds) (2022) *Ethology and Behavioral Ecology of Mysticetes*. Cham: Springer International Publishing (Ethology and Behavioral Ecology of Marine Mammals). Available at: https://doi.org/10.1007/978-3-030-98449-6_11.

Clark, E.H. (2021) ‘Introduction: Audibilities of colonialism and extractivism’, *The World of Music*, 10(2), pp. 5–20.

Clarke, E.F. (1999) ‘Rhythm and Timing in Music’, in D. Deutsch (ed.) *The Psychology of Music (Second Edition)*. San Diego: Academic Press (Cognition and Perception), pp. 473–500. Available at: https://doi.org/10.1016/B978-012213564-4/50014-7.

Clayton, M. (2000) *Time in Indian Music: Rhythm, Metre, and Form in North Indian Rag Performance*. Oxford University Press.

Clayton, M. (2016) ‘Aksak patterns and entrained interaction in Transylvanian village music’, *Empirical Musicology Review*, 10(4), pp. 292–301. Available at: https://doi.org/10.18061/emr.v10i4.4886.

Clayton, M. (2020) ‘Theory and practice of long-form non-isochronous meters: The case of the North Indian rupak tal’, *Music Theory Online*, 26(1), p. 26.1.2.

Clayton, M., Sager, R. and Will, U. (2005) ‘In time with the music: The concept of entrainment and its significance for ethnomusicology’, *European Meetings in Ethnomusicology*, 11, pp. 1–82.

Coleman, E.B., Coombe, R.J. and MacArailt, F. (2009) ‘A Broken Record: Subjecting “Music” to Cultural Rights’, in J.O. Young and C.G. Brunk (eds) *The Ethics of Cultural Appropriation*. Chichester, UK ; Malden, MA: John Wiley & Sons, Ltd, pp. 173–210. Available at: https://doi.org/10.1002/9781444311099.ch8.

Coleman, K. (2021) ‘Free jazz and the “New Thing”: Aesthetics, identity, and texture, 1960–1966’, *Journal of Musicology*, 38(3), pp. 261–295. Available at: https://doi.org/10.1525/jm.2021.38.3.261.

Collins, J. (1970a) *‘Farewell to Tarwathie’, Whales and Nightingales* [Vinyl]. New York: Elektra Records.

Collins, J. (1970b) *Whales and Nightingales* [Vinyl]. New York: Elektra Records.

Cook, P., Rouse, A., Wilson, M. and Reichmuth, C. (2013) ‘A California sea lion (Zalophus californianus) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic', *Journal of Comparative Psychology (Washington, D.C.: 1983),* 127(4), pp. 412–427. Available at: https://doi.org/10.1037/a0032345.

Cook, R.C. (2022) 'Heterospecific song and ecological politics: Whale song and Crumb's Vox Balaenae', *Contemporary Music Review,* 41(1), pp. 4–29. Available at: https://doi.org/10.1080/07494467.2022.2033568.

Cooper, G. and Meyer, L. (1960) *The Rhythmic Structure of Music*. Chicago, Illinois: University of Chicago Press.

Corcoran, C. and Frieler, K. (2021) 'Playing it straight: Analyzing jazz soloists' swing eighth-note distributions with the Weimar Jazz Database', *Music Perception*, 38(4), pp. 372–385. Available at: https://doi.org/10.1525/mp.2021.38.4.372.

Crist, E. (2013) 'On the poverty of our nomenclature', *Environmental Humanities*, 3(1), pp. 129–147. Available at: https://doi.org/10.1215/22011919-3611266.

Croll, D.A., Clark, C.W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J. and Urban, J. (2002) 'Only male fin whales sing loud songs', *Nature*, 417(6891), pp. 809–809. Available at: https://doi.org/10.1038/417809a.

D'Agostino, M.E. (2020) 'Reclaiming and preserving traditional music: Aesthetics, ethics and technology', *Organised Sound*, 25(1), pp. 106–115. Available at: https://doi.org/10.1017/S1355771819000505.

Damien, J., Adam, O., Cazau, D., White, P., Laitman, J.T. and Reidenberg, J.S. (2019) 'Anatomy and functional morphology of the mysticete rorqual whale larynx: Phonation positions of the U-fold', *The Anatomical Record*, 302(5), pp. 703–717. Available at: https://doi.org/10.1002/ar.24034.

Daniele, J.R. and Patel, A.D. (2013) 'An empirical study of historical patterns in musical rhythm analysis of German and Italian classical music using the nPVI equation', *Music Perception*, 31(1), pp. 10–18. Available at: https://doi.org/10.1525/mp.2013.31.1.10.

Darling, J.D., Jones, M.E. and Nicklin, C.P. (2006) 'Humpback whale songs: Do they organize males during the breeding season?', *Behaviour*, 143(9), pp. 1051–1101.

Datseris, G., Ziereis, A., Albrecht, T., Hagmayer, Y., Priesemann, V. and Geisel, T. (2019) ‘Microtiming deviations and swing feel in jazz’, *Scientific Reports*, 9(1), p. 19824. Available at: https://doi.org/10.1038/s41598-019-55981-3.

Davies, M., Madison, G., Silva, P. and Gouyon, F. (2013) ‘The effect of microtiming deviations on the perception of groove in short rhythms’, *Music Perception*, 30(5), pp. 497–510. Available at: https://doi.org/10.1525/mp.2013.30.5.497.

De Gregorio, C., Valente, D., Raimondi, T., Torti, V., Miaretsoa, L., Friard, O., Giacoma, C., Ravignani, A. and Gamba, M. (2021) ‘Categorical rhythms in a singing primate’, *Current Biology*, 31(20), pp. R1379–R1380. Available at: https://doi.org/10.1016/j.cub.2021.09.032.

De Jaegher, H. and Di Paolo, E. (2007) ‘Participatory sense-making: An enactive approach to social cognition’, *Phenomenology and the Cognitive Sciences*, 6(4), pp. 485–507. Available at: https://doi.org/10.1007/s11097-007-9076-9.

Dean, L.G., Vale, G.L., Laland, K.N., Flynn, E. and Kendal, R.L. (2014) ‘Human cumulative culture: A comparative perspective’, *Biological Reviews*, 89(2), pp. 284–301. Available at: https://doi.org/10.1111/brv.12053.

Demartsev, V., Haddas-Sasson, M., Ilany, A., Koren, L. and Geffen, E. (2022) ‘Male rock hyraxes that maintain an isochronous song rhythm achieve higher reproductive success’, *Journal of Animal Ecology* [Preprint]. Available at: https://doi.org/10.1111/1365-2656.13801.

Demartsev, V., Ilany, A., Kershenbaum, A., Geva, Y., Margalit, O., Schnitzer, I., Barocas, A., Bar-Ziv, E., Koren, L. and Geffen, E. (2017) ‘The progression pattern of male hyrax songs and the role of climactic ending’, *Scientific Reports*, 7(1), p. 2794. Available at: https://doi.org/10.1038/s41598-017-03035-x.

Demos, A.P., Lisboa, T., Begosh, K.T., Logan, T. and Chaffin, R. (2020) ‘A longitudinal study of the development of expressive timing’, *Psychology of Music*, 48(1), pp. 50–66. Available at: https://doi.org/10.1177/0305735618783563.

Despret, V. (2016) *What Would Animals Say If We Asked the Right Questions?* Translated by B. Buchanan. Minneapolis: University Of Minnesota Press.

Deutsch, D., Henthorn, T. and Lapidis, R. (2011) ‘Illusory transformation from speech to song’, *The Journal of the Acoustical Society of America,* 129(4), pp. 2245–2252. Available at: https://doi.org/10.1121/1.3562174.

Díaz-Báñez, J.-M., Farigu, G., Gomez, F., Rappaport, D. and Toussaint, G.T. (2004) ‘El Compas Flamenco: A Phylogenetic Analysis’, in *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*. *BRIDGES: Mathematical Connections in Art, Music and Science,* Southwestern College, Winfield, Kansas, pp. 61–70.

Diettrich, B., Freeman Moulin, J. and Webb, M. (2011) *Music in Pacific Island Cultures: Experiencing Music, Expressing Culture*. New York ; Oxford: Oxford University Press.

Dodson, A. (2020) ‘Visualizing the Rhythms of Performance’, in R. Hartenberger and R. McClelland (eds) *The Cambridge Companion to Rhythm*. Cambridge: Cambridge University Press (Cambridge Companions to Music), pp. 41–60. Available at: https://doi.org/10.1017/9781108631730.006.

Doolittle, E.L. (1999) *night black bird song* [Musical Score]. Canadian Music Centre.

Doolittle, E.L. (2007a) *Other Species’ Counterpoint: An Investigation of the Relationship Between Human Music and Animal Songs*. PhD Dissertation. Princeton University.

Doolittle, E.L. (2007b) *Social Sounds from Whales at Night* [Musical Score]. Chipping Norton UK: Composers Edition.

Doolittle, E.L. (2008) ‘Crickets in the concert hall: A history of animals in Western music’, *Trans. Revista Transcultural de Música,* 12.

Doolittle, E.L. (2009) *Utah, 1996* [Musical Score]. Chipping Norton UK: Composers Edition.

Doolittle, E.L. (2018a) *Reedbird* [Musical Score]. Chipping Norton UK: Composers Edition.

Doolittle, E.L. (2018b) *Woodwings* [Musical Score]. Chipping Norton UK: Composers Edition.

Doolittle, E.L. (2020a) *Gardenscape* [Musical Score]. Chipping Norton UK: Composers Edition.

Doolittle, E.L. (2020b) '"Hearken to the hermit-thrush": A case study in interdisciplinary listening', *Frontiers in Psychology*, 11. Available at: https://doi.org/10.3389/fpsyg.2020.613510.

Doolittle, E.L. (2020c) '"Scarce inferior to the nightingale": Hermit thrush song as a symbol of cultural identity in anglophone North America', *Ecomusicology Review*, 8. Available at: https://ecomusicology.info/doolittle20/.

Doolittle, E.L. (2022) *Luscinia* [Musical Score]. Available at: https://emilydoolittle.com/music/ (Accessed: 5 May 2022).

Doolittle, E.L. and Brumm, H. (2012) 'O Canto do Uirapuru: Consonant pitches and patterns in the song of the musician wren', *Journal of Interdisciplinary Music Studies*, 6(1), pp. 55–85.

Doolittle, E.L. and Gingras, B. (2015) 'Zoomusicology', *Current Biology*, 25(19), pp. R819–R820. Available at: https://doi.org/10.1016/j.cub.2015.06.039.

van Dooren, T., Kirksey, E. and Münster, U. (2016) 'Multispecies studies: Cultivating arts of attentiveness', *Environmental Humanities*, 8(1), pp. 1–23. Available at: https://doi.org/10.1215/22011919-3527695.

Dudzinski, K.M. and Gregg, J.D. (2018) 'Communication', in B. Würsig, J.G.M. Thewissen, and K.M. Kovacs (eds) *Encyclopedia of Marine Mammals (Third Edition)*. Academic Press, pp. 210–215. Available at: https://doi.org/10.1016/B978-0-12-804327-1.00062-5.

Dulau, V., Pinet, P., Geyer, Y., Fayan, J., Mongin, P., Cottarel, G., Zerbini, A. and Cerchio, S. (2017) 'Continuous movement behavior of humpback whales during the breeding season in the southwest Indian Ocean: on the road again!', *Movement Ecology*, 5, p. 11. Available at: https://doi.org/10.1186/s40462-017-0101-5.

Emmerson, S. (1986) ‘The Relation of Language to Materials’, in S. Emmerson (ed.) *The Language of Electroacoustic Music*. London: Palgrave Macmillan UK, pp. 17–39. Available at: https://doi.org/10.1007/978-1-349-18492-7_3.

Evans, R. (2014) ‘Cheltenham premieres review – planets, poems and pioneers’, *The Guardian*, 13 July. Available at: https://www.theguardian.com/music/2014/jul/13/cheltenham-premieres-review-planets-poems-pioneers (Accessed: 1 July 2023).

Fay, R. (1988) *Hearing in Vertebrates*. Winnetka, Illinois: Hill-Fay Associates. Available at: http://archive.org/details/hearinginvertebr0000rich_n2s4 (Accessed: 26 June 2023).

Feld, S. (1996) ‘Pygmy POP. A genealogy of schizophonic mimesis’, *Yearbook for Traditional Music*, 28, pp. 1–35. Available at: https://doi.org/10.2307/767805.

Feldman, A.C. (2021) ‘Minor listening, major influence: Revisiting Songs of the Humpback Whale’, *e-flux*. Available at: https://www.e-flux.com/journal/118/394434/minor-listening-major-influence-revisiting-songs-of-the-humpback-whale/ (Accessed: 7 July 2022).

Ferrari, L. (1970) *Presque rien No. 1: Le lever du jour au bord de la mer* [Sound Recording].

Ferraro, D.M., Miller, Z.D., Ferguson, L.A., Taff, B.D., Barber, J.R., Newman, P. and Francis, C.D. (2020) ‘The phantom chorus: Birdsong boosts human well-being in protected areas’, *Proceedings of the Royal Society B: Biological Sciences*, 287(1941), p. 20201811. Available at: https://doi.org/10.1098/rspb.2020.1811.

Fitch, W.T. (2006) ‘The biology and evolution of music: A comparative perspective’, *Cognition*, 100(1), pp. 173–215. Available at: https://doi.org/10.1016/j.cognition.2005.11.009.

Fitch, W.T. (2018) ‘Four Principles of Biomusicology’, in H. Honing (ed.) *The Origins of Musicality*. Cambridge, MA: MIT Press, pp. 23–48.

Fournet, M.E.H., Gabriele, C.M., Culp, D.C., Sharpe, F., Mellinger, D.K. and Klinck, H. (2018) ‘Some things never change: multi-decadal stability in humpback whale calling repertoire on Southeast Alaskan foraging grounds', *Scientific Reports*, 8(1), p. 13186. Available at: https://doi.org/10.1038/s41598-018-31527-x.

Frane, A.V. (2017) 'Swing rhythm in classic drum breaks from hip-hop's breakbeat canon', *Music Perception*, 34(3), pp. 291–302. Available at: https://doi.org/10.1525/mp.2017.34.3.291.

Frankel, A.S., Clark, C.W., Herman, L. and Gabriele, C. (1995) 'Spatial distribution, habitat utilization, and social interactions of humpback whales, Megaptera novaeangliae, off Hawai'i, determined using acoustic and visual techniques', *Canadian Journal of Zoology*, 73, pp. 1134–1146. Available at: https://doi.org/10.1139/z95-135.

Frantzis, A. and Alexiadou, P. (2008) 'Male sperm whale (Physeter macrocephalus) coda production and coda-type usage depend on the presence of conspecifics and the behavioural context', *Canadian Journal of Zoology*, 86, pp. 62–75. Available at: https://doi.org/10.1139/Z07-114.

Frühauf, J., Kopiez, R. and Platz, F. (2013) 'Music on the timing grid: The influence of microtiming on the perceived groove quality of a simple drum pattern performance', *Musicae Scientiae*, 17(2), pp. 246–260. Available at: https://doi.org/10.1177/1029864913486793.

Frumhoff, P.C. (1983) 'Aberrant songs of humpback whales (Megaptera novaeangliae): Clues to the structure of humpback song', in R.S. Payne (ed.) *Communication and Behavior of Whales*. Boulder, CO: Westview Press (AAAS Selected Symposium, 76), pp. 81–127.

Gales, N., Bannister, J., Findlay, K., Zerbini, A. and Donovan, G. (eds) (2011) 'Report of the workshop on the comprehensive assessment of southern hemisphere humpback whales', *Journal of Cetacean Research and Management*, Special Issue 3, pp. 1–50.

Garcia, M., Charrier, I., Rendall, D. and Iwaniuk, A.N. (2012) 'Temporal and spectral analyses reveal individual variation in a non-vocal acoustic display: The drumming display of the ruffed grouse (Bonasa umbellus, L.)', *Ethology*, 118(3), pp. 292–301. Available at: https://doi.org/10.1111/j.1439-0310.2011.02011.x.

Garland, E.C., Garrigue, C. and Noad, M.J. (2021) ‘When does cultural evolution become cumulative culture? A case study of humpback whale song’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1843), p. 20200313. Available at: https://doi.org/10.1098/rstb.2020.0313.

Garland, E.C., Gedamke, J., Rekdahl, M.L., Noad, M.J., Garrigue, C. and Gales, N. (2013) ‘Humpback whale song on the Southern Ocean feeding grounds: Implications for cultural transmission’, *PLOS ONE*, 8(11), p. e79422. Available at: https://doi.org/10.1371/journal.pone.0079422.

Garland, E.C., Goldizen, A.W., Lilley, M.S., Rekdahl, M.L., Garrigue, C., Constantine, R., Hauser, N.D., Poole, M.M., Robbins, J. and Noad, M.J. (2015) ‘Population structure of humpback whales in the western and central South Pacific Ocean as determined by vocal exchange among populations’, *Conservation Biology*, 29(4), pp. 1198–1207. Available at: https://doi.org/10.1111/cobi.12492.

Garland, E.C., Goldizen, A.W., Rekdahl, M.L., Constantine, R., Garrigue, C., Hauser, N.D., Poole, M.M., Robbins, J. and Noad, M.J. (2011) ‘Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale’, *Current Biology*, 21(8), pp. 687–691. Available at: https://doi.org/10.1016/j.cub.2011.03.019.

Garland, E.C., Lilley, M.S., Goldizen, A.W., Rekdahl, M.L., Garrigue, C. and Noad, M.J. (2012) ‘Improved versions of the Levenshtein distance method for comparing sequence information in animals’ vocalisations: Tests using humpback whale song’, *Behaviour*, 149(13–14), pp. 1413–1441. Available at: https://doi.org/10.1163/1568539X-00003032.

Garland, E.C. and McGregor, P.K. (2020) ‘Cultural transmission, evolution, and revolution in vocal displays: Insights from bird and whale song’, *Frontiers in Psychology*, 11. Available at: https://doi.org/10.3389/fpsyg.2020.544929.

Garland, E.C., Noad, M.J., Goldizen, A.W., Lilley, M.S., Rekdahl, M.L., Garrigue, C., Constantine, R., Daeschler Hauser, N., Poole, M.M. and Robbins, J. (2013) ‘Quantifying humpback whale song sequences to understand the dynamics of song exchange at the ocean basin scale’, *The Journal of the Acoustical Society of America*, 133(1), pp. 560–569. Available at: https://doi.org/10.1121/1.4770232.

Garland, E.C., Rendell, L., Lamoni, L., Poole, M.M. and Noad, M.J. (2017) ‘Song hybridization events during revolutionary song change provide insights into cultural transmission in humpback whales’, *Proceedings of the National Academy of Sciences*, 114(30), pp. 7822–7829. Available at: https://doi.org/10.1073/pnas.1621072114.

Gedamke, J., Costa, D.P. and Dunstan, A. (2001) ‘Localization and visual verification of a complex minke whale vocalization’, *The Journal of the Acoustical Society of America*, 109(6), pp. 3038–3047. Available at: https://doi.org/10.1121/1.1371763.

Gémard, C., Aubin, T., Reboud, E.L. and Bonadonna, F. (2021) ‘Call rate, fundamental frequency, and syntax determine male-call attractiveness in blue petrels Halobaena caerulea’, *Behavioral Ecology and Sociobiology*, 75(3), p. 55. Available at: https://doi.org/10.1007/s00265-021-02989-3.

Gero, S., Gordon, J. and Whitehead, H. (2015) ‘Individualized social preferences and long-term social fidelity between social units of sperm whales’, *Animal Behaviour*, 102, pp. 15–23. Available at: https://doi.org/10.1016/j.anbehav.2015.01.008.

Gero, S. and Rendell, L. (2015) ‘Oceanic Societies: Studying Cetaceans with a Social Networks Approach’, in J. Krause, R. James, D.W. Franks, and D.P. Croft (eds) *Animal Social Networks*. Oxford University Press, pp. 139–149.

Gero, S., Whitehead, H. and Rendell, L. (2016) ‘Individual, unit and vocal clan level identity cues in sperm whale codas’, *Royal Society Open Science*, 3(1), p. 150372. Available at: https://doi.org/10.1098/rsos.150372.

Girola, E., Noad, M.J., Dunlop, R.A. and Cato, D.H. (2019) ‘Source levels of humpback whales decrease with frequency suggesting an air-filled resonator is used in sound production’, *The Journal of the Acoustical Society of America*, 145(2), pp. 869–880. Available at: https://doi.org/10.1121/1.5090492.

Glasgow Improvisers Orchestra and Maggie Nichols (2019) *Energy Being* [CD]. Rayleigh, UK: FMR Records.

Glaze, C.M. and Troyer, T.W. (2006) ‘Temporal structure in zebra finch song: Implications for motor coding’, *The Journal of Neuroscience: The Official Journal of the Society* *for Neuroscience*, 26(3), pp. 991–1005. Available at: https://doi.org/10.1523/JNEUROSCI.3387-05.2006.

Glaze, C.M. and Troyer, T.W. (2012) ‘A generative model for measuring latent timing structure in motor sequences’, *PLOS ONE*, 7(7), p. e37616. Available at: https://doi.org/10.1371/journal.pone.0037616.

Glaze, C.M. and Troyer, T.W. (2013) ‘Development of temporal structure in zebra finch song’, *Journal of Neurophysiology*, 109(4), pp. 1025–1035. Available at: https://doi.org/10.1152/jn.00578.2012.

Glerean, E., Pan, R.K., Salmi, J., *et al.* (2016) ‘Reorganization of functionally connected brain subnetworks in high-functioning autism’, *Human Brain Mapping*, 37(3), pp. 1066–1079. Available at: https://doi.org/10.1002/hbm.23084.

Glockner, D.A. (1983) ‘Determining the sex of humpback whales (Megaptera novaeangliae) in their natural environment.’, in R.S. Payne (ed.) *Communication and Behavior of Whales*. Boulder, CO: Westview Press (AAAS Selected Symposium, 76), pp. 447–464.

Goldberg, D. (2015) ‘Timing variations in two Balkan percussion performances’, *Empirical Musicology Review*, 10(4), pp. 305–328. Available at: https://doi.org/10.18061/emr.v10i4.4884.

Gonneville, M. (1998) *Le cheminement de la baleine* [Musical Score]. Canadian Music Centre.

Grabe, E. and Low, E.L. (2002) ‘Durational Variability in Speech and the Rhythm Class Hypothesis’, in C. Gussenhoven and N. Warner (eds) *Papers in Laboratory Phonology VII*. De Gruyter Mouton, pp. 515–546.

Griffin, D.R. (1998) ‘From Cognition to Consciousness’, *Animal Cognition*, 1, pp. 3–16.

Grover Friedlander, M. (2020) ‘Whale wonder’, *Journal of Interdisciplinary Voice Studies*, 5(1), pp. 29–42. Available at: https://doi.org/10.1386/jivs_00014_1.

Gruen, L. (2021) ‘The Moral Status of Animals’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy (Summer 2021 Edition)*. Summer 2021. Metaphysics

Research Lab, Stanford University. Available at: https://plato.stanford.edu/archives/sum2021/entries/moral-animal/.

Guastavino, C., Gómez, F., Toussaint, G., Marandola, F. and Gómez, E. (2009) ‘Measuring similarity between Flamenco rhythmic patterns’, *Journal of New Music Research,* 38(2), pp. 129–138.

Guillot, G. and Rousset, F. (2013) ‘Dismantling the Mantel tests’, *Methods in Ecology and Evolution,* 4(4), pp. 336–344. Available at: https://doi.org/10.1111/2041-210x.12018.

Guinee, L.N., Chu, K.C. and Dorsey, E.M. (1983) ‘Changes over time in the songs of known individual humpback whales (Megaptera Novaeangliae)’, in R.S. Payne (ed.) *Communication and Behavior of Whales*. Boulder, CO: Westview Press (AAAS Selected Symposium, 76), pp. 59–80.

Guinee, L.N. and Payne, K.B. (1988) ‘Rhyme-like repetitions in songs of humpback whales’, *Ethology,* 79(4), pp. 295–306. Available at: https://doi.org/10.1111/j.1439-0310.1988.tb00718.x.

Guzman, H.M. and Felix, F. (2017) ‘Movements and Habitat Use by Southeast Pacific Humpback Whales (Megaptera novaeangliae) Satellite Tracked at Two Breeding Sites’, *Aquatic Mammals,* 43(2), pp. 139–155. Available at: https://doi.org/10.1578/AM.43.2.2017.139.

Guzman, H.M., Gomez, C.G., Guevara, C.A. and Kleivane, L. (2013) ‘Potential vessel collisions with Southern Hemisphere humpback whales wintering off Pacific Panama’, *Marine Mammal Science,* 29(4), pp. 629–642. Available at: https://doi.org/10.1111/j.1748-7692.2012.00605.x.

Hafner, G.W., Hamilton, C.L., Steiner, W.W., Thompson, T.J. and Winn, H.E. (1979) ‘Signature information in the song of the humpback whale’, *The Journal of the Acoustical Society of America,* 66(1), pp. 1–6. Available at: https://doi.org/10.1121/1.383072.

Hair, G. (2013) *‘Songs from the Turkish’, Music from 3 Continents* [CD]. Ravello Records (North Hampton, NH).

Handel, S., Todd, S.K. and Zoidis, A.M. (2009) ‘Rhythmic structure in humpback whale (Megaptera novaeangliae) songs: Preliminary implications for song production and perception’, *The Journal of the Acoustical Society of America*, 125(6), pp. EL225-230. Available at: https://doi.org/10.1121/1.3124712.

Handel, S., Todd, S.K. and Zoidis, A.M. (2012) ‘Hierarchical and rhythmic organization in the songs of humpback whales (Megaptera novaeangliae)’, *Bioacoustics*, 21(2), pp. 141–156. Available at: https://doi.org/10.1080/09524622.2012.668324.

Hansen, N.C., Sadakata, M. and Pearce, M. (2016) ‘Nonlinear changes in the rhythm of European art music: Quantitative support for historical musicology’, *Music Perception*, 33(4), pp. 414–431. Available at: https://doi.org/10.1525/mp.2016.33.4.414.

Haraway, D.J. (2008) *When Species Meet*. Minneapolis: University of Minnesota Press.

Haraway, D.J. (2016) *Staying with the Trouble: Making Kin in the Chthulucene*. Durham: Duke University Press.

Harding, B., Tremblay, C. and Cousineau, D. (2014) ‘Standard errors: A review and evaluation of standard error estimators using Monte Carlo simulations’, *The Quantitative Methods for Psychology*, 10(2), pp. 107–123. Available at: https://doi.org/10.20982/tqmp.10.2.p107.

Hasty, C.F. (1997) *Meter as Rhythm*. New York ; Oxford: Oxford University Press.

Hattori, Y. and Tomonaga, M. (2020) ‘Rhythmic swaying induced by sound in chimpanzees (Pan troglodytes)’, *Proceedings of the National Academy of Sciences*, 117(2), pp. 936–942. Available at: https://doi.org/10.1073/pnas.1910318116.

Hauser, M.D., Chomsky, N. and Fitch, W.T. (2002) ‘The faculty of language: What is it, who has it, and how did it evolve?’, *Science*, 298(5598), pp. 1569–1579. Available at: https://doi.org/10.1126/science.298.5598.1569.

Helble, T.A., Guazzo, R.A., Alongi, G.C., Martin, C.R., Martin, S.W. and Henderson, E.E. (2020) ‘Fin whale song patterns shift over time in the central North Pacific’, *Frontiers in Marine Science*, 7. Available at: https://doi.org/10.3389/fmars.2020.587110.

Hellmer, K. and Madison, G. (2015) 'Quantifying microtiming patterning and variability in drum kit recordings: A method and some data', *Music Perception*, 33(2), pp. 147–162. Available at: https://doi.org/10.1525/MP.2015.33.2.147.

Helweg, D.A., Cato, D.H., Jenkins, P.F., Garrigue, C. and McCauley, R.D. (1998) 'Geographic variation in South Pacific humpback whale songs', *Behaviour*, 135(1), pp. 1–27.

Helweg, D.A., Frankel, A.S., Mobley, J.R. and Herman, L.M. (1992) 'Humpback Whale Song: Our Current Understanding', in J.A. Thomas, R.A. Kastelein, and A.Ya. Supin (eds) *Marine Mammal Sensory Systems*. Boston, MA: Springer US, pp. 459–483. Available at: https://doi.org/10.1007/978-1-4615-3406-8_32.

Herman, L.M. (2002) 'Vocal, Social, and Self-Imitation by Bottlenosed Dolphins', in K. Dautenhahn and C.L. Nehaniv (eds) *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, pp. 63–108.

Herman, L.M. (2017) 'The multiple functions of male song within the humpback whale (Megaptera novaeangliae) mating system: Review, evaluation, and synthesis', *Biological Reviews of the Cambridge Philosophical Society*, 92(3), pp. 1795–1818. Available at: https://doi.org/10.1111/brv.12309.

Herman, L.M., Pack, A.A., Spitz, S.S., Herman, E.Y.K., Rose, K., Hakala, S. and Deakos, M.H. (2013) 'Humpback whale song: Who sings?', *Behavioral Ecology and Sociobiology*, 67(10), pp. 1653–1663.

Hersh, T.A., Gero, S., Rendell, L., *et al.* (2022) 'Evidence from sperm whale clans of symbolic marking in non-human cultures', *Proceedings of the National Academy of Sciences*, 119(37), p. e2201692119. Available at: https://doi.org/10.1073/pnas.2201692119.

Hoare, P. (2008) *Leviathan or, The Whale*. London: Fourth Estate.

Hoeschele, M. (2017) 'Animal pitch perception: Melodies and harmonies', *Comparative cognition & behavior reviews*, 12, pp. 5–18. Available at: https://doi.org/10.3819/CCBR.2017.120002.

Honing, H. (2012) ‘Without it no music: Beat induction as a fundamental musical trait’, *Annals of the New York Academy of Sciences*, 1252(1), pp. 85–91. Available at: https://doi.org/10.1111/j.1749-6632.2011.06402.x.

Horkheimer, M. and Adorno, T.W. (2002) *Dialectic of Enlightenment: Philosophical Fragments*. Edited by G. Schmid Noerr. Translated by E. Jephcott. Stanford, California: Stanford University Press.

Horner, A. (2014) ‘Watch - A Video Guide To Every Kate Bush Live Performance Ever’, *NME*, 21 March. Available at: https://www.nme.com/blogs/nme-blogs/watch-a-video-guide-to-every-kate-bush-live-performance-ever-20564 (Accessed: 27 October 2021).

Hovhaness, A. (1970) *And God Created Great Whales Op.229 No.1* [Musical Score]. Leipzig: Edition Peters.

Hu, J., Gao, J. and Principe, J.C. (2006) ‘Analysis of biomedical signals by the Lempel-Ziv complexity: The effect of finite data size’, *IEEE Transactions on Biomedical Engineering*, 53(12), pp. 2606–2609. Available at: https://doi.org/10.1109/TBME.2006.883825.

Huggan, G. (2004) ‘“Greening” postcolonialism: Ecocritical perspectives’, *MFS Modern Fiction Studies,* 50(3), pp. 701–733. Available at: https://doi.org/10.1353/mfs.2004.0067.

Humphreys, R. (2016) ‘Dignity and its violation examined within the context of animal ethics’, *Ethics and the Environment*, 21(2), pp. 143–162. Available at: https://doi.org/10.2979/ethicsenviro.21.2.06.

Huron, D. (2007) *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, Mass. ; London: MIT.

Hyland Bruno, J. and Tchernichovski, O. (2019) ‘Regularities in zebra finch song beyond the repeated motif’, *Behavioural Processes*, 163, pp. 53–59. Available at: https://doi.org/10.1016/j.beproc.2017.11.001.

Jacoby, N. and McDermott, J.H. (2017) ‘Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction’, *Current Biology*, 27(3), pp. 359–370. Available at: https://doi.org/10.1016/j.cub.2016.12.031.

Jairazbhoy, N.A. (1977) ‘The “objective” and subjective view in music transcription’, *Ethnomusicology*, 21(2), pp. 263–273. Available at: https://doi.org/10.2307/850947.

Jameson, F. (1990) *Late Marxism: Adorno, or, The Persistence of the Dialectic*. London: Verso.

Janik, V.M. (2014) ‘Cetacean vocal learning and communication’, *Current Opinion in Neurobiology*, 28, pp. 60–65. Available at: https://doi.org/10.1016/j.conb.2014.06.010.

Janik, V.M. and Slater, P.J.B. (1997) ‘Vocal Learning in Mammals’, in P.J.B. Slater, J.S. Rosenblatt, C.T. Snowdon, and M. Milinski (eds) *Advances in the Study of Behavior*. Academic Press, pp. 59–99. Available at: https://doi.org/10.1016/S0065-3454(08)60377-0.

Janney, E., Taylor, H., Scharff, C., Rothenberg, D., Parra, L.C. and Tchernichovski, O. (2016) ‘Temporal regularity increases with repertoire complexity in the Australian pied butcherbird’s song’, *Royal Society Open Science*, 3(9), pp. 1–15. Available at: https://doi.org/10.1098/rsos.160357.

Jennions, M.D., Bishop, P.J., Backwell, P.R.Y. and Passmore, N.I. (1995) ‘Call rate variability and female choice in the African frog, Hyperolius marmoratus’, *Behaviour*, 132(9–10), pp. 709–720. Available at: https://doi.org/10.1163/156853995X00108.

Johnson, H.D., Stafford, K.M., George, J.C., Ambrose, W.G. and Clark, C.W. (2015) ‘Song sharing and diversity in the Bering-Chukchi-Beaufort population of bowhead whales (Balaena mysticetus), spring 2011’, *Marine Mammal Science*, 31(3), pp. 902–922. Available at: https://doi.org/10.1111/mms.12196.

Jordania, J. (2015) *Choral Singing in Human Culture and Evolution*. Illustrated edition. LAP Lambert Academic Publishing.

K. Lisa Yang Centre for Conservation Bioacoustics at the Cornell Lab of Ornithology (2021) ‘Raven Pro: Interactive Sound Analysis Software (Version 1.6)’. Ithaca, New York. Available at: ravensoundsoftware.com.

Karlsson, F. (2012) ‘Critical anthropomorphism and animal ethics’, *Journal of Agricultural and Environmental Ethics*, 25(5), pp. 707–720. Available at: https://doi.org/10.1007/s10806-011-9349-8.

Kershenbaum, A. and Garland, E.C. (2015) ‘Quantifying similarity in animal vocal sequences: Which metric performs best?’, *Methods in Ecology and Evolution*, 6(12), pp. 1452–1461. Available at: https://doi.org/10.1111/2041-210X.12433.

Kivy, P. (1990) *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Ithaca, NY: Cornell University Press.

Kohn, E. (2013) *How Forests Think: Toward an Anthropology Beyond the Human*. Berkeley ; Los Angeles ; London: University of California Press.

Konishi, M. (2010) ‘From central pattern generator to sensory template in the evolution of birdsong’, *Brain and Language*, 115(1), pp. 18–20. Available at: https://doi.org/10.1016/j.bandl.2010.05.001.

de Kort, S.R., Eldermire, E.R.B., Valderrama, S., Botero, C.A. and Vehrencamp, S.L. (2009) ‘Trill consistency is an age-related assessment signal in banded wrens’, *Proceedings of the Royal Society B: Biological Sciences*, 276(1665), pp. 2315–2321. Available at: https://doi.org/10.1098/rspb.2009.0127.

Kotz, S.A., Ravignani, A. and Fitch, W.T. (2018) ‘The evolution of rhythm processing’, *Trends in Cognitive Sciences*, 22(10), pp. 896–910. Available at: https://doi.org/10.1016/j.tics.2018.08.002.

Kroher, N. and Díaz-Báñez, J.-M. (2019) ‘Modelling melodic variation and extracting melodic templates from flamenco singing performances’, *Journal of Mathematics and Music*, 13(2), pp. 150–170. Available at: https://doi.org/10.1080/17459737.2019.1610194.

Kvifte, T. (2007) ‘Categories and timing: On the perception of meter’, *Ethnomusicology,* 51(1), pp. 64–84.

Laland, K.N. and Hoppitt, W. (2003) ‘Do animals have culture?’, *Evolutionary Anthropology: Issues, News, and Reviews*, 12(3), pp. 150–159. Available at: https://doi.org/10.1002/evan.10111.

Lambrechts, M. and Dhondt, A.A. (1986) ‘Male quality, reproduction, and survival in the great tit (Parus major)’, *Behavioral Ecology and Sociobiology,* 19(1), pp. 57–63. Available at: https://doi.org/10.1007/BF00303843.

Lamoni, L., Garland, E.C., Allen, J.A., Coxon, J., Noad, M.J. and Rendell, L. (2023) ‘Variability in humpback whale songs reveals how individuals can be distinctive when sharing a complex vocal display’, *The Journal of the Acoustical Society of America,* 153(4), p. 2238. Available at: https://doi.org/10.1121/10.0017602.

Large, E.W. and Gray, P.M. (2015) ‘Spontaneous tempo and rhythmic entrainment in a bonobo (Pan paniscus)’, *Journal of Comparative Psychology,* 129(4), pp. 317–328. Available at: https://doi.org/10.1037/com0000011.

Lempel, A. and Ziv, J. (1976) ‘On the complexity of finite sequences’, *IEEE Transactions on Information Theory,* 22(1), pp. 75–81. Available at: https://doi.org/10.1109/TIT.1976.1055501.

Lengagne, T., Aubin, T., Lauga, J. and Jouventin, P. (1999) ‘How do king penguins (Aptenodytes patagonicus) apply the mathematical theory of information to communicate in windy conditions?’, *Proceedings of the Royal Society of London. Series B: Biological Sciences,* 266(1429), pp. 1623–1628. Available at: https://doi.org/10.1098/rspb.1999.0824.

Levenshtein, V.I. (1966) ‘Binary codes capable of correcting deletions, insertions, and reversals’, *Soviet Physics-Doklady,* 10(8), pp. 707–710.

Lewandowski, A. (2018) *Cetus: Life After Life* [Musical Score].

Liebermann, A. (2020) *A Bird in Quarantine* [Musical Score]. Available at: https://alexanderliebermann.com/.

Lien, M.E. and Pálsson, G. (2021) ‘Ethnography beyond the human: The “other-than-human” in ethnographic work’, *Ethnos,* 86(1), pp. 1–20. Available at: https://doi.org/10.1080/00141844.2019.1628796.

Lillis, A. and Mooney, T.A. (2022) ‘Sounds of a changing sea: Temperature drives acoustic output by dominant biological sound-producers in shallow water habitats’, *Frontiers in Marine Science*, 9.

Loch, G. (2018) ‘Between Szőke’s sound microscope and Messiaen’s organ: The cultural realities of blackcap song’, *Organised Sound,* 23(2), pp. 144–155. Available at: https://doi.org/10.1017/S135577181800002X.

London, J.M. (2012) *Hearing in Time: Psychological Aspects of Musical Meter*. 2nd Edition. New York ; Oxford: Oxford University Press.

London, J.M. and Jones, K. (2011) ‘Rhythmic refinements to the nPVI measure: A reanalysis of Patel & Daniele (2003a)’, *Music Perception,* 29(1), pp. 115–120. Available at: https://doi.org/10.1525/mp.2011.29.1.115.

Mâche, F.-B. (1992) *Music, Myth or Nature, or The Dolphins of Arion*. Translated by S. Delaney. Chur, Switzerland: Harwood Academic Publishers (Contemporary music studies, 6).

Macknight, F.L., Cato, D.H., Noad, M.J. and Grigg, G.C. (2001) ‘Qualitative and quantitative analyses of the song of the east Australian population of humpback whales’, *Memoirs of the Queensland Museum,* 47(2), pp. 525–537.

Mantel, N. (1967) ‘The detection of disease clustering and a generalized regression approach’, *Cancer Research,* 27(2 Part 1), pp. 209–220.

Marcoux, M., Whitehead, H. and Rendell, L. (2006) ‘Coda vocalizations recorded in breeding areas are almost entirely produced by mature female sperm whales (Physeter macrocephalus)’, *Canadian Journal of Zoology,* 84(4), pp. 609–614. Available at: https://doi.org/10.1139/z06-035.

Margulis, E.H. (2013) ‘Aesthetic responses to repetition in unfamiliar music’, *Empirical Studies of the Arts,* 31(1), pp. 45–57. Available at: https://doi.org/10.2190/EM.31.1.c.

Margulis, E.H. (2014) *On Repeat: How Music Plays the Mind*. New York, NY ; Oxford: Oxford University Press.

Marino, L. and Frohoff, T. (2011) 'Towards a new paradigm of non-captive research on cetacean cognition', *PLoS ONE*, 6(9), p. e24121. Available at: https://doi.org/10.1371/journal.pone.0024121.

Marler, P. (2004) 'Science and Birdsong: The Good Old Days', in P. Marler and H. Slabbekoorn (eds) *Nature's Music: The Science of Birdsong*. San Diego ; London: Elsevier Academic Press, pp. 1–38.

Martinelli, D. (1999) 'Il canto della megattera: Una musica bestiale', *Musica/Realtà*, 60, pp. 37–55.

Martinelli, D. (2009) *Of Birds, Whales, and Other Musicians: An Introduction to Zoomusicology*. Scranton ; London: University of Scranton Press.

Mathevon, N., Casey, C., Reichmuth, C. and Charrier, I. (2017) 'Northern elephant seals memorize the rhythm and timbre of their rivals' voices', *Current Biology*, 27(15), pp. 2352-2356.e2. Available at: https://doi.org/10.1016/j.cub.2017.06.035.

Mauch, M., MacCallum, R.M., Levy, M. and Leroi, A.M. (2015) 'The evolution of popular music: USA 1960–2010', *Royal Society Open Science*, 2(5), p. 150081. Available at: https://doi.org/10.1098/rsos.150081.

McAuley, J.D. (2010) 'Tempo and Rhythm', in M.R. Jones, R.R. Fay, and A.N. Popper (eds) *Music Perception*. New York, NY, US: Springer Science + Business Media (Springer handbook of auditory research), pp. 165–199. Available at: https://doi.org/10.1007/978-1-4419-6114-3_6.

McDonald, M.A., Hildebrand, J.A. and Wiggins, S.M. (2006) 'Increases in deep ocean ambient noise in the Northeast Pacific west of San Nicolas Island, California', *The Journal of the Acoustical Society of America*, 120(2), pp. 711–718. Available at: https://doi.org/10.1121/1.2216565.

McGowen, M.R., Spaulding, M. and Gatesy, J. (2009) 'Divergence date estimation and a comprehensive molecular tree of extant cetaceans', *Molecular Phylogenetics and*

Evolution, 53(3), pp. 891–906. Available at: https://doi.org/10.1016/j.ympev.2009.08.018.

McGregor, P.K., Holland, J. and Shepherd, M. (1997) ‘The Ecology of Corn Bunting (Miliaria calandra) Song Dialects and Their Potential Use in Conservation’, in P.F. Donald and N.J. Aebischer (eds) *The Ecology and Conservation of Corn Buntings Milaria calandra*. UK Nature Conservation 13, pp. 86–97.

McLean, P. (1977) *Beneath the Horizon I* [Musical Score]. Petersburgh NY: MLC Publications.

Mcloughlin, M., Lamoni, L., Garland, E.C., Ingram, S., Kirke, A., Noad, M.J., Rendell, L. and Miranda, E. (2018) ‘Using agent-based models to understand the role of individuals in the song evolution of humpback whales (Megaptera novaeangliae)’, *Music & Science*, 1. Available at: https://doi.org/10.1177/2059204318757021.

Merriam-Webster (no date) ‘Bioacoustics’, *Merriam-Webster.com dictionary*. Available at: https://www.merriam-webster.com/dictionary/bioacoustics (Accessed: 21 December 2021).

Meyer, L.B. (1956) *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

Mohamed-Tahrin, N., South, A.M., Newnham, D.A. and Jones, R.L. (2001) ‘A new accurate wavelength calibration for the ozone absorption cross section in the near-UV spectral region, and its effect on the retrieval of BrO from measurements of zenith-scattered sunlight’, *Journal of Geophysical Research: Atmospheres*, 106(D9), pp. 9897–9907. Available at: https://doi.org/10.1029/2000JD900743.

Mol, C., Chen, A., Kager, R.W.J. and ter Haar, S.M. (2017) ‘Prosody in birdsong: A review and perspective’, *Neuroscience & Biobehavioral Reviews*, 81, pp. 167–180. Available at: https://doi.org/10.1016/j.neubiorev.2017.02.016.

Moore, J.W. (2017) ‘The capitalocene, part I: On the nature and origins of our ecological crisis’, *The Journal of Peasant Studies*, 44(3), pp. 594–630. Available at: https://doi.org/10.1080/03066150.2016.1235036.

Moore, M.J. (2014) ‘How we all kill whales’, *ICES Journal of Marine Science*, 71(4), pp. 760–763. Available at: https://doi.org/10.1093/icesjms/fsu008.

Moore, M.J. (2021) *We Are All Whalers: The Plight of Whales and Our Responsibility*. Chicago: Chicago University Press.

Moore, M.J. and van der Hoop, J.M. (2012) ‘The painful side of trap and fixed net fisheries: Chronic entanglement of large whales’, *Journal of Marine Biology*, 2012, p. e230653. Available at: https://doi.org/10.1155/2012/230653.

Morizot, B. (2021) *On the Animal Trail*. Translated by A. Brown. Cambridge UK ; Medford MA: Polity Press.

Morton, T. (2018) *Being Ecological*. UK: Pelican Books.

Mundy, R. (2018) *Animal Musicalities: Birds, Beasts, and Evolutionary Listening*. Middletown, Connecticut: Wesleyan University Press.

Mundy, R. (forthcoming) *Music, Rights, and Animals after 1950 in the United States*.

Murray, A., Dunlop, R.A., Noad, M.J. and Goldizen, A.W. (2018) ‘Stereotypic and complex phrase types provide structural evidence for a multi-message display in humpback whales (Megaptera novaeangliae)’, *The Journal of the Acoustical Society of America*, 143(2), p. 980. Available at: https://doi.org/10.1121/1.5023680.

Nachtigall, P.E., Lemonds, D.W. and Roitblat, H.L. (2000) ‘Psychoacoustic Studies of Dolphin and Whale Hearing’, in W.W.L. Au, A.N. Popper, and R.R. Fay (eds) *Hearing by Whales and Dolphins*. New York: Springer-Verlag (Springer Handbook of Auditory Research), pp. 330–363. Available at: https://doi.org/10.1007/978-1-4612-1150-1.

Nagel, T. (1974) ‘What is it like to be a bat?’, *Philosophical Review*, 83(October), pp. 435–50.

Nakamura, E. and Kaneko, K. (2019) ‘Statistical evolutionary laws in music styles’, *Scientific Reports*, 9(1), pp. 1–11. Available at: https://doi.org/10.1038/s41598-019-52380-6.

National Research Council (US) Committee on Low-Frequency Sound and Marine Mammals (1994) *Low-Frequency Sound and Marine Mammals: Current Knowledge and Research Needs*. Washington, D.C.: National Academies Press (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK236686/ (Accessed: 1 August 2023).

Naveda, L., Gouyon, F., Guedes, C. and Leman, M. (2011) 'Microtiming patterns and interactions with musical properties in samba music', *Journal of New Music Research*, 40(3), pp. 225–238. Available at: https://doi.org/10.1080/09298215.2011.603833.

Naylor, S. (2014) 'Appropriation, culture and meaning in electroacoustic music: A composer's perspective', *Organised Sound*, 19(2), pp. 110–116. Available at: https://doi.org/10.1017/S1355771814000041.

Nishimura, T. (2020) 'Primate Vocal Anatomy and Physiology: Similarities and Differences Between Humans and Nonhuman Primates', in N. Masataka (ed.) *The Origins of Language Revisited: Differentiation from Music and the Emergence of Neurodiversity and Autism*. Singapore: Springer, pp. 25–53. Available at: https://doi.org/10.1007/978-981-15-4250-3_2.

Nishimura, T., Mikami, A., Suzuki, J. and Matsuzawa, T. (2007) 'Development of the Laryngeal Air Sac in Chimpanzees', *International Journal of Primatology*, 28(2), pp. 483–492. Available at: https://doi.org/10.1007/s10764-007-9127-7.

Noad, M.J., Cato, D.H., Bryden, M.M., Jenner, M.-N. and Jenner, K.C.S. (2000) 'Cultural revolution in whale songs', *Nature*, 408(6812), p. 537. Available at: https://doi.org/10.1038/35046199.

Nollman, J. (1999) *The Charged Border: Where Whales and Humans Meet*. New York: Henry Holt & Co.

Nollman, J. (2008) 'Getting into the groove', *Trans. Revista Transcultural de Música*, 12.

Norton, P. and Scharff, C. (2016) '"Bird song metronomics": Isochronous organization of zebra finch song rhythm', *Frontiers in Neuroscience*, 10. Available at: https://doi.org/10.3389/fnins.2016.00309.

Nussbaum, M.C. (2022) *Justice for Animals: Our Collective Responsibility*. New York: Simon and Schuster.

O’Connor, B. (2013) *Adorno*. Abingdon, Oxon ; New York: Routledge.

Odling-Smee, F.J., Laland, K.N. and Feldman, M.W. (1996) ‘Niche construction’, *The American Naturalist,* 147(4), pp. 641–648. Available at: https://doi.org/10.1086/285870.

Oleson, E.M., Calambokidis, J., Burgess, W.C., McDonald, M.A., LeDuc, C.A. and Hildebrand, J.A. (2007) ‘Behavioral context of call production by eastern North Pacific blue whales’, *Marine Ecology Progress Series,* 330, pp. 269–284.

Oliveira, C., Wahlberg, M., Silva, M.A., Johnson, M., Antunes, R., Wisniewska, D.M., Fais, A., Goncalves, J. and Madsen, P.T. (2016) ‘Sperm whale codas may encode individuality as well as clan identity’, *Journal of the Acoustical Society of America*, 139(5), pp. 2860–2869. Available at: https://doi.org/10.1121/1.4949478.

O’Neil, J.D. (2010) ‘Translator’s Introduction’, in Uexküll, J. von, *A Foray Into the Worlds of Animals and Humans; with, A Theory of Meaning*. Translated by J. D. O’Neil. Minneapolis, MN: University of Minnesota Press, pp. 35–38.

Osiejuk, T.S., Zbikowski, B., Wheeldon, A. and Budka, M. (2019) ‘Hey mister tambourine dove, sing a song for me: Simple but individually specific songs of Turtur tympanistria from Cameroon’, *Avian Research,* 10, p. 14. Available at: https://doi.org/10.1186/s40657-019-0153-x.

Otter, K.A., Mckenna, A., LaZerte, S.E. and Ramsay, S.M. (2020) ‘Continent-wide shifts in song dialects of white-throated sparrows’, *Current Biology,* 30, pp. 1–5. Available at: https://doi.org/10.1016/j.cub.2020.05.084.

Owen, C., Rendell, L., Constantine, R., Noad, M.J., Allen, J.A., Andrews, O., Garrigue, C., Poole, M.M., Donnelly, D., Hauser, N. and Garland, E.C. (2019) ‘Migratory convergence facilitates cultural transmission of humpback whale song’, *Royal Society Open Science*, 6(9), p. 190337. Available at: https://doi.org/10.1098/rsos.190337.

Pace, F., Benard, F., Glotin, H., Adam, O. and White, P. (2010) ‘Subunit definition and analysis for humpback whale call classification’, *Applied Acoustics*, 71(11), pp. 1107–1112. Available at: https://doi.org/10.1016/j.apacoust.2010.05.016.

Paddison, M. (1993) *Adorno's aesthetics of music*. Cambridge: Cambridge University Press.

Parks, S.E., Clark, C.W. and Tyack, P.L. (2007) ‘Short- and long-term changes in right whale calling behavior: The potential effects of noise on acoustic communication’, *The Journal of the Acoustical Society of America*, 122(6), pp. 3725–3731. Available at: https://doi.org/10.1121/1.2799904.

Patel, A.D. (2006) ‘Musical rhythm, linguistic rhythm, and human evolution’, *Music Perception*, 24(1), pp. 99–104. Available at: https://doi.org/10.1525/mp.2006.24.1.99.

Patel, A.D. (2008) *Music, Language, and the Brain*. Oxford: Oxford University Press.

Patel, A.D. (2014) ‘The evolutionary biology of musical rhythm: Was Darwin wrong?’, *PLOS Biology*, 12(3), p. e1001821. Available at: https://doi.org/10.1371/journal.pbio.1001821.

Patel, A.D. and Daniele, J.R. (2003) ‘An empirical comparison of rhythm in language and music’, *Cognition*, 87(1), pp. B35–B45. Available at: https://doi.org/10.1016/S0010-0277(02)00187-7.

Patel, A.D., Iversen, J.R., Bregman, M.R. and Schulz, I. (2009) ‘Experimental evidence for synchronization to a musical beat in a nonhuman animal’, *Current biology: CB*, 19(10), pp. 827–830. Available at: https://doi.org/10.1016/j.cub.2009.03.038.

Payne, K.B. (2000) ‘The Progressively Changing Songs of Humpback Whales: A Window on the Creative Process in a Wild Animal’, in N.L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music*. Cambridge, Mass: MIT Press, pp. 135–150.

Payne, K.B. and Payne, R.S. (1985) ‘Large scale changes over 19 years in songs of humpback whales in Bermuda’, *Zeitschrift für Tierpsychologie*, 68(2), pp. 89–114. Available at: https://doi.org/10.1111/j.1439-0310.1985.tb00118.x.

Payne, K.B., Tyack, P.L. and Payne, R.S. (1983) ‘Progressive changes in the songs of humpback whales (Megaptera novaeangliae): A detailed analysis of two seasons in

Hawaii', in R.S. Payne (ed.) *Communication and Behavior of Whales*. Boulder, CO: Westview Press (AAAS Selected Symposium, 76), pp. 9–57.

Payne, R.B. (1985) 'Behavioral continuity and change in local song populations of village indigobirds Vidua chalybeate', *Zeitschrift für Tierpsychologie*, 70(1), pp. 1–44. Available at: https://doi.org/10.1111/j.1439-0310.1985.tb00498.x.

Payne, R.S. (1970) *Songs of the Humpback Whale* [Vinyl]. Capital Records ST-620.

Payne, R.S. (1977) *Deep Voices (The Second Whale Record)* [Vinyl]. Capital Records ST-11598.

Payne, R.S. (1995) *Among Whales*. New York: Scribner.

Payne, R.S. (2013) 'The Night that Changed the Course of My Life', *Ocean Alliance,* 5 October. Available at: https://whale.org/the-rainy-night-that-changed-the-course-of-roger-paynes-life/ (Accessed: 10 January 2020).

Payne, R.S. and Guinee, L.N. (1983) 'Humpback whale (Megaptera novaeangliae) songs as an indicator of "stocks"', in R.S. Payne (ed.) *Communication and Behavior of Whales*. Boulder, CO: Westview Press (AAAS Selected Symposium, 76), pp. 333–358.

Payne, R.S. and McVay, S. (1971) 'Songs of humpback whales', *Science,* 173(3997), pp. 585–597. Available at: https://doi.org/10.1126/science.173.3997.585.

Percino, G., Klimek, P. and Thurner, S. (2014) 'Instrumentational complexity of music genres and why simplicity sells', *PLOS ONE,* 9(12), p. e115255. Available at: https://doi.org/10.1371/journal.pone.0115255.

Philharmonia Orchestra and Amos, D. (1989) *Music of Alan Hovhaness*. Sedro Wooley, WA: Crystal Records.

Plumwood, V. (1993) 'The politics of reason: Towards a feminist logic', *Australasian Journal of Philosophy,* 71(4), pp. 436–462. Available at: https://doi.org/10.1080/00048409312345432.

Plumwood, V. (2002) *Environmental Culture: The Ecological Crisis of Reason*. London ; New York: Routledge.

Plumwood, V. (2003) ‘Decolonizing Relationships with Nature’, in W.M. Adams and M. Mulligan (eds) *Decolonizing Nature: Strategies for Conservation in a Post-colonial Era*. London ; Sterling, VA: Earthscan Publications. Available at: https://doi.org/10.4324/9781849770927.

Plumwood, V. (2007) ‘Human exceptionalism and the limitations of animals: A review of Raimond Gaita’s “The Philosopher’s Dog”’, *Australian Humanities Review,* 42. Available at: http://australianhumanitiesreview.org/2007/08/01/human-exceptionalism-and-the-limitations-of-animals-a-review-of-raimond-gaitas-the-philosophers-dog/ (Accessed: 6 September 2021).

Podos, J. and Nowicki, S. (2004) ‘Performance Limits on Birdsong’, in P. Marler and H. Slabbekoorn (eds) *Nature’s Music: The Science of Birdsong*. San Diego ; London: Elsevier Academic Press, pp. 318–342.

Poesel, A., Foerster, K. and Kempenaers, B. (2001) ‘The dawn song of the blue tit Parus caeruleus and its role in sexual selection’, *Ethology,* 107(6), pp. 521–531. Available at: https://doi.org/10.1046/j.1439-0310.2001.00680.x.

Polak, R. (2010) ‘Rhythmic feel as meter: Non-isochronous beat subdivision in jembe music from Mali’, *Music Theory Online,* 16(4). Available at: https://mtosmt.org/issues/mto.10.16.4/mto.10.16.4.polak.html (Accessed: 19 March 2020).

Polak, R., London, J. and Jacoby, N. (2016) ‘Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: A corpus study of Malian jembe drumming’, *Frontiers in Neuroscience,* 10. Available at: https://doi.org/10.3389/fnins.2016.00285.

Post, O. and Toussaint, G. (2011) ‘The edit distance as a measure of perceived rhythmic similarity’, *Empirical Musicology Review,* 6(3), pp. 164–179.

Poster, R. and Reagan, R. (no date) *The Birdsong Project*. Available at: https://www.thebirdsongproject.com (Accessed: 20 June 2023).

Pound, J. (2014) ‘Wild at heart’, *BBC Music Magazine,* November.

Povey, I.M., South, A.M., de Roodenbeke, A. t’Kint, Hill, C., Freshwater, R.A. and Jones, R.L. (1998) ‘A broadband lidar for the measurement of tropospheric constituent profiles from the ground’, *Journal of Geophysical Research: Atmospheres,* 103(D3), pp. 3369–3380. Available at: https://doi.org/10.1029/97JD02969.

Powell, K. (1993) *Whale* [Musical Score]. Sounz. Available at: https://www.sounz.org.nz/works/15822.

Prögler, J.A. (1995) ‘Searching for swing: Participatory discrepancies in the jazz rhythm section’, *Ethnomusicology*, 39(1), pp. 21–54. Available at: https://doi.org/10.2307/852199.

Prum, R.O. (2018) *The Evolution of Beauty: How Darwin’s Forgotten Theory of Mate Choice Shapes the Animal World and Us*. New York: Anchor Books.

Raju, M., Asu, E.L. and Ross, J. (2010) ‘Comparison of rhythm in musical scores and performances as measured with the pairwise variability index’, *Musicae Scientiae,* 14(1), pp. 51–71. Available at: https://doi.org/10.1177/102986491001400102.

Ramp, C., Gaspard, D., Gavrilchuk, K., Unger, M., Schleimer, A., Delarue, J., Landry, S. and Sears, R. (2021) ‘Up in the air: drone images reveal underestimation of entanglement rates in large rorqual whales’, *Endangered Species Research*, 44, pp. 33–44. Available at: https://doi.org/10.3354/esr01084.

Ramp, C., Hagen, W., Palsbøll, P., Bérubé, M. and Sears, R. (2010) ‘Age-related multi-year associations in female humpback whales (Megaptera novaeangliae)’, *Behavioral Ecology and Sociobiology,* 64(10), pp. 1563–1576. Available at: https://doi.org/10.1007/s00265-010-0970-8.

Ravignani, A. (2017) ‘Visualizing and interpreting rhythmic patterns using phase space plots’, *Music Perception*, 34(5), pp. 557–568. Available at: https://doi.org/10.1525/mp.2017.34.5.557.

Ravignani, A., Bowling, D.L. and Fitch, W.T. (2014) ‘Chorusing, synchrony, and the evolutionary functions of rhythm’, *Frontiers in Psychology*, 5. Available at: https://doi.org/10.3389/fpsyg.2014.01118.

Ravignani, A., Delgado, T. and Kirby, S. (2016) ‘Musical evolution in the lab exhibits rhythmic universals’, *Nature Human Behaviour*, 1(1), pp. 1–7. Available at: https://doi.org/10.1038/s41562-016-0007.

Ravignani, A., Fitch, W.T., Hanke, F.D., Heinrich, T., Hurgitsch, B., Kotz, S.A., Scharff, C., Stoeger, A.S. and de Boer, B. (2016) ‘What pinnipeds have to say about human speech, music, and the evolution of rhythm’, *Frontiers in Neuroscience*, 10. Available at: https://doi.org/10.3389/fnins.2016.00274.

Ravignani, A., Kello, C.T., de Reus, K., Kotz, S.A., Dalla Bella, S., Méndez-Aróstegui, M., Rapado-Tamarit, B., Rubio-Garcia, A. and de Boer, B. (2019) ‘Ontogeny of vocal rhythms in harbor seal pups: An exploratory study’, *Current Zoology*, 65(1), pp. 107–120. Available at: https://doi.org/10.1093/cz/zoy055.

Ravignani, A. and Madison, G. (2017) ‘The paradox of isochrony in the evolution of human rhythm’, *Frontiers in Psychology*, 8. Available at: https://doi.org/10.3389/fpsyg.2017.01820.

Ravignani, A. and Norton, P. (2017) ‘Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations’, *Journal of Language Evolution*, 2(1), pp. 4–19. Available at: https://doi.org/10.1093/jole/lzx002.

Read, A.J., Drinker, P. and Northridge, S. (2006) ‘Bycatch of marine mammals in U.S. and global fisheries’, *Conservation Biology*, 20(1), pp. 163–169. Available at: https://doi.org/10.1111/j.1523-1739.2006.00338.x.

Reeves, R. (2015) ‘Whale-watching: Sustainable tourism and ecological management’, *Marine Biology Research*, 11(4), pp. 444–446. Available at: https://doi.org/10.1080/17451000.2014.983525.

Reidenberg, J.S. and Laitman, J.T. (2007) ‘Discovery of a low frequency sound source in Mysticeti (baleen whales): Anatomical establishment of a vocal fold homolog’, *The*

Anatomical Record, 290(6), pp. 745–759. Available at: https://doi.org/10.1002/ar.20544.

Reiss, D. and Marino, L. (2001) ‘Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence’, *Proceedings of the National Academy of Sciences,* 98(10), pp. 5937–5942. Available at: https://doi.org/10.1073/pnas.101086398.

Rendell, L. and Whitehead, H. (2003a) ‘Comparing repertoires of sperm whale codas: A multiple methods approach’, *Bioacoustics,* 14, pp. 61–81. Available at: https://doi.org/10.1080/09524622.2003.9753513.

Rendell, L. and Whitehead, H. (2003b) ‘Vocal clans in sperm whales (Physeter macrocephalus)’, *Proceedings of the Royal Society B: Biological Sciences,* 270(1512), pp. 225–231. Available at: https://doi.org/10.1098/rspb.2002.2239.

Rendell, L. and Whitehead, H. (2005) ‘Spatial and temporal variation in sperm whale coda vocalizations: Stable usage and local dialects’, *Animal Behaviour,* 70(1), pp. 191–198. Available at: https://doi.org/10.1016/j.anbehav.2005.03.001.

Repp, B.H. (1990) ‘Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists’, *The Journal of the Acoustical Society of America*, 88(2), pp. 622–641. Available at: https://doi.org/10.1121/1.399766.

RIAA (no date) *Gold & Platinum, Recording Industry Association of America*. Available at: https://www.riaa.com/gold-platinum/ (Accessed: 27 October 2021).

Rice, T. (2000) ‘Béla Bartók and Bulgarian Rhythm’, in E. Antokoletz, V. Fischer, and B. Suchoff (eds) *Bartók Perspectives: Man, Composer, and Ethnomusicologist*. Oxford University Press, pp. 196–212.

Rice, T. (2014) *Ethnomusicology: A Very Short Introduction*. Oxford, New York: Oxford University Press.

Riebel, K. and Slater, P.J.B. (2003) ‘Temporal variation in male chaffinch song depends on the singer and the song type’, *Behaviour,* 140(2), pp. 269–288.

Ritts, M. (2017) ‘Environmentalists abide: Listening to whale music – 1965–1985’, *Environment and Planning D: Society and Space*, 35(6), pp. 1096–1114. Available at: https://doi.org/10.1177/0263775817711706.

Robinson, D. (2020) *Hungry Listening: Resonant Theory for Indigenous Sound Studies*. Minneapolis, Minnesota: University of Minnesota Press.

Robisson, P., Aubin, T. and Bremond, J.-C. (1993) ‘Individuality in the voice of the emperor penguin Aptenodytes forsteri: Adaptation to a noisy environment’, *Ethology*, 94(4), pp. 279–290. Available at: https://doi.org/10.1111/j.1439-0310.1993.tb00445.x.

Roeske, T.C., Tchernichovski, O., Poeppel, D. and Jacoby, N. (2020) ‘Categorical rhythms are shared between songbirds and humans’, *Current Biology*, 30. Available at: https://doi.org/10.1016/j.cub.2020.06.072.

Rogers, T.L. (2017) ‘Calling underwater is a costly signal: Size-related differences in the call rates of Antarctic leopard seals’, *Current Zoology*, 63(4), pp. 433–443. Available at: https://doi.org/10.1093/cz/zox028.

Rolland, R.M., Parks, S.E., Hunt, K.E., Castellote, M., Corkeron, P.J., Nowacek, D.P., Wasser, S.K. and Kraus, S.D. (2012) ‘Evidence that ship noise increases stress in right whales’, *Proceedings of the Royal Society B: Biological Sciences*, 279(1737), pp. 2363–2368. Available at: https://doi.org/10.1098/rspb.2011.2429.

Roman, J., Estes, J.A., Morissette, L., Smith, C., Costa, D., McCarthy, J., Nation, J., Nicol, S., Pershing, A. and Smetacek, V. (2014) ‘Whales as marine ecosystem engineers’, *Frontiers in Ecology and the Environment*, 12(7), pp. 377–385. Available at: https://doi.org/10.1890/130220.

Rosch, E.H. (1973) ‘On the Internal Structure of Perceptual and Semantic Categories’, in T.E. Moore (ed.) *Cognitive Development and Acquisition of Language*. San Diego: Academic Press, pp. 111–144. Available at: https://doi.org/10.1016/B978-0-12-505850-6.50010-4.

Rosenthal, G.G. (2017) *Mate Choice: The Evolution of Sexual Decision Making from Microbes to Humans*. Princeton, N.J. ; Oxford: Princeton University Press.

Rothenberg, D. (2005) *Why Birds Sing: A Journey into the Mystery of Bird Song*. Paperback. New York: Basic Books.

Rothenberg, D. (2008a) *'Never Satisfied', Whale Music* [CD]. Terra Nova Music. Available at: https://www.terranovamusic.net/.

Rothenberg, D. (2008b) *Thousand Mile Song: Whale Music in a Sea of Sound*. New York: Basic Books.

Rothenberg, D. (2008c) 'To wail with a whale: Anatomy of an interspecies duet', *Trans. Revista Transcultural de Música*, 12.

Rothenberg, D. (2015) *New Songs of the Humpback Whale* [CD]. Important Records IMPREC433.

Rothenberg, D. (2023) *Whale Music: Thousand Mile Songs in a Sea of Sound*. A revised and updated version of Thousand Mile Song: Whale Music in a Sea of Sound. Terra Nova Press.

Rothenberg, D. and Deal, M. (2015) 'A new morphological notation for the music of humpback whales', *Art & Perception*, 3(3), pp. 347–358. Available at: https://doi.org/10.1163/22134913-00002040.

Rothenberg, D., Roeske, T.C., Voss, H.U., Naguib, M. and Tchernichovski, O. (2014) 'Investigation of musicality in birdsong', *Hearing research*, 308, pp. 71–83. Available at: https://doi.org/10.1016/j.heares.2013.08.016.

Ryan, M.J. (2018) *A Taste for the Beautiful: The Evolution of Attraction*. Princeton, N.J. ; Oxford: Princeton University Press.

Saar, S. and Mitra, P.P. (2008) 'A technique for characterizing the development of rhythms in bird song', *PloS One*, 3(1), p. e1461. Available at: https://doi.org/10.1371/journal.pone.0001461.

Sachs, C. (1953) *Rhythm and Tempo: A Study in Music History*. New York: W.W. Norton.

Sakata, J.T. and Vehrencamp, S.L. (2012) ‘Integrating perspectives on vocal performance and consistency’, *Journal of Experimental Biology*, 215, pp. 201–209. Available at: https://doi.org/doi:10.1242/jeb.056911.

Sanders, A. (2020) *Solo Humpback Whale* [Musical Score]. Available at: https://www.soundingwhales.com/.

Sanders, A. (2021) *Glacier Bay Nov 2020* [Musical Score]. Available at: https://www.soundingwhales.com/.

Sandoval, L., Méndez, C. and Mennill, D.J. (2014) ‘Individual distinctiveness in the fine structural features and repertoire characteristics of the songs of white-eared ground-sparrows’, *Ethology*, 120(3), pp. 275–286. Available at: https://doi.org/10.1111/eth.12206.

Savage, P.E. (2019) ‘Cultural evolution of music’, *Palgrave Communications*, 5(1), p. 16. Available at: https://doi.org/10.1057/s41599-019-0221-1.

Savage, P.E., Brown, S., Sakai, E. and Currie, T.E. (2015) ‘Statistical universals reveal the structures and functions of human music’, *Proceedings of the National Academy of Sciences*, 112(29), pp. 8987–8992. Available at: https://doi.org/10.1073/pnas.1414495112.

Savage, P.E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S. and Fitch, W.T. (2021) ‘Music as a coevolved system for social bonding’, *Behavioral and Brain Sciences*, 44(e59), pp. 1–22. Available at: https://doi.org/10.1017/S0140525X20000333.

Schneider, J.N. and Mercado III, E. (2019) ‘Characterizing the rhythm and tempo of sound production by singing whales’, *Bioacoustics*, 28(3), pp. 239–256. Available at: https://doi.org/10.1080/09524622.2018.1428827.

Schulze, J.N., Denkinger, J., Oña, J., Poole, M.M. and Garland, E.C. (2022) ‘Humpback whale song revolutions continue to spread from the central into the eastern South Pacific’, *Royal Society Open Science*, 9(8), p. 220158. Available at: https://doi.org/10.1098/rsos.220158.

Seashore, C.E. (1938) *Psychology of Music*. New York: McGraw-Hill.

Sebo, J. (2017) ‘Agency and moral status’, *Journal of Moral Philosophy*, 14(1), pp. 1–22. Available at: https://doi.org/10.1163/17455243-46810046.

Sek, A. and Moore, B.C.J. (1995) ‘Frequency discrimination as a function of frequency, measured in several ways’, *The Journal of the Acoustical Society of America*, 97(4), pp. 2479–2486. Available at: https://doi.org/10.1121/1.411968.

Semjen, A. and Ivry, R.B. (2001) ‘The coupled oscillator model of between-hand coordination in alternate-hand tapping: A reappraisal’, *Journal of Experimental Psychology. Human Perception and Performance*, 27(2), pp. 251–265. Available at: https://doi.org/10.1037//0096-1523.27.2.251.

Senn, O., Kilchenmann, L., Bechtold, T. and Hoesl, F. (2018) ‘Groove in drum patterns as a function of both rhythmic properties and listeners’ attitudes’, *PLOS ONE*, 13(6), p. e0199604. Available at: https://doi.org/10.1371/journal.pone.0199604.

Senn, O., Kilchenmann, L., von Georgi, R. and Bullerjahn, C. (2016) ‘The effect of expert performance microtiming on listeners’ experience of groove in swing or funk music’, *Frontiers in Psychology*, 7. Available at: https://www.frontiersin.org/article/10.3389/fpsyg.2016.01487 (Accessed: 1 May 2022).

Simchy-Gross, R. and Margulis, E.H. (2018) ‘The sound-to-music illusion: Repetition can musicalize nonspeech sounds’, *Music & Science*, 1, p. 2059204317731992. Available at: https://doi.org/10.1177/2059204317731992.

Simmons, L.W. and Zuk, M. (1992) ‘Variability in call structure and pairing success of male field crickets, Gryllus bimaculatus: The effects of age, size and parasite load’, *Animal Behaviour*, 44(6), pp. 1145–1152. Available at: https://doi.org/10.1016/S0003-3472(05)80326-4.

Simonis, A.E., Brownell, R.L., Thayre, B.J., Trickey, J.S., Oleson, E.M., Huntington, R. and Baumann-Pickering, S. (2020) ‘Co-occurrence of beaked whale strandings and naval sonar in the Mariana Islands, Western Pacific’, *Proceedings of the Royal Society B:*

Biological Sciences, 287(1921), p. 20200070. Available at: https://doi.org/10.1098/rspb.2020.0070.

Sinclair, N.C., Ursell, J., South, A. and Rendell, L. (2022) ‘From Beethoven to Beyoncé: Do changing aesthetic cultures amount to “cumulative cultural evolution”?’, *Frontiers in Psychology*, 12. Available at: https://doi.org/10.3389/fpsyg.2021.663397.

Širović, A., Oleson, E.M., Buccowich, J., Rice, A. and Bayless, A.R. (2017) ‘Fin whale song variability in southern California and the Gulf of California’, *Scientific Reports,* 7(1), p. 10126. Available at: https://doi.org/10.1038/s41598-017-09979-4.

Small, C. (1998) *Musicking: The Meanings of Performing and Listening*. Hanover: Wesleyan University Press.

Small, M. (2005) *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance*. World Scientific (World Scientific Series on Nonlinear Science Series A). Available at: https://doi.org/10.1142/5722.

Smalley, D. (1986) ‘Spectro-morphology and Structuring Processes’, in S. Emmerson (ed.) *The Language of Electroacoustic Music*. London: Palgrave Macmillan UK, pp. 61–93. Available at: https://doi.org/10.1007/978-1-349-18492-7_5.

Smith, J. (2020) ‘Self-Consciousness’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Summer 2020. Metaphysics Research Lab, Stanford University. Available at: https://plato.stanford.edu/archives/sum2020/entries/self-consciousness/ (Accessed: 5 August 2023).

Smith, J.N., Goldizen, A.W., Dunlop, R.A. and Noad, M.J. (2008) ‘Songs of male humpback whales, Megaptera novaeangliae, are involved in intersexual interactions’, *Animal Behaviour,* 76(2), pp. 467–477.

Sokal, R.R. and Rohlf, F.J. (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*. 3rd edn. New York: W.H. Freeman.

Sorce Keller, M. (2012) ‘Zoomusicology and ethnomusicology: A marriage to celebrate in heaven’, *Yearbook for Traditional Music,* 44, pp. 166–183. Available at: https://doi.org/10.5921/yeartradmusi.44.0166.

Sound of Yell (2014) *Brocken Spectre* [Vinyl]. Glasgow, UK: Chemikal Underground.

Sound of Yell (2020) *Leapling* [CD]. Glasgow, UK: Chemikal Underground.

Souster, T. (1972) *Spectral* [Musical Score]. Chipping Norton UK: Composers Edition.

South, A. (2011) ‘Two worlds? Gesture and speech in Thomas Reid and Maurice Merleau-Ponty’, *Journal of Scottish Thought,* 4(1), pp. 151–168. Available at: https://doi.org/10.57132/jst.83.

South, A. (2022) ‘Composing with cetaceans: Countering human exceptionalism through a practical zoömusicology’, *Journal of Interdisciplinary Voice Studies*, 7(1), pp. 51–67. Available at: https://doi.org/10.1386/jivs_00054_1.

South, A., Garland, E.C. and Rendell, L. (2022) ‘Cultural Behaviour in Cetaceans’, in T.M. Freeberg, A.R. Ridley, and P. d’Ettore (eds) *The Routledge International Handbook of Comparative Psychology*. Routledge, pp. 227–239. Available at: https://doi.org/10.4324/9781003091868-22.

South, A.M., Povey, I.M. and Jones, R.L. (1998) ‘Broadband lidar measurements of tropospheric water vapor profiles’, *Journal of Geophysical Research: Atmospheres*, 103(D23), pp. 31191–31202. Available at: https://doi.org/10.1029/98JD02852.

Stafford, K.M., Chapp, E., Bohnenstiel, D.R. and Tolstoy, M. (2011) ‘Seasonal detection of three types of “pygmy” blue whale calls in the Indian Ocean’, *Marine Mammal Science*, 27(4), pp. 828–840. Available at: https://doi.org/10.1111/j.1748-7692.2010.00437.x.

Stafford, K.M., Moore, S.E., Laidre, K.L. and Heide-Jørgensen, M.P. (2008) ‘Bowhead whale springtime song off West Greenland’, *The Journal of the Acoustical Society of America*, 124(5), pp. 3315–3323. Available at: https://doi.org/10.1121/1.2980443.

Stambaugh, L.A. (2011) ‘When repetition isn’t the best practice strategy: Effects of blocked and random practice schedules’, *Journal of Research in Music Education,* 58(4), pp. 368–383. Available at: https://doi.org/10.1177/0022429410385945.

Susam-Saraeva, Ş. (forthcoming) ‘Dolphin and Whale Communication in Arts and Music: Towards a New Ethics of Translation and Representation’, in E. De Francisci and C. Marinetti (eds) *Translation in the Performing Arts*. London & New York: Routledge.

Suzuki, R., Buck, J.R. and Tyack, P.L. (2006) ‘Information entropy of humpback whale songs’, *The Journal of the Acoustical Society of America*, 119(3), pp. 1849–1866. Available at: https://doi.org/10.1121/1.2161827.

Szőke, P., Gunn, W.W.H. and Filip, M. (1969) ‘The musical microcosm of the hermit thrush: From Athanasius Kircher’s naive experiments of musical transcription of bird voice to sound microscopy and the scientific musical representation of bird song’, *Studia Musicologica Academiae Scientiarum Hungaricae*, 11(1/4), pp. 423–438. Available at: https://doi.org/10.2307/901296.

Tanguiane, A.S. (1993) *Artificial Perception and Music Recognition*. Berlin: Springer-Verlag.

Tarr, B., Launay, J. and Dunbar, R.I.M. (2016) ‘Silent disco: Dancing in synchrony leads to elevated pain thresholds and social closeness’, *Evolution and human behavior : official journal of the Human Behavior and Evolution Society*, 37(5), pp. 343–349. Available at: https://doi.org/10.1016/j.evolhumbehav.2016.02.004.

Taylor, H. (2014) ‘Whose bird is it? Messiaen’s transcriptions of Australian songbirds’, *Twentieth-Century Music*, 11(1), pp. 63–100. Available at: https://doi.org/10.1017/S1478572213000194.

Taylor, H. (2017a) *Absolute Bird* [CD]. Thornton Heath, UK: ReR Megacorp.

Taylor, H. (2017b) *Is Birdsong Music? Outback Encounters with an Australian Songbird*. Bloomington, Indiana: Indiana University Press (Music, Nature, Place).

Temperley, D. (2017) ‘Rhythmic variability in European vocal music’, *Music Perception,* 35(2), pp. 193–199. Available at: https://doi.org/10.1525/mp.2017.35.2.193.

Templeton, C.N., Mann, N.I., Ríos-Chelén, A.A., Quiros-Guerrero, E., Macías Garcia, C. and Slater, P.J.B. (2013) ‘An experimental study of duet integration in the happy wren, Pheugopedius felix’, *Animal Behaviour*, 86(4), pp. 821–827. Available at: https://doi.org/10.1016/j.anbehav.2013.07.022.

Terleph, T.A., Malaivijitnond, S. and Reichard, U.H. (2017) ‘Male white-handed gibbons flexibly time duet contributions’, *Behavioral Ecology and Sociobiology*, 72(1), p. 16. Available at: https://doi.org/10.1007/s00265-017-2432-z.

Tervo, O.M., Parks, S.E., Christoffersen, M.F., Miller, L.A. and Kristensen, R.M. (2011) ‘Annual changes in the winter song of bowhead whales (Balaena mysticetus) in Disko Bay, Western Greenland’, *Marine Mammal Science*, 27(3), pp. E241–E252. Available at: https://doi.org/10.1111/j.1748-7692.2010.00451.x.

The MathWorks Inc. (2021) ‘MATLAB version: 9.10.0 (R2021a)’. Natick, Massachusetts.

The MathWorks Inc. (2023) *Dummy Variables - MATLAB & Simulink*. Available at: https://uk.mathworks.com/help/stats/dummy-indicator-variables.html (Accessed: 16 July 2023).

The One Ensemble (2017) *Saint Seven* [Vinyl]. Self-released.

The One Ensemble Orchestra (2009) *Other Thunders* [CD]. No-Fi.

Thompson, T.J. (1981) *Temporal Characteristics of Humpback Whale (Megaptera novaeangliae) Songs*. PhD Dissertation. University of Rhode Island. Available at: https://digitalcommons.uri.edu/dissertations/AAI8306423.

Thorpe, W.H. (1954) ‘The process of song-learning in the chaffinch as studied by means of the sound spectrograph’, *Nature*, 173, pp. 465–469.

Thul, E. and Toussaint, G.T. (2008a) ‘A Comparative Phylogenetic-Tree Analysis of African Timelines and North Indian Talas’, in *Proceedings of Bridges: Mathematics, Music, Art, Architecture, Culture*. *Bridges Leeuwarden: Mathematics, Music, Art, Architecture, Culture*, Tarquin Publications, pp. 187–194. Available at: http://archive.bridgesmathart.org/2008/bridges2008-187.html (Accessed: 4 October 2020).

Thul, E. and Toussaint, G.T. (2008b) ‘Rhythm complexity measures: A comparison of mathematical models of human perception and performance’, *ISMIR 2008: Ninth International Conference on Music Information Retrieval*, pp. 663–668.

Tinbergen, N. (1963) ‘On aims and methods of ethology’, *Zeitschrift für Tierpsychologie,* 20(4), pp. 410–433. Available at: https://doi.org/10.1111/j.1439-0310.1963.tb01161.x.

Tomlinson, G. (2015) *A Million Years of Music: The Emergence of Human Modernity*. First edition. New York: Zone Books.

Toussaint, G.T. (2006) ‘A comparison of rhythmic dissimilarity measures’, *Forma*, 21, pp. 129–149.

Toussaint, G.T. (2013) ‘The pairwise variability index as a measure of rhythm complexity’, *Analytical approaches to world music*, 2(2).

Toussaint, G.T. (2020) *The Geometry of Musical Rhythm: What Makes a ‘Good’ Rhythm Good?* 2nd Edition. CRC Press.

Toussaint, G.T., Campbell, M. and Brown, N. (2011) ‘Computational models of symbolic rhythm similarity: Correlation with human judgments’, *Analytical Approaches to World Music*, 1(2). Available at: http://www.aawmjournal.com/articles/2011b/Toussaint_AAWM_Vol_1_2.htm.

Toussaint, G.T., Matthews, L.J., Campbell, M. and Brown, N. (2012) ‘Measuring musical rhythm similarity: Transformation versus feature-based methods’, *Journal of Interdisciplinary Music Studies*, 6(1), pp. 23–53.

Toussaint, G.T. and Oh, S.M. (2016) ‘Measuring musical rhythm similarity: Further experiments with the many-to-many minimum-weight matching distance’, *Journal of Computer and Communications*, 4(15), pp. 117–125. Available at: https://doi.org/10.4236/jcc.2016.415011.

Trainer, J.M. (1989) ‘Cultural evolution in song dialects of yellow-rumped caciques in Panama’, *Ethology*, 80(1–4), pp. 190–204. Available at: https://doi.org/10.1111/j.1439-0310.1989.tb00739.x.

Turino, T. (2009) ‘Four fields of music making and sustainable living’, *The World of Music*, 51(1), pp. 95–117.

Tyack, P. (1981) ‘Interactions between singing Hawaiian humpback whales and conspecifics nearby’, *Behavioral Ecology and Sociobiology*, 8(2), pp. 105–116. Available at: https://doi.org/10.1007/BF00300822.

Tyack, P.L. (2009) ‘Human-generated sound and marine mammals’, *Physics Today*, 62(11), pp. 39–44. Available at: https://doi.org/10.1063/1.3265235.

Tyack, P.L. (2020) ‘A taxonomy for vocal learning’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1789), p. 20180406. Available at: https://doi.org/10.1098/rstb.2018.0406.

Tyack, P.L. (2022) ‘Social Organization of Baleen Whales’, in C.W. Clark and E.C. Garland (eds) *Ethology and Behavioral Ecology of Mysticetes*. Cham: Springer International Publishing (Ethology and Behavioral Ecology of Marine Mammals), pp. 147–175. Available at: https://doi.org/10.1007/978-3-030-98449-6_11.

Uexküll, J. von (2010) *A Foray Into the Worlds of Animals and Humans; with, A Theory of Meaning*. Translated by J.D. O’Neil. Minneapolis, MN: University of Minnesota Press.

Van Vugt, F.T., Jabusch, H.-C. and Altenmüller, E. (2013) ‘Individuality that is unheard of: Systematic temporal deviations in scale playing leave an inaudible pianistic fingerprint’, *Frontiers in Psychology*, 4, p. 134. Available at: https://doi.org/10.3389/fpsyg.2013.00134.

Vasquez, S. and Cojtí, A. (2020) *Cultural Appropriation: Another Form of Extractivism of Indigenous Communities, Cultural Survival*. Available at: https://www.culturalsurvival.org/news/cultural-appropriation-another-form-extractivism-indigenous-communities (Accessed: 29 June 2023).

de Waal, F.B.M. (1999) ‘Anthropomorphism and anthropodenial: Consistency in our thinking about humans and other animals’, *Philosophical Topics*, 27(1), pp. 255–280.

de Waal, F.B.M. (2016) *Are We Smart Enough to Know How Smart Animals are?* London: Granta Books.

Watkins, W.A., Tyack, P., Moore, K.E. and Bird, J.E. (1987) ‘The 20-Hz signals of finback whales (Balaenoptera physalus)’, *The Journal of the Acoustical Society of America*, 82(6), pp. 1901–1912. Available at: https://doi.org/10.1121/1.395685.

Weilgart, L. and Whitehead, H. (1993) ‘Coda communication by sperm whales (Physeter macrocephalus) off the Galapagos Islands’, *Canadian Journal of Zoology*, 71(4), pp. 744–752. Available at: https://doi.org/10.1139/z93-098.

Weinrich, M.T. (2011) ‘Stable social associations among humpback whales (Megaptera novaeangliae) in the southern Gulf of Maine’, *Canadian Journal of Zoology*, 69(12). Available at: https://doi.org/10.1139/z91-425.

White, F. (no date) *Composer and Material in Musique Concrète*. Available at: https://rosewhitemusic.com/composer-and-material-in-musique-concrete/ (Accessed: 6 June 2023).

White, T.I. (2007) *In Defense of Dolphins: The New Moral Frontier*. Malden, MA: John Wiley & Sons.

Whitehead, H., Dillon, M., Dufault, S., Weilgart, L. and Wright, J. (1998) ‘Non-geographically based population structure of South Pacific sperm whales: Dialects, fluke-markings and genetics’, *Journal of Animal Ecology*, 67(2), pp. 253–262. Available at: https://doi.org/10.1046/j.1365-2656.1998.00187.x.

Whitehead, H. and Rendell, L. (2015) *The Cultural Lives of Whales and Dolphins*. Chicago & London: Chicago University Press.

Whiten, A. (2019) ‘Cultural evolution in animals’, *Annual Review of Ecology, Evolution, and Systematics*, 50(1), pp. 27–48. Available at: https://doi.org/10.1146/annurev-ecolsys-110218-025040.

Whitmore, J. (2014) *Review: Richard Blackford The Great Animal Orchestra Symphony, MusicWeb International*. Available at: http://www.musicweb-international.com/classrev/2014/Dec14/Blackford_animal_NI6274.htm (Accessed: 30 June 2023).

Wichert, R.N. and Nussbaum, M.C. (2017) ‘Legal Protection for Whales: Capabilities, Entitlements, and Culture’, in L. Cordeiro-Rodrigues and L. Mitchell (eds) *Animals, Race, and Multiculturalism*. Cham: Springer International Publishing (The Palgrave Macmillan Animal Ethics Series), pp. 95–120.

Wichert, R.N. and Nussbaum, M.C. (2019) ‘The Legal Status of Whales and Dolphins: From Bentham to the Capabilities Approach’, in L. Keleher and S.J. Kosko (eds) *Agency and Democracy in Development Ethics*. Cambridge: Cambridge University Press, pp. 259–288.

Williams, H. and Lachlan, R.F. (2022) ‘Evidence for cumulative cultural evolution in bird song’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1843), p. 20200322. Available at: https://doi.org/10.1098/rstb.2020.0322.

Williams, H., Levin, I.I., Norris, D.R., Newman, A.E.M. and Wheelwright, N.T. (2013) ‘Three decades of cultural evolution in savannah sparrow songs’, *Animal Behaviour*, 85(1), pp. 213–223. Available at: https://doi.org/10.1016/j.anbehav.2012.10.028.

Wilson, M. and Cook, P.F. (2016) ‘Rhythmic entrainment: Why humans want to, fireflies can’t help it, pet birds try, and sea lions have to be bribed’, *Psychonomic Bulletin & Review*, 23(6), pp. 1647–1659. Available at: https://doi.org/10.3758/s13423-016-1013-x.

Wing, A.M. and Kristofferson, A.B. (1973) ‘Response delays and the timing of discrete motor responses’, *Perception & Psychophysics*, 14(1), pp. 5–12. Available at: https://doi.org/10.3758/BF03198607.

Winn, H.E. and Winn, L.K. (1978) ‘The song of the humpback whale Megaptera novaeangliae in the West Indies’, *Marine Biology*, 47(2), pp. 97–114. Available at: https://doi.org/10.1007/BF00395631.

Winter, P. (1978) *‘Ocean Dream’, Common Ground* [Vinyl]. Santa Monica, CA: A&M Records.

Winter, P. (1980) *'Lullaby from the Great Mother Whale for the Baby Seal Pups', Callings* [Vinyl]. Litchfield, CT: Living Music Records.

Winter, P. and Halley, P. (1987) *'Ocean Dream', Whales Alive* [Vinyl]. Sausalito, CA: Living Music Records.

Wood, M. and Širović, A. (2022) 'Characterization of fin whale song off the Western Antarctic Peninsula', *PloS one*, 17(3), p. e0264214. Available at: https://doi.org/10.1371/journal.pone.0264214.

Wray, J., Keen, E. and O'Mahony, É.N. (2021) 'Social survival: Humpback whales (Megaptera novaeangliae) use social structure to partition ecological niches within proposed critical habitat', *PLOS ONE*, 16(6), p. e0245409. Available at: https://doi.org/10.1371/journal.pone.0245409.

Xing, J., Sainburg, T., Taylor, H. and Gentner, T.Q. (2022) 'Syntactic modulation of rhythm in Australian pied butcherbird song', *Royal Society Open Science*, 9(9), p. 220704. Available at: https://doi.org/10.1098/rsos.220704.

Young, J.O. (2008) *Cultural Appropriation and the Arts*. Malden, MA: Blackwell Publishing.

Youngblood, M. (2019) 'Conformity bias in the cultural transmission of music sampling traditions', *Royal Society Open Science*, 6(9), p. 191149. Available at: https://doi.org/10.1098/rsos.191149.

Zandberg, L., Lachlan, R.F., Lamoni, L. and Garland, E.C. (2021) 'Global cultural evolutionary model of humpback whale song', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1836), p. 20200242. Available at: https://doi.org/10.1098/rstb.2020.0242.

Zhong, M., Castellote, M., Dodhia, R., Lavista Ferres, J., Keogh, M. and Brewer, A. (2020) 'Beluga whale acoustic signal classification using deep learning neural network models', *The Journal of the Acoustical Society of America*, 147(3), pp. 1834–1841. Available at: https://doi.org/10.1121/10.0000921.

Appendix A Whale Music

This inexhaustive survey of human music inspired by cetacean sounds took its lead from the second chapter of David Rothenberg's *Thousand Mile Song* (2008), and was substantially enriched by a Facebook discussion initiated by Emily Doolittle. It primarily covers genres of which I have a reasonable acquaintance: Western classical music (and its experimental underbelly), jazz and rock/pop. I have focused on music where the inspiration is sonic (e.g., as shown through the use of field recordings, or vocal/instrumental reference to the acoustic and other music-like aspects of whale calls and song) rather than linguistic (e.g., verbal reference to whale hunting, whale conservation or other narratives). Where composers or performers have explicitly stated that their work is connected to whale sounds, I have taken them at their word, even if the mode of connection is not obvious in the resulting work. Genres which are substantially under-represented include Traditional and New Age musics. This survey is divided into compositions where the principal output was a score for performance by others (Table A1), and compositions where the principal output was a recording (Table A2). When particular species of cetaceans are referenced in the piece of music, these are listed using the following abbreviations:

- Ba = *Balaenoptera acutorostrata* (Common Minke Whale)
- Bmus = *Balaenoptera musculus* (Blue Whale)
- Bmys= *Balaena mysticetus* (Bowhead Whale)
- Dl = *Delphinapterus leucas* (Beluga Whale, or White Whale)
- Ea = *Eubalaena australis* (Southern Right Whale)
- Ha= *Hyperoodon ampullatus* (Northern Bottlenosed Whale)
- Mn = *Megaptera novaeangliae* (Humpback Whale)
- Oo = *Orcinus orca* (Orca, or Killer Whale)
- Pm = *Physeter macrocephalus* (Sperm Whale)
- T=*Tursiops spp.* (Bottlenosed Dolphin)

Table A1 Works of music inspired by cetacean vocalizations, for which the principal output is a score for others to perform. Works are listed in date order. The incorporation of field recordings is indicated with (●) after species.

Composer	**Title**	**Year**	**Instrumentation**	**Species**	**Publisher/ Source/ Website**
Morrison, Frank	Tzinquaw	1950	Indigenous Opera	Oo	
Perry, Nina	The Whale	1959	Voices, Piano		OUP
Swann, Donald	The Whale Song (Mopy Dick)	1963	Voice, Piano	Ha	Warner Chappell Music
Tavener, John	The Whale - A Biblical Fantasy	1966	Chorus plus Orchestra		Chester Music
Lucier, Alvin	Quasimodo, the Great Lover	1970	Open	Mn	
Hovhaness, Alan	And God Created Great Whales	1970	Orchestra and Tape	Mn (●)	Edition Peters
Crumb, George	Vox Balaenae	1971	Electric Flute, Electric Cello, Amplified Piano	Mn	Edition Peters
Hodkinson, Sydney	Blue Whale (Sea Chanteys)	1971	2 x Mixed Chorus	Bmus	CMC
Farberman, Harold	The Blue Whale	1972	Mezzosoprano, Chamber Ensemble		
Souster, Tim	Spectral	1972	Solo Viola, Tape-delay system and Live Electronics	Mn	Composers Edition
Nam June Paik	Danger Music for Dick Higgins	1973	Open		

Wilder, Alec	Elegy For The Whale	1975	Solo Tuba and Orchestra		Music Sales Classical NMUSA
Lewis, Robert Hall	Nuances II 'Whale Lament'	1976	Orchestra and Tape	Mn (●)	NMUSA
Mclean, Priscilla	Beneath the Horizon I	1977	Tuba Quartet and Whale Ensemble	Mn (●)	NMUSA
Kok, Ronald	Killer Whales	1978	Viola and Four Double Basses		Donemus
Mikhashoff, Yvar Emilian	HWALC	1978	Cello and Tape		
Mclean, Priscilla	Beneath the Horizon III	1978	Tuba Solo and Whale Ensemble	Mn (●)	NMUSA
Cage, John	Litany for the Whale	1980	Vocal Duet		Edition Peters
Finney, Ross Lee	Lost Whale Calf	1980	Solo Piano		
Searle, Humphrey	The Apollonian Whale, Op.74	1980	Cello, Piano, Voice		Whalesound BMC
Perkinson, Coleridge-Taylor	The White Whale	1981	Chamber Orchestra with Baritone Soloist		NMUSA
Takemitsu, Toru	Toward the Sea	1981	Alto Flute and Guitar or Alto Flute, Harp and Strings or Alto Flute and Harp (1989)		Schott Music
Bodman, Christopher	Songs of the Humpbacked Whale	1982	Trombone Quartet	Mn	BMC
Holman, Derek	The Greenland Whale	1982	Mixed Chorus		CMC
Xenakis, Iannis	Pour les Baleines	1982	Large String Orchestra		Salabert IRCAM

Cresswell, Lyell	The Kauri Tree and the Whale	1983	Orchestra		Sounz
Ellias, Roddy	Whale Spirit Rising	1985	Baritone Saxophone and String Orchestra		CMC
Bruynèl, Ton	Save the Whale	1989 /91	(Contra)bass Clarinet and Two Soundtracks	Mn (●)	Composer
Knox, Garth	Jonah and the Whale	1990 /96	Viola, Tuba		CMC Ireland
Robb, Magnus	Delphi	1990	Orchestra		SMC
Werle, Lars Johan	Ännu sjunger valarna [Still the Whales Are Singing]	1992	S, Mez, T, B-Bar, str qt, db, perc		
Bolliger, Phillip	Benedictus Balaenarum (Benediction of the Whales)	1993	Trombone, Piano		Kookaburra Music AMC
Evans, Robert	Whale Song Dancing, A Cantata	1993	Solo voice, Chorus, Instrumental ensemble		CMC
Morawetz, Oskar	The Whale's Lament	1993	Solo Piano		
Powell, Kit	Whale	1993	SATB Choir and Tape	Mn (●)	Sounz
Powell, Kit	Whale	1993	Trombone and Tape	Mn (●)	Sounz
Grisey, Gerard	Vortex Temporum	1995	Flutes, Clarinet, Violin, Viola, Cello, Piano		Ricordi Ircam
Cresswell, Lyell	The Belly of the Whale	1997	Unaccompanied Choral		Composer BMC, Sounz
Gonneville, Michel	Le cheminement de la baleine	1998	Solo Clarinet, Ondes Martenot, 18 instruments	Mn	Composer CMC

Perrin, Ross	La baleine qui court dans le ciel du Cambodge	2001	Gamelan, Child's Voice, Clarinet and Harp		CMC
Andel, Laura	Whale Singing	2002	Solo Voice and Large Ensemble		NMUSA
Cionek, Edmund	Whale-tronica	2002	Saxophone Quartet and Fixed Media		NMUSA
Laporte, Jean-François	Le Chant des baleines	2002	Electric Guitar		IRCAM
Perkinson, Coleridge-Taylor	Piano Trio	2002	Violin, Cello, Piano		NMUSA
Cowan, Claire	Whaler	2003	Solo Cello		Sounz
Ball, Derek	Celtic Whale	2006	Cláirseach, Fiddle [electronically modified]		CMCI
Ritchie, Anthony	Whalesong	2006	Solo Double Bass and Orchestra	Mn	Sounz
Whitehead, Gillian	Puhake ki te rangi [Spouting to the skies]	2006	String Quartet and Taonga Puoro		Sounz
Uyeda, Leslie	Whales	2006	Soprano, Flute/Bass Flute, Piano and Pre-recorded Whale Sounds	Dl, Oo (●)	The Avondale Press CMC
Hirs, Rozalie	Kleine walvis en het ijs (Little Whale and the Ice)	2007	Large Ensemble (12 players)		Donemus
Powell, Kit	Whale Song	2007	Solo Piano		Sounz
Bright, Colin	The Last Whale	2007	Vocal Sextet, Chamber Ensemble, CD recording	Mn (●)	AMC

Doolittle, Emily	Social Sounds from Whales at Night	2007	Soprano or Solo Instrument, with Percussion and Tape	Mn (●)	Composers Edition
Blom, Diana	The Whale's Song	2008	Cello, Piano	Mn	AMC
Chapman, Stephen	Minke Whale (No.2 from Due East)	2008	SATB Chorus	Ba	CMC
Shapiro, Alex	Below	2008	Contrabass Flute and Electronic Soundscape	Mn (●)	Activist Music / NMUSA
Gorbos, Stephen	On the Whiteness of the Whale	2009	Solo Bass Clarinet		Composer Sequenza21
Fujikura, Dai	Dolphins	2010	Two Violas (also Two Cellos or Two Violins)	T	Ricordi
Hextall, Pieta	Planet Vandal	2010	Solo Piano		Sounz
Kirke, Alexis	Fast Travel	2011	Tenor Saxophone, Interactive Electronics	Bmus	
Svetlichny, Anton	The Whale (Total Loneliness)	2011	Violin, Cello, Double Bass		Babelscores
Shapiro, Alex	Immersion	2011	Symphony for Winds, Percussion and pre-recorded Soundscape	Mn (●)	Activist Music
Pécou, Thierry	Rorqual	2013	strings, harp, percussions, pre-recorded tapes and instrumental group ad libitum	Mn, Bmus, Dl, Pm (●)	Schott Music

Blackford, Richard (recordings from Bernie Krause)	The Great Animal Orchestra, Symphony for Orchestra and Wild Soundscapes	2014	Orchestra and CD recording	Mn (●)	Nimubus Music Publishing
Bozone, Judy	The Man and the Whale	2016	Clarinet, Piano		ISCM Soundcloud
Powell, Kit	Whale Fantasy	2016	Solo Piano	Mn	Sounz
Carter, Tristan	Tohoraha	2017	String Quintet and Taonga Pūoro		Sounz
Sharp, Barry	Even Here	2017	Piano and any other combination of instruments	Mn	
Pécou, Thierry	Méditation sur la fin de l'espèce	2017	Solo Cello, Six Instruments, Pre-recorded Whale Song	Pm, Mn, Oo, DI (●)	Schott Music, Gaudeamus
Miller, Jared	Leviathan	2018	Flute, Clarinet, Violin, Cello, Piano, Percussion, Electronics	Bmus	ISCM Soundcloud
Lewandowski, Annie	Cetus: Life After Life	2018	Whale Song and Chimes [Carillon]	Mn (●)	Composer
Byrne, Luke	Whale Pavane	2019	Cello, Piano		AMC Composer
Doolittle, Emily	Bowheads	2019	Piano Trio	Bmys	Composer
Redhead, Lauren	The Whale	2019	Open notation and fixed media		Composer
Danon, Nina	The Caress of the Sea	2020	Alto Saxophone, Piano	Pm	Composer
Danon, Nina	Aquatic Dance	2020	Alto Saxophone, Piano		Composer

Sanders, Abigail	Solo Humpback Whale	2020	French horn	Mn	Composer
Sanders, Abigail	Glacier Bay Nov 2020	2021	Horns, percussion, electronics	Mn	Composer
Sanders, Abigail	Glacier Bay May 2020	2022	French horn	Mn	Composer

Sources for Table A1 include

- Australian Music Centre (AMC)
- British Music Collection (BMC)
- Canadian Music Centre / Centre de Musique Canadienne (CMC)
- Contemporary Music Centre Ireland (CMCI)
- International Society for Contemporary Music (ISCM)
- New Music USA (NMUSA)
- Scottish Music Centre (SMC)

Table A2 Works of music inspired by cetacean vocalizations, for which the principal output is a recording (includes Electroacoustic music). Works are listed in date order. Genres were assigned following discogs.com. The incorporation of field recordings is indicated with (●) after species.

Artist (Performer)	Title (Album)	Year	Genre	Species	Label / Website
Neil, Fred	The Dolphins (Fred Neil)	1966	Rock	T	Capitol Records
Collins, Judy	Farewell to Tarwathie (Whales and Nightingales)	1970	Folk song	Mn (●)	Elektra
Robertson, Harry (Alex Hood)	Ballina Whalers (Whale Chasing Men: Songs of Whaling in Ice and Sun)	1971	Folk	Mn	MFP records
McDonald, Joe (Country Joe McDonald)	Save the Whales! (Paradise With An Ocean View)	1975	Rock	Mn (●)	Fantasy
Toop, David	The Divination of the Bowhead Whale (New and Rediscovered Musical Instruments)	1975	Experimental	Bmys	Obscure
Crosby, David (Crosby & Nash)	To the Last Whale (The Two of Us)	1975	Rock		ABC Records
Wendt, Larry	In the Beginning was the Whale	1976	Electroacoustic		
Interspecies Music	Whalescapes	1977	Experimental	Dl, Bmus, Oo	Music Gallery Editions
Bush, Kate	Moving (The Kick Inside)	1978	Electronic, Rock	Mn (●)	Amiga
Winter, Paul	Ocean Dream (Common Ground)	1978	Jazz	Mn (●)	A&M Records
Nollman, Jim	Orca Reggae	1979	Interspecies	Oo (●)	interspecies.com
Haden, Charlie (Old and New Dreams)	Song for the Whales (Old and New Dreams)	1979	Jazz		ECM Records

Winter, Paul	Lullaby from the Great Mother Whale for the Baby Seal Pups (Callings)	1980	New Age	Mn (●)	Living Music
Morris, David	Whale Song	1980	Electroacoustic		
Nollman, Jim	Playing Music with Animals: The Interspecies Communication of Jim Nollman with 300 Turkeys, 12 Wolves, 20 Orca Whales	1982	Field recording, Experimental	Oo (●)	Folkways Records
Douglas Ewart and Inventions / Clarinet Choir	Whale Song (Red Hills)	1983	Jazz		Arawak
Barnes, Andy (Sheena Wellington)	The Last Leviathan (Kerelaw)	1983 (1986)	Folk		Dunkeld Records
Jaffrennou, Pierre-Alain	Deuxième passage de la baleine	1984	Electroacoustic		Ircam
Jaffrennou, Pierre-Alain	Océane ou Troisième passage de la baleine	1984	Electroacoustic		Ircam
Nollman, Jim	Orca's Greatest Hits	1985	Field recording, Experimental	Oo (●)	Interspecies Communication
Dean, Roger (Lysis)	Wings of the Whale (The wings of the whale - You-yangs)	1985	Jazz		SOMA Records AMC
McMillan, Ann	Whale - Wail, In Peace, En Paix: For Voice and Tape Structures of Whale and Other Animal Sounds	1986	Electroacoustic	Mn,T (●)	Folkways Records
Winter, Paul (Paul Winter / Paul Halley, Leonard Nimoy)	Whales Alive	1986	New Age	Mn (●)	Living Music
Cocteau Twins	Whale's Tales (Victorialand)	1986	Rock		4AD

Terry Oldfield	Mystical Deep (Reverence)	1986	Electronic - New Age	Mn (●)	New World Cassettes
Chapin, Tom	Sing A Whale Song (Moonboat)	1989	Folk	Mn (●)	Sony Kids' Music
Jethro Tull	Whaler's Dues (Rock Island)	1989	Rock		Chrysalis Records
Reed, Lou	Last Great American Whale (New York)	1989	Rock		Sire
Michel Berger	Le Paradis Blanc	1990	Pop	Mn (●)	Apache
Cosmic Baby	Cosmikk Trigger 4.0 (Cosmikk Trigger)	1992	Electronic - Techno	Mn (●)	Time Out Of Mind
Rheostatics, George Blondheim	Whale Music - Music from the Motion Picture	1994	Rock, Pop		Sire Records
Gluck, Bob	Jonah under the Sea (Stories Hear and Retold)	1997	Electroacoustic	(●)	Electronic Music Foundation
Lawrie, Bunna	Whale Dance Song; Sound of the Whale (Rhythm of Nature/ Coloured Stone)	1998	Indigenous		CAAMA
Douglas Ewart and Inventions / Clarinet Choir	Migration of Whales (Angles of Entrance)	1998	Jazz		Arawak
Anderson, Laurie	One White Whale (Life on a String)	2001	Electronic		Nonesuch
Walker, Lisa	Grooved Whale	2001	Electronic, Ambient	Mn (●)	Earth Ear
Tom Waits / Kathleen Brennan	Starving in the Belly of a Whale (Blood Money)	2002	Rock		ANTI
Various	Belly of the Whale	2006	Various	Oo (●)	Important Records
Tarshito & Murray Burns (Bunna Lawrie)	Nullarbour Journey (Whaledreamers)	2006	Film soundtrack		tinyurl.com/ y5845lmw

Fahres, Michael	Cetacea	2007	Electroacoustic with video	T (●)	Composer
Rothenberg, David	Whale Music	2008	Jazz	Dl, Mn, Oo (●)	Terranova Music
Robertson, Harry (Faustus)	Ballina Whalers (Faustus)	2008	Folk	Mn	Navigator Records
Seeger, Pete	Song of the World's Last Whale (At 89)	2008	Folk		Appleseed Recordings
Rothenberg, David (Various)	Whale Music Remixed	2009	Electronic	Dl, Mn, Oo (●)	Terranova Music
Glass Wave	Balena; Moby Dick (Glass Wave)	2010	Rock	Mn (●)	Independent
Dirty Projectors with Björk	Mount Wittenberg Orca	2011	Electronic, Pop	Oo	Domino
Kings of the South Seas	Kings of the South Seas	2014	Folk		D. Wink
Danon, Nina	Pearls (The Dragon and the Golden Flower)	2016	Classical	Mn	Bandcamp
Lawrie, Bunna	Jeedara (White Whale Song)	2016	Indigenous	Ea	tinyurl.com/yxttfkaq
Various Artists	Pod Tune	2016	Electronic	Mn	brightcolors.com/podtune
Downes, Kit	Last Leviathan (Obsidian)	2018	Contemporary classical		ECM Records Sequenza21
Claquer with Catriona McKay	Blackfish (Remembrance Species 2018)	2018	Experimental	Oo	Bandcamp
Maartn	The Limit (Outro) (Technical Language)	2018	Hip Hop	Mn (●)	Dirty Beauty
Inner Child (Various Artists)	Canticum Megapterae - Song of the Humpback Whale	2019	Electronic, Ambient	Mn (●)	Bandcamp
Fraser, Al with Sam Leamy, Neil Johnstone	Rorqual; Whale Time (Panthalassa)	2019	Experimental	Mn (●)	Bandcamp

McCormack, Andrew	Belly of the Whale (Graviton - The Calling)	2019	Jazz		Ubuntu Music
Depeche Mode	Walking In My Shoes - Ambient Whale Mix (Songs of Faith and Devotion | The 12" Singles)	2020	Electronic, Pop	Mn (●)	Columbia
Lewandowski, Annie and Kyle McDonald	Siren: Composers of the Sea	2021	Electroacoustic	Mn (●)	
Ridley, David	Cetacean Song	2021			Bandcamp
Inner Child (Various Artists)	Canticum Megapterae II - The Evolution	2022	Electronic, Ambient	Mn (●)	Bandcamp
Sheldrake, Cosmo	Bathed in Sound (Wild Wet World)	2023	Experimental, Ambient	Mn+ (●)	Bandcamp

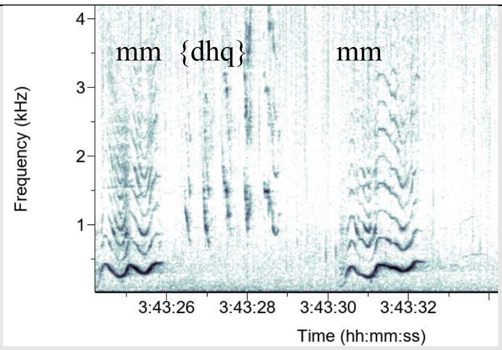
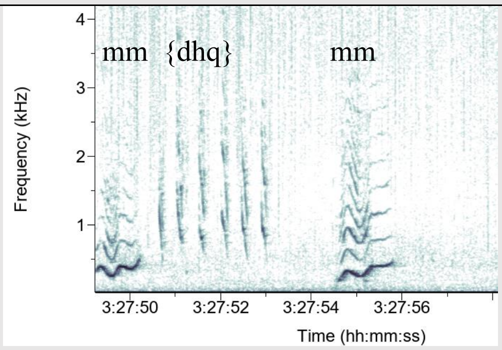
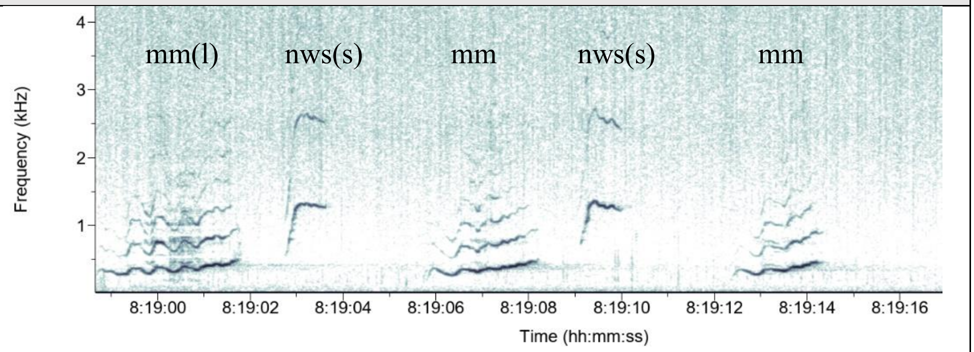
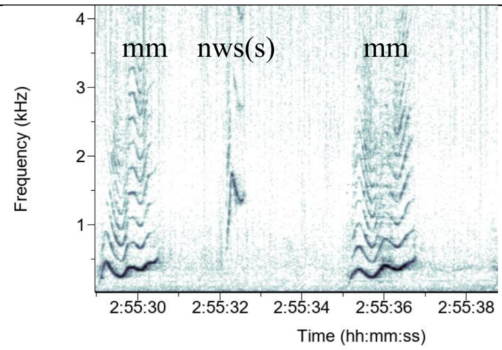
Appendix B Coded Song Spectrograms

Phrase variant (PV) codings (abbreviations in Table B1, following Garland *et al.* (2017)) and spectrograms of common PVs are provided below (Table B2).

Table B1 Song unit types and short codes. Groans, moans, cries, shrieks and whistles may be further qualified as long (l), short (s), and pulsative (pul). Hyphenated names indicate song units made up of multiple subunits.

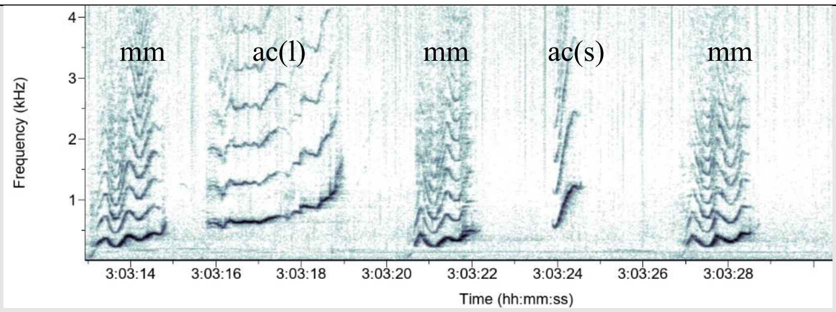
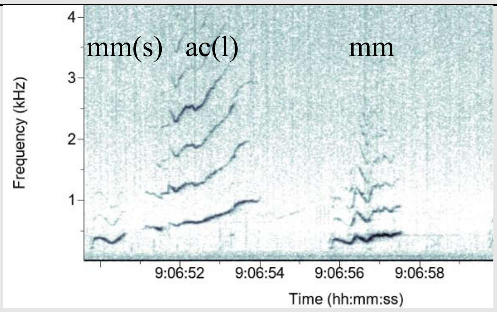
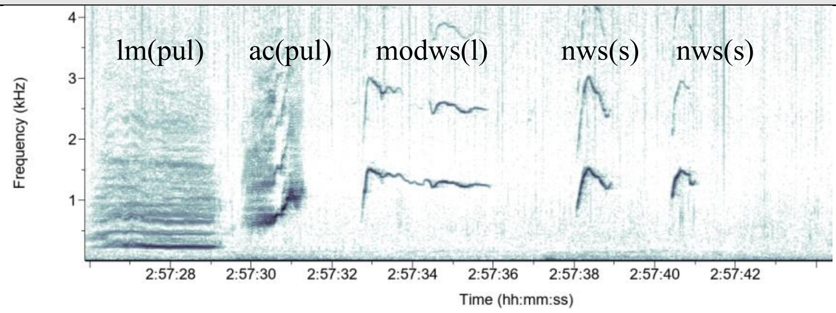
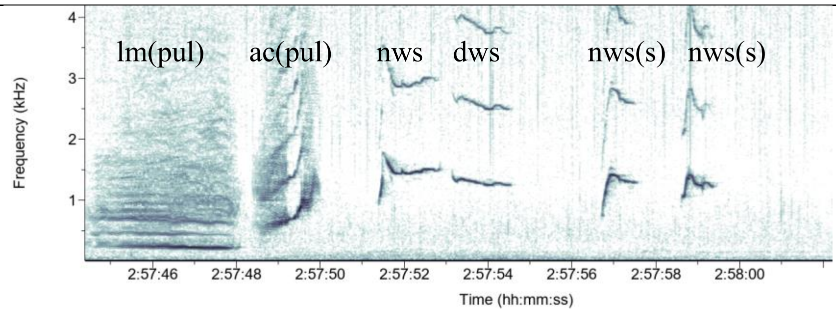
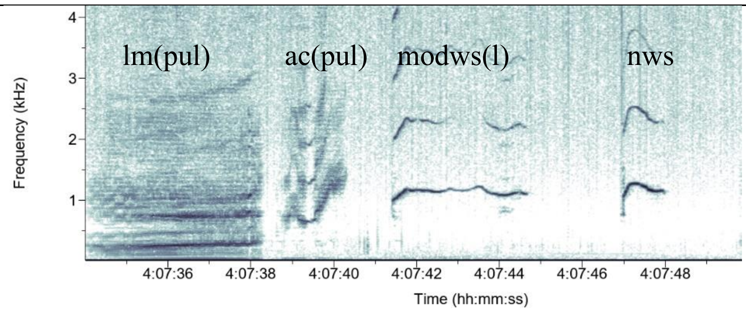
Class	**Name**	**Short Codes**
Groan	groan	gr
	ascending groan	agr
	descending groan	dgr
	n-shaped groan	ngr
	modulated groan	mgr
Moan	modulated moan	mm
	long moan	lm
Cry	ascending cry	ac
Shriek	high shriek	hs
	ascending shriek	as
	modulated high shriek	modhs
Whistle	ascending whistle	aws
	descending whistle	dws
	n-shaped whistle	nws
	modulated whistle	modws
Squeak	descending high squeak	dhq
	u-shaped high squeak	uhq
Pulsative	croak	c
	pulsative element	pe
	bark	ba
	long bark	lb
Other short	whoop	w

Table B2 Selected common PVs: typical spectrograms and codings. Themes are listed in normal order of occurrence. Spectrograms were generated in Raven Pro 1.6 (Hann window, 2048 samples, 75% overlap, 0.57 brightness and contrast). Unit name codes are grouped according to frequency.

PV	**Singer**	**Typical Spectrogram and Phrase Coding**
1Ba5	190924	
1Ba6	190924	
2Aa	190919	
2Ba	190924	

mm – modulated moan
mm(l) – long modulated moan

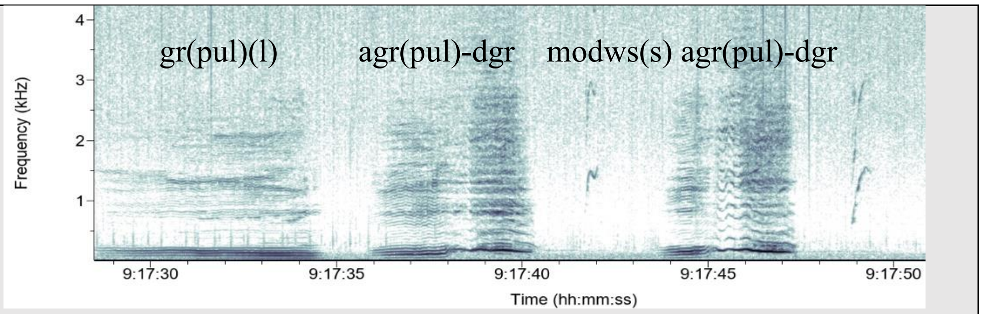
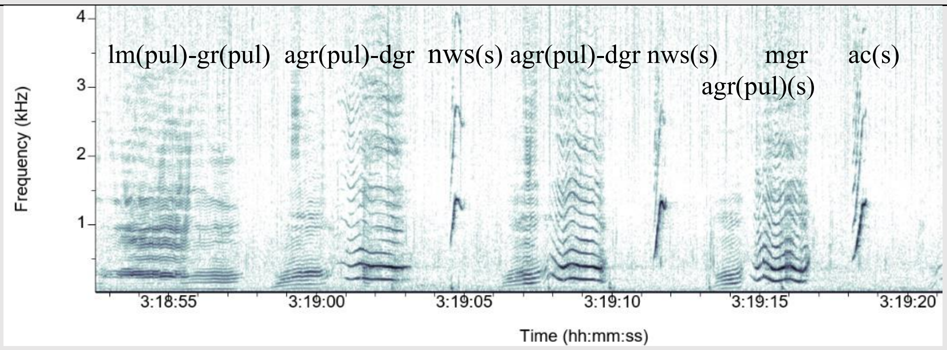
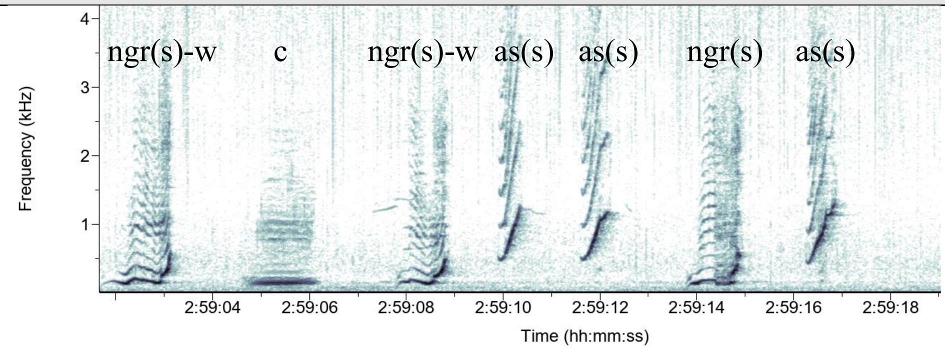
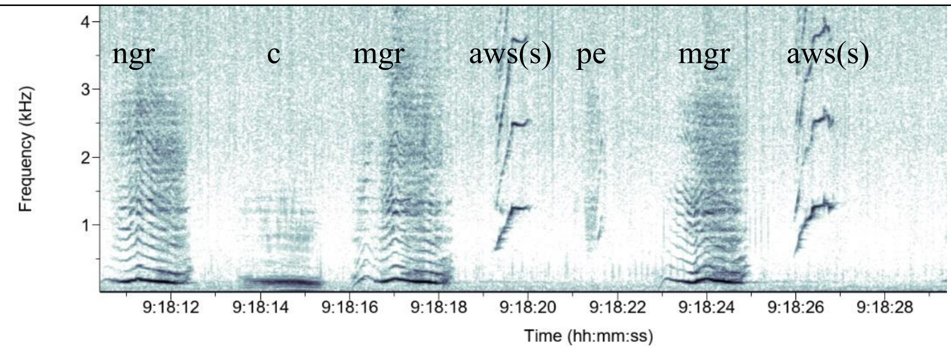
dhq – descending high squeak
nws(s) – short n-shaped whistle

7A°	190924	
7B°	190919	
5Aa	190924	
5Aa’	190924	
5Aa-	190925	

mm – modulated moan
mm(s) – short modulated moan
lm(pul) – pulsative long moan

ac(l) – long ascending cry
ac(s) – short ascending cry
ac(pul) – pulsative ascending cry

dws – descending whistle
modws(l) – long modulated whistle
nws – n-shaped whistle
nws(s) – short n-shaped whistle

6A°	190919	
6A'+	190924	
3A°	190924	
3A'	190919	

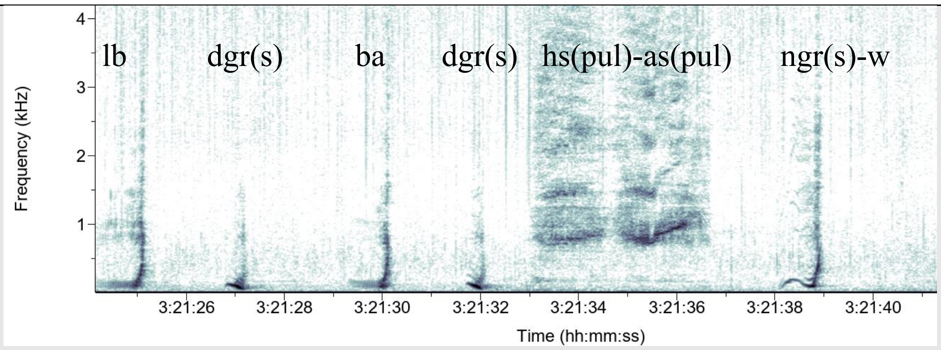
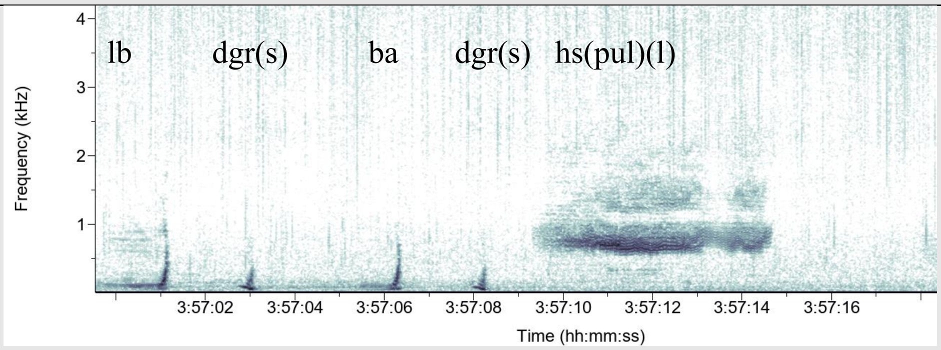
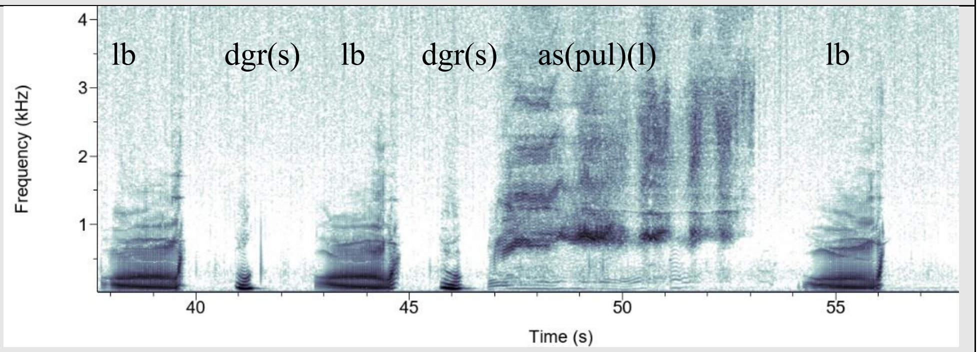
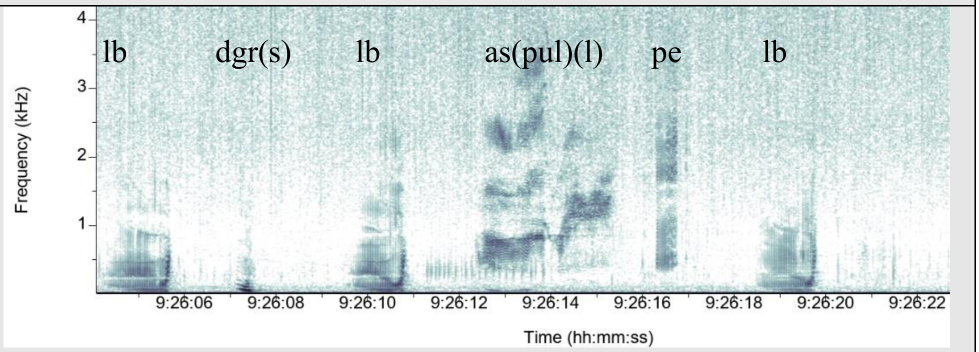
c – croak

agr(pul) – ascending pulsative groan
agr(pul)(s) – short ascending pulsative groan
dgr – descending groan
gr(pul) – pulsative groan
gr(pul)(l) – long pulsative groan
lm(pul) – pulsative long moan
mgr – modulated groan
ngr – n-shaped groan
ngr(s) – short n-shaped groan

ac(s) – short ascending cry
as(s) – short ascending shriek

pe – pulsative element
w – whoop

aws(s) – short ascending whistle
modws(s) – short modulated whistle
nws(s) – short n-shaped whistle
ws(s) – short whistle

4A°	190924	
4A-	190924	
4C°	191006	
4Da-4	190919	

ba – bark
lb – long bark

dgr(s) – short descending groan
ngr(s) – short n-shaped groan

as(pul) – pulsative ascending shriek
as(pul)(l) – long pulsative ascending shriek
hs(l) – pulsative high shriek
hs(pul)(l) – long pulsative high shriek

pe – pulsative element

Appendix C Programme Notes

‘(im)possible gestures’

In this work the four clarinettists stand outside the audience, facing inwards, creating an immersive acoustic environment for the listeners. They are playing a transcription of a humpback song recorded by Ellen Garland and M. Michael Poole off the north-west coast of Mo’orea (French Polynesia) in September 2015. As for a group of humpback whales, there is no synchronization of the song: individual players progress through the series of themes at their own speed and with individual variations. As the piece goes on, there is a gradual transition from the musical world of *Megaptera novaeangliae* to that of *Homo sapiens*, and a gradual convergence between players, in space both literal and harmonic. The title alludes to the difficulties involved in recreating humpback whale sounds, which display an extraordinary freedom of melodic contour and wide ranging timbral qualities.

‘The Path of the Unseen Whale’

In his book *Leviathan or, The Whale* (2008), Philip Hoare writes of the track left by the whale's flukes through the sea, “as still as a pond even in rough seas.” Both the early whalers and Inuit do not cross it, the latter out of respect, “seeing this *qaala* – ‘the path of the unseen whale’ – as the animal's mirror into our world, and our mirror into its own.” This piece makes use of a field recording of a humpback whale made in September 2015 by Ellen Garland (University of St Andrews) and M. Michael Poole (Marine Mammal Research Programme, Mo’orea, French Polynesia) off the island of Mo’orea in the South Pacific. The recording also contains the crackles and pops of multitudes of snapping shrimp. The cetacean soloist itself is, however, largely unheard, its voice excised from the recording and replaced with mimetic musical gestures sketched by the human performer and slowly transformed by electronics into an eerie clamour. The gaps on the field recording initially represent the still *qaala* on the surface of the ocean, but as they lengthen and fill with anthropogenic sound they mirror the reality of what has been happening in the deeps: the ever-increasing clattering, roaring and wailing of container ships, seismic mapping and military sonar. Such noise pollution is largely unreported by the media but causes chronic and acute harm to marine mammals and fish.

‘Underwater Rain (Farehau Humpback, 26.10.2019)’

Sounds of the sea by night: waves from above the surface, a tropical rainstorm from below, the crackle of snapping shrimp; somewhere a humpback whale begins to sing.

Listening suggestion (headphones are recommended)

In the first half, focus on the sounds of the coastal marine environment, attending to the rhythms of the waves, the rustling and sputtering of the rain. After the introductory bassline (drawn from the pitches and rhythms of the whale song), listen to the whale against its crackling chorus of snapping shrimp: can you follow the pattern of the song, describe the transition from one theme to the next, hear it as music? Phrases last around 50 seconds and are separated by several seconds of silence. You may want to use this silence to move from one breath pattern to another. In the second half of the piece the lowest sounds of the whale are accompanied by the sound of a bass clarinet. How do the sounds meld together, hold each other apart?

Compositional note

‘Underwater Rain’ was created in August 2020 for the University of St Andrews, to form part of a pack of mindfulness resources for students returning to a very different, post-pandemic, academic world. It showcases the particularly beautiful singing of a single humpback whale, recorded in October 2019 outside of the coral reef which surrounds the South Pacific island of Mo’orea. For this piece I selected two themes from the complete song and repeated them, adding bass clarinet accompaniment in the repetition. I transcribed pitches and rhythms from the first theme to create a sort of bass line (combining bass clarinet and piano recorded from the inside) to frame the whale song. To allow the fuller appreciation of its timbral and melodic detail, I slowed the humpback's song to half its original speed, lowering the pitch by an octave. I was especially struck by the expressiveness of the gradually descending high cries towards the end of the second theme. Electronic processing was kept to a minimum, primarily involving reverb and delays used to help create an immersive quality. The piece also includes my own field recordings of rain and waves, made above and below the surface of the South Pacific in Cook's Bay (Mo'orea).

'Entanglement' (Alex South with Katherine Wren)

This multi-stranded lament expresses our response to the growing threat posed to humpback whales worldwide by entanglement in fishing gear. It is based on the rhythmic structure and pitches transcribed from the song of a humpback whale recorded in October 2019 off the South Pacific island of Mo'orea, a whale belonging to the Endangered (IUCN Red List) Oceania subpopulation. This population is still recovering only slowly from intense Soviet whaling operations in the early 1960s.

'Whale, Bow, Echo' (Alex South with Sequoia)

'Whale, Bow, Echo' takes its form from three phrases sung by a humpback whale off the island of Mo'orea, French Polynesia, in September 2019 and translated into the musical gestures of violin and cello. Introduced by the crackling of snapping shrimp, the live instruments are placed in an imagined oceanic environment where their phrases reverberate and resonate, building up in waves determined by the slow tempo of the humpback's song.

'Submergence' (Jan Foote, realization of live part by Alex South)

AABBC (a phrase consisting of three types of sounds),

is heard / repeats (with variation) / repeats (with variation) /

DDEE (a phrase consisting of two different types of sounds),

is heard / repeats (with variation) / repeats (with variation) /

These words summarize the structure of Jan Foote's piece. They could easily describe part of the song of a humpback whale. Each of the complex sounds (**A**, **B**, etc.) in the piece portrays a humpback whale "song unit", without being based on any particular whale sound. Put together in a particular order, they make a "phrase". Repeats of a phrase make a "theme". The transition from the first theme to the second comes a little over halfway through the piece. Compared to the typical duration of humpback song, the two themes of 'Submergence' appear to have been put under an acoustic microscope, stretching them from two or three minutes into ten. A whale might have carried on repeating the material for ten hours.

'Submergence' was composed by Jan for Alex South and the Scottish Clarinet Quartet, originally intended as a work for four bass clarinets to be performed live with amplification and spatialization. Like so many other artistic plans, this one was modified after the onset of the COVID-19 pandemic, and 'Submergence' is performed today as a work for clarinet and fixed media. The pre-recorded material includes four bass clarinet parts pre-recorded by Alex, comprising mainly of multiphonics, that is, multi-pitched sounds achieved through the use of special fingerings. Rich assemblages of bass clarinet multiphonics represent humpback song units, slowly morphing as they repeat. In live performance the clarinet, in a line composed by Alex, mimetically shimmers through the acoustic texture, shadowing and animating it with colour trills, tremolos and further multiphonics.