9-2022

Experiencing Sonic Change: Acoustic Properties as Form- and Meter-Bearing Elements in Popular Music Vocals

Kristi D. Hardman
The Graduate Center, City University of New York

EXPERIENCING SONIC CHANGE: ACOUSTIC PROPERTIES AS FORM- AND METER-BEARING ELEMENTS IN POPULAR MUSIC VOCALS

by

KRISTI D. HARDMAN

A dissertation submitted to the Graduate Faculty in Music in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2022

Experiencing Sonic Change:

Acoustic Properties as Form- and Meter-Bearing Elements in Popular Music Vocals

by

Kristi D. Hardman

This manuscript has been read and accepted for the Graduate Faculty in Music in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 7, 2022 _____ Kofi Agawu
Date Chair of Examining Committee

July 7, 2022 _____ Norman Carey
Date Executive Officer

Supervisory Committee:

Johanna Devaney, Advisor

Mark Spicer, First Reader

William Rothstein

Kofi Agawu

THE CITY UNIVERSITY OF NEW YORK

Abstract

Experiencing Sonic Change:

Acoustic Properties as Form- and Meter-Bearing Elements in Popular Music Vocals

by

Kristi D. Hardman

Advisor: Johanna Devaney

This dissertation analyzes sound qualities in popular music vocals and examines the relationship between sound qualities and other musical structures, including beat, meter, and form. I use the term sound qualities to refer to those aspects of a sound that help us distinguish one sound from another. In this dissertation, I focus specifically on loudness, noisiness, brightness, and vowel quality. Listeners often attend to the vocals while singing along to their favorite popular song. Additionally, the voice is extremely malleable, which means that sound qualities regularly fluctuate in the vocals. This makes popular music vocals a good case study for considering the connections between changes in sound qualities and other musical structures. Through case studies of Tanya Tagaq's music and recent mainstream country music from the last ten years, I illustrate ways in which sound qualities influence our experience of beat, meter, and form in these two repertoires.

The analyses in this dissertation progress from small-scale to gradually larger-scale structures. Each chapter features slightly different approaches to analysis based on the differences between the music studied and the musical structures considered; however, the

general analytical approach developed throughout the dissertation emphasizes acoustic measurements and data visualization techniques. The acoustic measurements used to analyze sound qualities include RMS energy (related to loudness), periodicity (related to noisiness), spectral centroid (related to brightness), and formants (related to vowel quality). The acoustic measurements and data visualization techniques allow me to visualize and interpret changes in continuous data relating to loudness, noisiness, and brightness over time through a section of a song and over an entire piece. The analytical approaches I develop are informed by ethical approaches to analysis. I continuously critique my analytical approach and analytical assumptions in relation to the question of what an ethical analysis might look like, and I consider ways in which my approach to analysis can be more culturally appropriate and inclusive when analyzing not only Tagaq's music but also country music.

Acknowledgments

Reaching this level of university studies and completing this dissertation is due wholly to the support of my family and friends, and the guidance of teachers I have had along the way.

Thank you to Johanna Devaney, my dissertation supervisor and mentor, for regularly meeting with me, even when you were extremely busy. This document would have been very different without your guidance. You pushed me to think outside of the box with my analysis and taught me more than I ever imagined I would know about acoustic measurements and statistics. You helped me navigate academia and demonstrated what a good work/life balance looks like; I will take that with me throughout my career.

I also thank my committee members. Thank you to Mark Spicer for your support of my work over the course of my time at the Graduate Center and for sharing your copy-editing skills while thoroughly reading my dissertation. Thank you to William Rothstein for sharing your thoughts regarding the connection of poetic structure to musical meter, which highly influenced Chapter 5, and for pointing me to music theory literature that I would not have known about otherwise. Thank you to Kofi Agawu for your insightful questions and advice regarding the big picture aspects of my dissertation and for making me question the reasons surrounding how I write about topics.

This project has been greatly influenced by discussions that I have had with faculty at CUNY. Thank you to Poundie Burstein, Phil Ewell, Joseph Straus, and Jane Sugarman for your advice and support throughout my time at the Graduate Center: for taking the time to observe classes that I taught, for running professional development sessions that helped me understand

and navigate academia, and for holding paper previews during which I was able to test out many of the ideas found in this project in a safe space.

Thank you to Tonisha Alexander, Patrice Eaton, and Jackie Martelle for your support throughout my time at the Graduate Center, and for all that you do to keep the music department running smoothly. Tonisha, I will especially never forget how at ease you made me feel during the second exam; it was so appreciated. All that you three do made the music lounge a safe space for me within the busyness of NYC.

This work has also been greatly influenced by scholars that I have talked with over the years at various conferences and meetings. I am grateful to Janet Bourne and Antares Boyle for reading multiple drafts of most of these chapters. Your insightful comments and questions helped refine many of the ideas. Thank you to Charles Horton for being the first to suggest I apply for graduate studies in music theory and for your continued support of and interest in my work throughout the years. Thank you to Lawrence Shuster, Somangshu Mukherji, and Noé Dinnerstein for publishing material from this dissertation in their edited collection, *Trends in World Music Analysis: New Directions in World Music Analysis* (2022), and for your helpful suggestions during the editing process. Thank you to Taylor and Francis Group, LLC, for giving me permission to reproduce this material as part of Chapter 4 in this document.

Thank you to my colleagues and friends at the Graduate Center and NYC for making graduate school so enjoyable through our many conversations in the music lounge and at theory Wednesdays. Thank you to Michèle Duguay for being the best cohort mate I could have asked for, to Toru Momii for urging me to question my beliefs regarding a variety of topics, to Nathan Pell for sending me names of country songs that feature vocal backbeat, to Ruka Shironishi for our many kitchen talks about our projects and teaching, to Noel Torres-Rivera for sharing his

experiences with graduate school so freely and reassuring me that it would get easier while I was in my first year of the program, and to Xieyi (Abby) Z hang for our many conversations about research as we walked to a movie. Thank you also to Stephen Gomez-Peck, Tom Johnson, Erin Johnston, Megan Lavengood, Alex Martin, Tim Mastic, Demi Nicks, Simon Prosser, Molly Reid, Lina Tabak, and Alice Xue. You all influenced my work in some way and made my time at the Graduate Center a great experience.

Last, but certainly not least, thank you to my family. Thank you to my dad, Roger, for driving me to my music lessons, even when roads were terrible in the middle of winter, and patiently listening to me talk about my research; to my mom, Cathy, for helping me study for drop-the-needle music history tests and being my first editor; and to my brother, Scott, for confirming my suspicions about backbeat in recent country music vocals in one of our frequent music talks. Thank you to my in-laws, Ashok and Vaswati, for making it easier to focus on my dissertation in the last months by cooking fantastic meals and helping with cleaning. Thank you to my husband, Arko. Your unwavering support, especially in the last few months of working on this dissertation, has been a driving force in completing it. You make sure I never forget to enjoy life, including taking me to a Coldplay concert the weekend before my dissertation was due to my committee, and for that, I am so grateful.

This dissertation is dedicated to my grandparents, three of whom passed while working on it: Anna, Audrey, George, Helen, Marlean, Roger, and Sophia. While most are not here to celebrate this accomplishment, I have carried their encouragement and support with me as I worked to complete this degree. They always routed me on in whatever I chose to do, and that has meant the world to me.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The singing voice has the ability to draw our attention and guide us in our experiences of a song in multiple ways. The voice can guide our emotional reaction to a song, but it can also inform how we experience the structural elements of a song, such as the beat, meter, and form. I am arguing that people attend to, respond to, and engage with the changes in sound qualities and that these changes shape their experience of the structure of the music. "Sound qualities" is a term I use to refer to those aspects that help us distinguish between sounds, including loudness, noisiness, brightness, and vowel quality. I realized the importance of sound qualities to musical structure during two separate listening encounters, one while listening to Tanya Tagaq's *Retribution* (2016), and the other while listening to QX104 FM, a country music radio station in Winnipeg, Manitoba, Canada.

Tanya Tagaq (Inuk) is an improvisational, experimental vocalist and composer from Ikaluktutiak (Cambridge Bay), Nunavut, Canada, who regularly employs Inuit throat singing in her songs.[1] Many of the songs on *Retribution* have changes in sound qualities as the songs progress. For instance, "Nacreous" begins with fairly loud, noisy growling, but transitions to a very ethereal, high-pitched singing about midway through the song. Throughout the rest of the song, there is a back-and-forth interplay between growls and some screams with the softer, gentler melody. After listening to *Retribution*, I wondered whether Tagaq manipulates changes in sound qualities within individual songs as a type of formal device. I also wanted to figure out a

[1] Inuk is the singular of Inuit. Inuit are a group of Indigenous peoples from northern Canada, Greenland, and Alaska. It is convention at the time of writing this dissertation to name an Indigenous person's cultural group after their name in parentheses. The identity should be claimed by the person, however, when stylizing it in this manner. Tagaq is proud to be an Inuk and has spoken about this numerous times in interviews.

way to explain what I was hearing so I could share my experience with others, as I realized I did not have the analytical tools to articulate how I was hearing Tagaq's album at that time.

While listening to a radio station that plays contemporary country music (mostly music from the last ten years), I also noticed that I was focused more on sound qualities than pitch or even the specific rhythms used; however, I was attending to a different type of change in sound qualities compared to the way I was attending to sonic change in Tagaq's music. I was focusing on word changes in these country songs. Song after song after song, I noticed that the words in the lyrics that I was attending to most were often happening on beats 2 and 4. I realized again that I did not have the tools to describe or explain what I was hearing in these songs, but I wanted to find a way to do so.

These two repertoires, Tagaq's throat singing/experimental music and recent mainstream country music, may at first seem very disparate, but they have three important commonalities: 1) changes in sound qualities play a significant role in the composition, performance, and production of the music and how we experience it; 2) they are aurally transmitted (i.e., the recording, rather than a written score, is the primary source of the composition); and 3) they are not widely represented in music theory discourse. Because these repertoires feature regular changes in sound qualities, they can be used to study the ways in which sound qualities affect our experience of beat, meter, and form. Because they are aurally transmitted, listeners engage with the music through experiential means. They engage with the music by singing along, dancing, or tapping to the beat. I am positing that listeners also attend to changes in sound qualities while engaging with the music. Additionally, because they are aural styles of music, a psychoacoustic approach to analyzing timbre and loudness works well to illustrate the changes in sound qualities over time. Lastly, because these two repertoires are not widely represented in music theory, a

study of these repertoires provides a good opportunity to approach the study of sound qualities from a new perspective, and it contributes to the expansion of the analytical canon.

1.1 Analyzing Sound Qualities

This dissertation aims to answer the above questions by studying the ways in which sound qualities play a role in our experience of structural elements—specifically beat, meter, and form—in the music of Tanya Tagaq and recent mainstream country music from the last ten years. I define sound qualities as those aspects related to how we describe a sound, including timbre and loudness. In this dissertation, I focus on the specific sound qualities of loudness, noisiness, brightness, and vowel quality. I am arguing that changes in brightness, noisiness, and vowel quality, in conjunction with loudness, can affect the way that we experience beat, meter, and form in music. In this dissertation, I aim to contribute to the field of music theory in two ways: 1) I develop an approach to analyzing sound qualities in vocals throughout an excerpt or an entire song, and 2) I consider the intersection between sound qualities and other musical structures, including beat, meter, and form.

My approach to analyzing sound qualities is distinct from the way that most music theorists analyze timbre. In recent years, timbre has become an important topic of study for music theorists, especially popular music scholars (e.g., Blake 2012; Eidsheim 2019; Heidemann 2016; Lavengood 2020; Malawey 2020; Neal 2018; Peres 2016; Wallmark and Kendall 2018; Zagorski-Thomas 2018). One commonly employed methodological approach is spectrogram reading, but spectrograms can be difficult to accurately interpret. Spectrogram reading typically results in summative and qualitative descriptors of a sound across an entire piece rather than a dynamic, ever-changing interpretation of sound qualities. To investigate this intersection between sound qualities and musical structure, I instead focus on quantitative acoustic

measurements which provide continuous data regarding subtle changes in sound qualities throughout a piece of music.

Acoustic measurements are a quantitative method of extracting information from an audio signal. The ones used in this dissertation include root-mean square (RMS) energy (related to loudness), periodicity (related to noisiness), spectral centroid (related to brightness), and formant frequencies (related to the quality of vowels and voiced consonants). These quantitative acoustic measurements provide a link between the recorded audio and our psychoacoustic experience of sound. In this dissertation, I develop a new approach to understanding sound qualities in songs by Tanya Tagaq and recent mainstream country music, considering how loudness and timbre are form- and meter-bearing elements in these kinds of music.

As someone with no formal training in quantitative music research or statistics who, before embarking on the research for my dissertation, found a lot of the material published by quantitative music researchers to be intimidating, it is also a goal of mine to make this dissertation accessible to a wide range of readers. I hope that this dissertation is written in a way that is accessible to those with and without knowledge of the acoustic measurements that I use. I have chosen acoustic measurements which have specific, tangible connections to the way that musicians describe sound. Additionally, although I use acoustic measurements to visualize and confirm my hearing of sound qualities, my hearing takes priority. I relate the acoustic measurements to what I hear in the music as much as possible. I also detail the procedures that I used to create the graphs in this dissertation so that music theorists and analysts could replicate these techniques for themselves. In this way, I aim to demonstrate how acoustic measurements could be used for music analysis in a way that is useful to musicians and music theorists,

especially those who, like me, do not have much (or any) training in this style of analysis and are interested in analyzing individual pieces of music.

1.2 Structure of Dissertation and Chapter Outlines

Chapter 2 summarizes and critiques established techniques for analyzing timbre in music studies. It particularly considers two main ways of analyzing timbre: qualitative terms (Blake 2012; Wallmark and Kendall 2018) and spectrogram reading (Cogan 1984; Lavengood 2020; Peres 2016; Zagorski-Thomas). I also discuss some of the challenges that analyzing timbre presents and suggest reasons why my analytical approach, using acoustic measurements and graphs visualizing these measurements, might be able to help with these. To illustrate the differences in my approach compared to established methods of analyzing timbre, Chapter 2 introduces the techniques I use for analyzing sound qualities with some concrete, real-world examples. Lastly, in Chapter 2, I summarize the way that scholars have viewed the effect of sound qualities on meter and form, including work by Nicole Biamonte (2014), Matthew Butterfield (2010), and Fred Lerdahl (1987) for meter, and Asaf Peres (2016), Ciro Scotto (2017), and Mark Spicer (2004) for form. While pitch still dominates the field of music theory and affects how we analyze meter and form, these scholars have suggested ways in which timbre and loudness might contribute to our experience of meter and form. These suggestions have been mostly speculative studies thus far, and so I situate my contributions in this dissertation within this conversation.

In Chapter 3, I position myself as what I term an *armchair analyst*, someone who is an outsider to the music they analyze, as it pertains to my analysis of Tanya Tagaq's music, and I explain how this positionality relates to the choices that I have made regarding my analytical methodology, including the ways that my approach adopts Indigenous research methodologies.

In this chapter, I ground my analytical approaches in the work of leading scholars of Indigenous research methodologies, such as Margaret Kovach ([2010] 2021), Lester-Irabinna Rigney (1997), Linda Tuhiwai Smith (1999), and Shawn Wilson (2008). I also summarize ways in which Indigenous music scholars have incorporated Indigenous research methodologies in their work, including that by Dawn Avery (2014) and Dylan Robinson (2020). As part of this discussion, I establish and reflect on my listening positionality as a white, settler scholar and the ways that it affects the analyses in this dissertation.

The rest of the chapters are organized in such a way that in every successive chapter a larger-scale structure is considered in relation to various sound qualities. Chapter 4 examines the acoustic features of RMS energy, periodicity, and spectral centroid and how they can be used to describe sound, as well as the relationship of the sound qualities of loudness, noisiness, and brightness to beat structure. Chapter 5 considers how changes in vowel qualities through formant frequencies factor into my experience of meter. Chapter 6 focuses on the relationship of RMS energy, periodicity, and spectral centroid to the large-scale structure of form in a song's vocals. While acoustic measurements and a focus on vocals tie the dissertation chapters together, each of Chapters 4–6 features a particular approach relating to the specifics of how I analyze the interaction of sound qualities with my experience of beat, meter, and form, respectively. These are discussed in more detail below.

In the first part of Chapter 4, I present my analytical methodology, using the sounds that Tanya Tagaq uses while throat singing as an example. I situate her throat-singing sounds on continua of relative values for various sound qualities. The data derived from the acoustic measurements help illustrate differences in the sound qualities of her sounds, and, in turn, that allows for a nuanced comparison of them. These sounds are markedly different from each other

and provide a good introduction to my analytical approach. Although Chapters 5 and 6 develop slightly different analytical approaches to deal with the different repertoire in Chapter 5 and the larger-scale structure of form in Chapter 6, Chapter 4 introduces readers to the qualitative descriptors, quantitative acoustic measurements, and data visualization methods used for analyzing the songs found in subsequent chapters.

During the second part of Chapter 4, I present an analysis of beat structure in Tanya Tagaq's "Ilunikavi."[2] I compare the changes in data from the acoustic measurements for "Ilunikavi" to the beats indicated by Tagaq and myself in video recordings to gain an understanding of how the sound qualities affect our experiences of beat structure in this song. During this analysis, I also self-reflect and critique how and why I move the way that I do to music. To capture my experience of beat in the song, I recorded a video of myself moving to the performance.[3] I have combined this video with that of Tagaq singing "Ilunikavi" in a split-screen via iMovie. Next, I examine both Tagaq's movements while performing "Ilunikavi"—i.e., when and how she moves her hands, her head, or sways—and my own movements while listening to the song. I take our movements as indicators of where we are feeling the beat, but not as an indication of meter in the Western music theory sense (i.e., I do not assume any kind of grouping of the beats in the way that Tagaq and I move to the song). I then compare the acoustic measurements of RMS energy, periodicity, and spectral centroid from sound event to new sound event to see how our movements correspond to the changes in energy, noisiness, and brightness. I also make observations regarding the similarities and differences between the movements of

[2] In Chapters 4 and 6, I focus on a video recording of Tagaq's "Ilunikavi" because it features Tagaq singing alone and without accompaniment. In these chapters, I am interested in how vocals are experienced, so the song analyzed needs to feature only the voice.

[3] This video was recorded six months after initially watching the video recording of Tagaq's performance of "Ilunikavi" to ensure that my viewing of her performance would not cloud my movements.

Tagaq and myself. I ground my analyses in a discussion of the differences between Inuit ecological conceptions of time and Western mechanical conceptions of time, as well as how Western mechanical conceptions of time factor into notions of rhythm and meter in Western music theory.

In Chapter 5, I continue to use some aspects of the analytical approach established in Chapter 4, including the use of graphs to illustrate the acoustic measurements. This time, however, I use a different acoustic measurement, that of formant frequencies, when analyzing country songs. Because country singers mostly use one style of singing voice, the changes in sound qualities arise because of changes in vowel sounds, which are best measured using formants. In Chapter 5, I examine how changes in formant frequencies affect the experience of meter in recent mainstream country music vocals. More specifically, Chapter 5 explains how a backbeat can be heard in a vocal melody because of changes in formant structures of words in songs by Lil Nas X, Carrie Underwood, and Jason Aldean.[4] Typically, backbeat is considered a phenomenon of the instrumental parts of a pop song, but in recent mainstream country music, a noticeable, heavy backbeat appears in both the drums and the vocals because of changes in vowel quality. In many mainstream country music vocals from the last ten years, similar qualities of vowels and voiced consonants (as shown by formant data) in the lyrics emphasize timbral similarities on beats 2 and 4. This creates what I refer to as a *vocal backbeat* because phonetic accents and/or rhyming words appear on beats 2 and 4 of a 4/4 meter and these are most prominent and stressed within a song's melody, like a snare backbeat within a drum pattern. Rhyming words or assonances feature similar vowel sounds, and thus similar formant

[4] Because I focus solely on vocals in this chapter, the songs I have chosen to study in detail are ones for which I have been able to locate remix stems online at www.remixpacks.ru, or I have produced "vocals only" tracks via www.voxremover.com.

frequencies. Country singers will also often change the pronunciation of certain words to make them rhyme with other words, which changes the formant frequencies. This work contributes to current meter scholarship by expanding the parameters that are generally understood to contribute to our experience of rhythm and meter.

In Chapter 6, I return to the acoustic measurements of RMS energy, periodicity, and spectral centroid used in Chapter 4, but this time to consider the relationship between sound qualities and form. In this chapter, I examine changes in energy, noisiness, and brightness in Tagaq's "Ilunikavi," proposing ways in which these changes suggest formal interpretations of the song. By tracking changes in various sound qualities (energy, noisiness, and brightness), which I visualize using line graphs and moving average trendlines of the acoustic measurements of a large section of a song, I create multiple readings of the form of a single song, showing the complexities at work in Tagaq's music. When I take into consideration changes in sound qualities, I can experience the form of Tagaq's music in several ways that are different from the formal analysis that I experience when only pitch is considered. I also experiment with an alternative way of presenting my analysis through storytelling in Chapter 6 (inspired by Robinson 2020). Using the changes in acoustic measurements as a basis for the analysis of "Sivulivinivut," I present this analysis through a fictional story meant to represent what I hear and imagine as I listen to the song. This method of presentation is more open to interpretation than what is typically found in scholarship and aims to retain some of the affective state of the music itself. This chapter contributes to current scholarship not only by expanding the parameters that help determine formal aspects of a song, but also by showing multiple possible formal narratives in a single song, rather than the "one correct" analysis that has traditionally pervaded music theory scholarship.

In Chapter 7, I summarize the main findings and detail some of the technical and analytical limitations of my approach. I also outline future research avenues that build on the topics found in this dissertation but were ultimately outside of its scope. The future research avenues are related to two broad categories: sound qualities and repertoire. I end with a personal reflection on the writing of this dissertation, critiquing the steps I have taken toward more ethical approaches to analysis and contemplating the ways forward.

Chapter 2: The Intersection of Sound Qualities with the Experience of Beat, Meter, and Form

Throughout this dissertation, I am arguing that we experience music in time holistically, listening to the melody, the harmonies, the texture, the timbre, the form, and the rhythms and meter all at once. In my experience of listening to various popular music genres, changes in loudness and timbre often affect the way that I understand the form and meter of songs just as much as changes in pitch and harmony. Many scholars have argued that timbre is an important feature of popular music (Blake 2012; Eidsheim 2019; Heidemann 2016; Lavengood 2020; Malawey 2020; Neal 2018; Peres 2016; Wallmark and Kendall 2018; Zagorski-Thomas 2018; among others) but have not extensively brought timbre into dialogue with other features of the music, such as beat, meter, and form. This dissertation aims to make significant contributions to timbre research by working at the intersection of timbre research and theories of meter and form and by presenting some ways in which changes in sound qualities affect our experience of beat, meter, and form.

2.1 Timbre + Loudness = Sound Qualities

The American National Standards Institute ([1960] 1994, 35) defines timbre as "that attribute of auditory sensation which enables a listener to judge that two nonidentical sounds, similarly presented and having the same loudness and pitch, are dissimilar." In this dissertation, I use the term sound qualities to refer to timbre and loudness. I include not only timbre-related descriptors in my definition of sound qualities but also loudness because, while the ANSI definition of timbre is widely accepted as the standard, it negates real-world settings. It is rare that a song features different sounds with the same loudness and pitch. Throughout this dissertation, I am arguing that loudness equally affects the way that we identify sounds. So, we

must also consider loudness in a study of sound qualities with actual music to get a fuller impression of the sounds.

The sound qualities of acoustic instruments, especially the voice, are extremely malleable and are affected by a number of factors. It is valuable to look at all of these aspects together because listeners experience sound in a holistic manner; timbre and loudness can affect our perception of prominence, especially in the voice (Sluijter and van Heuven 1996; Streefkerk 1997; Tamburini 2003). As a clarinetist, I can play the same pitch with a variety of subtle timbral differences.[5] I can affect the timbre of a pitch by changing my embouchure and the dynamic level. When playing the lowest note on the clarinet (a concert D3), if I play the pitch with a loud dynamic, the pitch will have a buzzy, noisy sound, whereas if I play it in a very soft dynamic, I can make the pitch have an airy, whisper-like quality. Likewise, if I change my embouchure, loosening my lower jaw, I can create a squawking sound while playing the concert D3. Singers are also able to change the sound qualities of a vocal line through changes in dynamic and the manipulation of their vocal tract, including the position of their tongue, teeth, lips, etc., and whether the sound is voiced or unvoiced (Sundberg 1987). These changes in vocal quality can occur over the course of a single event—that is, a single sung note can have different sound qualities as it progresses from the onset to the sustained part to the eventual decay. It is, therefore, valuable to take into consideration timbral and loudness descriptors to get a more complete sense of how changes in sound qualities affect our understanding of beat, meter, and form.

[5] Slawson (1981) also discusses the difference in sound qualities of a clarinet, stating that register also plays a role in the sound quality of the instrument.

2.2 Approaches to Studying Timbre in Music Studies

Not long ago, scholars lamented the lack of timbre studies within music scholarship (Blake 2012; Lavengood 2017). Although this can no longer be stated given the wealth of timbre research that has been published in recent years, including two edited volumes of essays (Dolan and Rehding 2018; Fink et al. 2018), scholars still generally agree that timbre remains poorly understood and an area for study (Blake 2020; Fink et al. 2018; Lavengood 2020). While timbre is one of the most accessible aspects of music, it is also one of the most difficult to analyze (Cogan and Escot 1976; Fales 2002). As Cornelia Fales (2002, 57) states, "we hear it, we use it … but we have no language to describe it. With no domain-specific adjectives, timbre must be described in metaphor or by analogy to other senses." Timbre is something that we use regularly in our everyday lives to distinguish between sounds—for instance, a stranger's voice from a loved one's voice—but timbre is a multi-dimensional phenomenon that in scholarship is difficult to analyze.

Music scholars often engage with timbre through single, summary, qualitative terms (bright, dull, pure, noisy, rich, hollow, etc.) to describe their experience of listening to different sounds. These descriptions give us a tool for communicating our understanding of a sound and for distinguishing between sounds. David Blake (2012), for instance, uses four adjectives to describe sound in indie songs: full, distorted, digestible, and homogeneous. Zachary Wallmark and Roger Kendall (2018) take a cognitive linguistic approach to timbre analysis and examine the connections that listeners need to make to map timbral qualities onto verbal descriptors. They also provide a useful chart of qualitative terms used in timbre research dating back to Hermann von Helmholtz in 1863. Like Wallmark and Kendall, many music scholars have found the need to draw on other tools when using qualitative terms to bolster their analyses.

The next most common approach to analyzing timbre is probably the use of spectrograms in conjunction with these qualitative terms. Scholars use spectrograms to visualize timbre and qualitative terms to describe the spectrograms. One of the seminal writings on the use of spectrograms in the analysis of timbre is Robert Cogan's *New Images of Musical Sound* (1984), in which he developed opposition charts as a way of visually summarizing timbral qualities of sounds after interpreting the spectrograms. Modelling his opposition charts on those used in phonology (particularly Jakobson and Waugh 1979), Cogan suggested using thirteen antonym pairs, based on the qualitative descriptors. In Cogan's method, a sonic event is analyzed according to these thirteen antonyms based on information located in the spectrogram, and is given either a positive, negative, mixed, or neutral value. The score is then totalled at the bottom of the table, which is used to compare fluctuations in the quality of sound events throughout a piece of music. I provide a sample opposition chart comparing familiar sounds, a wooden flute and snare drum, in Table 2.1. The descriptors on the left of the slash in the first column are represented with a negative sign (-) in the subsequent columns, the descriptors on the right of the slash are represented with a positive sign (+) in the subsequent columns, and those sounds that exhibit both descriptors are assigned a mixed sign (-/+). The flute and snare drum are relatively high-pitched instruments. They can also be very soft or very loud. The distinguishing factors of these sounds are noisiness and brightness. A wooden flute is pure and dull sounding, especially if the pitch played is in the lower register, while a snare drum is noisy and bright sounding.[6] This approach is still in use today to analyze timbre, most notably by Megan Lavengood (2017; 2020), who adopts Cogan's approach, including his use of spectrograms and opposition charts,

[6] The timbre of an instrument depends on multiple factors, including the material that the instrument is made of (wood, silver, gold, etc.), the way that the sound is produced (through striking the instrument, blowing into the instrument, etc.), and register.

for her analysis of 1980s synthesizer pop music. Lavengood uses descriptors that are more widely used among musicians, producers, and composers, and considers the cultural implications of spectrogram reading.

Table 2.1 Timbral opposition chart for wooden flute and snare drum.

Opposition (-/+)	Wooden Flute	Snare Drum
Low/High	+	+
Soft/Loud	-/+	-/+
Noisy/Pure	+	-
Dull/Bright	-	+

Although I draw on Cogan's descriptors (as well as those used in Wallmark and Kendall 2018 and Lavengood 2020), I find the antonym pairs too limiting when used to analyze timbre in relation to beat, meter, and form and instead build on them, emphasizing comparatives and superlatives rather than binary oppositions. We understand the world through oppositions, but in addition to binaries, we also use comparatives (for instance, "the coffee is warmer than the tea" or "the snare drum is noisier than the flute") and superlatives ("the coffee is the warmest" or "the snare drum is the noisiest"). In this dissertation I develop a method of analyzing timbre that emphasizes comparatives and superlatives, which is necessary to understand the subtle differences in Tanya Tagaq's throat singing and country music vocals, as well as how subtle changes in sound qualities affect our experience of a song's beat, meter, and form.

Other scholars who use a mixture of spectrograms and qualitative terms include Denis Smalley, Asaf Peres, and Simon Zagorski-Thomas. Smalley (1997) establishes another approach to analyzing timbre through spectrograms that he refers to as spectromorphology, with which he analyzes frequency content over time in electroacoustic compositions. Peres (2016) and Zagorski-Thomas (2018) have adopted Smalley's method for studies of contemporary popular

music that is recorded. Peres (2016), for instance, traces changes in spectrograms to show how three major formal functions—setup, buildup, and peak—in contemporary popular music are governed by sonic density, relative brightness, and filter sweeps. Yet other scholars use spectrograms in a less systematic way to simply visually illustrate what they are hearing. Cornelia Fales (2002) uses spectrograms to visually illustrate timbre in such varied instruments as overtone singing, didjeridoo, Ghanaian balofon, and sitar. Victoria Malawey (2020) uses spectrograms to illustrate a variety of singing styles, such as rasp, breathiness, and nasality, which affect vocal quality in popular music.

Despite being one of the few ways of visualizing timbre, spectrograms have not been widely adopted in the study of timbre by music theorists and musicologists. Just three out of fifteen authors in *The Relentless Pursuit of Tone: Timbre in Popular Music* (2018) employ quantitative visualizations in their analysis of timbre (Neal, Wallmark and Kendall, and Zagorski-Thomas). Simon Zagorski-Thomas is the only one to use spectrograms, while Jocelyn Neal uses plots of fast Fourier transform; Zachary Wallmark and Roger Kendall use spectra graphs. Likewise, of the essays published in *The Oxford Handbook of Timbre* (2018), only five out of twenty-eight essays (those by Levin and Süzükei; Klotz; Hui; Goodchild and McAdams; and Hasegawa) show visual representations of the timbre of a sound, and few use spectrograms. Instead, many scholars continue to rely on intuitions, focusing on describing more subjective experiences of sound through qualitative descriptors alone.

An overreliance on qualitative methods in timbre studies can be problematic. Fales (2002) suggests that humans imperfectly identify sound sources and make associations with timbre based on the presumed source regardless of whether that perceived source is accurate. Our lack of ability to identify ambiguous sounds, such as a high singing voice that belongs to a

man rather than a woman (Eidsheim 2019), or the difference between "ba" and "da" (the McGurk effect), are prime examples of how our ears (and eyes) can be biased. This suggests that we cannot rely on our intuitions alone to correctly identify sound qualities or, at the very least, that a more objective means of analyzing timbre and loudness through computer-assisted methods might help us hear otherwise. I find it difficult to interpret spectrograms reliably, so in addition to qualitative descriptors, I use an alternative method to visualize sound qualities. I graph data from acoustic measurements—RMS energy, periodicity, spectral centroid, and formants—and employ descriptive statistics to interpret these data. The quantitative acoustic measurements employed in this dissertation are not widely used within the fields of music theory and musicology but are valuable because they provide a link between changes in sound qualities in recorded audio and our experience of the music in relation to beat, meter, and form. The quantitative approach taken here is more tangible than qualitative approaches alone and more reliable than reading data from spectrograms. It is a method with which others can come up with the same measurements and a similar analysis (though the interpretation of the data may still be different).

2.3 Sound Qualities and the Voice

Often scholars studying vocal quality approach their research through an interdisciplinary lens. Some voice scholars tend to concern themselves with the relationship of timbre to identity. For example, Bauer and Devaney (2022) posit that the voice is curated in audio in such a way that it influences our conception of gender identity. Nina Sun Eidsheim (2019) examines race in connection to timbre. Eidsheim critiques and makes known the biases in the way that people experience sound with regard to race. Jocelyn Neal (2018) theorizes the cultural and political implications of using twang in country music. Others have connected timbre to embodiment.

Kate Heidemann (2016) grounds her discussion of vocal timbre in the physiology of the voice, as well as the perception and embodiment of timbre, in which timbre is something that is not only heard but also felt. Zachary Wallmark (2014) argues that timbre is a mediator of musical embodiment and affect. Following this scholarship, I take an interdisciplinary approach to analyzing timbre, drawing on the physiology of the voice in Chapters 4 and 5, and Indigenous epistemologies in Chapters 3, 4, and 6. I base my analytical approach in topics that I had nearly no experience with prior to the writing of this dissertation: acoustic measurements, data visualization, and statistics.

In this dissertation, I analyze sound qualities of the voice through acoustic measurements. I show that this approach of analyzing sound qualities can be effectively adopted to analyze and describe sound qualities in popular music vocals; however, the specific type of acoustic measurement chosen is based on the types of singing style being studied, as I discuss in further detail in the next two sections.

2.3.1 Analyzing Vocal Sound Qualities Using Acoustic Measurements

For the chapters that focus on Tanya Tagaq's vocal style, I use three acoustic measurements—RMS energy, periodicity, and spectral centroid—to analyze changes in sound qualities.[7] *Root-mean square* (RMS) *energy* is a low-level feature that measures the energy in the signal. It is related to loudness, but it is not a strong indicator of the perceived loudness of a sound.[8] Figure 2.1 shows that Tagaq's Exhaling High sound (similar to the type of singing

[7] For more information on audio tools for music analysis see Devaney (in press). For more detail on these descriptors and how they are calculated, see Peeters et al. (2011).
[8] As Devaney (in press) states, "like pitch, loudness is a perceptual construct, one that is based on both amplitude and frequency information within the signal." RMS energy is a low-level acoustic measurement rather than a psychoacoustic measurement (loudness). A loudness measure such as phons would consider loudness curves (how the same amplitude has different perceived loudness at different frequencies).

featured in most popular music) has more energy than Exhaling Deep (a kind of growl), and Breath In is essentially silence with the data points hovering just above 0. *Periodicity* is a calculation of the time it takes to complete one cycle of a waveform pattern and an indicator of the noisiness of the sound (Mauch and Dixon 2014). Periodicity is calculated in tandem with the f0 estimation (a measurement used to determine the fundamental frequency of a sound). If a sound is pure, the periodicity value will be close to 1. Conversely, the closer the periodicity value is to 0, the noisier the sound is. In melodic singing, such as in country music vocals, the periodicity of a vocal sound will be nearly 1 consistently. In contrast, because of the percussive manner of Tagaq's throat singing, the periodicity of her vocals changes rather significantly from sound to sound and over an entire song. In Figure 2.1, Tagaq's Exhaling High sounds have a periodicity that is nearly a value of 1, while the data from Exhaling Deep sounds range from 0.3 to 0.9. The *spectral centroid* calculates the average of the spectral energy in a signal in Hertz, indicating where the bulk of the energy is located. Perceptually, spectral centroid is robustly connected to the brightness of a sound (Grey and Gordon 1978). The higher the spectral centroid, the brighter the sound. The Exhaling High sounds have a duller, smoother spectral centroid measurement that hovers around 1000 Hz compared to the spectral centroid data for the Exhaling Deep sounds that vary quite substantially and range from 1500–3000 Hz. Because Tagaq uses her voice in a percussive manner and drastically changes the sounds that she produces to create her songs, these three measurements provide a useful summary of the sound qualities of her singing. Additionally, because the measurements are continuous, they can be used to consider the ways in which changes in sound qualities over time shape our experience of beat and form in her music.

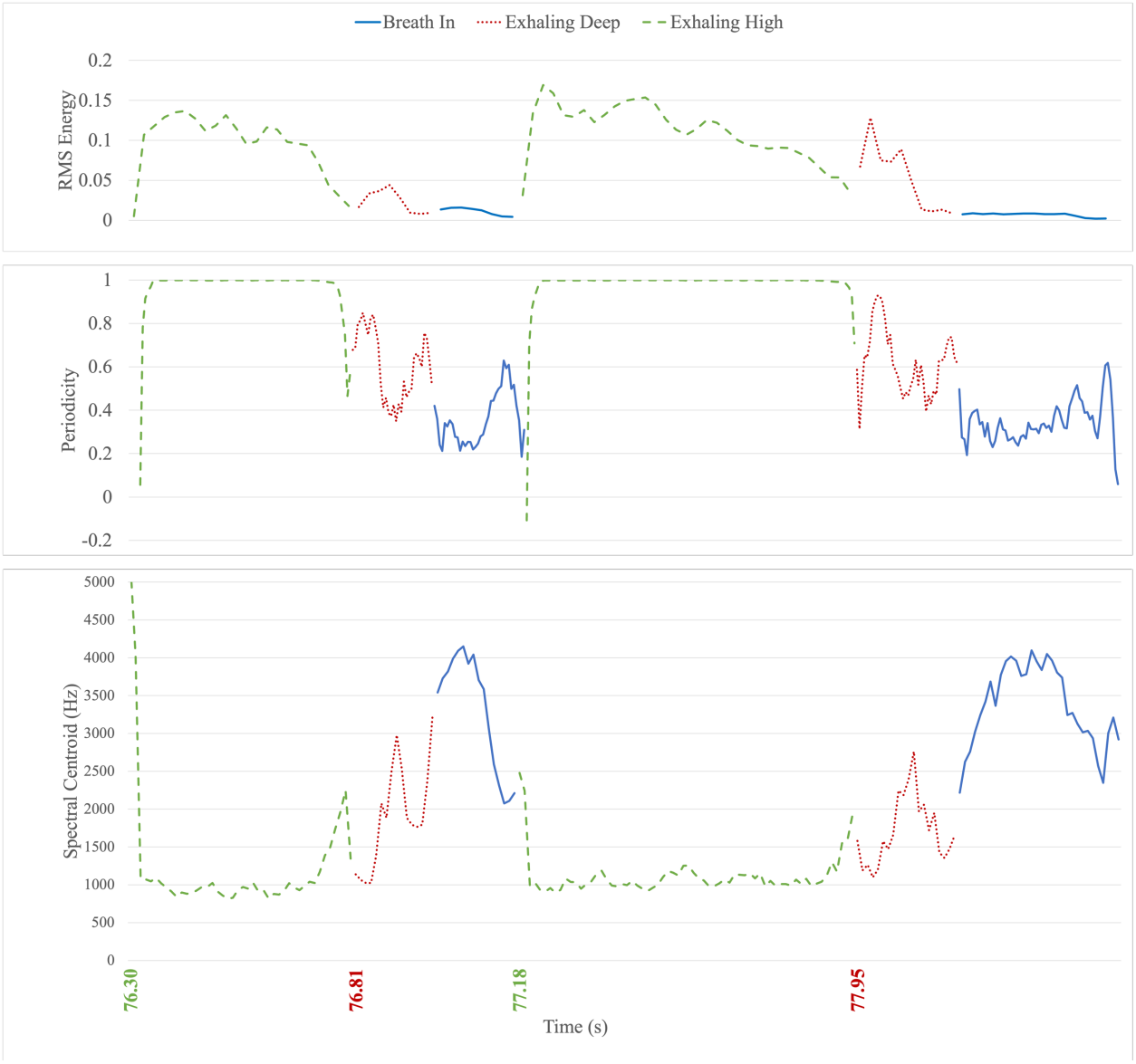


Figure 2.1 Acoustic measurements from a video performance of Tanya Tagaq's "Ilunikavi," showing differences between acoustic properties of two sounds that Tagaq uses in her throat singing. Top: RMS Energy; Middle: Periodicity; Bottom: Spectral Centroid.[9]

2.3.2 Analyzing Vocal Sound Qualities Using Formant Frequencies

For Chapter 5 featuring country music vocals, I interpret and compare the characteristics of formants. Formants are peaks in sound pressure levels at particular frequencies that help us

[9] To hear the version of "Ilunikavi" from her video series with the Open University, consult the following link: https://www.youtube.com/watch?v=vJlrZNrbMfc

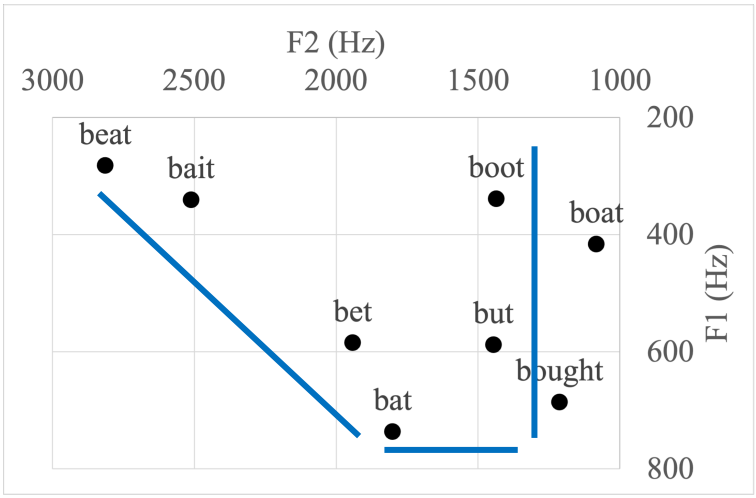
identify spoken or sung phonemes. In everyday speech, listeners attend to subtle changes in sound qualities, particularly formants, to determine the sounds that they hear and the prosodic structure. Vowels, particularly, are distinguished by formants as each vowel (and voiced consonant) sound has a unique formant structure. By changing the shape of your vocal tract, the shape of your lips and/or position of your tongue as you pronounce a word, you can change the formant frequencies of the sound produced. Thus, formants influence the sound quality of the voice. As Johan Sundberg states, the "vocal tract…gives each sound its final characteristic spectral shape and thus its timbral identity" (1999, 172). I examine formants in country music because changes in vowels are a feature in this style of music.

The lowest two formants (F1 and F2) determine the sound quality of vowels, the third formant (F3) affects our perception of voiced consonants, and higher formants indicate differences between the voice quality of individuals (Sundberg 1991, 124).[10] The frequency of the first formant is typically between 150 and 1200 Hz, the second between 800 and 3000 Hz, the third between 2000 and 3500 Hz, and the fourth is typically over 2500 Hz.[11] In this research, I focus on the first three formants because I am interested in how the words used in country lyrics affect our experience of meter in vocal melodies, rather than the differences between singers. Figure 2.2 shows a scatter plot of mean frequencies of F1 and F2 from a recording of myself speaking words that begin with the consonant "b" and end with the consonant "t." The vowels for each word are different, and as such, have different formant structures. The formant

[10] According to Sundberg (1991, 124), higher formants (F3, F4, and F5) do not vary as much and do not aid in the identification of vowels. The higher formants, on the other hand, contribute to an individual's vocal quality, or the overall sound of one's voice. The higher formants are the reason we can identify a loved one's voice from a stranger's or distinguish between favorite singers.

[11] Sundberg (1991, 122) has stated that "by means of articulation an adult male can vary the frequency of the first formant between approximately 150 and 900 Hz, the second between 500 and 3,000 Hz, and the third between 1,500 and 4,500 Hz." The numbers that I provide here are based on the maximum and minimum frequencies of formants from the songs analyzed in this dissertation.

values of these words correspond to the International Phonetic Alphabet vowel chart. The vowel chart indicates average differences in how the sounds are produced. For instance, the vowel sound in “beat” is produced with the lips almost closed and the tongue at the front of the mouth; while the vowel sound in “bought” is produced with open lips and the tongue at the back of the mouth. A sound produced with closed lips and the tongue at the front of the mouth will have a much lower F1 and higher F2 than a sound produced with open lips and the tongue at the back of the mouth. Figure 2.2 shows how formant frequencies correspond to the IPA vowel chart when words are spoken in a neutral manner. The formant values indicated on the IPA chart are generalized and will deviate slightly from speaker to speaker. For instance, my pronunciation of “beat” has an F2 that is quite a bit higher than the generalized highest frequency on the IPA chart. The important aspect to note, however, is that each vowel sound has its own unique formant structure.



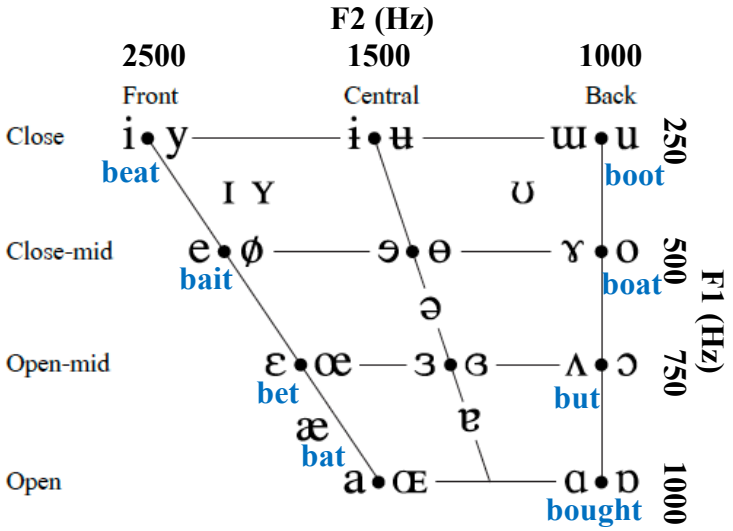


Figure 2.2 Top: Scatter plot of mean values of F1 and F2 of words spoken by me. Bottom: International Phonetic Alphabet vowel chart with the generalized corresponding F1 and F2 values of the vowels.

While a speaker can alter formants by manipulating their vocal tract when pronouncing words to make words sound more alike than they might otherwise be, they are less likely to do so than a singer. Country singers, in particular, quite often manipulate their pronunciation of words to make them sound like they rhyme. A blatant example of this happens in Billy Currington's "People Are Crazy" (2009). Currington makes "wars" rhyme with "divorced" by altering the

pronunciation of "wars" (1:08–1:15). Another example in the same song comes during the rhyme between "A.M." and "him" (1:59–2:10). Currington clearly alters the pronunciation of "him" to rhyme with "A.M." because his normal pronunciation of "him" is heard just seconds later in the last line of the bridge and is substantially different. When this happens, the formant structure of vowel sounds changes, and it deviates from the standard placement of the vowels on the IPA chart.

Singers will often alter the formant structure of words to make them more prominent. In Chapter 5, I observe that country singers use vowel quality and change formant structures to make certain words sound more alike, which in turn makes these moments prominent in the melody. This concept stems from work in perception, in which experiments have shown that singers change formant structures in particular ways so that words may be perceived as prominent (Sundberg 1999; Ammirante and Copelli 2019).

2.4 Sound Qualities as a Structure-Bearing Element

According to Leonard Meyer (1989), timbre is a "secondary parameter" in music theory and analysis. Meyer labeled pitch and rhythm as primary parameters because they were quantifiable and syntactic aspects of music, and he labeled timbre and dynamics as secondary parameters because they, in contrast, were non-syntactic and statistical. Despite Meyer's warning against placing value judgments on these terms, recently Emily Dolan (2013, 18 and 54) has suggested that by labeling instrumentation and tone-color as secondary parameters, timbre has taken a secondary, "inessential" role in musical works and, by extension, in scholarship. Questions of tone and timbre are often acknowledged by music theorists as important but are then set aside and largely ignored. Pitch centricity, which still dominates the field of music theory, leads to particular understandings of beat, meter, and form. Secondary parameters, as

Leonard Meyer suggests, can "play important roles, reinforcing or undermining the processes generated by the primary parameters" (1989, 16). Using acoustic measurements as a way to quantify and visualize loudness and timbre, this dissertation considers the ways in which sound qualities affect our experiences of small- and large-scale structures of a song and lead to alternative understandings of these structures in popular music vocals. This analytical approach could have implications for the study of other musics, including Western art music.

2.4.1 Sound Qualities, Beat, and Meter

Traditional theories of beat and meter do not discuss how sound qualities affect our experience of meter. In the three most oft-cited books on meter—those by Fred Lerdahl and Ray Jackendoff (1983), Christopher Hasty (1997), and Justin London (2004)—there are only brief mentions of timbre. Hasty distinguishes between timbre and rhythm, saying that timbre, pitch, and volume may remain relatively fixed after the attack with no becoming, whereas duration cannot remain fixed and cannot be determinate (1997, 93). Hasty speaks of rhythm and timbre as separate entities, ones that do not influence each other. In this dissertation, I show through case studies in Chapters 4 and 5 that timbre is not fixed, and it influences our experience of beat and meter. While I disagree with Hasty's position regarding timbre's influence on meter, I find Hasty's method of annotation useful for illustrating my understanding of meter in country songs, so I adopt and adapt his analytical symbols in Chapter 5. In contrast to Hasty, London indicates that timbral differences may help to distinguish different rhythmic layers, which can help in perceiving metric dissonance, but he does not go into detail as to which timbral qualities affect our understanding of meter and how (2004, 80). In Chapter 5, I build on London's theory that timbre plays a role in perceiving metric dissonance, suggesting that the rhythmic layer in the voice and drums is different from that found in the guitars and keyboards. Decades after its

publication, Lerdahl and Jackendoff's *A Generative Theory of Tonal Music* continues to inform new work on rhythm and meter, yet the book also does not give much attention to timbre. Lerdahl and Jackendoff (1983, 17) classify a change in timbre as a phenomenal accent—an "event at the musical surface that gives emphasis or stress to a moment in the musical flow," which also includes "attack points of pitch-events, [and] local stresses such as sforzandi, sudden changes in dynamics, long notes, leaps to relatively high or low notes, [and] harmonic changes." While they postulate that beats in a metrical structure will likely coincide with attack points of pitch-events (MPR 3), stress (MPR 4), and long duration (MPR 5), they do not identify metric preference rules for changes in timbre. Additionally, in these theories, beat 1 or an early event is typically considered strong (MPR 2 in Lerdahl and Jackendoff 1983) because the aspects considered reinforce the downbeat of a measure. In my opinion, this is a result of not looking at timbre in relation to meter. In Chapters 4 and 5, I provide examples where the first beat is not necessarily strong in certain contexts when timbre is considered a determining factor in understanding rhythm and meter.

A few years after *A Generative Theory of Tonal Music* was published, Lerdahl (1987) published an article specifically addressing timbre. In the article, he posits a "timbral prolongational structure" in which he establishes a hierarchical structure of timbre with consonance and dissonance that is similar to pitch-related consonance and dissonance. He states that timbres described as more tense are dissonant, such as a bright sound or that with a sharp attack. In the prolongational trees presented in this article, Lerdahl suggests that timbral consonance is stable/strong and dissonance is unstable/weak. It can be inferred from this that timbral consonances help a listener to determine strong beats within the musical meter. Lerdahl, however, does not make this correlation explicit, nor does he examine actual music in this article,

instead suggesting this as an area of possible future research. I draw on this theory in Chapter 4 with my analysis of beat structure in Tagaq's "Ilunikavi" to explain the way in which I experience meter in relation to the changing qualities of Tagaq's vocal sounds. There are significant correlations between Lerdahl's theory and my understanding of beat. I most often marked a beat with timbrally consonant sounds that are loud, pure, and dull. Based on my study of Tagaq's "Ilunikavi," I do not, however, agree with Lerdahl that timbrally consonant sounds are preferably strong; in fact, the way that Tagaq moves to "Ilunikavi" contradicts this theory.

Although music theorists have yet to base analyses of rhythm and meter on timbral changes, some scholars studying popular musics have noted timbral changes as a reason for experiencing a rhythmic pattern in a particular way. Matthew Butterfield (2010, par. 16, emphasis mine) suggests that, in a performance of Miles Davis' "Tune Up" by drummer Steve Davis, "the salient *timbre* of the snare drum generates a distinct rhythmic layer that stands out against the jazz waltz figure played on the ride cymbal." Butterfield explains that this generates a grouping dissonance between the rhythmic layers because the snare is displaced from the 4/4 meter and the ride cymbal. In an earlier article from 2006, while discussing a basic two-beat rock drumbeat, Butterfield also states that "the *timbre* of the snare, its higher frequency, and apparent brevity relative to the booming bass drum, causes it to be heard as anacrusis; thus beat 4, like beat 2, is directed toward the emergence of a new beginning" (2006, par. 40, emphasis mine). For Butterfield, in these particular statements, timbre is simple and one-dimensional. What is it about the timbre of the snare, which can include how noisy, bright, full, and rich it is in comparison to the bass drum, that causes it to be heard as anacrusis? Nicole Biamonte (2014, par. 6.1, emphasis mine) agrees with this statement from Butterfield, adding that "the snare drum is higher in pitch, has a sharper *timbre*, and … is louder than the bass drum—thus it is registrally,

timbrally, and dynamically more prominent." She goes on to say that "this creates a clear distinction within the pattern between odd-numbered and even-numbered beats, and places several types of phenomenal accent on what are traditionally the weak beats in a 4/4 pattern" (2014, par. 6.1). Unlike Butterfield, Biamonte understands the snare as "reactive rebounds" that are continuational rather than anacrustic, but like Butterfield, Biamonte similarly refers to the "timbre" of the snare drum, not going into further detail regarding what aspects of the timbre cause it to be heard as continuational. Mark Butler (2006, 93) suggests that phenomenal accents in electronic dance music are created by timbre, and that they draw attention to textural and metrical layers. Butler also suggests that sudden timbral changes can be rhythmically significant, but does not go into further detail (2006, 228). I draw on much of this literature when discussing the timbre of a standard rock drum pattern with backbeat in Chapter 5 to set up my explanation of what I term vocal backbeat. These studies, however, are rather cursory in terms of the influence of timbre on our experience of meter. To be clear, timbre was not the main area of interest in any of this scholarship, and so there is much room to expand on these studies. I contribute to these discussions through a detailed study of beat in Tagaq's "Ilunikavi" in Chapter 4 and meter in recent country songs in Chapter 5. Additionally, while these studies assume that the qualities of a sound remain consistent throughout its duration, the acoustic measurements that I use throughout this dissertation show that sound qualities can change over time.

While few theorists have explicitly discussed the influence of vowel quality on musical meter, many have considered the influence of prosodic stresses on meter. Raymond Monelle (1989) and Fred Lerdahl (2006) experiment with setting poetry to music. Both maintain that the nuclear stress in a phrase should fall on a metrically strong beat. Yonatan Malin (2010) makes connections between poetic rhythm and musical rhythm and meter in German Lieder and

explains the metric conflict that sometimes arises due to the joining of these two types of rhythm. William Rothstein (2011), in his examination of Verdi's operas, argues that national poetic traditions should affect the way that we hear and count musical meter, and this requires different metrical strategies. Rothstein (2022) notably suggests that for some music, such as nineteenth-century Italian opera, we should prefer a metrical structure in which the strongest metrical beat aligns with the strongest syllable in the poetry (that of the *accento comune* in Italian poetry). Because the *accento comune* is found at the end of a poetic line, the strongest beat in Italian opera typically happens near the end of a melodic group. Kyle Adams (2009) examines the metrical placement of rhyming syllables and accented syllables in rap music, noting that while some rappers place their lyrical stresses in predictable locations, others place them unpredictably. Nancy Murphy (2022) examines metric flexibility in 1960s and 1970s singer-songwriter music, where a regular metric pulse is not present but stresses in the singing give rise to a metric hierarchy. These scholars all indicate the importance of considering lyrical stresses when analyzing musical meter in texted music because lyrical stresses have an impact on how we hear/experience meter. I build on these theories by giving preference to the stresses in the lyrics of country songs in Chapter 5. I contribute to these theories by showing that vowel quality, in addition to prosodic structure, influences experiences of musical meter in country vocals through an examination of formants.

2.4.2 Sound Qualities and Form

In theories of form in popular music, most analysts focus on melody, chord changes, lyrics, and sometimes texture (Covach 2005; Summach 2011; de Clerq 2017), rather than changes in timbre. They also prioritize the identification of boundaries between sections. John Covach (2005) outlines the primary formal types used in rock music, which are largely based on

repetitions and contrasts in harmonic patterns and lyrics. While Jay Summach (2011) mentions the possibility of timbral changes as a momentum-building device in prechoruses in popular music, nothing else is said on this topic in the article. Instead, Summach argues that prechoruses developed out of an expansion of the Statement, Restatement, Departure, and Conclusion (SRDC) format (a phrase structure first discussed by Walter Everett in his 1999 book *The Beatles as Musicians:* Revolver *through* Anthology) and its convergence with verse-chorus and strophic forms. He also details the ways in which lyrics, instrumentation, and harmonic progressions contribute to the establishment of a prechorus. Trevor de Clerq (2017) prioritizes lyrics, melody, and harmonic progressions in his analysis of pop-rock music, but he suggests that there is often ambiguity in the boundaries of sections and/or labeling of these sections.

On the other hand, Stephen McAdams (1989) has long suggested that timbre is a form-bearing element; notably, a few scholars and composers have begun to consider the connection of sound quality changes to our experience of form in music. Western art music composers of the twentieth century, such as Kaija Saariaho (1987), use timbral changes as a way of creating formal structures in her compositions. Saariaho writes about playing with the differences between pure and noisy sounds to create feelings of dissonance and consonance throughout her works. Mark Spicer (2004) was one of the first scholars to discuss the role of recording technology on the large-scale formal structures of popular music, suggesting that multitracking allows for thicker textures to be built as a song progresses. While texture is different than timbre, it is related in the sense that usually new textures bring new sound qualities. Brad Osborn (2013) also suggests that in certain types of popular music there is a syntactical climax, a kind of hyperchorus, in which there might be a textural or timbral change, particularly the addition of new vocals. He also notes that in rock music, there is often a shift from clean guitar sounds in the

verses to distorted guitar sounds in the choruses. Likewise, Ciro Scotto (2017) studies the role of distortion in the form of hard rock and heavy metal songs. Scotto suggests that the use of distortion in hard rock and heavy metal morphs the traditional verse-chorus form into a "progressive directional form" in which the sections of a song progressively become more and more distorted until the end. In his 2016 dissertation, "The Sonic Dimension as Dramatic Driver in 21st-Century Pop Music," Asaf Peres argues that recent pop music is governed by timbre, texture, and spatialization. He proposes that popular songs contain three major formal functions—setup, buildup, and peak—based on his observations of spectrograms, and observes that these three formal functions are governed by changes in sonic density, relative brightness, and filter sweeps. In Chapter 6, I build on these scholars' work by considering ways to understand form based on changes in energy, noisiness, and brightness in two songs by Tanya Tagaq. I focus specifically on the vocals without any additional instruments, in contrast to these scholars who looked at the mix as a whole. In contrast to much of the literature discussed here, which assigns a single timbral quality to a section and compares it to other sections, I aim to show how sound qualities can change throughout a section of a song or an entire song through my analyses of form in Tagaq's "Ilunikavi" and "Sivulivinivut," respectively.

2.5 Summary

Although many scholars recognize the importance of sound qualities in a variety of popular genres, we continue to have difficulty when formally analyzing and describing timbre; possibly because of this, timbre has not been put into dialogue with form and meter as much as pitch has. David Blake (2012, par. 2.1) states that "timbre is especially frustrating for analytic description, at once the most apparent and least systematizable musical parameter." It is the most apparent because we all experience it as we listen; it is the way that we distinguish a clarinet

from a violin, the whistle of a kettle from the whistle of the wind, and a stranger's voice from that of a loved one. Timbre is also, however, the least systematizable based on the methods of analysis that are currently available; this was the reason that Leonard Meyer deemed it a secondary parameter. My intuitions, and the acknowledgements by scholars that timbre is an integral part of popular music, led me to search for a way to analyze this repertoire that allows connections to be made between various sound qualities and form, as well as rhythm and meter. In this dissertation, I establish an approach using acoustic measurements for analyzing sound qualities that makes this possible. In the next chapter, I discuss the ethics involved in creating a new analytical approach and explain my positionality as it relates to the music featured in this dissertation.

Chapter 3: Considering Ethics of Analysis as an Armchair Analyst

The primary focus of my research is the influence of sound qualities on the experience of musical structure, particularly in relation to beat, meter, and form. As summarized in Figure 3.1, this analytical focus on sound qualities contributes to two other large-scale goals, which include expanding the analytical canon and diversifying analytical approaches. I focus on popular music, particularly vocals in popular repertoires that have largely been understudied in music theory. As mentioned in previous chapters, I specifically focus on music by Tanya Tagaq (Inuk) and recent mainstream country music in this dissertation. For my analytical approach, I visualize the results of acoustic measurements using graphs. Through interpretation of these graphs, I draw connections between various literature on timbre and theories of meter and form that are based in Western popular music and art music.

These broad goals of expanding the analytical canon and diversifying analytical approaches go hand in hand while researching sound qualities, which is why they are connected with a line in Figure 3.1. When expanding the analytical canon, there is a need to also diversify analytical approaches because established analytical techniques will not necessarily work for the new music that is being analyzed. Most established music theories are rooted in Western art music and Western popular music, and they cannot necessarily be applied to other styles of music effectively. When these theories are applied to music outside of Western art and popular musics, sometimes important aspects of the music get lost or the theories themselves require adaptation (Clarke 2017; Temperley 2000). Even if one is studying Western art music, one might also consider diversifying analytical approaches to gain alternative perspectives on music within the existing canon. While the study of sound qualities in and of themselves does not necessitate a consideration of the ethics associated with analysis, when expanding the canon and developing

appropriate analytical approaches for the repertoire at hand, it is imperative to consider the ethics involved in one's choices.

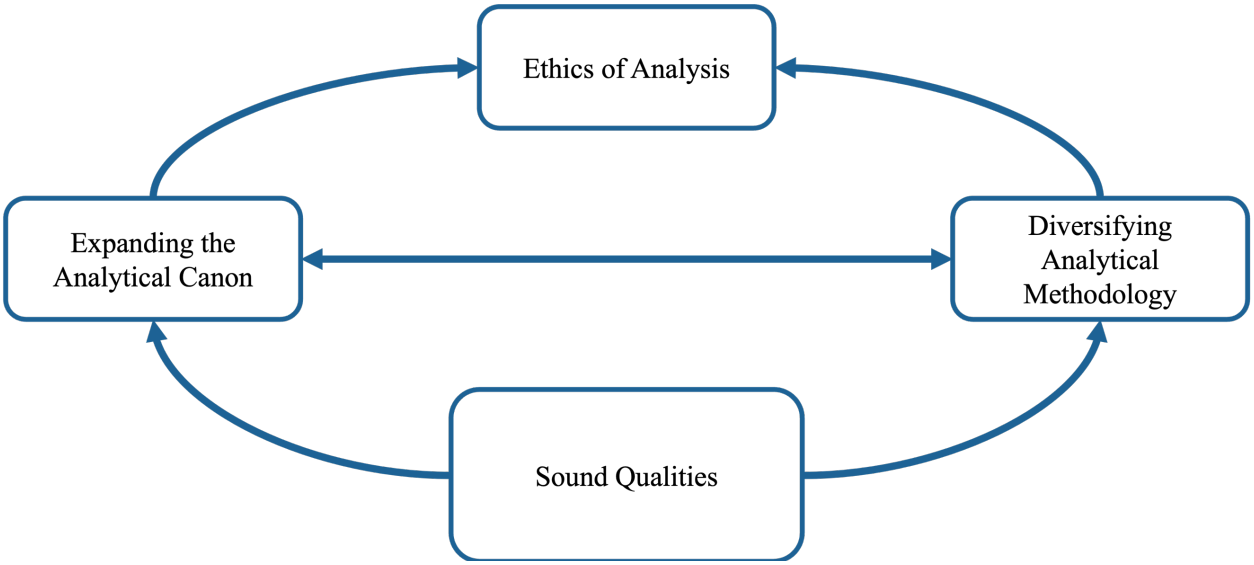


Figure 3.1 The broad goals of my research. Each informs the others.

There has been a move to consider ethics in music scholarship in recent years.[12] I began contemplating my ethics of analysis while working on a talk for a 2020 Society for Ethnomusicology roundtable titled "Decolonization and Analysis." During that time, I started reading Dylan Robinson's *Hungry Listening* (2020), which greatly inspired me to reflect on how I approach my research and seek alternative research methodologies for writing about Indigenous music. This is particularly important since I am writing about Tagaq's music, which features Inuit throat singing. In this chapter, I describe my role as an analyst, one that will likely be familiar to most music theorists reading this dissertation. Later, I explain steps that I have taken to uncolonize my analyses by drawing on Indigenous research methodologies.[13] Finally, I

[12] See for example Agawu (2003), Attas (2019), Ewell (2020), Luong (2019), and Robinson (2020).
[13] I use the term uncolonize rather than decolonize because decolonize needs to be used in a specific context. As Tuck and Yang (2012) explain, "decolonization is not a metaphor." It is a return of Indigenous land and sovereignty to Indigenous peoples and a removal of settler society from these lands. To use the term for anything else is to make

identify and critique my listening positionality, explaining ways in which it biases what follows in this dissertation.

3.1 Armchair Analyst

Here's a scene that, in my opinion, is fairly common in music theory:

> *An armchair analyst opens their computer, goes to Spotify (or whatever streaming service they prefer), starts listening to a "Discover" playlist, and stumbles upon the music of Tanya Tagaq. The armchair analyst comes to discover that she incorporates Inuit throat singing into her music. In curiosity, the armchair analyst does what they know best, analyzes a song or two the way that they know based on their training in a university music program.*

The term I use in the above description, armchair analyst, is akin to "armchair analysis"—a term coined by Alan Merriam in his *Anthropology of Music* (1964) as an opposition to that of fieldwork-based ethnomusicology. Merriam identified two types of armchair analysis, which he considered inappropriate: "failure to take theories to the empiric test of the field materials, and the analysis of materials collected by others" (1964, 39). This criticism of armchair analysis was first used to critique the work of comparative musicologists who approached the study of music as objectively as possible, while researching from their desk, transcribing music into Western notation, listening to field recordings, and reading the documentation of others. I use this term, armchair analyst, in a narrow sense to distinguish between analysts who have interacted with the artists/musicians/composers/producers of whose music they write and those who haven't.[14] During the writing of this dissertation I have been an armchair analyst, but armchair analysts also include many other music theorists. They might be scholars without time and resources to build relationships with performers/composers for a given

a metaphor of it. To indicate that I understand that what I am doing here with this dissertation is not decolonizing work, I adopt the term uncolonizing instead.

[14] This is also a take on the phrase "armchair expert" as it relates to music analysis. An armchair expert is someone who claims to know a lot about a subject but has no real experience with the matter or understanding of it.

project, or they might be scholars who could build relationships with performers/composers but choose not to. Whatever the reason for working independently, the armchair analyst often chooses to do what is easy, drawing upon what they know, often using analytical tools that were originally designed for Western art music. It is not only easy but comfortable to use established analytical tools.

It may surprise some that I openly refer to myself as an armchair analyst since this term has previously had derogatory connotations. Furthermore, the concept of an armchair analyst as originally defined by Merriam is somewhat of a myth. As Kofi Agawu notes, "all ethnomusicology is armchair ethnomusicology, because no matter how long you spend in the field, getting out is almost always a condition for producing your ethnography" (2003, 180). Agawu also goes on to say that not all armchair analysts are alike; the quality of the analysis largely depends on the questions asked and the observations made. Agawu (2003) cites Erich von Hornbostel as a noteworthy comparative musicologist whose work continues to influence African music research, even though he did not conduct fieldwork of his own. Like Agawu, I am not advocating for armchair analysts to step away from the armchair in order to do ethical work. Sometimes doing fieldwork does more harm to the people a scholar works with than remaining in the armchair. I am advocating, however, for armchair analysts, such as myself, to consider ethics in relation to the research conducted and to foreground ethics in the process of analysis regardless of whether the analyst directly consults with the people who make the music. Determining what is considered ethical in music analysis is a matter for the individual scholar. In this chapter, I consider my ethical stance toward analyzing music, and what I felt was appropriate ways of approaching the analysis of Tagaq's music (and country music).

The problem with being an armchair analyst is that most of us are likely to practice *hungry listening*, a term used by Dylan Robinson (2020) in his book of the same name. The practice of hungry listening prioritizes "the capture and certainty of information over the affective feel, timbre, touch, and texture of sound" (38).[15] This practice, while often unmarked in music scholarship, dominates the field.[16] For instance, hungry listening pervades typical university music theory curricula, in which we prioritize formalism over the affective aspects of music and attempt to include non-Western musics only when doing so does not interfere with the established pedagogy. In some ways I have remained a quintessential armchair analyst during the writing of this dissertation. My analyses are structural, and in this way, very similar to what remains prominent in the field of music theory. In other ways, however, I have actively made steps to broaden the ways in which I research and write about music in an attempt to make my work more ethical and inclusive. Those choices are detailed below in Section 3.2 in conjunction with a discussion of Indigenous research methodologies.

3.2 Indigenous Research Methodologies; Some Applications for an Armchair Analyst

As I thought about my position in my research, I also began to question how, as an armchair analyst, I might approach analyzing music outside of (and even inside of) my own cultural background in a way that would be more appropriate for the type of music and more in line with the cultural practices of the musicians/composers. In order to ethically approach the analysis of Tagaq's music, I needed to develop analytical research approaches that incorporate

[15] This type of listening relates to Carolyn Abbate's distinction between drastic and gnostic ways of listening to and understanding music. Abbate defines drastic listening as "physicality, but also desperation and peril, involving a category of knowledge that flows from drastic actions or experiences and not from verbally mediated reasoning"; whereas, the goal of gnostic listening is to make "the opaque transparent, knowledge based on semiosis and disclosed secrets, reserved for the elite and hidden from others" (2004, 510).

[16] Although critiques of settler colonial forms of perception in music scholarship are rare, Philip Ewell (2020) has recently critiqued music theory for being rooted in the "white racial frame" while discussing the impacts of anti-Black racism in the United States on the field of music theory.

Indigenous research methodologies. In the second edition of her book *Indigenous Methodologies: Characteristics, Conversations, and Contexts* (2010, 2nd edition 2021), Margaret Kovach (Sakewew p'sim iskwew) suggests that there are four core foundations of an Indigenous conceptual framework within an Indigenous research methodology: epistemology, ethics, community, and self. These foundations are listed in the leftmost column of Table 3.1. I list the main features of Indigenous research design in the middle column. These characteristics are a summary of those proposed by Margaret Kovach ([2010] 2021), Lester-Irabinna Rigney (1999), Linda Tuhiwai Smith (1999), and Shawn Wilson (2008). Nado Aveling (2013) provides a useful and concise summary of these features in an article titled "'Don't talk about what you don't know': on (not) conducting research with/in Indigenous contexts." Reading this article was actually my first introduction to the concept of Indigenous research methodologies; from this article I learned of the numerous substantial works in this field.

These works highlight the importance of insiders of Indigenous culture to implement Indigenous research design in their methodologies. The work of Rigney (Narungga) and Smith (Māori, Ngāti Awa, and Ngāti Porou iwi) are seminal writings on Indigenous research methodologies. Rigney (1999) explicitly states that Indigenous research is research by Indigenous people, with Indigenous people, for Indigenous people, but he builds his concept of Indigenist Research from feminist scholarship and research methodologies. Rigney presents three fundamental principles of Indigenist research: resistance, political integrity, and privileging Indigenous voices. Concurrently, Smith published her influential book *Decolonizing Methodologies: Research and Indigenous Peoples* (1999, 3rd edition 2021), in which she critiques the history and legacy of colonial research and presents possibilities for a decolonizing research paradigm with examples from her own research experiences as a Māori woman working

with Indigenous peoples. Wilson (Opaskwayak Cree) emphasizes the importance of relationship in research. He explains that *Research is Ceremony* (2008), the ceremony of maintaining accountability to these relationships. As Rigney explains, Indigenous researchers tend to hold themselves more accountable than settler scholars because they are responsible to their communities. I, however, am by no means an insider of Inuit culture, nor have I built relationships with these musicians. So how can someone like me, an armchair analyst, incorporate the tenets of Indigenous research design into their research if they have not built a relationship with and worked with the Indigenous musicians about whom they write? I have struggled with this throughout the writing of my dissertation. What I have come up with is by no means a perfect solution, but I would like to take a moment to discuss ways that I have attempted to incorporate Indigenous research methodologies in this dissertation. To the right of the columns summarizing the principles of Indigenous research methodologies in Table 3.1, I outline some possible applications of Indigenous research design for music-theoretical research, some of which have been inspired by the work of Avery (2012) and Robinson (2020). Below, I go into further detail about the ways in which this is manifested in my dissertation.

Table 3.1 Summary of Indigenous research methodologies and some possible applications in the study of music analysis.

Core Foundations of Indigenous Research (Kovach [2010] 2021)	**Features of Indigenous Research** (Summarized in Aveling 2013, pp. 206–7)	**Applications in Analyzing Music as an Armchair Analyst**
Epistemology	• Research is grounded in an Indigenous epistemology (Kovach [2010] 2021; Rigney 1999) • Research privileges Indigenous perspectives (Rigney 1999; Wilson 2008) • Research design demonstrates an explicit decolonizing aim (Smith 1999) • Researchers utilize Indigenous methods, such as storytelling (Kovach [2010] 2021; Wilson 2008)	• Engage with published Indigenous literature (Avery 2012; Robinson 2020) • Ground analysis in published interviews with the musicians, focusing on aspects of the music that are important to the musician • Graphic representations of music, rather than use of Western art music notation • Storytelling as a way of presenting material and as analysis itself (Robinson 2020) • Present multiple analyses of the same song
Ethics	• Researchers observe cultural protocols (Wilson 2008) • Research honours and respects sacred knowledges (Kovach [2010] 2021; Wilson 2008)	• No analysis of sacred music • Treat Indigenous knowledges the same as scholarly writing, citing and giving the same credit
Community	• Research emphasizes collaborative research and should benefit Indigenous peoples (Kovach [2010] 2021; Smith 1999; Wilson 2008) • Research gives back to the communities (Kovach [2010] 2021; Smith 1999)	• Make research accessible to people outside of academia, both in content and availability of publication
Self	• Research locates self through identification and critique of positionality and reflection (Kovach [2010] 2021; Wilson 2008)	• Make known one's listening positionality and biases (Robinson 2020) • Reflect on process of analysis

Rigney (1999) argues that research should be grounded in Indigenous epistemologies and should privilege Indigenous perspectives (refer to the first row of Table 3.1).[17] This research highlights the experiences, ideas, traditions, and aspirations of Indigenous people. This feature of Indigenous research design counters the colonial citational practices of academia, in which white scholars are cited as though they are the authority on a subject while excluding people of color and non-scholars. In his 2007 article "To Cite or not to Cite?", Kofi Agawu ponders the appropriate citational practices for writing about African music. He calls on scholars to critique these colonial citational practices and reject them. Like scholarship on African music, scholarship on Indigenous music has largely been a Euro-American archive of knowledge written in European languages and aimed at a European audience (Agawu 2007, 256).[18] Because of this, care needs to be taken in citational practices when writing about Inuit music. Indigenous music scholars, such as Dawn Avery (2012) and Dylan Robinson (2020), have shown ways of grounding research in Indigenous epistemology. Avery (2012, 130) bases her study of classical Native music in the concept of twinness, explaining that the "Kanienkeha or Mohawk word for two is tékeni," which means "the balance between two beings or particles of spirit working together." Avery also develops a theoretical model to illustrate the concept of twinness reminiscent of the Plains Indian medicine wheel, which is an important symbol in many Indigenous cultures in North America. Robinson's term hungry listening (defined above) derives from "two Halq'eméylem words: shxwelítemelh (the adjective for settler or white person's methods/things) and xwélalà:m (the word for listening). shxwelítemelh comes from the word xwelítem (white settler) and more precisely means 'starving person'" (2020, 2). This definition

[17] Kovach ([2010] 2021) and Wilson (2008) share similar suggestions in their scholarship.
[18] I am particularly thinking of the writings of Franz Boas and other comparative musicologists from the late nineteenth and early twentieth centuries. These writings, while not widely cited today, still have an influence in some part on the way that settler scholars write about Indigenous music.

corresponds to the first experiences that Stó:lō have of settlers during the gold rush of the late 1800s, in which the newcomers were starving for both sustenance and gold (Robinson 2020, 48). For Chapters 4 and 6, the ones discussing Tanya Tagaq's music, I draw on numerous published interviews by Inuit throat singers, including Tagaq herself, and Tiffany Kuliktana Ayalik and Kayley Inuksuk Mackay (of PIQSIQ), to center Tagaq, as well as the perspectives of other Inuit musicians, in my analyses of Tagaq's music. I also incorporate Indigenous concepts in my analyses. For instance, I ground my discussion of beat structure in Tagaq's "Ilunikavi" in Chapter 4 in both Western and Inuit conceptions of time, as discussed by Gombay 2009, Keyes 2009, and Robinson 2020. I maintain this practice in Chapter 5, on country music, by quoting the words of country songwriters from published interviews. Regardless of the type of music, it is important to highlight the voices of the people who make it in one's analyses. This ensures that other perspectives are included in the scholarship beyond that of the analyst and other academic scholars who presumably think somewhat alike.

Smith (1999) emphasizes that research should demonstrate an explicit decolonizing aim; in a similar sentiment, Wilson (2008) suggests that researchers utilize Indigenous methods, such as storytelling (refer to the first row of Table 3.1). Philip Ewell (2020) has argued that music theory is grounded in the *white racial frame* (Feagin [2009] 2013), as work in this field typically amplifies Western art music (WAM). Even when writing about music outside of WAM, theorists often draw from analytical approaches originally intended for WAM to analyze these other styles. Music theorists tend to listen from one perspective, the *ideal listener* who is a trained musician in Western art music, producing a singular analysis of a piece of music, without regard for other experiences. In this dissertation, I decenter Eurocentric approaches to analyzing music by focusing on sound qualities, avoiding Western staff notation, and developing graphic

visualizations. This project engages with Tagaq's music in a way that amplifies the rich sonic aspects of the music that are erased when it is reduced to Western staff notation. I do this in Chapters 4 and 6 by focusing on developing graphic visualizations comprised of data from the acoustic properties of audio recordings. Additionally, in Chapter 6, I visualize Tagaq's "Ilunikavi" through multiple graphic representations—one for each of the acoustic descriptors. Each of these three graphs tells a slightly different story about the music.

In an effort to decenter the Western academic writing style and to push myself out of my comfort zone with analysis, I have also drawn inspiration from Robinson's concept of *sensory-formalist analysis*, in which "the writer seeks to extend the form and structure of the listener/viewer/reader's sensory engagement through their writing" (94). Robinson takes an intersectional approach to analysis, adopting performative writing techniques inspired by the scholarship of feminist, Indigenous, Black, and queer scholars (Cusick 1994; hooks 1992; Koestenbaum 1993; Lorde 1981; McClary 1991; Morin 2016; Phelan 1997; among others). Presented as what Robinson terms *event scores*, these analyses are, like the music, open to interpretation.[19] They are poetic and a form of art, expressing Robinson's understanding of a piece in a way that retains the affective qualities of the music. Following these ideas, in Chapter 6 I present an analysis of Tagaq's "Sivulivinivut" in the form of a story, one that is open to interpretation and aims to maintain the affective feeling of the song.

As summarized in the second row of Table 3.1, when analyzing music by Indigenous composers and musicians, analysts need to consider the ethics of their work. Wilson (2008) argues that researchers need to observe cultural protocols, including honoring and respecting sacred knowledges; for this reason, it is inappropriate to research sacred Indigenous songs. In

[19] Likewise, Marion Guck (2006, 191) advocates for music analysis as interpretation, as something "open to interpretation" or "something that might be taken in a number of ways."

this dissertation, I consciously chose to focus on the music of Tanya Tagaq because she is a public figure who composes and sings secular popular music while using Inuit throat singing as a primary singing style. She composes, performs, records, and tours her own music and profits from these ventures. She has agency, and she regularly protests injustices, especially with regard to Indigenous peoples, in interviews and through Twitter. I also show that I value the knowledges of Tagaq and other Inuit singers by grounding my analyses in their published interviews and by citing them in the same way that I cite academic scholarship.[20] I use these interviews both as a basis for my analysis—as in Chapter 4, where I use Tagaq's discussion of her throat singing as a basis for the way I categorize the sounds—and as a way to push myself to think otherwise, as in the discussion of oppositions by the members of PIQSIQ (also in Chapter 4). While I did not work directly with Tagaq on this project, this practice centers Tagaq in the analyses that I have created. Furthermore, I suggest that it is important to consider this when analyzing all types of music. In Chapter 5, I focus on recent mainstream country music, which continues to be a predominantly white genre. While I could ignore the whiteness and racism in country music, to do so would be unethical. Instead, I confront it by foregrounding Lil Nas X's "Old Town Road" in the chapter, discussing the reasons that it was removed from country radio and the ways in which it matches the characteristics of current country music hits.

Community and relationships are key to Indigenous research methods. It is the third foundation of Indigenous research methodologies listed in Table 3.1. Kovach ([2010] 2021), Smith (1999), and Wilson (2008) argue that research needs to emphasize collaboration and

[20] Robin Attas (2019) suggests that decolonization has two parts: "uncovering Eurocentric elements in the specific context where decolonization will happen, and second, finding ways to put Indigenous and settler knowledge on an equal footing in terms of both content and pedagogy." By grounding my analyses in Tagaq's sentiments about her own music, I put Tagaq's knowledge on equal footing with the music theory scholars whom I also cite.

should benefit Indigenous peoples in a way that ultimately gives back to communities being researched. This point is difficult to address if, as an armchair analyst, I am not in direct communication with the people who write and perform the music about which I write. To address this aspect of Indigenous research methodology, I have attempted to make my research known to Tagaq by posting conference papers I have presented via Twitter and tagging her. I will also provide a copy of my dissertation to Tagaq via her management so that she can be aware of my work and determine for herself if it is of any benefit or not.

It must also be acknowledged that I have already unquestionably benefited from working on Tagaq's music: I have presented at multiple conferences, published an article and a book review, and obtained a job because of my work on Tagaq's music. This, in itself, is an ethical conundrum, and one with which I am frankly struggling. I do not have any good solutions. In response to this ethical question, Aveling (2013), a white scholar who used to write about Indigenous culture, wrote an article titled "'Don't talk about what you don't know': on (not) conducting research with/in Indigenous contexts." Aveling argues that non-Indigenous scholars working with/in Indigenous communities either need to conduct research in ways that are culturally appropriate and inclusive, or to give up the role of researcher within Indigenous communities and make way for Indigenous researchers. By the end of the article, she determines that it is not appropriate for her to do research in Indigenous communities and mentions alternative ways that she has and plans to continue being an ally with Indigenous researchers. I agree with much of what Aveling says in this article. Music scholars need to develop ways of engaging with music outside of their cultural background in ways that are appropriate and inclusive or to not analyze this music at all. Like Aveling, I could pledge to no longer write about Tagaq's music or any music made by Indigenous composers/musicians, but to not write about

Tagaq's music or other Indigenous artists' music would be taking the easy way out. It would allow me to return to a place of comfort, one in which I would get to continue with my (inherited) practice of hungry listening without concerning myself with these difficult questions. To move away from hungry listening practices and toward culturally appropriate ways of doing analysis is a journey that takes time. I will inevitably make mistakes when writing about Tagaq's music. (Given some time, I hope to return to this document with a fresh perspective to identify things that I could have done better ethically.) On the other hand, I feel that it is a worthwhile task if it can in some small way contribute to making the field of music theory more inclusive and equitable.

Lastly, as summarized in the fourth row of Table 3.1, Kovach ([2010] 2021) and Wilson (2008) argue that researchers need to make room for self-reflection and -identification in their scholarship. Robinson defines critical listening positionality as an act that "involves a self-reflexive questioning of how race, class, gender, sexuality, ability, and cultural background intersect and influence the way we are able to hear sound, music, and the world around us" (Robinson 2020, 10). Furthermore, identifying and critiquing listening positionalities, such as settler-colonial listening positionalities, turns "a critical lens toward individual inheritances of (hetero)normative, ableist, colonial, and racializing formations for compositional and presentational practice" (Robinson 2020, 254), a move that addresses issues relating to diversity, equity, and inclusion in music disciplines. Articulating this is an important step in music analysis because it gives readers insight into my analytical assumptions and biases—namely, my positionality affects how and why I hear and analyze music in the way that I do. While I will reflect throughout my dissertation on choices that I make, I take a moment now to situate myself in relation to my research.

3.3 My Listening Positionality

I am a white, cis woman from a middle-class family. I was born and raised in Thompson, Manitoba, Canada, which is located in Treaty 5 territory on the lands of the Néhiyaw.[21] I learned about Néhiyaw (which at that time, I knew as Cree) and Anishinaabe music, dance, and cultural values through school. My high school years were spent in Treaty 1 territory on the lands of the Anishinaabe peoples in a small town in southern Manitoba. I didn't learn any more Indigenous cultural teachings until one course at the University of Manitoba. While writing this dissertation, I lived in Manhattan, the past and present land of the Lenape peoples. In more than ten years of university training and three music degrees, I only took five courses that featured non-Western music, and three of those I audited rather than took for credit because they did not fit into my program of study.

To state that I am a white, Canadian, cis woman from a middle-class family, and a music theorist is easy. They are facts that are a part of my identity, but what does this mean for my ways of listening and for my scholarship? When listening for fun, I listen to the music as a whole, moving and singing along with it; I listen for how it makes me feel, taking in whatever emotion the music raises in me. While some may wonder if I listen this way because I am a woman, I doubt that would be the only reason. I think it is because I love music, and I retain that love while listening for the enjoyment of music. I have had similar conversations regarding

[21] There were eleven numbered treaties made between the Crown and First Nations from 1871 to 1921. (The Crown is a term used to refer to the state and its government in Canada as it remains a constitutional monarchy. First Nations refers to Indigenous peoples in Canada who are not Métis or Inuit.) These treaties ensured that the Crown claimed land for settlement and development, and in exchange, the government promised special rights to treaty lands for First Nations, as well as the distribution of cash payments, hunting and fishing tools, farming supplies, and more. These treaties are controversial and continue to be contested. Many believe that the written treaties misrepresent or neglect to include certain aspects discussed and agreed upon during the negotiations. Treaty 1 was signed in 1871 and Treaty 5 was signed in 1875. These treaties are still in effect today, and new ones are being created. For more on treaties in Canada, see Michelle Filice (2016) and Taylor MacLean's "6 Common Myths about Treaties in Canada."

listening practices with my brother and father, who equally love music. In contrast, when listening with scholarly goals in mind, I listen structurally and formally, often without any emotional attachment, seeking to identify aspects of the music in a way that makes sense to me based on what I have learned in my formal music training. My listening is based on my prior experiences with music, both informally and formally; however, in my scholarship, the formal side often takes over and the emotional relationship between myself and the music is removed. In this dissertation, I attempt to retain both parts of my listening habits while presenting my analyses of both Tagaq's music and country songs. The analyses in this dissertation are structural at various levels, focusing on beat, meter, and form, but I also attempt to show the emotional side of my listening habit through regular self-reflection and by considering narrative within the songs discussed.

I take this moment to situate my identity and my positionality because they factor into all of the analyses and the words that I write in this dissertation. In other words, I engage with the music from a settler Canadian perspective, but I have tried to bring in others' opinions to make my research less biased. I, however, cannot change my listening positionality, and since this dissertation is from my perspective, my listening positionality should be kept in mind and critiqued while reading this document. I will continue to critique my ways of writing about music throughout this dissertation, and I hope you, the reader, will do the same.

Chapter 4: A Comparison of Two Experiences of Beat in Tanya Tagaq's "Ilunikavi"

In recent years, Tanya Tagaq (Inuk), an experimental, improvisational singer and avant-garde composer who grew up in Ikaluktutiak (Cambridge Bay, Nunavut), has achieved international success, performing at major festivals and venues all over the world. Her complex and individual vocal style—blending pop, punk rock, heavy metal, electronica, and Inuit throat singing—mixed with her lively stage presence has become an interest of numerous scholars, though they have focused more on her public persona and activism than her musical style.[22] In describing her music, Tanya Tagaq states, "It's breath, it's rhythm. To be very, well, pompous about it, it's like the sushi of sound. When you hear it, you either love it or you hate it" (Perry 2011). This quote suggests that Tagaq considers rhythm to be a primary element of her music, but here I am suggesting that timbre can also be regarded as an important element. This chapter attempts to answer the question: How do rhythm and timbre interact to affect our experience of the beat structure in Tagaq's music?

This chapter develops a method for analyzing Tagaq's vocal style through an examination of the qualities of individual sound types and the ways in which changes in sound qualities affect our experience of beat. I contend that loudness, noisiness, and brightness affect the experience of beat in Tagaq's song "Ilunikavi." Section 4.1 compares generalized conceptions of Inuit and Western time, as well as the connection between the two cultures' understandings of time and musical time. In Section 4.2, I detail some of my concerns with

[22] Among those paying attention to Tagaq of late are Galloway (2020), Stévance (2010, 2016, and 2017), Woloshyn (2017) and Taylor-Neu (2018). Galloway makes connections between Tagaq's use of songs, live performance, and audiovisual media and her ecological activism. Stévance examines the meaning in her bodily gestures during performances. Woloshyn and Taylor-Neu make connections between her artistic identity and activism, suggesting that her persona challenges Indigenous stereotypes and continues to play an important role in reconciliation in Canada.

writing a chapter such as this, the ethics involved, and the assumptions established in order to make possible a comparison of my feeling of a beat with Tagaq's feeling of a beat for the same song. Section 4.3 presents an introduction to Tagaq's sounds and their sound qualities. Using Tagaq's own descriptions of these sounds from her throat-singing demonstrations in "The Sounds of Throat Singing" (Tagaq 2009), I develop a taxonomy of her throat-singing sounds. Using computer-aided acoustic measurements, I study the ways in which music information retrieval techniques may be employed both in the analysis of individual vocal sounds and for the conceptualization of timbre. In Section 4.4, I present the analytical approach used in this chapter to identify the timing of an ictus, which I associate with feeling a beat, in video recordings of Tagaq's "Ilunikavi."[23] In Section 4.5, I make observations regarding the similarities and differences between Tagaq's movements and my own to the same performance of "Ilunikavi" in relation to the onsets of the sounds and the changes in sound qualities. In the last part, Section 4.6, I speculate on possible reasons for the differences between our movements in relation to our cultural understandings of time. Research for this chapter began with a hunch that I listen to music with a focus on timbral changes and that these changes affect how I understand beat. I also assumed that if timbre affects my way of understanding beat, then maybe it does for others too. I was surprised to find that the way that I mark beats aligns with Lerdahl and Jackendoff's Metric Preference Rules—and Western conceptions of mechanical time—while Tagaq's movements during this performance did not.

[23] An ictus is the moment when a beat occurs. In a Western conducting pattern, for instance, the ictus of the first beat is the lowest point that the hand/baton reaches before changing direction.

4.1 Inuit and Western Conceptions of Time

Western mechanical conceptions of time are drastically different from Inuit ecological conceptions of time. Gombay (2009) suggests that Western mechanical conceptions of time are absolute, linear, and quantitative, governed by calendars and clocks, while Inuit may view time as relational, contingent, and ecological, governed by natural cycles such as the placement of the sun in the sky, lunar months, the migration of animals, etc. This Western conception of time is rigid, following the mechanical clock time, while Inuit ecological conceptions of time are flexible, following changes in the environment (Gombay 2009). These differences were felt by Inuit children forced into residential schools and mandated to abide by Western "technical time," which was used as a weapon for colonization (Smith 2007; Keyes 2009; Robinson 2020). Sara Keyes explains that "bells ordered students' lives, dictating when to sleep, rise, learn, pray, and eat. On their way to the dining hall, students marched in time to the sound of a bell. Upon their arrival, two bells rang; one to direct students to pull out their chairs and the other to indicate that they could sit down" (2009, 36).

These simplified conceptions of time relate to our most accepted conceptions of musical time in music theory discourse.[24] Like the Western conception of time described above, Western meter theory is mechanical. For example, the belief that "every beat at a given level must also be a beat at all smaller levels" and "each metrical level has evenly spaced beats" (Lerdahl and Jackendoff 1983, 69–70) connects to Western conceptions of time as hierarchical and evenly spaced. Within a day there are 24 hours, within an hour there are 60 evenly spaced minutes, within a minute there are 60 evenly spaced seconds, etc. Similarly, within a 4/4 measure, there are four evenly spaced quarter-note beats, within a quarter-note beat there are two eighth-note

[24] Grant (2014) and Scherzinger (2019) discuss the historical events which led to the association of musical time with clock time.

subdivisions, within an eighth note there are two sixteenth-note subdivisions, etc. Johann Forkel was one of the first to theorize the possible connection of musical rhythm to the division of days into hours, but it was Maelzel's metronome that cemented the relationship between musical time and clock time, as it measured and indicated the number of beats per minute (Grant 2014). In Western cultures, time is regulated in a mechanical sense, as is the Western experience of musical time (rhythm, beats, and meter). Western musical time is so regulated that when composers or musicians want to deviate from this type of time, they often develop alternative ways of notating music. Furthermore, computer programs, such as music notation software and Digital Audio Workstations, are built in such a way that suggests music must be created using steady, even pulses and recognized metric groupings.

Our conceptions of musical time can become, and have been, colonizing forces in themselves. Robinson notes that "in the inclusion of Indigenous voices and musicians into Western composition: the imposition of Western temporality through 'clock-defined time,' regimented through time signatures and meter" has become common practice (119). In these supposed collaborations, Indigenous musicians are often forced to comply with certain practices held by Western musicians. Western art musicians, for instance, play music from written notation and prefer to keep a steady, even pulse and to group beats into hierarchical groupings to establish meter. This relates to the power dynamic in the relationship between Western art musicians and Indigenous musicians; with more Western art musicians involved in the collaboration, the default is to continue with the Western status quo and make the Indigenous musicians adapt. This conception of time in Western art music, as Robinson goes on to explain in *Hungry Listening*, is antithetical to Indigenous musical practices. Robinson argues that "while the regularity of the drum may seem a defining feature of much Indigenous music, to force

Indigenous performance into unvarying meter misconstrues the life-giving force of the heartbeat as having its own rhythm and relative variation from moments of exertion to rest" (2020, 141). Indigenous music tends to be more flexible in time than Western art (and even popular) music, and this may be due to Indigenous conceptions of time. Inuit conceptions of time, as something that is contingent and relational, correspond to the basic premise of Inuit throat singing as "play and flexibility of time in play" (Robinson 2020, 141).

This comparison of Inuit and Western conceptions of time, including musical time, is admittedly reductionist.[25] This distinction that I draw between Western mechanical time and Inuit ecological time is intentionally simplified and generalized, for the sake of convenience. Roger Matthew Grant (2014) and Martin Scherzinger (2019) historicize temporality much more eloquently and thoroughly than I do here. Western mechanical time is an invention of the early modern era and has only been in use widely for the last 200 years or so. Furthermore, the distinction between these conceptions of time are colonial constructs. Ethnographers suggested that mechanical time was modern and superior to the time of their interlocutors, which was considered non-modern (Scherzinger 2019). As Scherzinger (2019) points out, however, the preference rules of Lerdahl and Jackendoff that are based on Western mechanical conceptions of time are not universal and so some difference in understanding musical time among different people must be considered possible. Even if there is some truth to the distinctions between Western and Inuit conceptions of time described above, other issues arise in the simplicity of the characterization of them. It should also be acknowledged that Inuit culture, traditions, dialects, and beliefs are regional, based on Inuit relationships to people, place, and land, and conceptions

[25] In *Beyond Settler Time: Temporal Sovereignty and Indigenous Self-Determination*, Mark Rifkin (2017) makes a much more nuanced comparison of Indigenous and settler temporal frameworks.

of time may differ among ethnic groups.[26] Additionally, even if at one time Inuit relied upon ecological conceptions of time, the clear-cut distinction between Western and Inuit conceptions of time presented above likely no longer exists due to colonization and changes in ways of living (many of which were violently forced on Inuit by the Government of Canada and the churches). Many Inuit likely navigate between (or blend) Inuit conceptions of time and the Western mechanical conception of time, or might even feel more comfortable with (or know nothing but) the Western conception of time.

As I explain later in this chapter, despite the issues related to drawing this dichotomy between Western and Inuit conceptions of time, my Western understanding of time undoubtedly influences the way in which I feel and locate beat in music based on observations I make regarding how I mark beat while listening to "Ilunikavi." For instance, I mostly maintain equally spaced beats, following Lerdahl and Jackendoff (1983) and the characteristics of clock time. Interestingly, Tagaq's way of marking the beat, as observed in a single performance of "Ilunikavi," also corresponds in some ways to the generalized Inuit conception of time described above. I describe in more detail these speculative connections between conceptions of time and musical time at the end of this chapter, after I explain my observations regarding Tagaq's and my own experience of beat.

[26] The traditional homelands of the Inuit are in northern Canada, Greenland, and parts of Alaska. As indicated in *The Canadian Encyclopedia*, "there are eight main Inuit ethnic groups: the Labradormiut (Labrador), Nunavimmiut (Ungava), Baffin Island, Iglulingmuit (Iglulik), Kivallirmiut (Caribou), Netsilingmiut (Netsilik), Inuinnait (Copper), and Inuvialuit or Western Arctic Inuit" (Freeman 2020). The Inuit language is Inuktitut. However, there are five main dialects in Canada: Inuvialuktun (Inuvialuit region in the Northwest Territories); Inuinnaqtun (western Nunavut); Inuktitut (eastern Nunavut dialect used west of Hudson's Bay); Inuktitut (Nunavik dialect used east of Hudson's Bay in Quebec); and Nunatsiavumiuttut (Nunatsiavut) (Freeman 2020). There are likely differences among these groups in terms of conceptions of time, as well.

4.2 Ethics of Analyzing Inuit Popular Music and Meter

In an interview with Light Echo Productions (2014), Tagaq recounts a time when she went to Pond Inlet, Nunavut and was "travelling across the sea ice…and the guides caught a yearly seal." She mentions that she has always loved color, particularly the color of the bright red, fresh blood on the snow. She states in the interview that it is one of her favorite visuals. She goes on to explain that "food is life," and asks the listener to "imagine what it would be like living in 24-hour darkness in the middle of winter, -40, -50, no food, with your child hungry and someone brings a seal to your doorstep, someone coming to your door with life" (Light Echo Productions 2014). For Tagaq, the sight of bright red fresh blood is comforting. It means that she and others will have fresh meat to eat. For me, on the other hand, bright red fresh blood reminds me of horror movies or pain after cutting myself or falling and scraping my knee. This is just one example to illustrate that my perspective is different from Tagaq's, just as my perspective is likely also different from yours, the person reading this dissertation.

As I discussed in Chapter 3, identifying and naming my positionality matters in analyzing music because it affects how and why I hear and analyze music in the way that I do. It gives the reader insight into my analytical assumptions. One's identities and personal experiences affect their listening in ways that are inaccessible to others. As Dylan Robinson pithily states, "unlike iPhone photo filters, one cannot simply select and add noncolonial, feminist, queer, or black listening filters in order to listen otherwise" (2020, 51). I cannot choose to listen as someone else does. The way that each of us listens is wrapped up in prior experiences, both musical and extra-musical. My hearing, like everyone's hearing, is biased. I, as I have stated in Chapter 3, am a white settler/colonialist who is trained in Western art music, but I spend most of my time listening to and playing various types of American and Canadian Anglophone popular music. I

cannot listen as Tanya Tagaq listens. In addition to the differences between our past experiences and cultural backgrounds, we also have different roles. Tagaq is the performer/composer of the song, while I am the listener/consumer (to be discussed in further detail below). This gives Tagaq a different insight to her music than I have. For this reason, I am careful to make observations based on the movement data and sound quality data available to me. I also put more of a focus on analyzing myself and the reasons that I move to the music in the way that I do because I am able to question myself, but I am not able to question Tagaq.

In this chapter, I consciously chose to focus on *beat* analysis, rather than *metrical* analysis. As stated in the previous section, most Western music theorists define meter as being, in some way, hierarchical. Lerdahl and Jackendoff, for instance, define meter as a "regular, hierarchical pattern of beats to which the listener relates musical events" (1983, 17). We entrain to these regularly recurring events, feeling a musical meter (London 2004, 4). The videos of our movements, however, do not give enough information to speculate about how Tagaq and I distinguish beats hierarchically. Therefore, I focus on how and when we appear to feel a beat in relation to changes in sound qualities. Furthermore, this is not a study of how Tagaq and I understand beat in general, as I only make observations about when and how we move during a single listening/performance. It is very plausible that, while performing "Ilunikavi" in a different setting and at a different time, Tagaq would move differently than she does in the video performance examined in this chapter. Likewise, listening to "Ilunikavi" at a later date, I may have moved differently to the song than I did in the video I recorded of myself.

In sections 4.6 and 4.7, I make some speculations that when and how we embody beats while listening to music or organized sound is connected to our positionalities as performer vs. listener and Inuk vs. Euro-Canadian. This is based on the observations made regarding Tagaq's

and my movement to the same performance. Stronger claims can be made when comparing the differences between a listener and performer, in this case, than when comparing the differences between coming from an Inuit background and a Euro-Canadian background. The distinction between listener and performer here is clear. Tagaq is performing "Ilunikavi" while also moving to it. She can predict the next sounds (if she chooses) as she is the one making the music. I, on the other hand, am a listener to "Ilunikavi" and not privy to knowing what comes next. I have to guess and assume based on what has come before where I should indicate a beat.[27] The distinction between our cultural backgrounds, however, is much less clear. I speculate that our cultural conceptualizations of time, as well as our prior experiences with music, affect how we feel beat. It must be noted, though, that these distinctions, which I laid out so cleanly above, are not as clearly defined in practice. While Tagaq is Inuk and I am Euro-Canadian, these labels are neither our only identity markers nor the only contributing factors to how we understand the beat of music. We have had many more experiences that contribute to the ways we feel a beat. For instance, Tagaq did not grow up listening to Inuit throat singing as a child because it was banned.[28] Likewise, I did not grow up with Western art music. We both started engaging with these respective styles when we were in university. Tagaq has mentioned in interviews that she listened mostly to classic rock growing up, including Jimi Hendrix, Leonard Cohen, the Doors, the Beatles, and Bob Marley (Dickie 2014), all of which I grew up listening to as well. Tagaq's music itself is not one-dimensional; she incorporates punk, metal, and electronica to create a style of music that is all her own. While I attempt to ascribe reasons for our movements to

[27] While I had listened to "Ilunikavi" numerous times prior, I recorded myself moving to Tagaq's live performance of "Ilunikavi" more than six months after last seeing or hearing the recording of her performance to ensure that my movements would be minimally influenced by Tagaq's.

[28] The Government of Canada did many things to try to supress Inuit people and culture, including forcing Inuit to relocate into communities established by the government, killing the sled dogs, giving Inuit identification numbers, and banning throat singing.

"Ilunikavi" to differences in our cultural backgrounds, this is extremely speculative and not as clearly defined as the performer/listener dichotomy. Because I cannot speak for Tagaq, during this speculative part of the chapter, I focus more on how my movements embody a beat of a song in a Western sense of time. This study is intended to be an analytical self-reflection of how and why I moved to "Ilunikavi" as I did.

Reflection and critique are important aspects of the continual process of decentering Western music theory and centering the voices that composed, created, perform, and/or own the music that is being analyzed. With this chapter, I hope to provide one example of a path that a theorist might take for self-reflection, while presenting my struggles with this process openly. It is also my hope that others will use this as a springboard for even more contextually informed and culturally appropriate approaches for analyzing music.

4.3 An Introduction to Tagaq's Throat-Singing Sounds[29]

Tagaq's vocal styling draws on a variety of sounds, including pure high tones, deep growls, and audible breaths—sounds regularly used in Inuit throat singing.[30] Mixed in are screams, cries, and moans, sounds not typically heard in throat singing. Here, I focus on the qualities of individual sounds in order to develop a nuanced and systematic way of describing the very important, though subtle, differences between them.[31]

---

[29] Much of Sections 4.3 and 4.4 appear in Hardman 2022, though this material has been substantially revised. (Copyright 2022 from *Trends in World Music Analysis: New Directions in World Music Analysis*, edited by Lawrence Beaumont Shuster, Somangshu Mukherji, and Noé Dinnerstein. Reproduced by permission of Taylor and Francis Group, LLC, a division of Informa plc.)

[30] In Inuktitut, throat games (or throat singing) are referred to as katajjait (pl.) and katajjaq (sing.). Because I am, frankly, not sure of when it is appropriate to use the plural vs. the singular, I have chosen to mostly use "throat singing," as Tagaq typically does in interviews.

[31] While Tagaq has taken inspiration from traditional Inuit throat singing, often using the main sounds of throat singing in her music, she is quick to note in interviews that what she does is very different from traditional throat singing in that she typically sings alone, her music features instruments, and she incorporates a variety of popular music styles. Traditional throat singing involves two women, face to face, where one takes a leader role and the other follows. The follower mimics the sound of the first person a split second later, "fitting together like puzzle pieces," to create a hocket (Red Bull Music Academy 2016). The roles of the two people may switch throughout the

In an educational video entitled "The Sounds of Throat Singing" from a course on "Words and Music" from the Open University (2009), Tagaq explains and demonstrates the throat-singing sounds that she frequently uses in her own recordings and improvised performances.[32] Tagaq differentiates her sounds in two ways, each of which is a binary opposition. The first is how she breathes while producing the sound: exhaling (E) or inhaling (I). The second is the overall effect of the sound: a low, deep growl (D) or a pure, high pitch (H).[33] Thus there are four types of sounds: 1) *Inhaling High* (IH); 2) *Exhaling High* (EH); 3) *Inhaling Deep* (ID); and *Exhaling Deep* (ED). These sounds are summarized in Figure 4.1.[34] The words in the triangles are the four elements that can be combined to create the sounds Tagaq uses. I indicate the sounds themselves by the abbreviations on the diagonal lines. The high pitch is a voiced, pure, head tone, whereas the low, deep sounds are unvoiced. Rather than using her vocal cords, Tagaq creates the deep sound by allowing her epiglottis to vibrate. And while Tagaq does not explicitly define audible *Breath In* as a "sound," it is included in the following discussions, since it is crucial to the overall effect of Tagaq's singing.[35]

---

performance—in other words, the follower may become the leader. Although Tagaq has since learned traditional songs and has participated in this style of singing, she was originally self-taught, listening to tapes of throat-singing singers from her hometown while she was away at university and trying to reproduce the sounds she heard on the recordings. This unconventional method of learning throat singing and her interest in a wide variety of musical styles has led to the distinctive, original sound of her music.

[32] "The Sounds of Throat Singing" video may be located in Apple Podcasts or on the Open University website: https://www.open.edu/openlearn/history-the-arts/music/inuit-throat-singing?trackno=5.

[33] Low may be the more obvious opposite of high, but I identify the low, deep sounds with a "D" because in the video describing these sounds, Tagaq primarily refers to these as the "deep sounds."

[34] Although there is a stipulation on the website that houses "The sounds of throat singing" video, stating that the material was published twelve years ago and information provided may be out of date or inaccurate and any views or opinions expressed may no longer be relevant, many people have referred to throat-singing sounds in similar ways. Beaudry (1978) differentiates the sounds of throat singing using inspiration/expiration and voiced/unvoiced. Here I use the terms that Tagaq herself uses to describe her sound production, though they are analogous to the terms used by Beaudry. Beaudry's voiced sounds map onto Tagaq's high sounds, and unvoiced equates to deep for Tagaq. Tagaq puts an emphasis on the pitch of the sound, whereas Beaudry puts an emphasis on how the sound is produced, whether it is produced by the vocal cords or the vibration of the epiglottis. Stévance (2010, 86), like Beaudry, uses the terms "voiced" and "unvoiced," but she also refers to the opposition as "high sounding" and "dull."

[35] Beaudry (1978) and Nattiez (1999) also note the importance of breathing in throat singing.

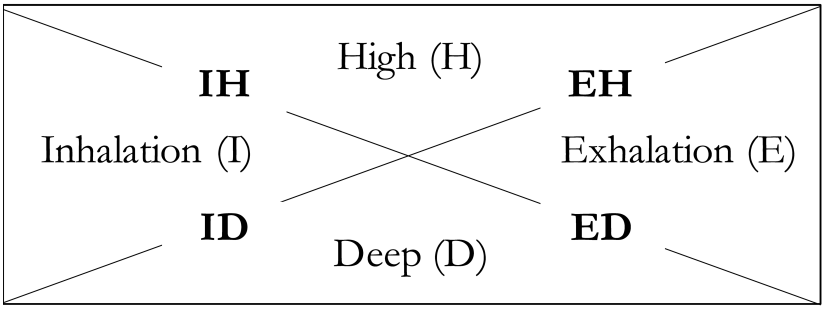


Figure 4.1 The four sounds of Tanya Tagaq's throat-singing–inspired music.[36]

Figure 4.2 provides three spectrograms of excerpts from "The Sounds of Throat Singing." In the video, Tagaq demonstrates the four main sounds by performing them cyclically in pairs—what I will refer to here as *sound patterns*—with minimal changes in pitch and dynamics.[37] The frequency bands with the most energy appear darker than those with less energy.[38] Examination of the spectrograms from the three demonstrations shows that one can parse the spectrogram into the different sounds. The difference between the low, deep growls (D) and the pure, high pitches (H) is particularly noticeable, with more energy distributed in the lower frequencies for the deep sounds than for the high sounds. In addition, the deep sounds typically feature more lower frequencies than the high sounds, albeit less clearly defined. Conversely, the differences between EH and IH are less noticeable—as are the differences between ED and ID—when examining the spectrogram.

[36] This figure is adapted from one in Beaudry (1978, 271).
[37] The original interview is a video recording that was published in mp4 format. Thus, the audio files used for this study are lossy. *Lossy* compression formats use acoustics to remove the parts of sound that have the least effect on perceived quality. Therefore, the range of frequencies in this audio file only goes to about 16kHz, as that is the frequency threshold with which there is no loss in timbral quality according to acoustic measurements. However, when converting the video audio into a strict audio format, I used FLAC files to prevent any further loss of data.
[38] The spectrograms were created in Praat. The view range is from 0 to 5000 Hz, the window length is 0.01s, and the dynamic range is 70dB.

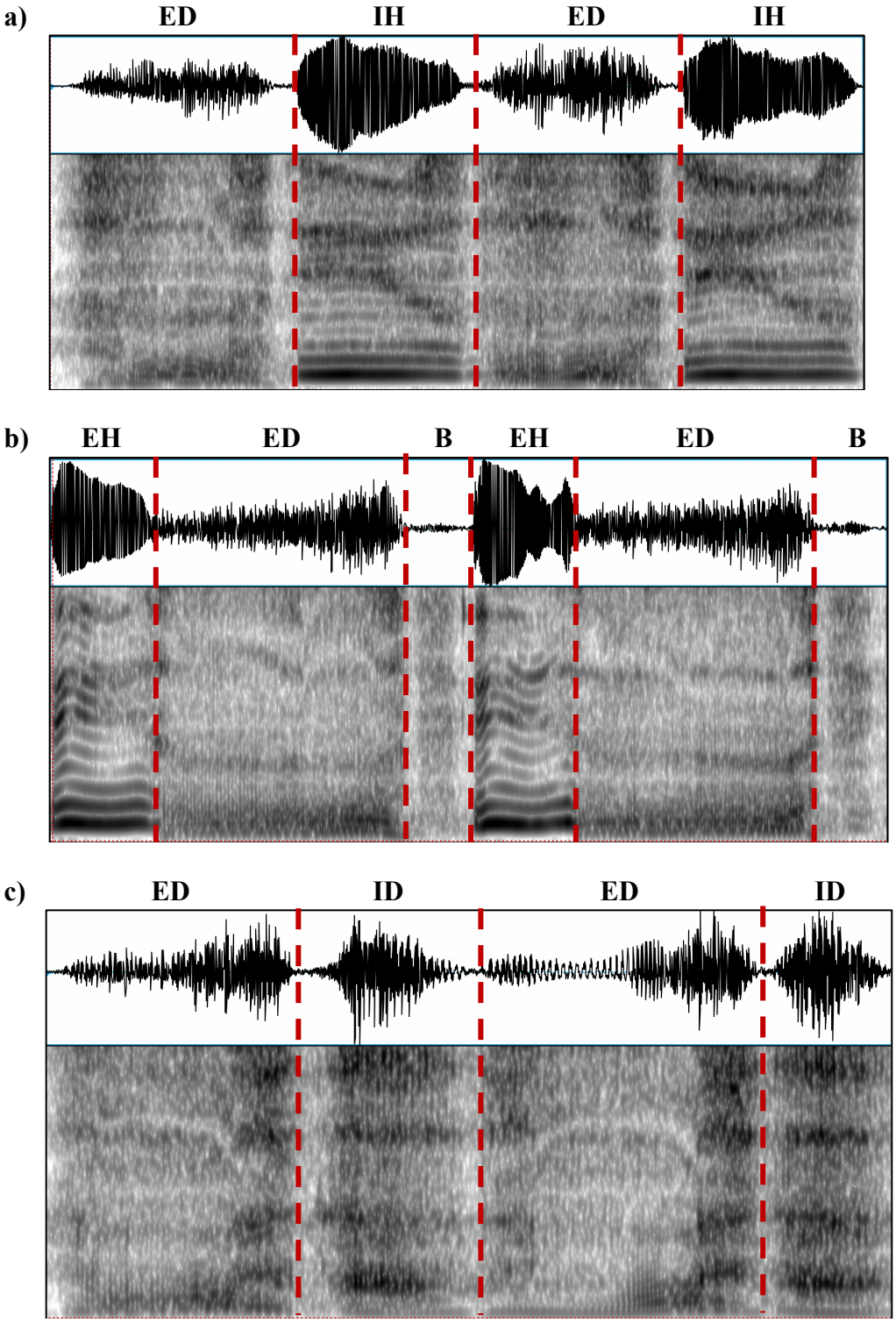


Figure 4.2 Spectrograms of Tagaq's throat-singing demonstrations of a) exhaling deep, then inhaling high; b) exhaling high followed by three exhaling deep sounds; c) exhaling deep, then inhaling deep.

In Tagaq's music (and in traditional Inuit throat singing), the individual sounds I've described here are never performed alone; they are always combined with one or more sounds and repeated cyclically until a new sound pattern is introduced. Nevertheless, an individual examination of each sound shows that there are nuanced distinctions in their timbre. (More detailed examinations of sound patterns appear later in this chapter and in Chapter 6.)

4.4 Analyzing Sound Qualities: Timbral Descriptors and Acoustic Descriptors

4.4.1 Timbral Descriptors and Opposition Chart

The timbral descriptors used here are borrowed from Lavengood (2020), who has updated the terms from Cogan (1984) to better match those used by audio engineers and the average listener. Table 4.1 illustrates a timbral opposition chart using spectrogram analysis as per Cogan and Lavengood for comparing the sounds of Tagaq's throat singing.[39] These oppositions are assigned a contextual value of either a negative sign (-) or a positive sign (+).[40] The negative sign is used when the sound corresponds to the descriptor to the left of the slash, and a positive sign when it corresponds to the descriptor on the right. Cogan allows for more nuance within these binaries with two additional symbols: 1) if the sound seems to have both properties, a mixed symbol may be used (+/-); and 2) if the sound does not display either descriptor, it may be assigned a null value (ø).

[39] Cogan and Lavengood use many more oppositions than I do here. I focus on these five oppositions because they correlate to acoustic measurements easily computed using VAMP plugins in Sonic Visualiser.
[40] These signs are arbitrary, though I have deliberately chosen descriptor words that, to my mind, bring an association with either + or -. Thus, in Table 4.1, loud sounds generally have more energy and are indicated by a positive sign, while soft sounds have less energy and are indicated by a negative sign.

Table 4.1 Timbral opposition chart for the sounds used in Tagaq's demonstrations of throat singing, based on what is seen in the spectrogram and heard while listening to her demonstrations.

Opposition (-/+)	**Breath In**	**Exhaling Deep**	**Inhaling Deep**	**Inhaling High**	**Exhaling High**
Low/High	ø	-	-	+	+
Unvoiced/Voiced	-	-	-	+	+
Soft/Loud	-	+	+	+	+
Noisy/Pure	-	-	-	+	+
Dull/Bright	+	-	-	+/-	+/-

In Table 4.1, I assign these contextual values to five opposition categories for the six different sounds. The categorization of Tagaq's throat-singing sounds is first and foremost based on the type of pitch produced—that is, whether it is a high pitch or a low pitch. Second, a sound may be categorized based on whether it is voiced or unvoiced. One produces a voiced sound when creating the sound using the vibration of the vocal cords. All vowels are voiced in the English language, but consonants may be voiced or unvoiced. For example, the "g" in "gate" is voiced, while the "K" in "Kate" is unvoiced. This plays a major factor in the classification of Tagaq's sounds. Tagaq produces the high sounds with her vocal cords—i.e., they are voiced—but she produces the deep sounds by allowing her air stream to vibrate her epiglottis—i.e., they are unvoiced.[41]

Tagaq's sounds can also be classified based on the general dynamic of the sound, i.e., whether the sound is relatively loud or soft. Based on the spectrograms in Figure 4.2 and in listening to the demonstrations, I have indicated that the four main sounds are loud, in comparison to Breath In. Although these three opposition categories are not timbral indicators

[41] For more information on the physiology and ways of describing the timbre of a singing voice, see Heidemann 2016.

per se, they can affect the quality of the sounds and may influence how they are perceived (Lavengood 2020). I group them under the broader umbrella of sound qualities.

In addition, Table 4.1 illustrates two spectral oppositions of a sustained sound: pure/noisy and bright/dull. The pure/noisy opposition refers to the difference in frequencies present in the sound, as observed in Figure 4.2. A sound is *pure* if the frequency bands are distinct in the spectrogram, but *noisy* if the bands are less distinct or more closely grouped together. For example, a flute produces a pure sound while a snare drum produces a noisy sound. In Tagaq's vocal style, the high sounds have a pure quality, while the deep sounds have a noisy quality.[42] For the bright/dull opposition, the number of higher frequencies affects the timbral quality. If there are numerous higher frequencies above the fundamental frequency, the sound will be *bright*, but if there are fewer higher frequencies heard above the fundamental, the sound will be *dull* (Grey 1978). In Table 4.1, Breath In is bright, while the deep sounds are dull, and the high sounds have a mixed quality.

In his seminal 1984 book *New Images of Musical Sound*, Cogan drew heavily on the ideas of linguists, particularly Jakobson and Waugh, and on philosophers such as C.S. Peirce, who argue that we understand the world through binaries (for example, hot/cold). Cogan modeled his table of oppositions on those used in phonology and adapted them for musical timbre. Although the opposition chart in Table 4.1 shows major differences between the deep growls and high pitches, it does not indicate much of a difference between ED and ID sounds, even though the distinction is audible in Tagaq's demonstration, nor does it indicate a distinction between IH and EH sounds.[43] This is a problem, not only because the difference is audible, but

[42] With the voice, if a sound is pure, it will also be harmonic; if it is noisy, it will be inharmonic. Therefore, the harmonic/inharmonic category will not be discussed here.
[43] Cogan (1984) has more steps to his opposition-chart method of analyzing timbre than I have used here. After completing the chart, he totals up a score based on the positives, negatives, and mixed symbols for each sonic

also because throat singers understand the differences between exhalation and inhalation to be elemental to the sound of throat singing (Beaudry 1978, 266). So, while we understand the world through oppositions, we likewise do so by means of comparatives (for instance, "the coffee is warmer than the tea") and superlatives ("the coffee is the warmest"). A method of analyzing sound qualities that incorporates comparatives and superlatives is necessary to understand the subtle differences between Tagaq's throat-singing sounds. Acoustic measurements of the individual sounds from the demonstration video, which can be obtained using Sonic Visualiser and various VAMP plugins, assist in creating a more robust picture of the sonic distinctions and allow for comparative descriptions of the various sounds.

4.4.2 Western and Inuit Conceptions of Oppositions

In Western society, we likely will associate "pure" with "good" and "noisy" with "bad." It might also be noted from the above discussion that the high sounds, which are more familiar tones for Western ears, are labelled as "purer" than the deep growls, which are perhaps less familiar to Western ears and might even be considered as "not as nice sounding." But these value judgments are ones that Euro-Canadians and -Americans ascribe to these labels. Based on an interview conducted with the members of PIQSIQ, an Inuit style throat singing duo, Inuit would not likely make these value judgments. When asked about their album *Taaqtuq Ubluriaq: Dark Star* (2020), Kayley Inuksuk Mackay and Tiffany Kuliktana Ayalik describe why strict oppositions and associating oppositions with good/bad are not as prevalent in Inuit culture:

> **Mackay:** With this story, we were looking to challenge the really strict dichotomies between light and dark. So, growing up in the Arctic, we have some extremes, one of them being weather, the other being daylight. So, in the summer we have 24-hour sunlight, and in the winter we have nearly 24-hour darkness, and in some of the more

moment. He then graphs these sums, showing the "directed spectral flow." These additional steps, however, will not elucidate differences between the low growls or between the high, pure pitches, since they have been analyzed identically for each similarly pitched sound.

> northerly communities, total darkness for a couple of months in the winter. So, we're used to navigating extremes and not putting a value on them, they just are. And so, I think that's a little bit of the culture clash between Inuit culture and settler Christianity coming over. You know, there's, in settler culture, especially being rooted in Christianity, there's quite the dichotomy between light and dark and light is good and dark is bad. And in the Arctic, we just can't afford to see it that way.
>
> **Ayalik:** Like half—more than half our life is bad [*laughs*].
>
> **Mackay:** And so, we learned to navigate extremes and to navigate darkness and to appreciate light, but also to see beauty in the dark…
>
> **Ayalik:** And with a lot of Inuit storytelling, it's hard to find the hero story. It's hard to find the classic kind of Aristotelian model of narrative. It's not, you know, a story rising action, climax, and then everything is ta da at the end. The heroes in one story, in the next are the villain. The people are complicated and complex and, you know, are hypocritical at times, if you're looking at it from a really simplistic lens, and there's not this obsession with the sort of hero's journey and triumphing over evil. It's at the end of the story a lot of times if—sometimes even if you're asking an elder a question, they might tell you a story and at the end you might be like, ah, well jeez, I know less now than I did before, because it's like just tell me what—and that's sometimes not the answer, you know. There's a lot of teaching through story and I think that being brought up with some of those stories really helps to highlight that things are complicated and that there isn't this super clear, you know, light and dark, black and white; everything is like shades of complexity that you have to navigate and like trust your intuition and be comfortable with. (Library of Congress 2021, 20:43–23:39)

With this in mind, I implore anyone reading this dissertation to suspend any value judgments that might be brought up by the pure/noisy descriptors (or the bright/dull descriptors for spectral centroid) used throughout. Try to see the beauty in both the purer, high head tones and the noisier, deep growls.

4.4.3 Acoustic Descriptors and Continua

When other forms of comparison, including comparatives and superlatives, are considered, it is necessary to present a range of values across a variety of descriptors. The acoustic measurements of Tagaq's throat-singing sounds provide continuous quantitative data, and through experimenting with these data, I investigate my intuitions that timbre is too complex

a phenomenon to be represented with binary oppositions.[44] Therefore, instead of binary oppositions, I present these data over a series of continua of relative values that visually convey these degrees of nuance. While I continue to use the binary descriptors of Cogan (1984) and Lavengood (2020), this is a departure from the method of timbral analysis discussed above that focuses on binary oppositions—what might be called either-or comparisons.

To enable a more specific comparison between Tagaq's sounds, I use acoustic measurements, specifically RMS energy (loudness), periodicity (noisiness), and spectral centroid (brightness). To obtain the acoustic measurements, I first imported the audio into Sonic Visualiser (Cannam et al. 2006) and applied one VAMP plugin for each of the descriptors.[45] Second, I identified the onset timings of the sounds, using the combination of a note onset–detection plugin, the eye, and ear. I found the note onset detectors helpful in giving clues to where potential onsets occurred, but these plugins were rather unreliable in terms of accuracy. To augment the data provided by the plugins, I also visually looked for changes in the waveform and spectrogram to identify note onsets, and then checked my timings by moving the starting position of the audio and listened for where the best onset location seemed to be. Third, I exported the data into a spreadsheet and parsed it based on the onset timings identified in step two. Fourth, I repeated the first three steps for all of Tagaq's demonstrations. Fifth, I combined the data into a single spreadsheet and created two types of graphs: 1) box and whisker graphs to compare the central tendencies of different individual sounds; and 2) line graphs showing all of the data points over time, which I use to visualize the connections between changes in sound

[44]Several of the measurements that can be used for timbral analysis, including spectral centroid, are based on what we know about how listeners perceive sounds (Devaney, in press).
[45] See Salamon and Gómez 2014 for information regarding the MIR.EDU plugins used for RMS energy and spectral centroid. See Mauch and Dixon 2014 for information regarding the plugins used for periodicity. For each plugin, I used the default settings.

qualities with the beat structure and form of Tagaq's songs (form is discussed in Chapter 6). Sixth, and finally, based on these graphs, I constructed continua that allow for a comparative analysis of the sound qualities for Tagaq's throat-singing sounds. This systematic procedure resulted in graphs that illustrate the differences between the sounds that I could hear but did not previously have a way to illustrate.

4.4.3.1 RMS Energy (Loudness)

Figure 4.3 gives an example of the two types of graphs used throughout this chapter.[46] Above the graphs is a waveform representation of Tagaq's demonstration of Exhaling Deep followed by Inhaling High. This demonstration is also illustrated with a spectrogram in Figure 4.2a. The line graph in Figure 4.3 plots the RMS energy values provided by Sonic Visualiser and the MIR.EDU plugin for RMS energy for the ED and IH sounds in this demonstration.[47] This line graph signifies two aspects of the sounds: first, how much energy each sound contains; and second, the consistency of energy across the sound. The energy level correlates to the amplitude of the waveform. The higher the amplitude on the waveform, the larger the RMS energy value at that moment.[48] ED's RMS has a more jagged shape, which corresponds to a more rapid variation in energy levels, whereas IH has its peak energy level near the beginning of the sound and its RMS is variable overall.

The lower graph in Figure 4.3 summarizes this information by extracting the central tendency of the data points of the two sounds in the line graph. The boxes indicate the value of

[46] Chapter 5 also features box and whisker graphs, and Chapter 6 features line graphs.
[47] RMS energy was measured using the default settings of the MIR.EDU plugin by Justin Salamon. This plugin has no adjustable parameters. The mean of the source channels was used. The audio frames per block and window increment were both set at 1024.
[48] Panagiotakis and Tziritas (2005) use RMS energy to distinguish between silence, speech, and music signals. They note that RMS energy is nearly 0 for silence, more consistently high for music, and rather variable for speech.

the 25th percentile and the 75th percentile. The line inside the box indicates the median value, while the "X" indicates the mean value. The lines extending out from the boxes indicate the fifth and 95th percentile of the values. The plots show that IH has more RMS variability overall than ED because it has a longer box and thus a wider range of values, confirming what was observed in the continuous line plots. The sounds can be positioned on a continuum of energy, as shown to the left of the lower graph, to summarize this difference; however, due to variability of RMS across instances of both sound types, there is not a significance difference in RMS between ED and IH. IH might have slightly more energy overall compared to ED, since the median line of IH lies just outside the upper line of ED's box. For that reason, IH is placed slightly higher on the continuum for RMS energy than ED.

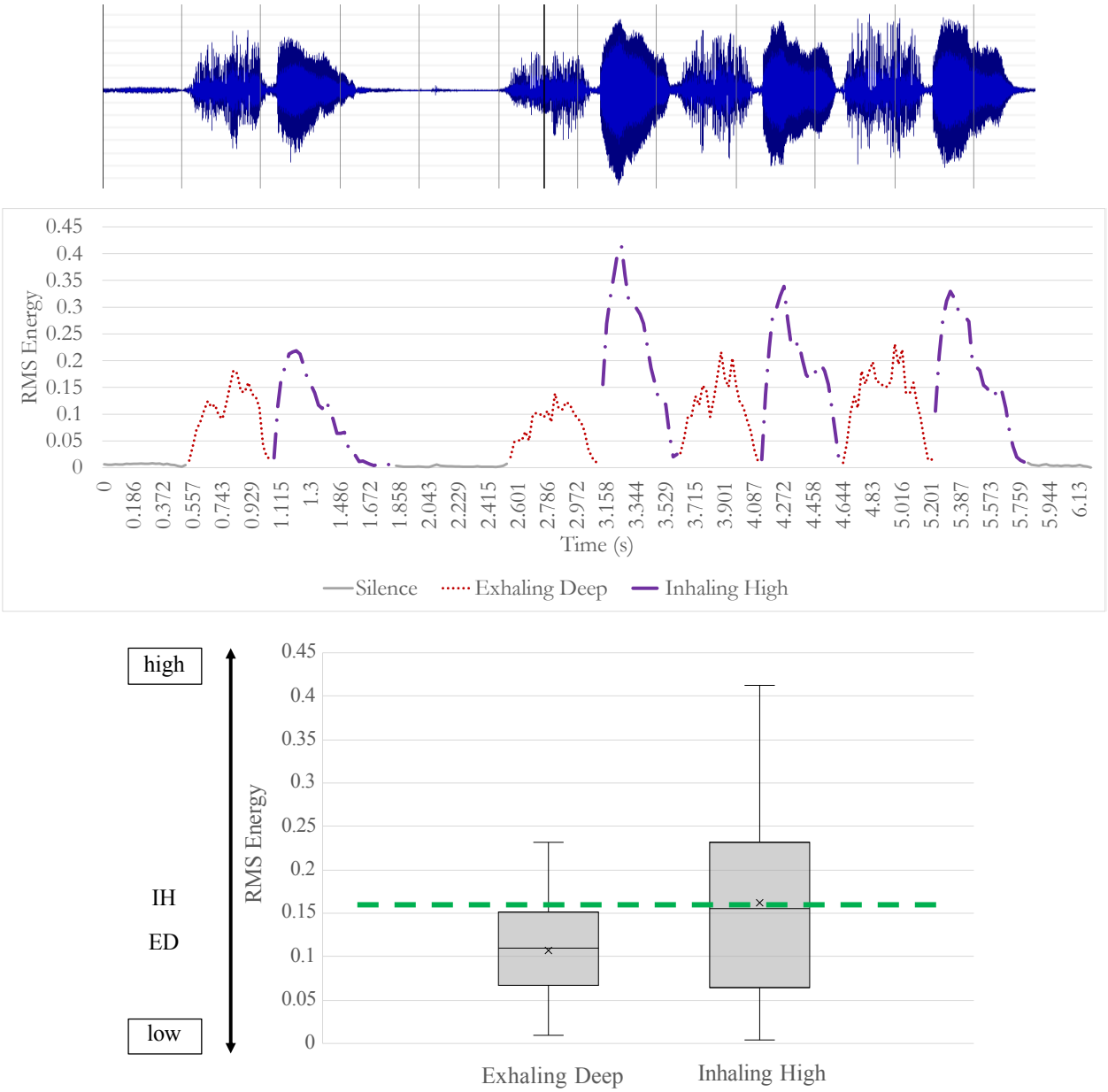


Figure 4.3 Top: Waveform of audio with a sound pattern of Exhaling Deep–Inhaling High. Middle: RMS energy plot, with data obtained using the MIR.EDU RMS energy plugin. Bottom: The central tendency of RMS energy values for Exhaling Deep and Inhaling High sounds, corresponding to the top graph. The green dashed line shows that the median value of IH lies outside the box of ED.

Figure 4.3 shows the results of the RMS measurement in the middle plot for a demonstration of the sound pattern Exhaling Deep–Inhaling High, and a statistical analysis in the lowest plot. Similar plots summarizing data obtained from five of Tagaq's demonstrations from

"The Sounds of Throat Singing" were created and replicated twice more for each of the other descriptors. The amount of information provided by the plugins is extremely unwieldy, and so the focus for the remainder of this section will be on the central tendency values for each of the different sounds within each of the different descriptors across the five demonstrations of sound patterns, as illustrated in the three graphs in Figure 4.4.

Figure 4.4a graphically illustrates the central tendency of RMS energy values for all five sounds. Contrary to the earlier opposition chart, which showed that Tagaq's four throat-singing sounds are all loud, this graphic of RMS energy shows some differences in loudness among the sounds. Breath In (BI) has the least amount of energy, and Exhaling High (EH) has the most. These sounds are placed at the edges of the energy continuum to the left of the box and whisker plot. Exhaling Deep (ED) also clearly has a lower amount of energy than EH, but a higher energy level than BI. This is seen in the graph because the median line of ED lies outside the boxes of BI and EH, sitting above the BI box and below the EH box. The difference between Exhaling Deep, Inhaling Deep (ID), and Inhaling High (IH) is less clear, since the median values of these sounds overlap with the other boxes. All three sounds have less energy than EH and more energy than B, so they are placed in the middle of the energy continuum. Although ED, ID, and IH effectively have similar energy levels, there are some small differences in the way that they relate to each other that I use to determine their placement on the energy continuum. The median line of ED lies within the box for IH. Likewise, the median line of IH lies within the box of ED, indicating that these sounds are not significantly different in overall energy level. This is highlighted in the graph with a red dotted line. The median line of ID lies within the box of IH but outside the box of ED, which indicates that there may be a difference between ED and ID in terms of overall energy. There is no difference, however, between ID and IH in terms of energy.

The more significant difference between these three sounds is the distribution of values. ID and IH both have wide distributions and variable data, while ED has a smaller distribution, indicating that ED has more consistency among values.[49] Taking all of these observations into consideration, I have placed these three sounds on the continuum in the order of their mean values from lowest to highest—ED, IH, then ID. Because of the likely difference in distribution between ID, IH, and ED, I have grouped together ID and IH in a box on the continuum.

[49] The dots above the 95th percentile of B and ED represent outliers.

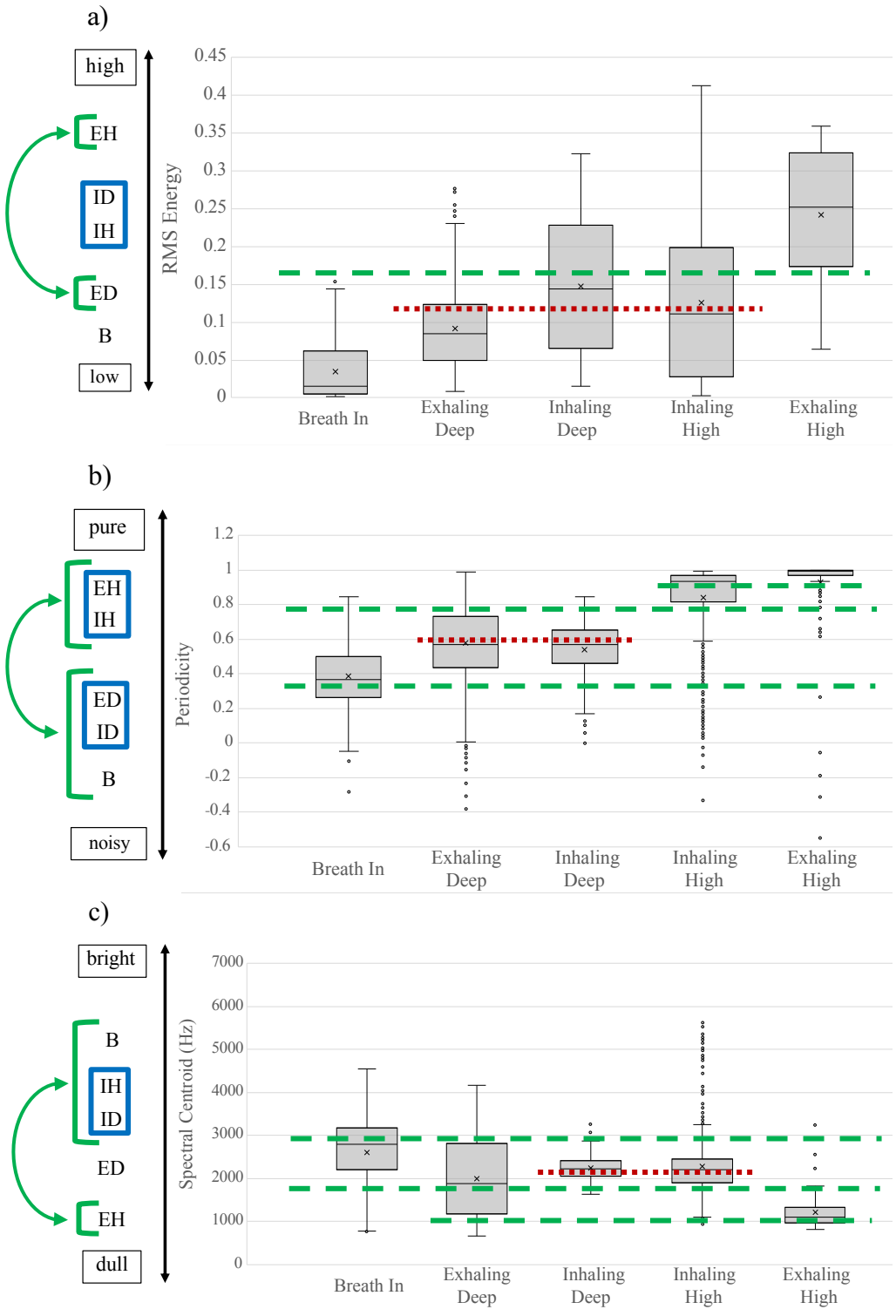


Figure 4.4 Central tendency of the a) RMS energy, b) periodicity, and c) spectral centroid values for the *five* different sounds heard in Tagaq's demonstrations. Green lines show that median values lie outside the boxes of other sounds. The red lines show overlap between median values and boxes. The green brackets and arrows connect sounds that are definitively different.

4.4.3.2 Periodicity (Noisiness)

In Figure 4.4b, there is a noticeable difference between the high sounds and the deep sounds in terms of periodicity, but less of a difference between sounds with similar pitches.[50] The high sounds are purer than the deep sounds, and Breath In is noisier than Tagaq's four throat-singing sounds. With this in mind, I placed the high sounds near the pure end of the periodicity continuum and the deep sounds near the noisy end, with BI below the deep sounds. The median value of IH lies outside the box of EH as well, indicating that it likely has a lower overall periodicity than EH. I place EH closer to pure on the continuum than IH because the median value of EH is slightly higher than that of IH. It is also notable that EH has a much narrower distribution than IH, indicating less variability in the EH values and that it is likely purer than IH. While the differences between the deep sounds cannot be determined by examining the median values of these sounds (indicated with the red dotted line), the distribution of values indicates some difference. ID has a slightly smaller distribution than ED. Even though these sounds are essentially the same in periodicity, I place ED higher in the box with ID on the periodicity continuum, because the values of ED in the box reach closer to the pure end of the continuum than those of ID.[51]

4.4.3.3 Spectral Centroid (Brightness)

In Figure 4.4c, there is a noticeable difference between the inhaling sounds and the exhaling sounds in terms of spectral centroid, but less of a difference between sounds that are

[50] Periodicity was measured using the default settings of the Yin plugin by Mauch. The Yin threshold was set at 0.1500. Audio frames per block were 2048 and the window increment was 256.

[51] Zero-crossing is another indicator of noisiness. The zero-crossing rate (ZCR) is "defined as the number of times the audio signal waveform crosses the zero-amplitude level during a one second interval, which provides a rough estimator of the dominant frequency component of the signal" (Alías et al. 2016, 9). If the ZCR is high, the sound is noisy, whereas if it is low, the sound is tonal (Panagiotakis and Tziritas 2005). If the sound signal is periodic, the ZCR is unnecessary since it is roughly related to the fundamental frequency. Thus, the continuum of noisiness using the ZCR will mirror the continuum of noisiness established in Figure 4.4b with the measurement of periodicity.

produced through inhalation and between those produced via exhalation.[52] Breath In has a higher median value than all of Tagaq's throat-singing sounds, and it lies outside the boxes of the other sounds. For this reason, I place BI closest to bright on the spectral centroid continuum. The inhaling sounds are located just below BI on the continuum, because the median values of these sounds lie between those of BI and the exhaling sounds. They are also boxed on the continuum because they are essentially the same in terms of spectral centroid. Their median values are the same and lie within the box of the other sound, as indicated with the red dotted line on the spectral centroid plot. The distribution of values for ID is slightly narrower than that of IH, indicating more consistency in the values of ID. Because of this, I placed IH higher within the box on the continuum than ID. IH is more variable and might be heard as slightly brighter in certain instances than ID. While the median lines of the inhaling sounds lie within the box of ED, the ED median value lies below the boxes of the inhaling sounds. The median value of EH is significantly lower than the boxes of the inhaling sounds, but only slightly below that of ED. Additionally, the distribution of values of EH is much narrower than ED, indicating a consistency in the EH sounds toward a duller sound quality. I, therefore, place EH closer to the dull end of the continuum than ED.

Contrary to the periodicity continuum, where similar pitches were grouped together, the brightness continuum shows a similarity among sounds produced via inhalation compared to those produced via exhalation. Bader (2013) notes that both spectral centroid and pitch contribute to the perceived brightness of a sound. If a pure sinusoidal tone is played in a low register, it will sound less bright than in a higher register (Bader 2013, 355). Contrary to this

[52] Spectral centroid was measured using the default settings of the MIR.EDU plugin by Justin Salamon. This plugin has no adjustable parameters. The mean of the source channels was used. The Hann window shape was used with 1024 audio frames per block and a window increment of 512.

perspective, based on this plot of spectral centroid, the amount of breath heard in the sound seems to be a greater factor in the perceived brightness of a sound than pitch. The inhaling sounds, regardless of pitch, are brighter than their exhaling counterparts. The more air heard in the sound, the brighter the sound will be.

4.4.4 Summary of Sound Qualities

As the reader will recall, the opposition table (Table 4.1) suggests that there are noticeable differences between the two pitches—that is, the low growl and the high, pure tone; but there is no difference in that analysis between the exhaling and inhaling sounds with similar pitches. In the alternative analytical method that I have just presented, the three continua illustrate subtle, but important and audible, differences between the four sounds of Tagaq's throat-singing passages (plus BI). For loudness, the greatest difference in energy occurs between EH and ED. The ID and IH sounds are essentially the same. This is contrary to the opposition chart for loudness, which indicates that all sounds are equally loud. For noisiness, the continuum presented corresponds well to the opposition chart. Both show that pitch plays a large role in our perception of the noisiness of a sound. On the other hand, the inhaling versions of sounds might be said to be slightly noisier than their exhaling counterparts, as indicated by the periodicity continuum. Additionally, the brightness continuum contrasts significantly with the analysis in the opposition chart. Rather than grouping sounds based on pitch, the spectral centroid values indicate that the inhaling sounds are brighter than the exhaling sounds overall. As mentioned above, the amount of breath heard and the physical exertion required to produce the inhaling sounds seems to contribute to their overall brightness. For inhaled sounds, Tagaq draws in breath while also performing the sound, which takes more effort than exhaled sounds and gives the sound an airy quality that is less full than the exhaled variant. The airiness gives these sounds a

noisier and brighter effect than their exhaling counterparts. The acoustic measurements and continua give a fuller impression of the qualities of Tagaq's throat-singing sounds than can be shown in an opposition chart. The nuances in the qualities are possible only once the sounds are compared to one another on continua, rather than in opposition.

4.5 Analyzing Movement as Feeling Beats: The Methodology

From the beginning of this study, there were two objectives I had in analyzing beats in "Ilunikavi." Tagaq's perspective is extremely important to bring into any discussion of her music, so I attempt to bring her voice into the analysis by analyzing her movements. Another aspect that I wanted to address included avoiding analyzing beats as a purely intellectual exercise. I wanted to see how I *felt* the beats. With these objectives in mind, I decided to analyze the movements that Tagaq and I make while performing or listening to "Ilunikavi." I therefore used a video recording of Tagaq performing "Ilunikavi" to observe how she feels the beat of the song. I also recorded a video of myself moving to the same performance of "Ilunikavi" to observe how I feel the beat of the song.

This analytical approach was inspired by the methodology described in Rosa Abrahams's 2019 article "Mimicry as Analysis." While working with Jewish communities, Abrahams developed a method to assist in analyzing live music and movement that cannot be recorded for cultural reasons. She developed this method because she was working in a religious setting, and it was inappropriate to record (video or audio) the participants in such a setting. Through body mimicry, i.e., learning to move as the participants move, Abrahams gained an empathic understanding of the participants' movements. She then used these observations to analyze meter in chants from Jewish worship settings. Abrahams demonstrates an approach to analysis that is ethical and respectful of the people with whom she works and about whom she writes. I use

Abrahams (2019) as inspiration for this study because she looks more broadly at improvised movement to music, rather than choreographed dance.

There are some key differences in my approach compared to Abrahams's. In contrast to Abrahams, who used mimicry to get a sense of how participants moved to the music, I actively tried to move to Tagaq's music without having recently observed her movements. I recorded myself moving to Tagaq's live performance of "Ilunikavi" more than six months after last seeing or hearing the recording of the performance. This was to ensure that my movements would be minimally influenced by Tagaq's movements. I then used the application Final Cut Pro to create a split-screen video of the recording with my movements next to Tagaq's visual performance.[53] This application also allowed me to complete a micro-timing analysis of the movements by slowing down the video while starting and stopping at moments where beats were indicated. The data were then entered into an Excel spreadsheet and compared to the acoustic measurement data obtained in Sonic Visualiser.

In analyzing movement as it relates to music, we must first ask: what types of movement constitute demarcating a beat? Abrahams confronts this issue in her 2019 article. She identifies three main aspects of motion and effort in movement: the locus, the ictus, and the trajectory. For the purposes of this study, I focus on the ictus. The ictus, as Abrahams defines it, is "where a participant's movement reaches its outermost point before reversing or ceasing" (2019, 50).[54] This is the point at which I have assumed a beat to be indicated. For my movements, that was the point at which my hand was lowest before reversing direction. This is illustrated in screenshots provided in Figure 4.5. In some cases, my ictus was understood to be the moment when I

[53] This video can be located at the following link: https://youtu.be/pcOvqwmP6EU
[54] John Paul Ito (2020) has developed a similar concept of *focal impulses*, which are bodily motions that indicate felt beats at a chosen main beat level.

snapped my fingers, which might or might not align with reversing the direction of my hand movement. For Tagaq, I also assumed that moments where she reversed direction were icti. However, identifying Tagaq's icti was generally more challenging. Tagaq's movements, like most performers, were more gestural than a clear demarcation of a beat. Most of the time, this entailed looking for moments when Tagaq reversed the direction of her hand movement from left to right or from down to up. One such motion is illustrated in Figure 4.5. Notably, Tagaq's motion up and down with her right hand is much subtler than mine. The middle picture, boxed in red, is the moment of ictus in Tagaq's right hand. In moments where she made a circular motion with her hands, I focused on the moment where her movement ceased briefly before continuing with the circular motion. I sometimes also noted the ictus in her head movements when she would move her head from left to right. Her movements were less exaggerated than mine; sometimes the movement right before she reversed or ceased motion was unclear. If I was at all unsure of an ictus, I did not record a time marker for Tagaq's movement. This led to a large section where I have almost no data for Tagaq's icti.[55]

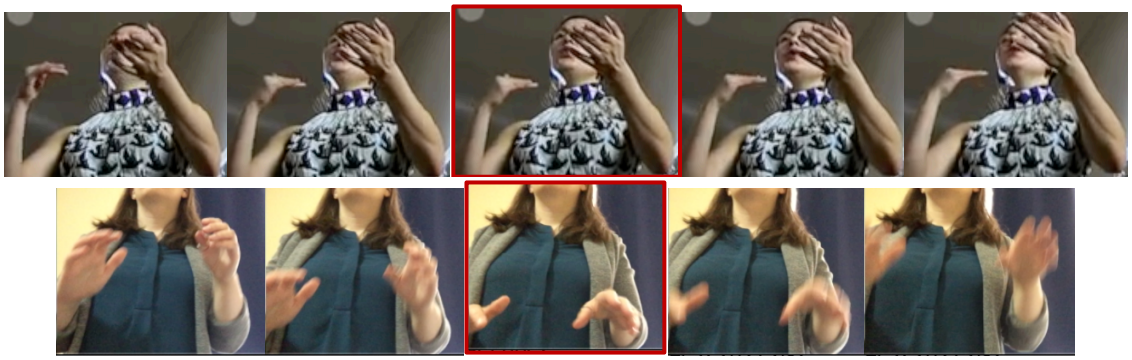


Figure 4.5 Screenshots detailing movements to and from the ictus for Tagaq (top row) and myself (bottom row). The box outlined in red is the ictus.

[55] Due to the lack of data, this section of the song (approximately 103.5–118.5s) is not discussed in this chapter.

Differences abound in the way Tagaq and I move to "Ilunikavi," which made comparing our movements challenging. While I was actively trying to indicate a beat through an up/down motion with my hand and snapping my fingers at moments where I was feeling a beat, Tagaq was performing the song, and she may or may not have been actively attempting to indicate a beat.[56] Her movements, like most performers, are more gestural than beat-oriented in many parts of the performance. Tagaq changes her movements regularly throughout the performance, moving her head from side to side, her hands circularly and from side to side, and her fingers at different points. Also, the editing of the video obscures Tagaq's movements in some parts, cutting to a different camera angle during what is likely an ictus. As such, our movements cannot be compared exactly. This should be taken into consideration while reading the following section. Observations can be made regarding similarities and differences only when there are clear icti by Tagaq. In the following analysis, I focus on passages where both of our markings of icti were clearly visible in the videos.

4.6 Observations: Beat Analyses of Tanya Tagaq's "Ilunikavi"

As alluded to earlier, in this section I make observations about the way that Tagaq and I experience beat during the middle part of "Ilunikavi," which features throat singing. The recording used for this study comes from a video published by the Open University (2009) in which Tagaq performs live and solo.[57] Here, the discussion focuses on how Tagaq and I mark beats during "Ilunikavi" through our icti. In comparing our icti, there are some striking similarities with regard to how frequently we align icti with onsets of sounds; but there are also

[56] When I play clarinet, I often mark the beat with the bell of my clarinet (sometimes unknowingly, I have been told), but other times I try to show the dynamic changes or other aspects of the music with my movements as I play. It is likely that Tagaq would do similarly while performing.
[57] To hear the version of "Ilunikavi" from her video series with the Open University, consult the following link: https://www.youtube.com/watch?v=vJlrZNrbMfc

differences, particularly in terms of duration between icti and the alignment of icti with the music. Because it was difficult for me to identify Tagaq's icti for a portion of the throat-singing part of "Ilunikavi," I focus on two sections, from approximately 76–104s and 118–160s.

While moving to "Ilunikavi," Tagaq and I both regularly aligned our icti with onsets, though we were rarely exactly in sync with the onsets. If using a threshold of within 70ms of the identified onset of a sound, 72% of Tagaq's movements fell within that threshold, while 84% of my movements did. If a 90ms threshold was used, 86% of Tagaq's fell within that threshold, while 93% of my movements did. The larger the discrepancy threshold used, the more often that icti can be considered to fall on the onset of a sound, so choosing a larger threshold than 90ms might skew data.[58] I was actively trying to mark a beat, so it is not surprising that the majority of my icti align with the onsets of Tagaq's sounds. Tagaq was performing, which means that she may or may not have been indicating a beat. The frequency with which Tagaq aligns an ictus with an onset of a sound, however, seems to indicate that she is marking some sort of beat. In sum, regardless of the potential discrepancy between onset timings and icti, Tagaq and I tend to feel the beat slightly differently than when a note onset actually happens. This needs to be considered when reading the observations below. As a general rule in what follows, I consider an ictus to align with whichever note onset is closest.

Though we both aligned our icti with the onset of sounds, failing to do so only a small portion of the time (if leeway is given to what counts as aligning with an onset), we differed

[58] As discussed above, it is very difficult to precisely identify the onset of a sound for various reasons, including how gradual the onset is and potential blending with a previous sound. While 70ms and 90ms are arbitrary thresholds, they allow me to take into consideration this discrepancy in the timing of sound onsets. It also shows that, for the most part, regardless of the sounds with which icti correspond, both Tagaq and I placed icti at the onsets of sounds, indicating that we were both marking a beat of some kind. This leeway is necessary because if I only consider those icti that have the exact same time markings as the onsets of sounds that I identified, only 11% of both Tagaq's and my icti would be considered to align with sounds. Certainly, both of us are much more capable of carrying a beat than that.

fairly substantially on how we missed onsets. I, more often than Tagaq, marked an ictus slightly after the identified onset of a sound. Nearly 75% of my movements occurred after the onset of a sound, while only 50% of Tagaq's movements occurred after the onset of a sound. That is, I regularly felt a beat after the onset of a sound, while Tagaq felt beats equally before and after the onset. This may have to do with the roles that we have—Tagaq as the performer and me as the listener. Tagaq is likely able to predict what is coming since she is the performer, whereas I take a more passive role as the listener. This difference between our ictus placements is discussed in further detail below and illustrated in Figures 4.7–4.10.

These observations correspond to Justin London's studies of entrainment, in which subjects tap along with a metronome at varying tempos. These studies have shown that "our synchronization systematically deviates from the stimulus onset, as people tend to tap 20–60ms before the metronome click" (London 2004, 12). The discrepancy of when people tap compared to the metronome click aligns with when Tagaq and I mark the beat during "Ilunikavi," although I expanded the threshold to 90ms. I used 90ms as a threshold because, in some cases, our icti were slightly farther away from the onset of the sounds than is suggested by these studies of entrainment. Tagaq's ictus placement also aligns with London's findings that people will often tap before the onset of a sound, whereas I often placed my ictus after the onset, in contrast to London's findings. It is also notable that these studies are based on observations from a controlled environment with a steady click of a metronome. Here, Tagaq and I are moving to actual music, which is not strictly metronomic. Despite the difficulty associated with marking a beat to music with a slightly fluctuating tempo, the majority of Tagaq's and my icti occur relatively closely, if not always exactly, at moments of note onset.

Not only did we differ in terms of the placement of our icti in relation to note onsets, but our beat lengths also differed. Figure 4.6 illustrates this observation by showing box and whisker graphs summarizing the interonset intervals between icti for Tagaq's movements and my own during the periods from 76–104s and 118–160s in the video recording. The length of the boxes in these graphs indicates the variability in the durations between our icti. While I maintained fairly consistent durations between icti, the durations between Tagaq's icti varied significantly. As the left graph in Figure 4.6 shows, I maintained a beat length of approximately 0.5s at the beginning of the middle (throat-singing) section of "Ilunikavi" (76–104s). During the last portion of the throat-singing section (118–160s), I doubled the time between icti to approximately 1s on average, as shown in the right graph in Figure 4.6. I was slightly less successful in maintaining a steady beat during this time, as shown by the slightly larger box in the right graph compared to the left graph, but the durations between Tagaq's icti were even more variable, with the majority of interonset intervals ranging from 0.5s to 1.4s throughout the parts of the throat-singing section examined. Even with this wide range in durations between icti, Tagaq's average alignment with onsets of sounds was similar to my own.

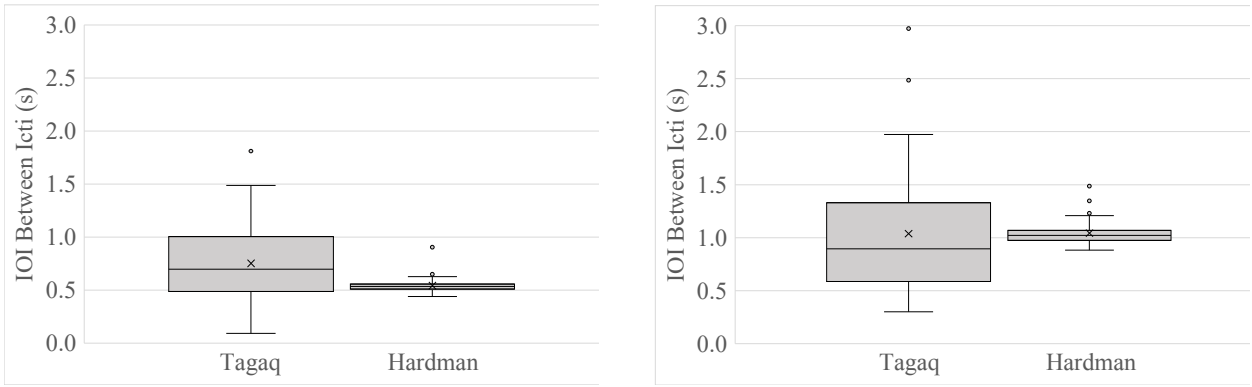


Figure 4.6 Comparison of durations between Tagaq's and my icti. Left: The first part of the throat-singing section in "Ilunikavi," 76–104s. Right: The last part of the throat-singing section in "Ilunikavi,"118–160s.

Further observations regarding the placement of icti can be made when considering the acoustic features of the recording. Figure 4.7 illustrates the timings of our icti and the corresponding RMS energy, periodicity, and spectral centroid data extracted from the audio from 76.30–80.78s. The timings indicated on the x-axis of the lowest graph are the onset timings for Exhaling High (EH) and Exhaling Deep (ED). These timings can be read vertically up the page for the other graphs. The colors help to orient the eye. Green timings correspond to the onset of EH, which is color-coded green and with long dots, while red timings correspond to the onset of ED, which is color-coded red and with short dots. Later examples have a similar structure.

At the beginning of the throat-singing section (76.30s), Tagaq and I placed icti with every note onset, though I did not align my icti with note onsets as well as Tagaq did. This is likely due to the difference between performer and listener. Tagaq plays with the time at this moment of the song, as she shifts from the texted melody of the first part of "Ilunikavi" to this throat-singing section. Because Tagaq is the performer, she knew when she was going to begin a new sound, and this would have enabled her to align her movements more closely with onsets of sounds. In contrast to Tagaq and her ability to mark the beginning of sounds with icti, I was a listener trying to guess the new tempo and had more difficulty. As noted above, I tried to maintain a steady pulse with equally spaced icti throughout the song, but this was a moment where the durations between my icti vary rather substantially, and I failed at aligning them properly with note onsets. Often, my ictus was later than the onset of the sound and even later than the peak energy level.

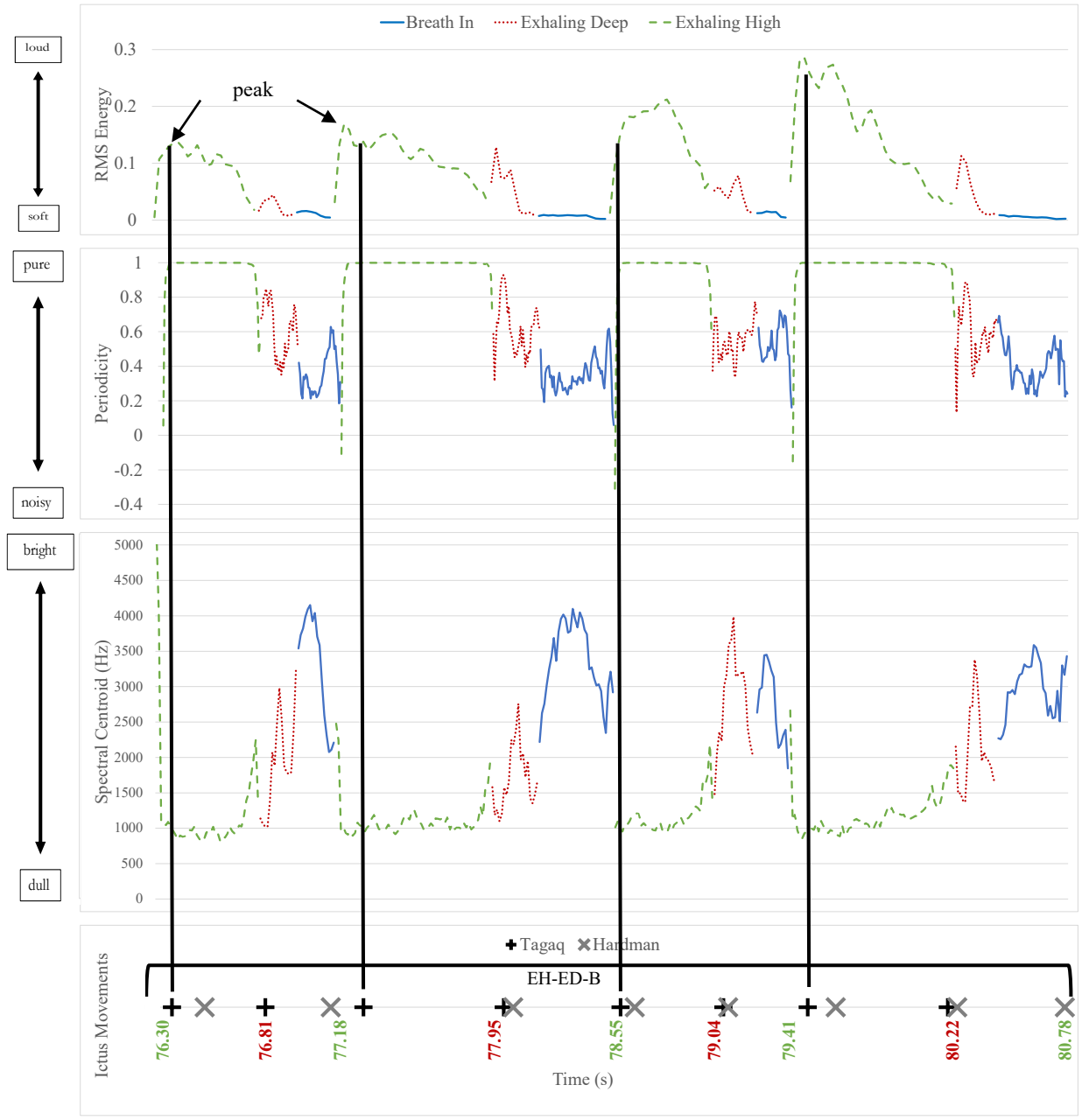


Figure 4.7 Ictus placement and corresponding RMS energy, periodicity, and spectral centroid data of audio for 76.30–80.78s of "Ilunikavi."

As I mentioned above, my hypothesis for this study was that loudness, noisiness, and brightness affect our experience of beat in "Ilunikavi." At the beginning of the throat-singing section (76s), however, this theory fails, at least for my icti. Both Tagaq and I marked an ictus with every onset of a sound, as illustrated at the bottom of Figure 4.7. The tempo fluctuations

made it so that I just attempted to place an ictus on each onset, without regard for the sound I was hearing. While Tagaq also marked icti with every note onset, her icti align with certain aspects of the sound. Tagaq's icti sometimes fall later than the identified onsets in this part of the song, but her ictus placements align regularly with the peak energy level of the EH sounds, as well as the beginning of the purest and dullest parts of the sound. Later, once the tempo stabilizes and a more predictable rhythm ensues, loudness, noisiness, and brightness seem to play more of a factor in the placement of both our icti.

At approximately 85s into the song, Tagaq and I begin to differentiate between the sounds used, marking icti only with certain sounds, rather than with the onsets of all sounds. I aligned my icti only with the Exhaling High (EH) sounds; Tagaq more often aligned her icti with the Exhaling Deep (ED) sounds. Figure 4.8 illustrates the placement of our icti, along with the corresponding RMS energy, periodicity, and spectral centroid data for the audio excerpt from 85.29–91.58s, in a similar manner as Figure 4.7.

In 1987, Fred Lerdahl established a hierarchical structure of timbre, in which he posited that pure, dull sounds are consonant, and noisy, bright sounds are dissonant. This theory correlates with the way I seemingly approached locating a beat while listening to Tagaq's "Ilunikavi." According to Lerdahl's definition of timbral consonance and dissonance, the EH sounds are timbrally consonant because they are pure and dull while the ED sounds are timbrally dissonant because they are bright and noisy. Lerdahl also argues that timbral consonance is stable, and therefore preferably strong metrically, compared to the unstable and preferably weaker timbral dissonance. Timbral consonances, therefore, are more likely to fall on a beat than timbral dissonances—at least for Western art music, which would have likely been the type of music Lerdahl had in mind while writing this article. Additionally, EH sounds are also more

familiar to me as a listener of mostly Western Top 40 popular music, since most singers in this style use only EH sounds. These are likely the reasons why I attend to the EH sounds throughout Tagaq's throat-singing section.

While Lerdahl's theory of timbral hierarchy corresponds to the way that I, a WAM-trained music theorist, mark beats while listening to "Ilunikavi," Tagaq's icti dispute this theory. Tagaq aligns her icti with the supposedly timbrally consonant EH and timbrally dissonant ED sounds at different points in this excerpt. She does not prefer one over the other but attends to both. From 80.22–86.59s (part of this excerpt is not shown in the example), she aligns her icti with the ED sounds before shifting to align with EH sounds for two icti at approximately 87.38s. Starting at 88.68s, Tagaq aligns her icti with both sounds in what is loosely a short–long durational pattern. Overall, Tagaq's movements more often correspond to sounds with a lower periodicity (a noisier sound) and a higher spectral centroid (a brighter sound) in the portion shown in Figure 4.8. Unlike myself, Tagaq seems more willing to attend to both sounds, sometimes moving with the pure, dull, high sounds, and sometimes moving with noisy, bright, deep growls. This shows a willingness to change her listening approach that is absent in mine. The data from later parts of the throat-singing sections emphasize these observations: my fixation on the more timbrally stable (purer and duller) sounds and Tagaq's flexibility in her placement of icti.

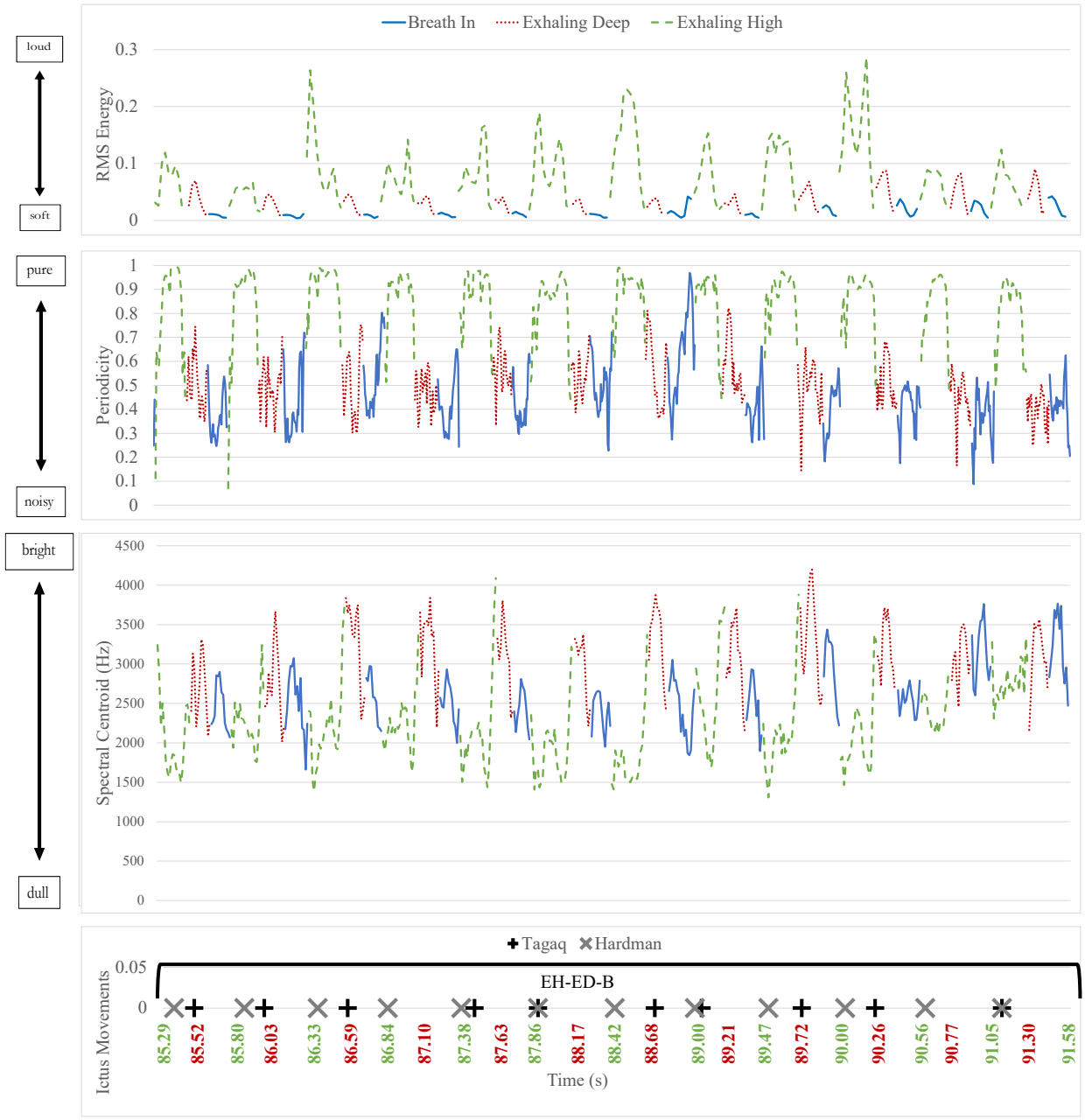


Figure 4.8 Ictus placement and corresponding RMS energy, periodicity, and spectral centroid data of audio for 85.29–91.58s of "Ilunikavi."

The excerpt illustrated in Figure 4.9 confirms my fixation on high sounds. Starting at 112.5s (not shown in Figure 4.9), I doubled the duration between my icti so that I now only identified every second EH as a beat. My attention remained, however, on the pure, dull, high sounds. Based on this observation, I seemingly identified the EH sounds as the beginning of a

grouping made up of two EH–ED–B sound patterns.[59] (Exhaling High was, after all, the first sound heard at the beginning of this throat-singing section.) When the pattern switches from EH–ED–B to ED–IH at 126.97s, I mark the first ED sound (the first sound of the new sound pattern) as a beat. This moment, at 126.97s, is approximately 1s from the last beat I marked with an ictus. Because I tried my best to maintain a steady, evenly spaced pulse while listening to "Ilunikavi," it is not surprising that I marked an ictus at this moment. Rather than continuing to align with ED sounds thereafter, however, my next ictus aligns once again with a high sound, though this time with IH at 128.29s. To align with this IH sound, I elongated the duration between my icti. This indicates that aligning my ictus with a high sound was more important to me than maintaining equally spaced beats. The last ED (128.13s) of the ED–IH part was just over 1s from my last ictus and would have been a better choice if maintaining my pulse was the ultimate goal. With the return of the EH–ED–B pattern at 128.8s, I once again align my icti with every second EH sound, though for a while after the change in my beat length, my feeling of the beat was regularly behind the onset of Tagaq's EH sounds. Instead of aligning with the initial onset, I aligned my icti with the peak energy level of these sounds.

Tagaq's icti in the excerpt shown in Figure 4.9 also confirm the previous observation that she is willing to adapt to the given scenario. There is no discernible pattern to the durations between Tagaq's icti in this excerpt.[60] At most, it can be observed that there are often a few shorter durations followed by one or two longer durations. Again, in contrast to my focus on high

[59] This grouping of sound patterns aligns with the sound patterns identified during Tagaq's demonstrations in "The Sounds of Throat Singing" (2009). In this video, Tagaq describes, then demonstrates individual sound patterns. I have assumed that the way that Tagaq begins a demonstration in the video indicates the beginning of a sound pattern.

[60] My data may also not be completely accurate, as sometimes it was difficult to discern moments of ictus in the video. Sometimes this was due to the way in which Tagaq was moving (there may have been more circular motions, which made it difficult to posit a moment of ictus) or camera angle changes, which broke up the flow of motion by Tagaq. This observation (the lack of discernible patterns in Tagaq's icti) should be read with this caveat in mind.

sounds, Tagaq frequently marks the ED sounds with an ictus. The most notable moment happens during the ED–IH sound pattern: when I presumably try to place an ictus with the IH sound, Tagaq aligns her icti more closely with the ED sounds.

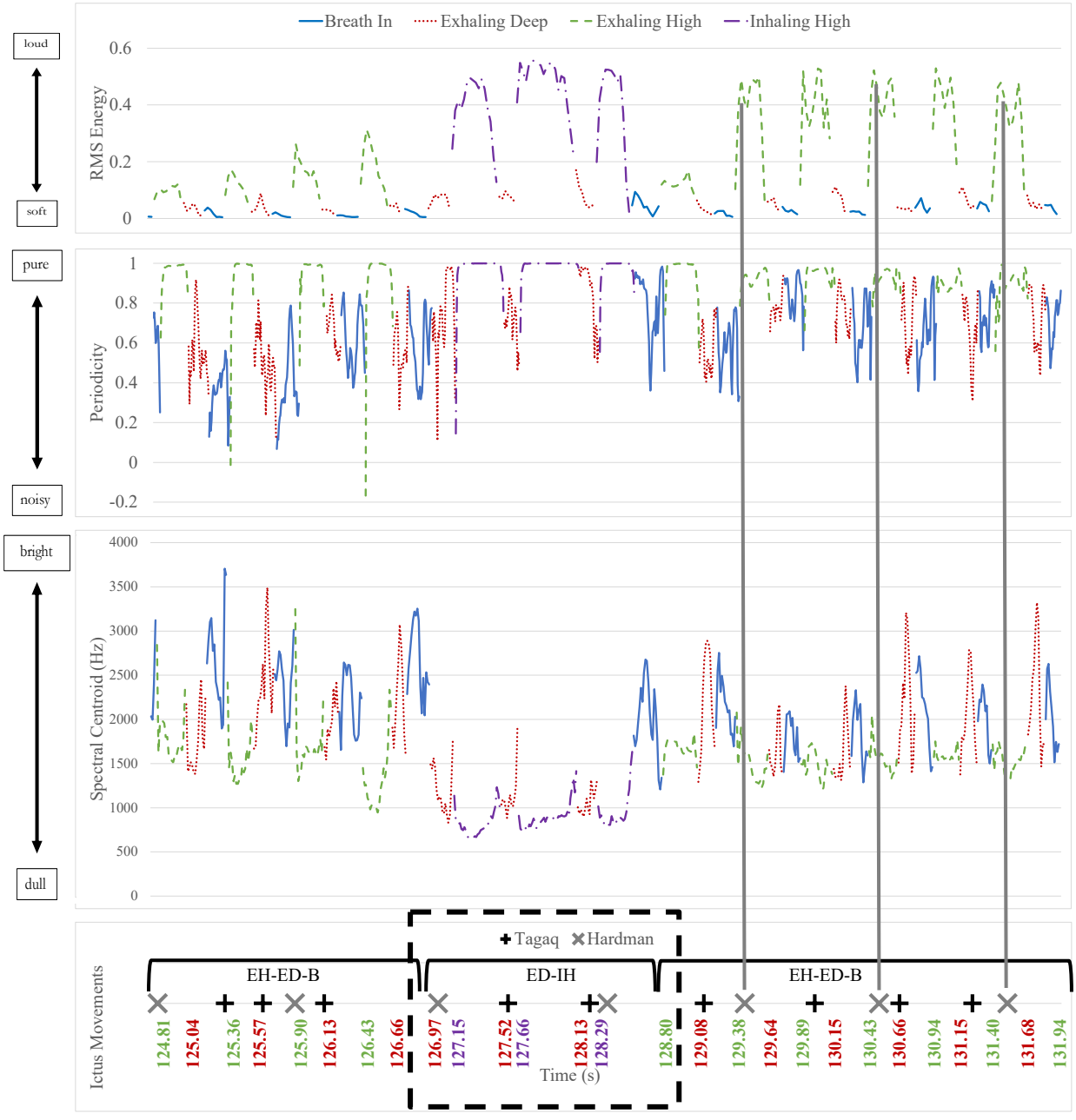


Figure 4.9 Ictus placement and corresponding RMS energy, periodicity, and spectral centroid data of audio for 124.81–131.94s of "Ilunikavi."

Toward the end of the middle section of "Ilunikavi," the sound patterns change more frequently. As Figure 4.10 indicates, each sound pattern (EH–ED–B, ED–B, and ED–ID) is sung only twice (thrice for the last ED–B) before changing to a different sound pattern. The placement of my icti occurs every two iterations of a sound pattern, at what I deemed the beginning of a new pattern. There were few moments to align my icti with an EH sound, so it seems that I might have listened for changes in sound patterns, grouping similar patterns together and feeling a new beat once a new pattern was started, regardless of the sound type at the beginning of the pattern. Notably, however, I also seem to have wanted to maintain the evenly spaced pulse of 1s that I had established earlier, and these moments often happened to coincide with the beginning of a new sound pattern, at least at the beginning of this excerpt. When Tagaq changes the number of times she sings a sound pattern with the last ED–B pattern at 156.43, my ictus no longer aligns with the start of a new pattern. This suggests that I was actually more concerned with maintaining a steady, evenly spaced duration between beats than aligning with the beginning of a sound pattern. Tagaq, on the other hand, continued to most often align her icti with the deep growls (ED and ID sounds) at the end of the throat-singing section, as illustrated in Figure 4.10.[61] She regularly placed her icti at moments of peak energy in the deep sounds that were also moments of noisiness and brightness. Additionally, in this excerpt, most of Tagaq's icti align with the second sound in the sound patterns that I identified. This may indicate that the first sound in a sound pattern is not the most important for Tagaq to attend to. Based on these observations, it seems that Tagaq may place little weight on Lerdahl and Jackendoff's metric preference rule 2, which states that listeners typically prefer to hear a strong beat early in a

[61] Approximately 64% of Tagaq's icti throughout the middle (throat-singing) section of "Ilunikavi" were closer in time to the deep sounds than the high sounds. Only 23% of her icti were closer in time to high sounds, while approximately 12% were equidistant from a high and a deep sound. In contrast, approximately 85% of my icti were closer in time to high sounds.

grouping.[62] Moreover, it does not seem that changes in sound qualities played a role in my experience of beat in this excerpt, but it might have in Tagaq's.

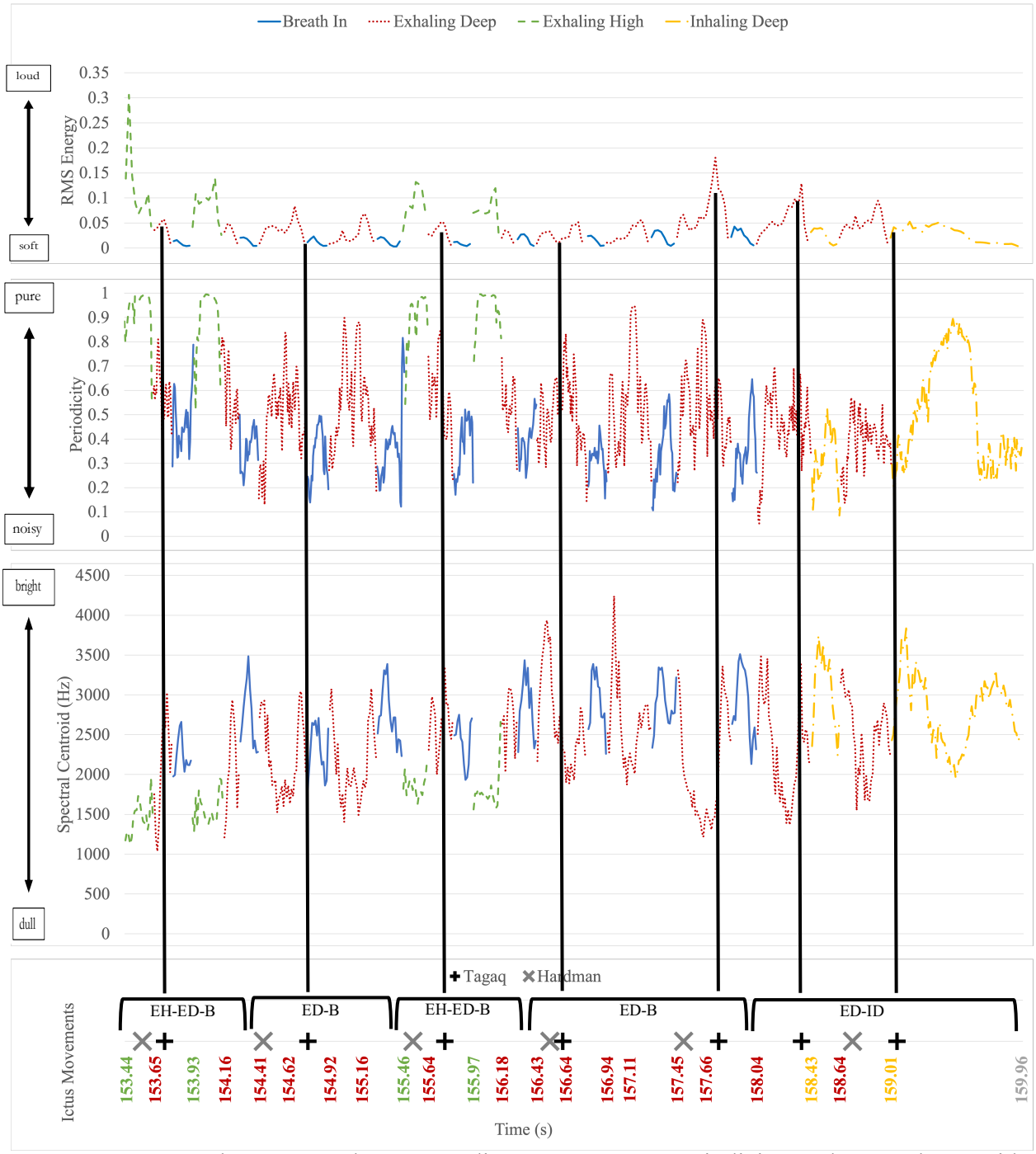


Figure 4.10 Ictus placement and corresponding RMS energy, periodicity, and spectral centroid data of audio for 153.44–159.96s of "Ilunikavi." The onsets of sounds are marked by the time in the bottom graph.

In sum, the main observations from this study are that 1) Tagaq and I both generally align our icti with the onsets of sounds; 2) I aligned my icti with different sound qualities from Tagaq much of the time, attending to high sounds while Tagaq often attended to deep sounds; and 3) I maintained a steady beat, while the durations between Tagaq's icti fluctuated. These differences between the placement of our icti are revealing of differences in our conception of beat structure in music and may be at least partially due to our cultural understandings of time. In the next section, I return to the discussion from earlier that compared Inuit ecological conceptions to Western mechanical conceptions of time, explaining how these conceptions of time correspond to the way that we each move to the same performance of the same song.

4.7 Speculations: Relationship between Experiences of Beat and Conceptions of Time

It is widely argued that we experience rhythm, beats, and meter based on our prior experience with music (Clayton et al. 2005; London 2004). Justin London has theorized that "metric behaviors are…learned—they are rehearsed and practiced…So we fit, so to speak, patterns of events in the world to patterns of time we have in our minds" (2004, 4). London clarifies that "these entrainment behaviors are highly practiced: from early childhood we are steeped in a musical environment, and have many opportunities to develop and hone our attentional habits relative to particular musical styles" (2004, 6). I contend that our conceptions of time based, at least in part, on cultural teachings also plays a factor in how we entrain to music.[63] As discussed above, there are substantial differences between Inuit ecological

[62] This observation is based on my assumption of where a sound pattern begins.
[63] Jean-Jacques Nattiez (1983) questions whether there is a connection between Inuit conceptions of time and the musical structures of some genres of Inuit music. In the same article, he casts doubt on this potential linkage, giving numerous reasons for why one should not try to make the connection. It should be noted, however, that Nattiez was interested in the form of traditional katajjait rather than a beat structure of a particular song and the specific movement of the performers/listeners.

conceptions of time and Western mechanical conceptions of time. While descriptions of these conceptions are reductionist, these distinctions can nevertheless be used as a point of departure for considering reasons why Tagaq and I move to "Ilunikavi" as we do. Based on my own past experiences with listening to a variety of styles of music and the observations made in the preceding section, it is my hypothesis that our cultural conceptions of time play an important factor in our understanding of time in music.[64] As a white, settler Canadian, I can make stronger claims for the connection between Western mechanical conceptions of time and Western conceptions of beat, and only speculative claims regarding the connection between Inuit ecological conceptions of time and Inuit conceptions of beat. Besides, Tagaq's listening habits include Western popular music genres, and how she moves are surely also influenced by these, making any connections between the observations about her ictus placements and Inuit conceptions of time and musical beat precarious.

Not only do different conceptions of musical time correlate to the abstract understandings of time in Western and Inuit music, but they also correspond (coincidentally?) to the movements of Tagaq and me during "Ilunikavi." Tagaq was more flexible with beat duration, rarely maintaining the same duration between icti. She also regularly changed the placement of her icti based on what was currently happening in the song, rarely aligning with one sound type for an extended period. As discussed above in Section 4.1, this flexibility is important in Indigenous performance practice, and it corresponds to Inuit conception of time.[65] I, on the other hand, was much more regimented and quantitative, attempting to maintain equal durations between icti and

[64] This connections have been historicized in Grant (2014) and Scherzinger (2018).

[65] Flexibility of time in music is also important in Western styles of music such as Gregorian chant, recitative, Baroque keyboard fantasies, certain genres of Roma music, certain genres or passages in jazz, singer-songwriter music of the 1960s and 1970s, and other musics that are partially or wholly aligned with Western culture. Not all Western music is based on equally spaced beats, though often Western theories of meter privilege musics with equally spaced beats.

the placement of my icti on a particular sound type. Tagaq's spacing of icti was contingent upon what was happening in the music at that moment, whereas my spacing remained as consistent as I could make it. Following the expectations of Western mechanical conceptions of time, I attempted to control the recorded music by maintaining my beat duration, even when my icti did not align with onsets in the music.

Tagaq's movements while performing "Ilunikavi" are an interpretive choice, which aligns (whether intentionally or not) with certain Inuit ecological (and anti-colonial) conceptions of time and could be seen as an act of resistance to Western mechanical conceptions. Not only is Tagaq known for being an improvisational singer, avant-garde composer, best-selling author, and visual artist; she is also known for her social, political, and environmental activism. Tagaq regularly uses embodied performance and visuals to enact activism (Galloway 2020; Stévance 2017; Taylor-Neu 2018; Woloshyn 2017). Galloway (2020, 123) argues that "Tagaq works with language, the voice and the body as an act of activism that challenges the generally commonly accepted views of Indigenous peoples (e.g., all Indigenous people live on reserves), their lived experiences, human–non-human relationships, and the Circumpolar North." For example, during her 2014 Polaris Prize gala performance, Tagaq dressed in red and had a screen behind her with the names of murdered or missing Indigenous women on it to amplify discussion of these women and bring more awareness to the plight of Indigenous women in Canada. She also wore sealskin cuffs around her wrists, in a protest against PETA and in support of Indigenous sovereignty. Tagaq's movements while performing are also a form of activism, as can be seen in her music videos such as "Retribution" and live performances. She often evokes animals, such as the tusks of walruses, the capture of an animal during the hunt, the claw of a wolverine, etc., emphasizing Inuit connections to the land and human–non-human relationships (Stévance 2016). Although

Tagaq's movements during this performance of "Ilunikavi" are extremely subdued compared to her usual live show, I wonder if, based on her activism in other performances, her flexibility with regard to beat placement might be a way to assert Inuit ways of engaging with time and space.[66]

This section has examined differences between Western mechanical conceptions of time and Inuit ecological conceptions of time to postulate possible reasons why Tagaq and I move to "Ilunikavi" in the way that we do. Although this section is speculative, particularly when connecting the reasons for Tagaq's movements to generalized Inuit ecological conceptions of time, the stronger claim that there is a connection between Western mechanical conceptions of time and Western theories of beat gives reason for further consideration of this question. Much more research would need to be done to see if my idea that our cultural conceptions of time influence how we entrain to music actually carries weight.

4.8 Conclusion

I began this chapter by describing the differences between Inuit ecological and Western mechanical conceptions of time and how these may relate to conceptions of musical time. I then considered the ethical implications and issues with a study such as this, which aims to identify beat structure from two perspectives and for a music of which I am an outsider. I built on this to establish the approach that I used in this chapter, specifically to analyze the general sound qualities of Tagaq's throat singing and the ways in which beat structure can be identified using pre-recorded video. The bulk of the chapter compares the sound qualities to the beat structures of Tagaq and me to draw correlations between the two. At the end of the chapter, I returned to the

[66] In other live, recorded performances, such as when Tagaq performs "Centre" with Shad, she moves to the drumbeat, which features equally spaced beats in a 4/4 meter.

discussion of Inuit ecological and Western mechanical conceptions of time to speculate about the ways that these may have impacted the way Tagaq and I felt the beat of "Ilunikavi."

In this chapter, I introduced the four main sounds that Tanya Tagaq employs in her throat-singing–inspired passages, as she describes in a published interview: Exhaling Deep, Inhaling Deep, Exhaling High, and Inhaling High. In examining the qualities of these sounds, I realized that the oppositional chart often used to analyze timbre is not effective in explaining the differences between them, because it does not explain the differences between the sounds that I can aurally identify as different. By using acoustic measurements and placing these sounds on continua of three acoustic qualities related to energy, noisiness, and brightness, it became possible to produce a more nuanced comparison of these sounds. Contrary to an oppositional analysis, which suggests pitch as a primary factor in the categorization of sound quality, the acoustic measurements (RMS energy, periodicity, and spectral centroid) suggest that, in some cases, the type of breathing affects the experience of sound quality more than pitch does. The continua put forth in this chapter allow for comparison of one sound to the others, providing a way of describing the subtle yet important differences in the sound qualities of Tagaq's throat-singing sounds. The continua elucidate a different way of describing sounds that spectrogram reading and qualitative descriptors do not enable; however, the continua are complicated to interpret (like spectrogram readings), and they require some statistics in order to make sense of them.

Although loudness, noisiness, and brightness are certainly not the only aspects of a voice that affect how we move to music, based on observations made in this chapter, it seems that changes in these three sound qualities may play roles in our experience of beat. The basis of these observations comes from Tagaq's movements while performing her song "Ilunikavi" and

my movements while listening to the song using pre-recorded video. While this method did not allow for a complete comparison of our movements, as there was sometimes difficulty in determining Tagaq's icti, certain observations in connection to changes in sound qualities could be made. There were some similarities between how Tagaq and I marked a beat, but there were also three notable differences: 1) in moments of uncertainty in the tempo, Tagaq was better able to more closely align her icti with onsets of sounds; 2) I mostly maintained a steady pulse, while Tagaq frequently changed durations between icti; and 3) I more often aligned my icti with pure, high sounds, while Tagaq more often aligned hers with deep growls.

In this study, I initially set out to only answer the question of whether changes in sound qualities—specifically loudness, noisiness, and brightness—affect our feeling of a beat while performing or listening to music; but it grew to include much larger questions about how to analyze beat structure from multiple perspectives and the meanings that can arise from doing analysis in this way. I ultimately chose to make myself the subject of the study, in addition to Tagaq. By recording myself listening to "Ilunikavi," I removed some of the bias that might have arisen if I had listened to the song repeatedly, taking portions of it out of context, and only *thinking* about how I might feel the beat in the music. Here, the recording *shows* how I felt the beat during one listening experience, similar to how the recording of Tagaq performing "Ilunikavi" shows how she moved to it in that instance. While this study is surely not without flaws, I hope that it can serve as a point of departure for a way of analyzing beat that takes into consideration two or more individual perspectives.

In the next two chapters, I use the analytical approach established in this chapter to consider how changes in sound qualities affect our understanding of meter in recent mainstream country music, and how they affect the experience of form in two of Tagaq's songs, "Ilunikavi"

and “Sivulivinivut.” Although the approach will vary slightly when dealing with music with English lyrics, these chapters visualize acoustic measurements using similar statistical and data visualization techniques. They demonstrate the wide applicability of this approach for visualizing and analyzing sound qualities.

Chapter 5: Phenomenal Accents, Formants, and the Vocal Backbeat in Country Music Vocals from 2010 to 2020

In recent years, much country music, especially what has been termed *bro-country*, has been dismissed for sounding too alike.[67] About six years ago, Sir Mashalot, a Nashville songwriter and producer also known as Greg Todd, released a video in which he mashed together six mainstream country songs to prove that they all sound the same.[68] Indeed, these songs are very similar in many ways. They use similar drum patterns and instrumentation, the same basic four chords, and have a walking tempo of about 76 beats per minute. Another important aspect that makes these songs so similar are the lyrics and how they are rhythmically set within the meter. All six of the songs in Sir Mashalot's video feature a heavy emphasis in the vocals on beats 2 and 4. I refer to this style of text setting as "vocal backbeat." The perceived emphasis with words on the backbeats in recent mainstream country vocals is due to four factors: duration, loudness, pitch, and sound quality. The focus of this chapter is to study ways in which sound qualities interact with prosodic accentuation and meter to affect the ways in which we experience the vocals of recent country music.

While much of this dissertation focuses on the music of Tanya Tagaq, this chapter focuses on selected mainstream country songs from the last decade. The choice to focus on country music in this chapter was one of practicality and ethics. Country music often has a clear metrical structure, making it possible to relate changes in timbre to our experience of meter.

---

[67] Bro-country is a term used to describe a sub-genre of mainstream country music from the 2010s that is influenced by hip-hop, hard rock, and electronica. The term was coined by music critic Jody Rosen in 2013. Bro-country songs are often upbeat songs that discuss attractive young women, partying, and pick-up trucks. Some of the main artists associated with this style of country music include Florida Georgia Line, Luke Bryan, Jason Aldean, and Jake Owen.

[68] The mashup by Sir Mashalot (2014) can be located at the following link: https://www.youtube.com/watch?v=FY8SwIvxj8o

Additionally, this music is designed for dance, whether it is a two-step or line dancing, and counting for these dance styles is very metrical. On the other hand, to fit Tagaq's music into a hierarchical metric structure could enact violence toward the music and Tagaq. I, personally, do not know enough about Inuit throat singing or Tagaq's music to suggest a metric structure for her songs. But I also feel that it is important to consider connections between sound qualities and meter because so much popular music is governed by meter, especially 4/4 meter. By focusing on country music in this chapter, I have the opportunity to consider the relationship between changes in sound qualities and meter, a slightly larger-scale structural aspect of music than was discussed in Chapter 4.

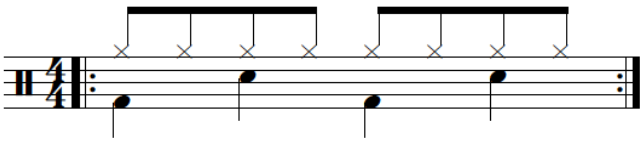
In this chapter, I begin by defining instrumental and vocal backbeat, outlining the characteristics of these phenomena. I then explain the approach used in this chapter to analyze sound qualities in vocals with English lyrics. While the approach resembles that used in previous and subsequent chapters, this chapter focuses on formants, rather than the acoustic measurements of RMS energy, periodicity, and spectral centroid, to analyze vowel quality. Next, I propose some novel ways of interpreting the meter of songs with vocal backbeat and a theory of how backbeat in country music affects our understanding of the overall meter in such songs. Later, I present in-depth analyses of vocal backbeat through three case studies: Lil Nas X's "Old Town Road" (2019), Carrie Underwood's "Church Bells" (2016), and Jason Aldean's "The Only Way I Know" (2012). While examining these three songs in detail, I also contrast the conventional way of analyzing melody with my novel way of understanding vocal backbeat, which is informed by phenomenal accents (including phonetic accents) and sound qualities. Finally, I consider the narrative context in which songwriters employ vocal backbeat within a country song. While some songwriters of recent mainstream country music use vocal backbeat throughout an entire

song, others use it sparingly as a text-painting device to create a sense of tension which furthers the narrative of the song.

5.1 Definitions of the Instrumental Backbeat and the Vocal Backbeat

The first thing that comes to mind when we think of the term "backbeat" is probably the ever-prominent snare drum on beats 2 and 4 in one of our favorite popular songs. As illustrated in Figure 5.1a, the basic beat of rock music features a quarter-note pattern in 4/4 with the kick drum and snare drum in alternation, while the hi-hat (or potentially the ride cymbal) has continuous eighths. Walter Everett (2009), Garry Tamlyn (1998), and others have noted that backbeats are also found in accompanimental rhythms. A common backbeat pattern for the guitars and piano is rest–quarter–rest–quarter, as shown in Figure 5.1b. Both types of backbeat can be heard during the first verse of the Temptations' "My Girl" (1965).

a)



b)

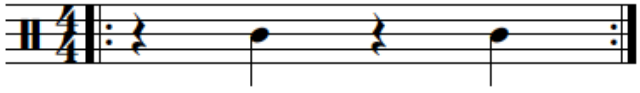


Figure 5.1 Examples of instrumental backbeats: a) the basic beat of rock/pop music with snare backbeat (bass or kick drum indicated on first space, snare indicated on third space, and hi-hat indicated by x above staff); and b) a guitar/piano backbeat.

In both of the most common instrumental backbeat patterns, there are a few crucial elements. The instrumental backbeat is characterized by phenomenal accents on the normally weak beats 2 and 4 (Biamonte 2014). Additionally, there is a change to a drum with noticeably

different sound qualities on the backbeats, which emphasizes the salience of these beats. Alternatively, there may be a rest on beat 1, which de-emphasize the strength of the odd-numbered beats, giving more emphasis to the weaker beats 2 and 4. These characteristics are also found in vocal backbeat.

Vocal backbeat occurs when phonetic accents, rhyming words, and/or regular repetitions of vowel or consonant sounds regularly appear on (or in relation to, as in syncopated against) beats 2 and 4 of a 4/4 meter.[69] The accented word may not appear directly on beat 2 or 4, but it will be heard in relation to the backbeats as syncopation.[70] Essentially, vocal backbeat arises in country songs because words/syllables are perceived as being prominent on beats 2 and 4. Multiple factors contribute to perceiving a word as prominent within a phrase, including duration, loudness, pitch, and spectral balance.[71] These four descriptors are what determine prominence and prosodic structure in everyday speech, and they also contribute to our experience of the vocal backbeat. Therefore, in a vocal backbeat, words on beats 2 and 4 are performatively emphasized by the singer through phenomenal accents in the melody, including durational, dynamic, and pitch accents, as well as spectral changes.[72] A vocal backbeat is perceived when prosodically stressed syllables on beats 2 and 4 are longer in duration than those that precede them, louder than neighboring words/syllables, higher in pitch, and/or have a spectral profile similar to other backbeat words. Additionally, to be identified as vocal backbeat, I stipulate that the effect must recur throughout a section of the song, not just for one bar, like

[69] Songwriters of country music are known to use similar-sounding words on strong beats. Aaron Fox (2004, 217) describes that "honky-tonk songs…conventionally use alliteration on strong beats to produce a strong rhythmic 'feel' in the vocal line."
[70] I conceive of syncopation much the same way as Temperley (1999), for whom syncopations are considered to exist in relation to metrical beats.
[71] Sluijter and van Heuven (1996) found that the best determinants of whether a sound is stressed are duration and spectral balance.
[72] According to Lerdahl and Jackendoff (1983, 17), phenomenal accents refer to "any event at the musical surface that gives emphasis or stress to a moment in the musical flow."

instrumental backbeats. Lastly, I have observed that songs with vocal backbeat typically have a "walking pace" of 80–86 beats per minute. While this may be coincidental, it is worth noting as a characteristic of songs with vocal backbeat.

Vocal backbeat arises in country songs in two ways: 1) the *shifted anacrusis* backbeat and 2) the *amped up 2 and 4* backbeat. In the shifted anacrusis type, the strength of beat 1 is deemphasized in the vocals with a rest, making the "and" of beat 1 an anacrusis to the performatively stronger beat 2. This type of vocal backbeat is found in Eric Church's "Stick That in Your Country Song." Figure 5.2a contains the requisite portion of the lyrics from the chorus.[73] The chorus features a rest on beat one, emphasis on the word "that" on beat 2, and emphasis on the words "song," "one," and "along" on beat 4. The words on beat 4 share a similar sound, that of "on." This, along with the deemphasis of beat 1 with a rest in the vocals, contributes to the perception of vocal backbeat in the chorus.

In the amped up 2 and 4 type, beats 1 and 3 are strongly accented, but beats 2 and 4 receive equal or greater weight through phenomenal accents—including durational, dynamic, pitch, or phonetic accents—and/or sound qualities. The amped up 2 and 4 vocal backbeat can be heard in the chorus of Blake Shelton's "I'll Name the Dogs," the lyrics of which are provided in Figure 5.2b. The lyrics feature rhymes at the ends of lines, including "money"/"funny," "kisses"/"business," "stars"/"guitar." Furthermore, our attention is drawn to the words on the backbeats because they are the only words in a line that change significantly. For each line, the general structure is "You'll [insert verb] [insert noun] and I'll [insert same verb] [insert different noun]." While much of the chorus features vocal backbeat, there is an abrupt shift back to a

[73] I have purposely not included a transcription of the vocal melody because I want the reader to engage with the music aurally. Vocal backbeat is an aural effect used in recent mainstream country songs, one that can be easily heard while listening to the music. Please take a moment to listen to each example discussed in this chapter.

focus on beats 1 and 3 near the end of the chorus, with the title line "you'll name the babies and I'll name the dogs." The last two lines rhyme ("frog"/"dog"), just as the previous lines did, but this time the rhyme happens on beats 1 and 3. The shift in placement of the rhyme draws our attention to the differences between the vocal backbeat and the more normal placement of stressed words on odd-numbered beats within the melody. This type of vocal backbeat, where all beats are emphasized but the backbeats receive equal or greater stress due to phenomenal accents, aligns with Tamlyn's (1998) argument that a snare backbeat can be heard in an instrumental backbeat even if there are four snare hits per bar as long as the second and fourth are accented performatively. The unusual vocal accentuation of beats 2 and 4 through phenomenal accent creates a sensation of vocal backbeat.

a) "Shifted Anacrusis" Backbeat, Eric Church's "Stick That in Your Country Song" (2021), chorus, starting at 0:29.[74]

Stick **that** in your country **song**, yeah
Take **that** one to number **one**, yeah (number one)
And get the **whole** world singing **along**, yeah
Stick **that** in your country **song**, yeah

b) "Amped Up 2 and 4" Backbeat, Blake Shelton's "I'll Name the Dogs" (2017), chorus, starting at 0:36.[75]

You'll find the **spot** and I'll find the **money**
You'll be the **pretty** and I'll be the **funny**
You'll plant the **flowers**, I'll plant the **kisses**
Baby, let's **get** right down to **business**

I'll hang the **pictures**, you'll hang the **stars**
You'll pick the **paint**, I'll pick a **guitar**
Sing you a **song** out there with the crickets and the ***frogs***
You'll name the ***babies*** and I'll name the ***dogs***, yeah

Figure 5.2 Comparison of vocal backbeat types. The green, italicized words are heard on or in relation to beats 1 and 3; the red, underlined words are heard on or in relation to beats 2 and 4.

In these examples of vocal backbeat, I pointed to instances where the words rhyme. This is a common feature of vocal backbeat in recent mainstream country songs, though it can sometimes appear with a twist of sorts. In discussing Blake Shelton's "God's Country," Michael Hardy (one of the songwriters) mentions that the songwriters were trying to find "sound" rhymes while writing the song, and they changed one of the lines from "the devil went *down* to Georgia, but he didn't stay *long*" to "The devil went *down* to Georgia, but he didn't stick *around*" (Dukes 2019, emphasis mine). Based on the example Hardy provides, "sound" rhymes are those more commonly referred to as "imperfect rhymes" (also known as slant, oblique, near, or partial

[74] Eric Church's "Stick That in Your Country Song" can be located at the following link: https://www.youtube.com/watch?v=V8KRC5KH7HU
[75] Blake Shelton's "I'll Name the Dogs" can be located at the following link: https://www.youtube.com/watch?v=pAsN4lSa7O0

rhymes), in which two words share either a vowel sound or a consonant sound. “Down” and “around” share the same vowel sound (aʊ) and the n that directly follows it. This change from “long” to “around” also makes the line rhyme with the others in the chorus, which features words such as “sound,” “found,” “wind,” “bound.” All these rhyming words fall on the backbeats. Rhyming words have similar formant structures because they typically share vowel sounds (Sunberg 1991). Based on the following case studies, it seems that words that do not rhyme can also have similar formant structures if the singer alters their vocal tract and pronunciation of the word while singing.

5.2 Metric Interpretations of the Backbeat

Theorists such as Attas (2015), Biamonte (2014), Butterfield (2006), and Hughes and Vuvan (forthcoming) continue to debate the best ways to interpret or understand backbeat, but many of these music scholars turn to Christopher Hasty’s (1997) metrical functions when analyzing backbeat. Each of Hasty’s functions has a corresponding symbol: beginning (|), anacrusis (/), and continuation (\). Figure 5.3 summarizes the ways in which theorists typically analyze instrumental backbeats. Matthew Butterfield (2006) prefers to understand the backbeat as anacrustic. Butterfield (2006) states that “the *timbre* of the snare, its higher frequency, and apparent brevity relative to the booming bass drum, causes it to be heard as *anacrusis*; thus beat 4, like beat 2, is directed toward the emergence of a new beginning” (2006, par. 40, emphasis mine). In contrast, Nicole Biamonte (2014) prefers to understand the backbeat as functionally continuational. Still others, like Robin Attas (2015), are open to either possibility shown in Figure 5.3, depending on the context and microtiming shifts in performance.

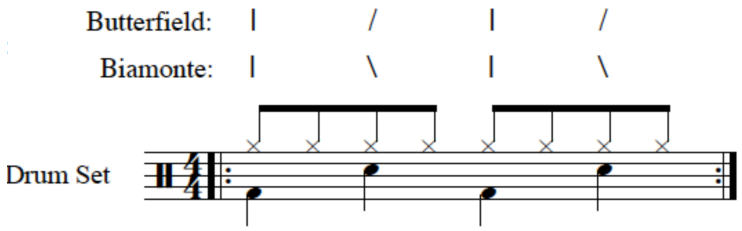


Figure 5.3 Analyses of drum backbeats

While these theorists disagree as to whether backbeats are anacrustic or continuational, they all seem to agree that it is weak compared to the traditionally strong beats 1 and 3, despite believing that "the snare drum…is registrally, timbrally, and dynamically more prominent. This…places several types of phenomenal accent on what are traditionally the weak beats in a 4/4 pattern" (Biamonte 2014, par. 6.1). Waveform representations of the basic beat show that both the bass drum and snare drum have more energy than the hi-hats alone. The snare drum, however, often appears to have slightly more energy than the bass drum in these representations, given that it is the only sound to clip the boundaries of amplitude. This places phenomenal accent on the weak beats of a 4/4 bar.

Theorists continue to analyze backbeats using continuation and anacrusis symbols, which effectively represent weak beats because they maintain a separation of metric and phenomenal accents. Lerdahl and Jackendoff (1983) identify three kinds of accent: phenomenal, structural, and metrical.[76] Phenomenal accents refer to "any event at the musical surface that gives emphasis or stress to a moment in the musical flow," including attack points, dynamic changes, timbral changes, long durations, large leaps in pitch and harmonic changes, while metrical accents refer to "any beat that is relatively strong in its metrical context" (Lerdahl and Jackendoff 1987, 17).

[76] Structural accents refer to accents "caused by the melodic/harmonic points of gravity in a phrase or section—especially by the cadence, the goal of tonal motion" (17). Structural accents, at least as described by Lerdahl and Jackendoff are not common in popular music, and therefore will not be discussed further in this chapter.

Lerdahl and Jackendoff carefully outline the rules for how to determine metric accent in a piece of music but choose not to delve into many of the aspects of phenomenal accent, especially if they do not pertain, in their opinion, to metric accent. Often, phenomenal and structural accents align with metric accents in Western art music (Lerdahl and Jackendoff 1983), but phenomenal accents do not as regularly align with metric accents in Western popular music.[77] Often, in popular music, the use of syncopation, or the shifting of an unstressed syllable to a metric accent, weakens the downbeat and strengthens the backbeat. While I agree with Biamonte that "there is still an accent on beats 1 and 3, as demonstrated by the propensity of musically untrained listeners to clap on beats 1 and 3 of a backbeat rather than 2 and 4" (2014, par. 6.2), the metrical accent, in my opinion, does not come from the drum beat in most popular songs, but rather from the other instruments—the guitars, keyboards, and usually the vocals—which continue to emphasize beat 1 as the strongest beat of the 4/4 bar with a change of harmony.

5.3 Sound Qualities and the Backbeat

5.3.1 Instrumental Backbeat and Acoustic Measurements

A primary characteristic of instrumental backbeat is the change in sound qualities on beats 2 and 4 of a standard rock drumbeat. As Nicole Biamonte (2014, par. 6.1, emphasis mine) describes, "the snare drum is higher in pitch, has a sharper *timbre*, and … is louder than the bass drum—thus it is registrally, *timbrally*, and dynamically more prominent." This description, however, does not tell us much about the specific changes in timbral quality that occur on the backbeats. What makes the snare drum timbrally more prominent, and in what way does the snare drum have a sharper timbre?

[77] Scholars such as Fink (1990), Lerdahl (2006), and Malin (2010), also argue that stresses in poetry should align with metric stresses.

Because the sound qualities of the snare backbeats are significantly different from the bass/kick drum, the sound qualities of a standard rock drumbeat can be analyzed effectively using acoustic measurements, much the way that acoustic measurements were used to analyze Tanya Tagaq's vocals in previous chapters. When I analyze a standard drum beat of popular music using the acoustic measurements from the previous chapters (RMS energy, periodicity, and spectral centroid), the data, as provided in Figure 5.4, shows that the snare and bass drums are equally loud, the snare is less variable in periodicity and noisier than the bass drum, and the snare drum is brighter than the bass drum. I equate the *noisier* and *brighter* sound of the snare drum to Biamonte's statement that the snare has a *sharper* timbre than the bass drum. To say that the snare is brighter and noisier than the bass drum is just a more specific way of describing the timbre of these instruments. Biamonte states that the snare drum is louder than the bass drum in the basic beat, but that is not the case with the recording analyzed here. In other recordings of a standard rock drumbeat, we might find that the snare drum has more energy overall than the bass drum. The brighter and noisier qualities of the snare contribute to its prominence within the basic beat pattern, which may make use perceive the snare are louder. Changes in sound qualities are used in a similar manner in the vocals of country songs, though instead of changing the brightness and noisiness from beat to beat, the vowel and consonant sounds change.

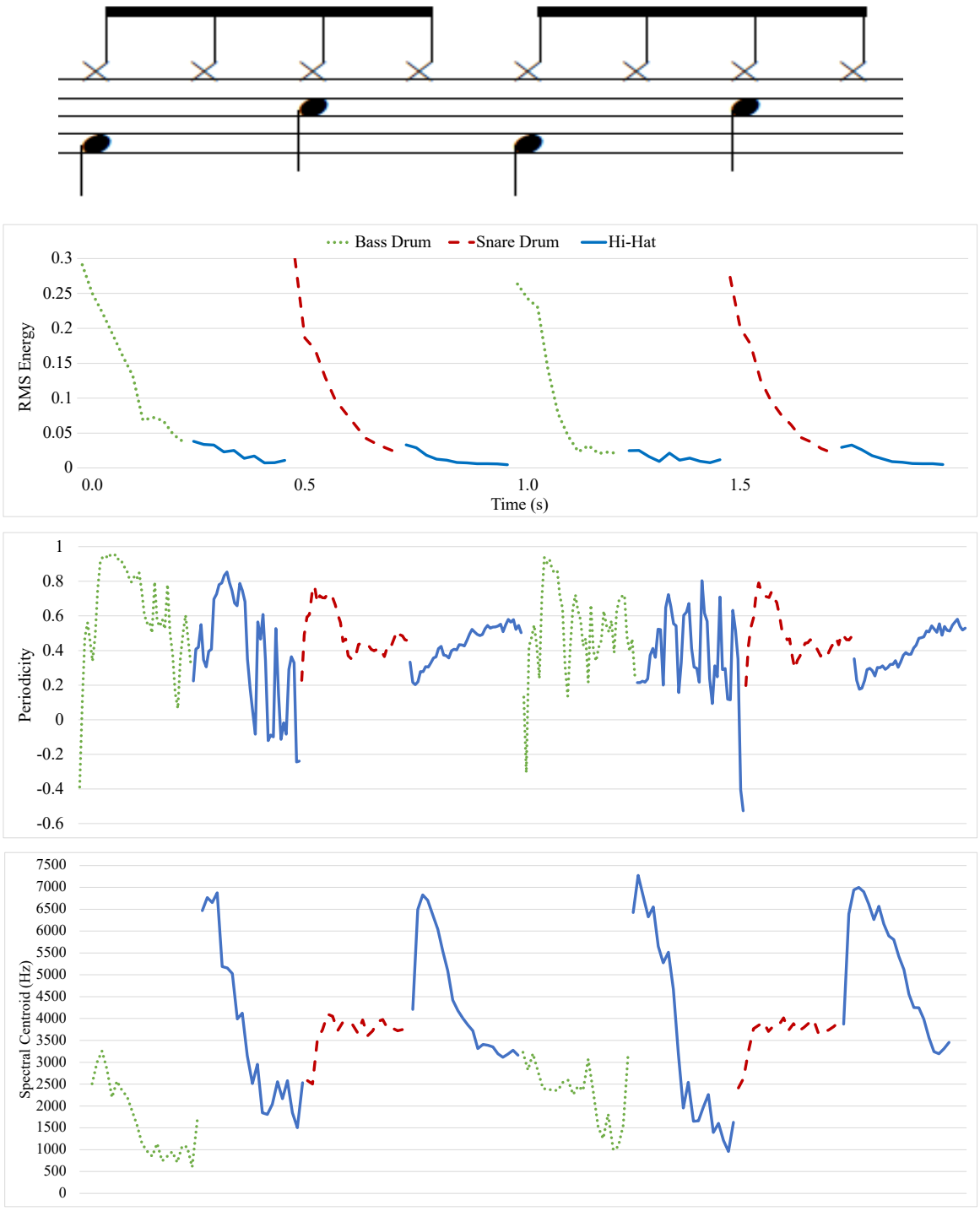


Figure 5.4 Sound qualities of a basic beat of rock music.

5.3.2 Vocal Backbeat and Formants

Like instrumental backbeats, sound qualities play a large role in our perception of a vocal backbeat. As in instrumental backbeat, vocal backbeat occurs when the voice is "registrally, timbrally, or dynamically more prominent" on beats 2 and 4 of a 4/4 meter; however, the differences are much subtler and related to the pronunciation of lyrics. Changes in sound qualities in vocals with lyrics cannot be measured in acoustic properties, but rather with formants. RMS energy and periodicity show few differences from word to word. Because the style of singing in country music uses only exhaling, pitched sounds, the entire melody features a very periodic, pure sound quality. This is different than Tagaq's throat-singing style, which is the focus of Chapters 4 and 6. Throat singing is a rather percussive style of singing and features both voiced and unvoiced sounds. Spectral centroid (which is related to formants) may show some differences from beat to beat because it is a measure of the average of frequencies present in the sound signal, but spectral centroid by itself is too general to be able to give a robust understanding of the vowel qualities in the vocals of country songs. For these reasons, this chapter focuses on formants, while the other chapters in this dissertation focus on RMS energy, periodicity, and spectral centroid.

Formants are peaks in sound pressure levels at particular frequencies that help us identify spoken or sung words. In everyday speech, listeners attend to subtle changes in vocal qualities, particularly formants, to determine the sounds that they hear and the prosodic structure. Vowels, specifically, are distinguished by formants. By changing the shape of your vocal tract, the shape of your lips and/or position of your tongue as you pronounce a word, you can change the frequencies of the formants produced. Think of the differences between the words "bat" and "bait," for instance, and the way that the mouth is shaped in order to produce these words. If the

syllables are produced similarly, they will have similar formants and similar sound quality. The formants in Figure 5.5 correspond to the words "black," "back," "match," and "attached."[78] While small differences occur between these words, the formants happen at similar frequencies because these words all feature the "æ" vowel sound. Figure 5.5 indicates four formants for these words: F1, F2, F3, and F4. The first formant occurs around 700 Hz; the second around 1700 Hz; the third around 2500 Hz; and the fourth around 3500 Hz. Throughout this chapter, the focus will be on the first three formants because these are the ones that vary the most from word to word.[79]

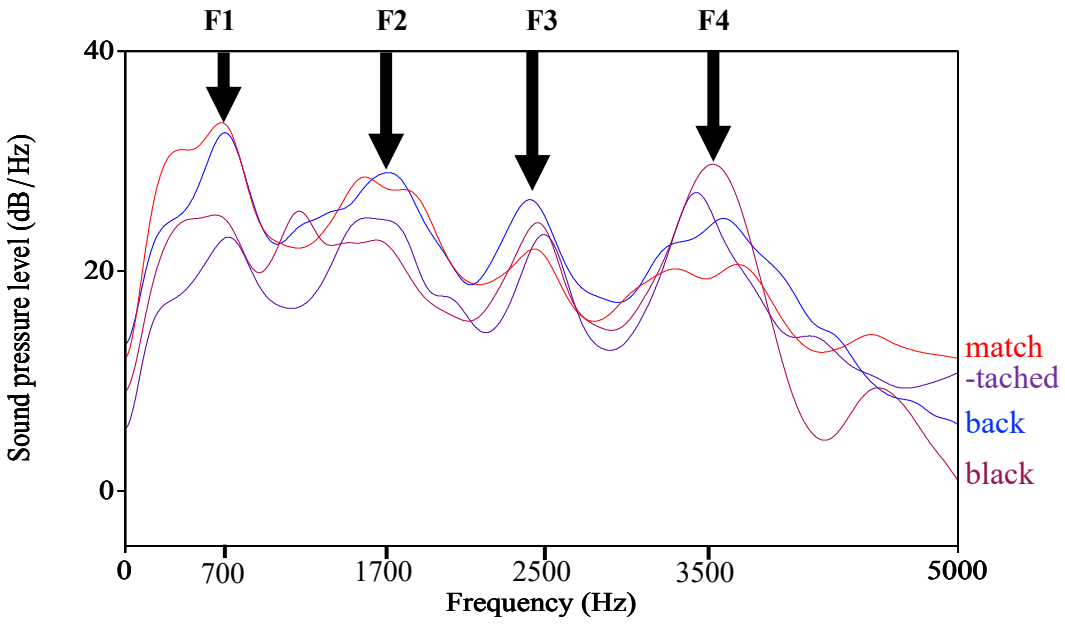


Figure 5.5 Spectral comparison of the first four formants (F1, F2, F3, and F4) of the words "black," "back," "match," and "attached." The frequencies indicated on the x-axis give an approximate value of the formants of the words.

Singers can manipulate their vocal tracts to increase the energy of particular frequencies to make certain moments of their vocals stand out from the rest. Opera singers, for instance,

[78] This graph was created in Praat, a computer application intended for the study of linguistics.
[79] Formants higher than the third formant are better used when comparing the vocal quality of singers (Sundberg 1991).

often adjust their vocal tracts to increase the energy of frequencies around 3000 Hz (Sundberg 1991). This aids in projecting their voices above the low-frequency energy of the orchestra. Sundberg refers to this as the "singer's formant." In contrast, country singers show consistency between their spoken and sung formant values. Country singers, like good speakers (e.g., actors or radio announcers), tend to have a strong fourth formant, which occurs between 3 and 4 kHz. This feature has been referred to as a "speaker's formant" (Cleveland et al. 2001). The formant structures shown in Figure 5.5 of Lil Nas X singing the words "black," "back," "match," and "-tached" correlate to the "speaker's formant." The fourth formant is either stronger in sound pressure than the third formant, as in "black" and "-tached," or similar in sound pressure to the third formant, as in "back" and "match." In country music, singing is closer to speaking. In this way, country music vocals are similar to hip-hop vocals. In a study concerning the correlation between vowel formants and metric positions of words in hip-hop lyrics, Ammirante and Copelli (2019) argue that vowels with higher F2 are less likely to be masked by percussion and are therefore favored for on-beat words. Words with lower F2 vowels, which may be masked, were more likely to deviate from the beat. In this chapter, I suggest that prominent words will have more consistency among formant structures than those that are not heard as prominent. In Figure 5.5, the formant structures have a lot of consistency among the four words graphed, and these words are prominent in the melody of Lil Nas X's "Old Town Road," which will be discussed in further detail below. The data also suggest that country singers might manipulate certain words to make the formant frequencies of words on the backbeats more similar.

5.3.3 Approach to Measuring Formants

In this chapter, I use a method of analyzing sound qualities related to that presented in Chapter 4, though it differs in some important ways. In Chapter 4, I used the acoustic

measurements of RMS energy, periodicity, and spectral centroid. I used box and whisker plots to compare the central tendencies of data relating to Tagaq's individual sounds. Here I use box and whisker graphs in a similar manner, though now I use them to compare the central tendency of mean formant frequencies of words on the odd-numbered beats, even-numbered beats, and offbeats in individual country songs to see if there are significant differences in the formant structures of words that occur in different places in the musical meter. I also add one other type of graph, 3D scatter plots, to my analytical approach. I use 3D scatter graphs to plot the first three mean formant frequencies of all words to be considered in the analysis to see if there is clustering among any sets of words. If there is clustering, it means that the words are similar in formant frequencies and thus similar in vowel qualities.

I obtain the formant frequencies in a program called Praat, which was developed by Paul Boersma and David Weenik for phonetic speech analysis. As in Chapter 4, I used multiple systematic steps to achieve the graphs used below. First I opened the audio file in Praat, and I added a "TextGrid" to the list of objects. This enabled me to annotate the beginning and end of syllables within the program. Figure 5.6 presents a screenshot of the "View & Edit" mode with both the sound and TextGrid objects. The formants are visible in the spectrogram; the dark lines in the spectrogram are formants. To further aid the analyst, Praat has a feature that highlights the formants in red. Both versions, with and without the formants highlighted, are provided in Figure 5.6. While it is interesting to see the formants on the spectrogram, I found the mean formant frequencies more beneficial for analyzing country songs (at least at this time). So, after locating the beginning and endings of syllables, I highlighted the word for which I wanted a formant measurement. After highlighting the time span of interest, I went to the "Formant" column and

selected "Get…Formant."[80] Next, I copied and pasted the values for F1, F2, F3, and F4 for every word that I wanted to analyze into a spreadsheet in Excel. There was less change among the fourth formant values within a single song, so I focus here on the first three formants. To plot all three formants on a single graph, I used Plotly Chart Studio, a free and open-source graphing library. I also created box and whisker graphs, similar to those in Chapter 4, through Excel to determine differences in the dispersion of formant values between syllables on odd-numbered beats, even-numbered beats and offbeats. Based on the visualizations of these two types of graphs, I have found that songs with a vocal backbeat tend to have consistency among the formant values of words heard on or in relation to the backbeats.

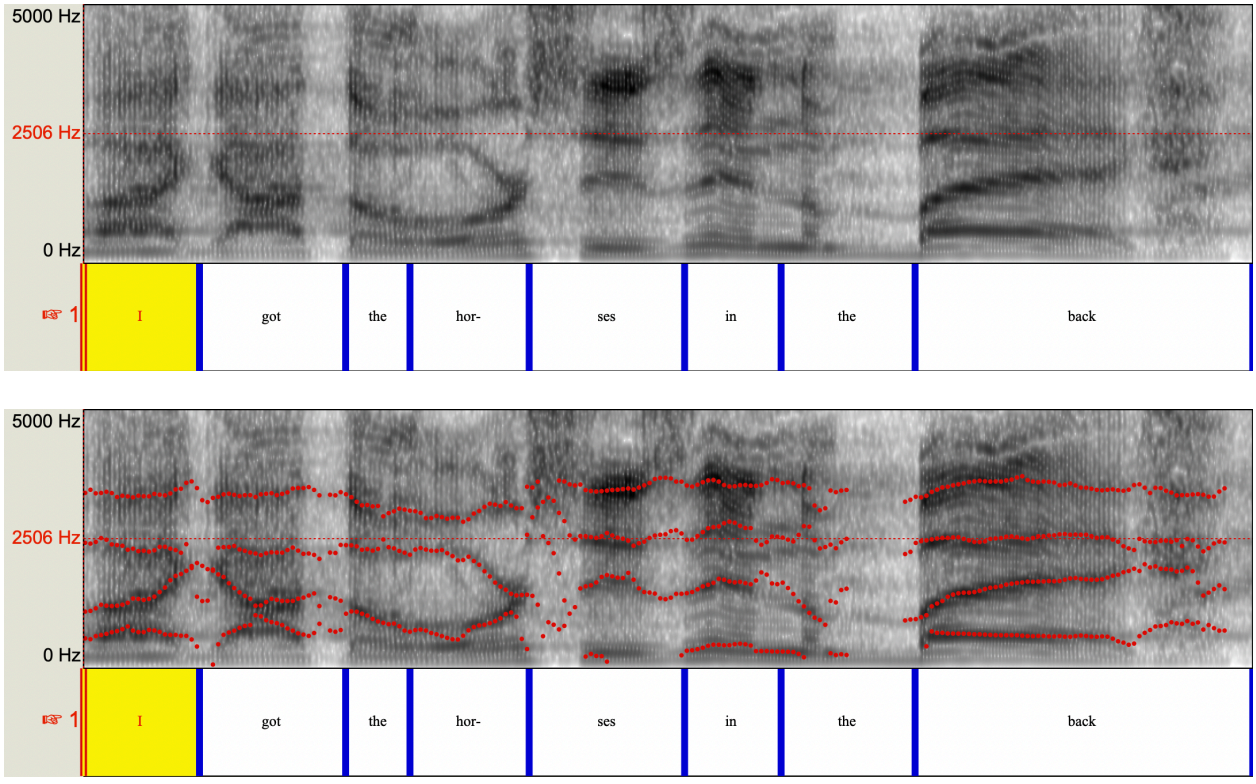


Figure 5.6 Spectrogram without and with formants highlighted in red.

[80] When measuring formants, I set the maximum formant frequency to 4000 Hz in "Formant Settings" for the male singers and 4500 Hz for Carrie Underwood's song. For all songs in this chapter, I set the number of formants to 4, the window length to 0.04s and the dynamic range to 20dB.

5.4 A Novel Interpretation of Backbeat Based on Vowel Quality

Playing with rhythms and meter has long been a feature of country music. In his article on old-time country music, Joti Rockwell (2011) notes that country musicians would often add or delete a beat to make a "crooked tune." Jocelyn Neal (2000, 113) has noted that more recent country hits also often contain hypermetrical irregularities, and "the manipulations of a song's metric structure contribute to the 'sound' of an individual singer." The vocal backbeat in recent mainstream country music is just another metric effect that country songwriters employ. I postulate that this effect changes our experience of meter in country music.

Contrary to the aforementioned ways of analyzing backbeat, I propose analyzing backbeat with a symbol that represents a stressed moment in the vocals. I suggest that, within the singer's part, the backbeats—beats 2 and 4 of 4/4 meter—feel like the strongest of the four beats, subverting our expectations of 4/4 meter. In the following analyses I use a modified version of Christopher Hasty's analytical technique from *Meter as Rhythm* (1997).[81] I adopt Hasty's symbols for analyzing meter: a beginning (|) represents a strong beat at the beginning of a group of beats; an anacrusis (/) leads into the following strong beat; while a continuation (\) continues the duration initiated by the previous strong beat.[82] To distinguish between the end-accented backbeat in the vocals and the conventional beginning-accented (|) meter in the guitars and keyboards, I use a downward arrow (↓) for the backbeats, as shown in Figure 5.7.

[81] I do not use the dot notation of Krebs (1999) and Lerdahl and Jackendoff (1983) because it does not explain the complexities of how surface rhythms interact. The dots do not suggest relations to one another. Hasty's beginning, anacrusis and continuation allow for a deeper discussion of surface rhythms, as will be demonstrated in this chapter.
[82] Although Hasty (1997) is careful to explain that beginnings are not accents in his theory, beginnings will inevitably attract some sort of ontological marking, and therefore build in some sort of accent into Hasty's markings.

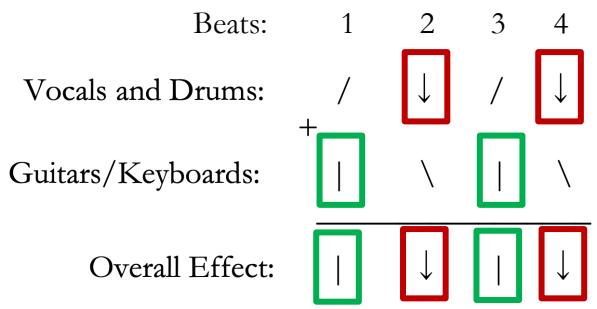


Figure 5.7 The effect of combining vocal and drum backbeats with the metrically normative guitars and/or keyboards in a country song.

The vocal backbeat in these songs influences how listeners engage with the music. Building on John Covach's (2020) concept of positional listening, in which a performer increases their focus on certain elements while casting aside others within the musical texture to make performance easier, I suggest that the average listener of country music likely engages more with the drums and vocals, and that this can alter their experience of meter in a song. For the average listener, the drums and vocals are the most accessible and most compelling parts of a song to engage with physically, whether by dancing, tapping out the drumbeat, or singing along. Hughes and Vuvan (forthcoming) confirmed in an experiment that listeners readily attend to the backbeat and argue that the backbeat is a second attentional peak to which a listener entrains. The combined strength of the vocal and drum backbeats puts even more weight on the traditionally weak beats for the listener. The phenomenal accents on 2 and 4 marked with an end-accent (boxed in red in Figure 5.7) do not negate the beginning-accents on 1 and 3 (boxed in green)—the other instruments emphasize beats 1 and 3 with their changing harmonies. However, the combination of these contrasting accentual patterns has an effect of making all of the beats equal in strength, similar to that of four-on-the-floor.[83] In live performances of songs featuring vocal

[83] Four-on-the-floor is a rhythmic pattern used in popular music, particularly disco and electronic dance music, in which a steady, uniformly accented bass drum hit is heard on all four beats of a 4/4 bar.

backbeat posted on YouTube, audience members regularly clap or bob to every single beat during the performance, showing that they are not distinguishing between strong and weak beats.[84] Singers such as Jason Aldean sometimes prompt the audience to clap on each beat.[85] In these performances, the singer and the audience members seem to feel all of the beats as equally strong, which is the overall effect of combining the vocal and snare backbeat with the metrically normative guitars.

This conception of meter in country music has precedents in scholarship on Western art vocal music, particularly opera. In his analysis of the first two lines of "Columbine" from Schoenberg's *Pierrot Lunaire*, David Lewin (2006) rewrites the vocal melody so that the stresses in the poetry more closely match the metrical accents of the ¾ meter that Schoenberg chose. Lewin indicates that there is a difference between "vocal meter," which takes into consideration the natural poetic stresses, and the "written meter," which the violin fits to perfection (346). Reinhard Strohm (1986), following Thrasybulos Georgiades, uses the term *Periodenbeginn* to refer to a metric phenomenon where the stresses in the melody and the instrumental accompaniment do not coincide (293). In these cases, the instrumental accompaniment is typically beginning-accented, while the vocal melody is end-accented. Willian Rothstein (2022) refers to a similar phenomenon in Italian opera as *accentual equilibrium*, in which accents on odd- and even-numbered measures alternate in a systematic and predictable way." Frank Samarotto (1999) has referred to a similar feature in as *shadow meter*. In effect, this is the phenomenon that I hear happening in recent country music with vocal backbeat. The vocal melody has stresses one beat later than the instrumental accompaniment. For this reason, I

[84] See the following videos of Carrie Underwood performing "Church Bells": https://www.youtube.com/watch?v=0qCRQoVEtG0; https://www.youtube.com/watch?v=N_d6nomoDG0
[85] One such performance can be viewed at the following link: https://www.youtube.com/watch?v=EMg_0J2_U5w

analyze the vocal melody using end-accents, which I show with a downward arrow. I have simply taken it one step further to suggest that the cumulative effect of this for a listener is to feel all beats as equally strong.

5.5 Case Study 1: Lil Nas X's "Old Town Road"

One of the clearest examples of vocal backbeat happens in Lil Nas X's "Old Town Road" (2019), a song that was declared "not country enough" by many in the country music industry.[86] While this song contains many elements of recent country music, including vocal backbeat, "Old Town Road" sparked controversial discussions about what makes a song country. In March 2019, Lil Nas X achieved a rare feat with this song, charting on both the *Billboard* "Hot Country Songs" and "Hot R&B/Hip-Hop Songs" charts before being removed from the country chart after reaching number 19. Representatives from *Billboard* argued that "Old Town Road" "does not embrace enough elements of today's country music to chart in its current version" (Leight 2019).[87] Lil Nas X dubbed the song "country trap," a fusion of country and trap—a subgenre of hip-hop. Rap and Hip-Hop are not foreign styles to country radio. Recently, bro country artists have collaborated with rappers for their singles (e.g., Florida Georgia Line's "Cruise," featuring Nelly) and, in some cases, artists have even done the rapping themselves (e.g., Colt Ford's "Workin' On"). Despite featuring elements of trap including a drumbeat that is characterized by quick subdivisions in the hi-hat and a hip-hop vocal, the introduction features a prominent banjo sound, which is a timbral marker of country music; the lyrics discuss common country imagery

[86] Even though "Old Town Road" was "not country enough" for some, I consider it a mainstream country music hit, since it made it onto the *Billboard* "Hot Country Songs" chart. For me, "mainstream music" is broadly that which appears on radio, *Billboard* charts, or in Top Hits playlists.

[87] Many pointed to country music's racist history as the true reason that "Old Town Road" was removed from the country charts. For more on the complex history of country music, see Olivia Carter Mather's "Race in Country Music Scholarship" (2017) or Diane Pecknold's *Hidden in the Mix: The African American Presence in Country Music* (2013).

(horses, tractors, and cowboy hats); and the vocals feature vocal backbeat throughout, a vocal effect that was used in approximately 60% of the number one songs on *Billboard*'s "Country Airplay" chart in the three years prior to this song's release.[88]

Lil Nas X's "Old Town Road" has a text setting that is rhythmically extremely similar to the majority of songs featured on *Billboard*'s country charts, with an amped up 2 and 4 vocal backbeat heard throughout the verses; the first two bars are provided in Figure 5.8. This rhythm is especially similar to Jason Aldean's rapping on "The Only Way I Know," a song to be discussed in detail later in this chapter. The last two bars of Aldean's verse from "The Only Way I Know" are also provided in Figure 5.8 for comparison. Four sixteenths in both songs are regularly followed by a durational accent (circled in dotted red). The words with durational accents also rhyme or share ending consonants (boxed), further drawing the listener's attention to the backbeat. Additionally, Lil Nas X performs the words on the backbeats with slightly more energy than those on the offbeats or on beats 1 and 3, giving these words more emphasis with a subtle dynamic accent. In sum, "Old Town Road" features words on the backbeats that have durational accents, slightly more energy, and rhyme, all of which contribute to the perceived emphasis of these syllables.

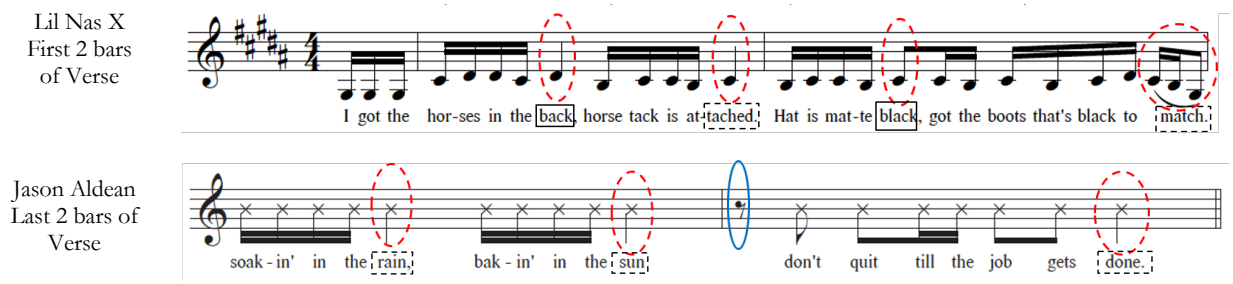


Figure 5.8 Comparison of a) the verse of Lil Nas X's "Old Town Road" and b) the last two bars of the verse of Aldean's "The Only Way I Know"

[88] In 1998, 20 years earlier, only 30% of the number one songs on *Billboard*'s Country chart contained vocal backbeat.

In terms of vowel quality, there is more consistency among formants for syllables on the backbeats in "Old Town Road" than those on odd-numbered beats or offbeats. I assume that consistency among mean formant values indicates that the syllables share a similar vowel quality. I suggest that the consistent use of similar vowel qualities on the backbeats is also a contributing factor to the perceived emphasis of these syllables. Figure 5.9 shows a 3D space in which the axes represent the mean frequencies of the first three formants (F1, F2, and F3) in Hz of the words in Verse 1 of "Old Town Road." The green words with square markers are those on beats 1 and 3, the red words with X markers are those heard on beats 2 and 4, and the black circles represent words in the "and" space of a beat (the offbeats). There is a clear difference between the formants of words on even-numbered beats (the backbeats), odd-numbered beats (the traditional strong beats), and between beats (the offbeats). The words on the backbeats cluster together in this 3D mapping of frequency space in two areas, which are circled in Figure 5.9. The F1, F2, and F3 values of "black," "back," "match," and "-tached" (the words/syllables on backbeats in the first four lines of Verse 1) are clustered together, and the F1, F2, and F3 values of "porch," "Porsche," and "horse" (the words on backbeats of the last four lines of Verse 1) are clustered together. The words on odd-numbered beats and the words on offbeats are, in contrast, much more spread out. The significant aspect of this text setting is that only the words on the backbeats consistently share a vowel sound and similar formants.

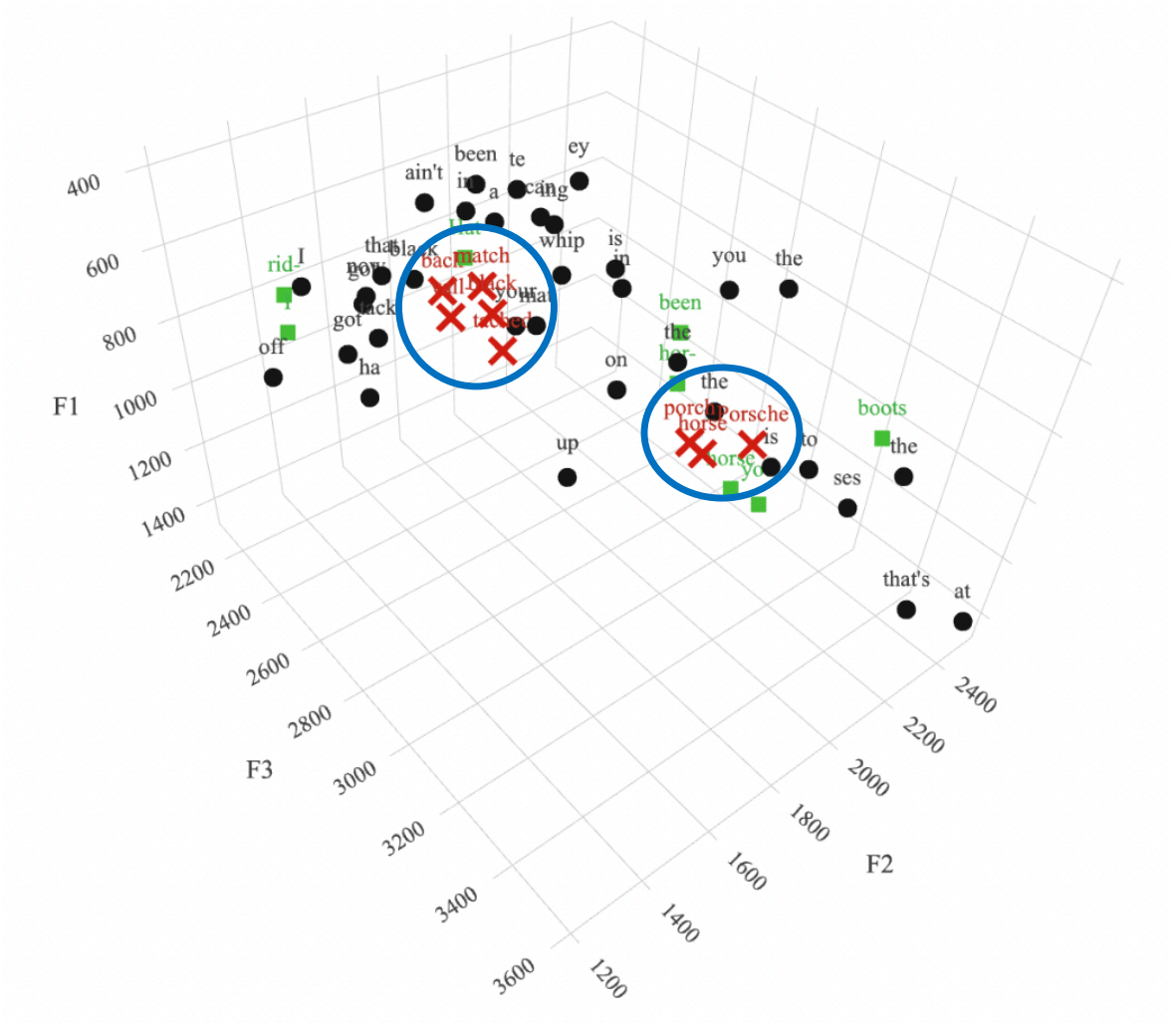


Figure 5.9 Mean F1, F2, and F3 values for words in the first verse of "Old Town Road." Words on odd-numbered beats are indicated with a green square, words on even-numbered beats are indicated with a red X, and words on offbeats are indicated with a black circle. The backbeat words are clustered together in this 3D mapping of frequency space. For an interactive 3D graph, click here.

The consistency among the formants of the backbeat words is further illustrated in Figure 5.10 with box and whisker graphs summarizing the distributions of the mean frequencies of the first three formants for the words on even-numbered beats, odd-numbered beats and offbeats. The graphs on the left show the dispersion of formant values for the first four lines of Verse 1, i.e., those with the "æ" vowel sound. The graphs on the right show the dispersion of formant values for the last four lines of Verse 1, i.e., those with the "ɔ" vowel sound. The 25th to 75th

quartile boxes in Figure 5.10 are noticeably smaller for the backbeats across the three formants compared to those on beats 1 and 3 and the offbeats. For the first four lines (the graphs on the left), the backbeat words have F1values ranging from 636 to 757 Hz with a standard deviation of 52 Hz, whereas the F1 values of the words on odd-numbered beats range from 597 to 997 Hz with a standard deviation of 189 Hz, indicating that the dispersion of words on beats 1 and 3 is much greater.[89] Likewise, the F2 values of the backbeat words range from 1672 to 1720 Hz with a standard deviation of 21 Hz, whereas the F2 values of the words on odd-numbered beats range from 1721 to 2432 Hz with a standard deviation of 296 Hz. The F3 values of the backbeat words for the first four lines range from 2503 to 2687 Hz with a standard deviation of 83 Hz, whereas the F3 values of the words on odd-numbered beats range from 2524 to 3259 Hz with a standard deviation of 329 Hz. The dispersion of the offbeat words is even greater than that of the words on odd-numbered beats for the first four lines of Verse 1. The last four lines also show a similar difference in the dispersion of values. The backbeat words are noticeably closer in range than those on odd-numbered beats or offbeats. The remarkable aspect about Lil Nas X's singing and word choice on "Old Town Road" is that the backbeat words consistently share similar values for the first three formants compared to the non-backbeat words. While the other songs to be discussed in this chapter also maintain consistency of formants for words on the backbeats, they do not have it as regularly across multiple formants as Lil Nas X exhibits in "Old Town Road."

[89] A smaller standard deviation means that data are less spread out and there is a smaller distribution of values.

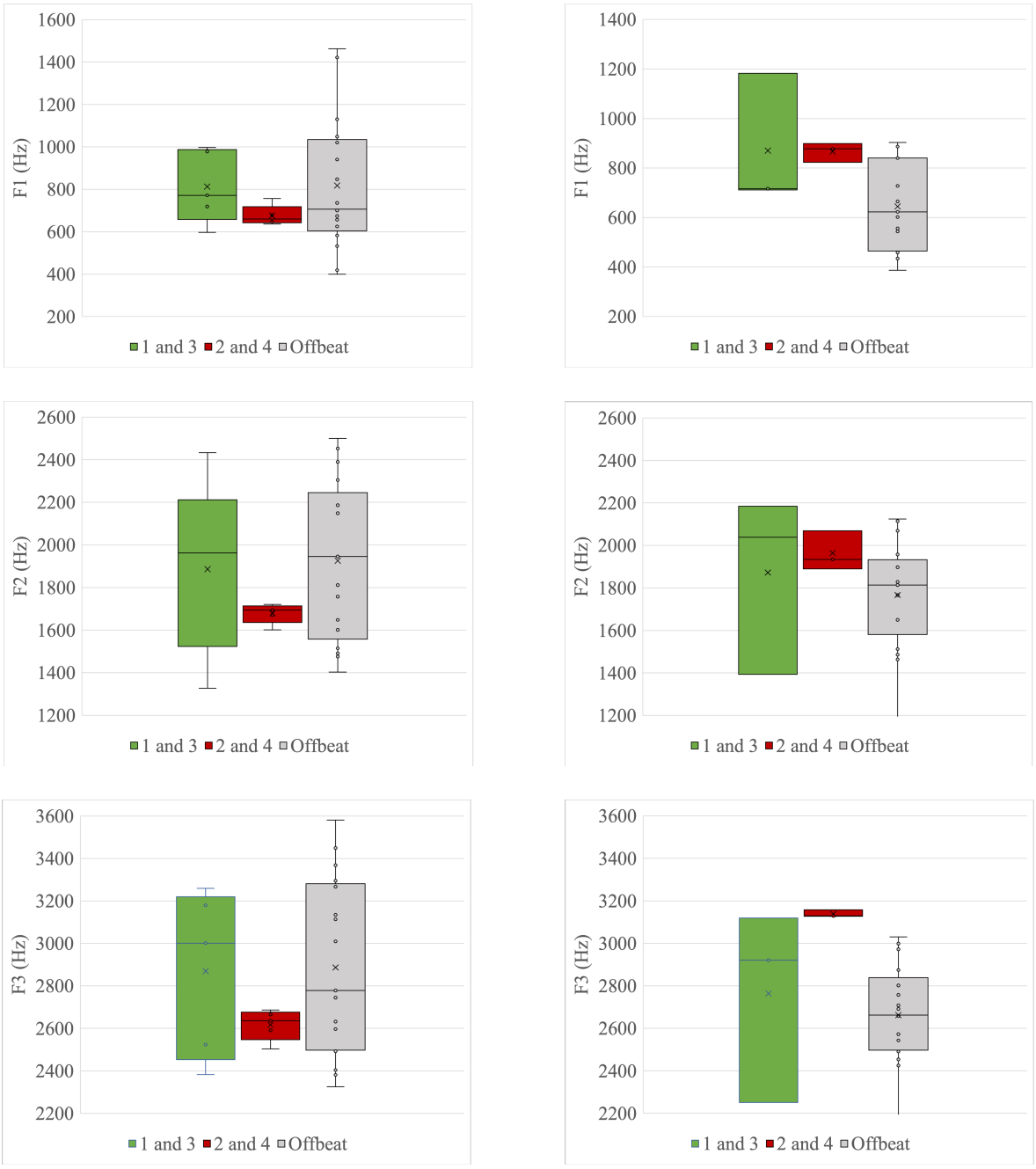


Figure 5.10 Left: The central tendency of mean formant frequencies of the words on odd-numbered beats, even-numbered beats, and offbeats for lines 1–4 and 7 of Verse 1 of "Old Town Road." These lines feature words with the "æ" vowel sound on the backbeats. Right: The central tendency of mean formant values of the words on odd-numbered beats, even-numbered beats, and offbeats for lines 5, 6, and 8 of Verse 1 of "Old Town Road." These lines feature words with the "ɔ" vowel sound on the backbeats.

Words with vowels that are produced in a similar way with the throat, tongue, and lips are expected to share similar formants (Sundberg 1991). Because the words on the backbeats in the first four lines feature the "æ" vowel sound and the backbeat words in the last four lines feature the "ɔ" vowel sound, they cluster in their respective areas on the chart. Yet, in other case studies later in this chapter, there are instances when the words on the backbeats do not share vowel or consonant sounds, yet they still have similar formants. This consistency among the mean frequencies of formants indicates a similarity in the way the sound is produced, as well as the vowel quality, of the backbeat words. This similarity of vowel quality draws our attention to these words in the vocals, making them prominent in the texture, much like the snare drum in the standard beat shown earlier.

The focus above has been on vowels since vowels are best measured through formants; however, the consistent use of similar consonants on particular beats can also draw a listener's attention to these moments. Many of the consonant sounds that Lil Nas X uses on the backbeats mimic the snare drum sound, especially the consonant sounds of "ck" and "ch." The "ck" sound, as in "back" and "black," is a voiceless velar plosive, meaning that it is produced without the use of vocal cords, but with the back of the tongue against the roof of the mouth and a full obstruction of airflow. The "ch" sound, as in "match" or "-tached," is a voiceless postalveolar affricate, meaning that it is similarly produced without the use of vocal cords. The "ch" sound is articulated with the tip of the tongue at the alveolar ridge (the point where the teeth and gums meet on the roof of the mouth) and by first stopping the airflow, then allowing air to flow through a small opening between the lips. Both of these sounds, to my mind, are onomatopoeic for the snare drum that is also heard on these beats.[90] Ammirante and Copelli (2019) note that

[90] For more on the ways in which using similar vowels and consonants can alter our perception of words and sound quality, see Noriko Manabe's "Abe Road: Kuwata Keisuke's Beatles Parody" (2021). Manabe discusses the ways in

music producers will use "boom bap" as onomatopoeia for the sounds of the kick drum and snare, respectively. Some other onomatopoeia for the sound of a snare drum might be "rum pum" (as in the lyrics from the song "Little Drummer Boy," which is meant to imitate a snare drum), as well as "chick," "crack," "blatt," and "cack." All these examples of onomatopoeia feature plosive or fricative consonants. The consonants are created by, at some point, stopping the airflow and using turbulence to create the sound, giving the sound a similarly noisy quality to the snare drum. While "Old Town Road" consistently features plosive and fricative consonants on the backbeats, the two other country songs discussed below in this chapter use other techniques to emphasize the words on the backbeats.

5.6 Case Study 2: Carrie Underwood's "Church Bells"

Carrie Underwood's "Church Bells" (2016) features the amped up 2 and 4 vocal backbeat, in which beats 1 and 3 are accented as usual, but the listener clearly hears performative emphasis on beats 2 and 4. This song has a cumulative texture. It begins with a solo banjo for two measures (Figure 5.11). Then a bass drum, snare drum and a gliss from the bass guitar are heard on the downbeat. When Underwood begins with the verse, she is accompanied by the banjo riff from earlier, bass guitar glisses on the downbeat, and a four-on-the-floor kick drum. In the fifth measure of the verse, the drums switch to the basic beat and the guitars enter. This texture continues throughout the rest of the song with a prechorus and chorus following each verse. The banjo introduction features a two-note motive (C#–D#) that is heard throughout Underwood's vocals in the verse and prechorus (Figure 5.12). The metric placement of this motive, in addition to the phenomenal accents, informs my analysis of this song.

---

which Kuwata Keisuke chose Japanese lyrics with similar vowels and consonants when covering songs by the Beatles to mask his critiques of Japanese politicians.

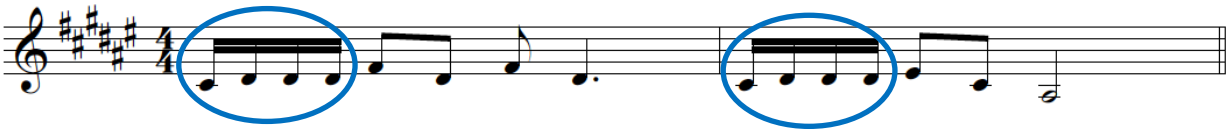


Figure 5.11 The introduction, played on banjo, of Carrie Underwood's "Church Bells," featuring the two-note motive (C#–D#) heard throughout her vocals.

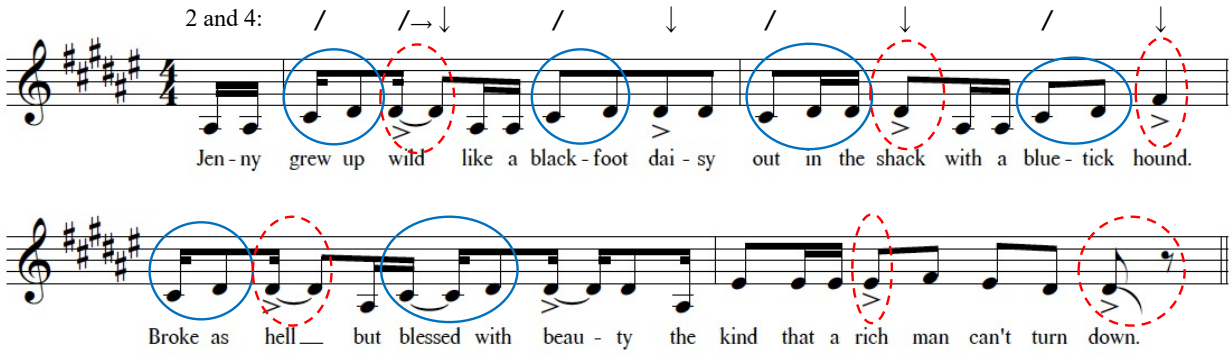


Figure 5.12 The first half of the verse. The two-note motive from earlier occurs on beats 1 and 3 followed by textual accents and performative emphasis on beats 2 and 4.

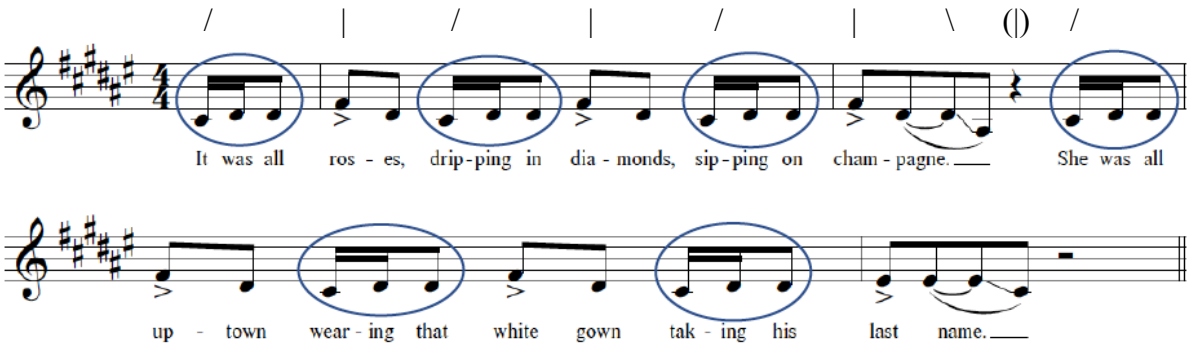


Figure 5.13 The prechorus of Carrie Underwood's "Church Bells." The two-note motive appears on beats 2 and 4 with text accents and phenomenal accents occurring on beats 1 and 3.

In the verse (Figure 5.12), Underwood retains the position of the C#–D# motive on beats 1 and 3 from the banjo introduction but adds phenomenal accentuation to beats 2 and 4 through durational accents and a slight dynamic change. Similar to David Temperley, who in his "Syncopation in Rock: A Perceptual Perspective" (1999) argues that poetic stresses in the lyrics are perceived in relation to strong beats in the musical meter, I consider syncopations to exist in

relation to a beat. Thus, I consider "wild" and "hell" in the verse to be "on" beat 2. If analyzed with beats 1 and 3 as beginnings, beats 2 and 4 would probably be considered continuations. I prefer, however, the reading of beats 2 and 4 as end-accents and beats 1 and 3 as anacruses. The reason has to do with the similarities between the verse and the prechorus.

Significantly, in the prechorus (Figure 5.13), the two-note C#–D# motive from beat 1 of the banjo introduction and verse melody appears on beats 2 and 4, and stressed syllables appear on beats 1 and 3 with pitch accents. The similarities between the prechorus and the verse suggest that the verse is best analyzed using end-accents on the backbeats. So, locally, we can analyze Underwood's verse melody as anacrustic on beats 1 and 3 and end-accented (↓) on beats 2 and 4.

In live performances, Underwood herself physically shows that beat 4 is emphasized in the verse. In her performance of "Church Bells" at the 2016 Academy of Country Music Awards, Underwood marks beat 4 by raising or dropping her hand or shifting her head as she comes down the stairs.[91] These movements switch to beats 1 and 3 when the prechorus starts and Underwood moves the stressed syllables to beats 1 and 3. Krahmer and Swerts (2007) found that marking beats visually has a significant effect on the perceived prominence of words both by the speaker and the person observing the speaker. If a word is marked visually, it is more likely to be produced with more emphasis by the speaker and to be perceived as prominent by the observer. Underwood's marking of beat 4 in the first verse and beats 1 and 3 in the prechorus during her performance makes the effect of vocal backbeat in her singing even more pronounced.

In many recent country music songs featuring vocal backbeat, there are rhyming words, assonances, and/or recurring consonants regularly on the backbeats. "Old Town Road" is one such example discussed above. In Underwood's "Church Bells," however, use of end-rhyme,

[91] A video of Underwood's 2016 ACM performance can be located at the following link: https://www.youtube.com/watch?v=4vSmKbu20Fs

head-rhyme, and assonance happens less frequently on the backbeats. In the first four lines of Verse 1, "hound" and "down" share assonance, but they are the only words that do. It is notable, however, that these two words are the only ones to share assonance because they are the words that end the two-bar phrases. The similarity of vowel sounds helps to draw a listener's attention to these words. The other words in the first verse, however, do not have similarities of vowel or consonant sounds. Despite this, Figure 5.14, a 3D graph of the mean formant frequencies of the words in the first four lyric lines from Verse 1, shows consistency among formants for more words appearing on the backbeats. While "hound" and "down" share similar formants (as is to be expected due to the similarity of vowel sounds), "shack" also has a similar formant structure to these two words. Additionally, "hell" and "rich" share similar mean formant frequencies. These two words appear on beat 2 of the third and fourth measures of the verse as transcribed in Figure 5.12. To make "rich" and "hell" have similar formants, Underwood likely sings these words slightly differently than they would be spoken to affect the formant frequencies. Changing the pronunciation of words to make them sound more similar is a common tactic in country singing. In an interview with Billy Dukes for Taste of Country, Devin Dawson states that "dog's running" and "God's country" is a perfect rhyme. As Dukes notes, "phonetically they're not a *perfect* rhyme, but this is country music where accents make up the difference" (Dukes 2019, emphasis in original). These are imperfect rhymes similar to those discussed above; they share a vowel sound and a consonant sound. These types of rhymes are very common in country music, and often pronunciation plays a role in making the words sound more similar. In Chapter 2, I provided the example of Billy Currington's singing on "People Are Crazy" (2009), in which he rhymes "wars" with "divorced" and "him" with "A.M." by altering his pronunciation of one of

the words. Here, the formant structure indicates that Carrie Underwood likely alters the pronunciation of "rich" and/or "hell" to make the words sound more similar.

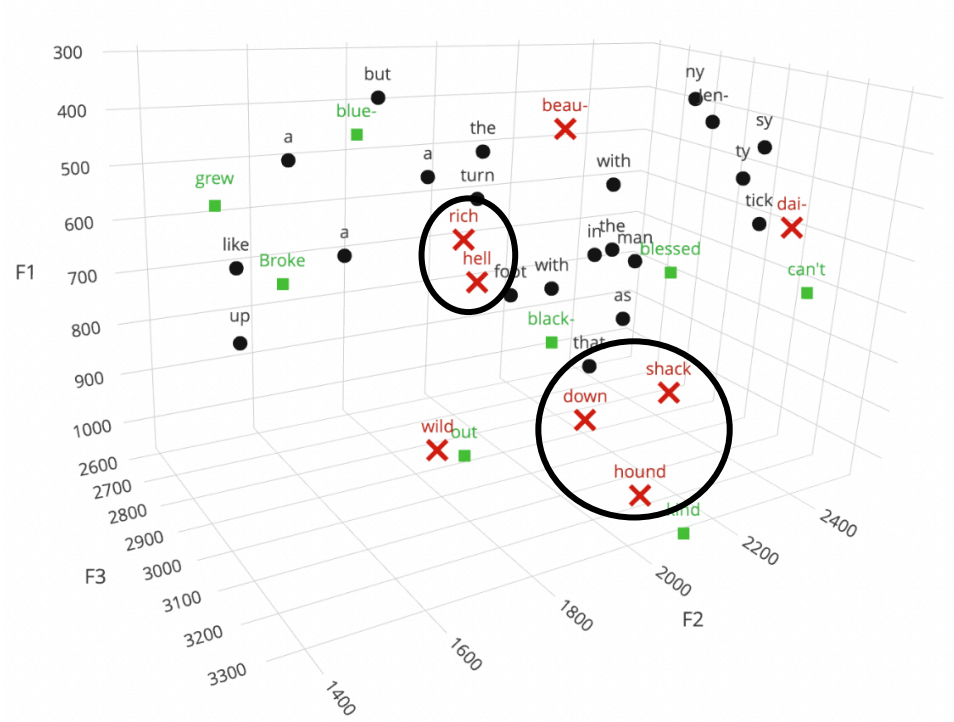


Figure 5.14 Mean formant values for "Church Bells" verse. Words on odd-numbered beats are indicated with a green square, backbeat words with a red X, and offbeat words with a black circle. For an interactive 3D graph, click here.

Figure 5.15 shows the distribution of mean F2 values for the first four lines of verse 1 of "Church Bells." Similar to that in "Old Town Road," the 25th to 75th quartile box for the backbeat words is significantly smaller compared to those of the other words in the first four lines of verse 1. In other words, the F2 values of words on the backbeats are more similar than those on beats 1 and 3 or the offbeats. Most of the backbeat words appear within the range from 1700 to 2100 Hz for F2 with a standard deviation of 235 Hz, while the words on odd-numbered beats range from 1400 to 2300 Hz with a larger standard deviation of 311 Hz. In contrast, in the prechorus, where the words on beats 1 and 3 are perceived as more emphasized, the F2 values of words on odd-

numbered beats have a lower standard deviation than those on the backbeats (187 Hz for words on odd-numbered beats compared to 232 Hz for backbeat words).

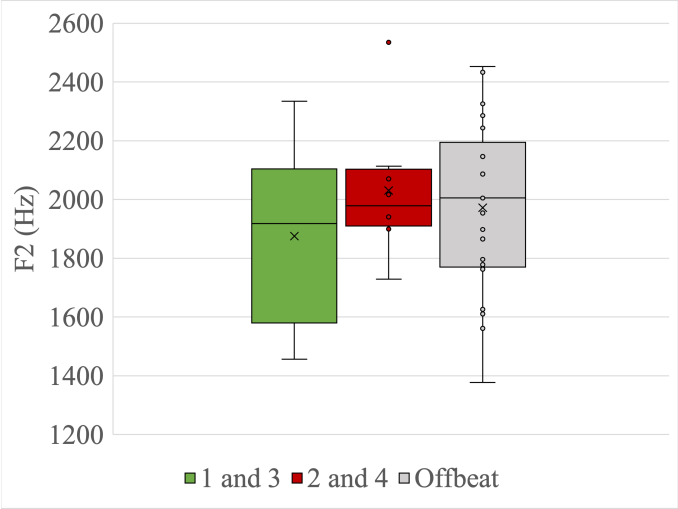


Figure 5.15 The central tendency of mean F2 frequencies for words on odd-numbered beats, backbeats, and offbeats in the first four lines of verse 1 of "Church Bells."

5.7 Case Study 3: Jason Aldean's "The Only Way I Know"

The last case study, Jason Aldean's 2012 hit single "The Only Way I Know," features both types of vocal backbeat. In some measures, the downbeat is deemphasized with a rest (circled in solid blue in Figure 5.16), putting the stress on beat 2 and, because of the alternation of strong and weak beats, beat 4. Other measures maintain a metric accent on beats 1 and 3, but through phenomenal accents including longer durations (circled in dotted red) and repetitions of ending consonances (boxed), beats 2 and 4 are also accented.

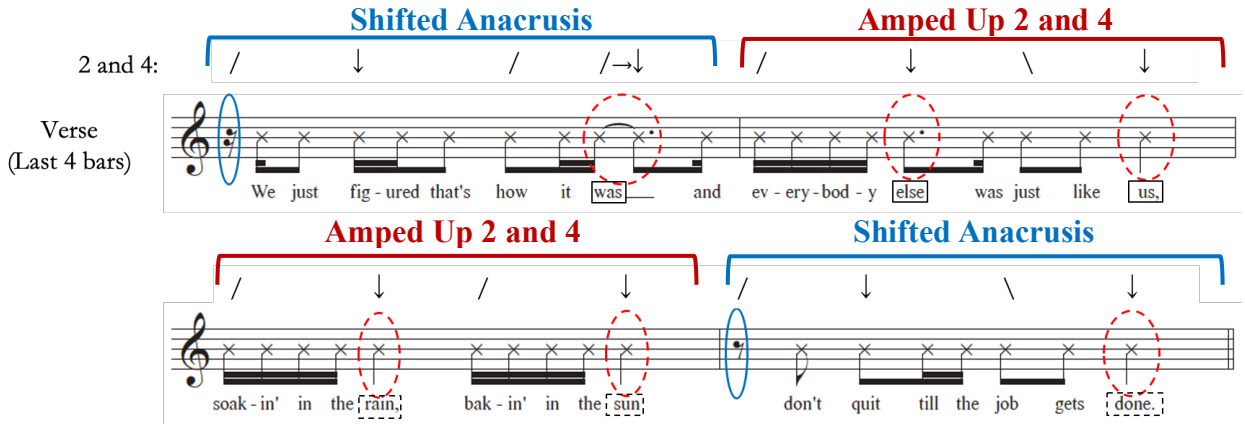


Figure 5.16 Verse of Jason Aldean's "The Only Way I Know" (2012), with both the shifted anacrusis backbeat and amped up 2 and 4 backbeat.

The formants of words in the last four bars of verse 1 show some similarities to the formant structures of the previous songs examined, but also some differences. The words "was," "else," and "us" group in a similar area of the graph in Figure 5.17, indicating a consistency among the formants, which in turn indicates a consistency in vowel quality. The first syllable in "figured" and "rain" also share similar formant frequencies; however, these words do not share vowel sounds or even initial consonants. This indicates that Aldean likely manipulated the way in which these words are sung in order to make them sound more alike, similar to the way in which Underwood changed the usual pronunciation of "rich" and "hell" to make them sound more similar. It is significant that the first syllable of "figured" and "rain" share a similar formant structure, because they are the first backbeat words of the two-measure phrases. While Aldean's vocals share this aspect of manipulation of the vocal tract to produce words with similar sound qualities with Underwood's vocals, his song differs from Underwood's when examining the dispersion of formant frequencies. In contrast to Underwood's "Church Bells," where the mean F2 values of the backbeats were the most similar, for Aldean's "The Only Way I

Know" the mean F1 values are significantly less dispersed compared to the words on odd-numbered beats and offbeats in the last four bars of the verse, as illustrated in Figure 5.18.

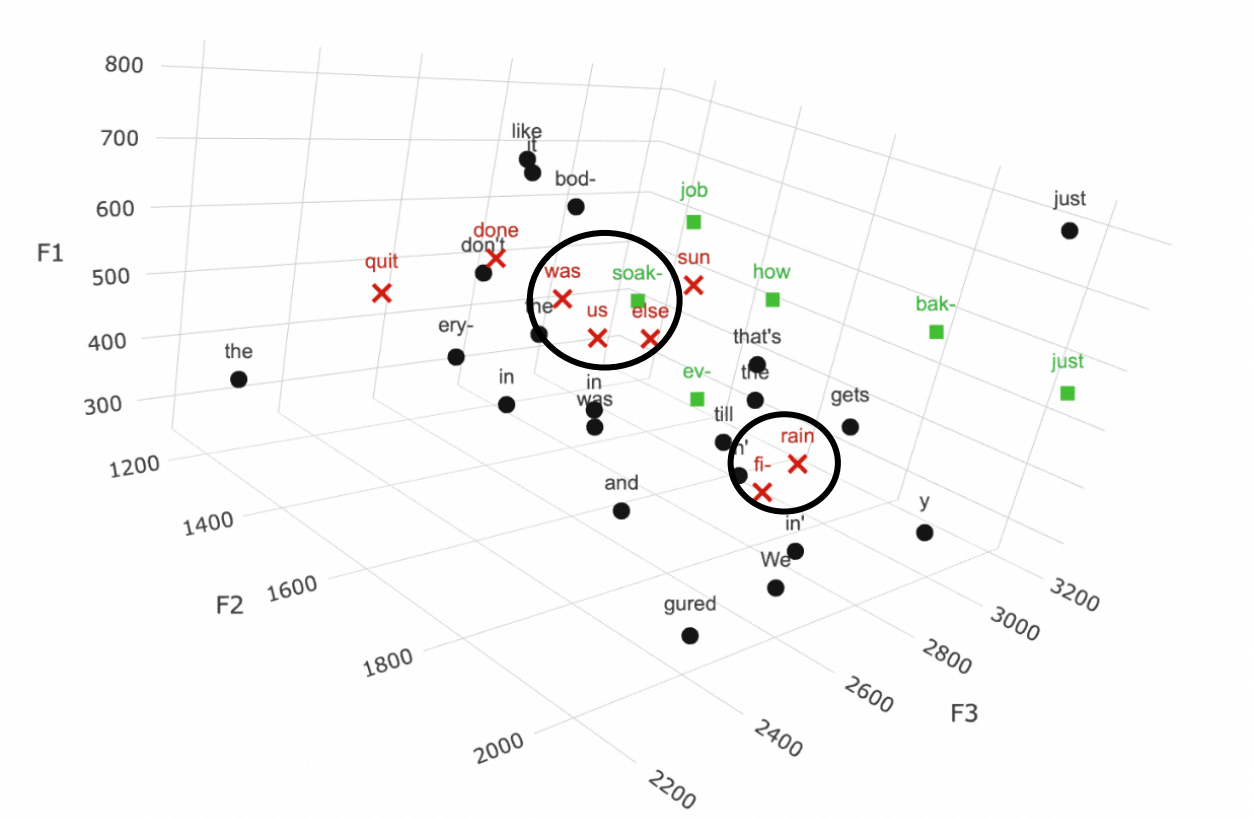


Figure 5.17 Mean formant values for "The Only Way I Know" verse. Words on odd-numbered beats are indicated with a green square, backbeat words with a red X, and offbeat words with a black circle. For an interactive 3D graph, click here.

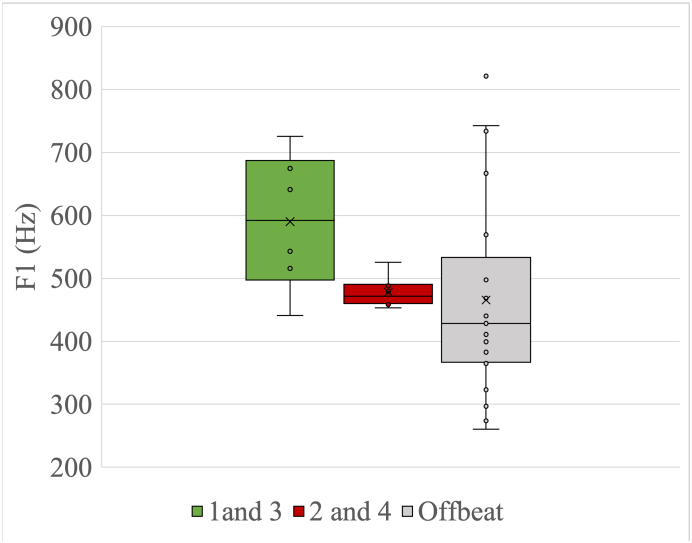


Figure 5.18 The central tendency of mean F1 frequencies for words on odd-numbered beats, backbeats, and offbeats in the last four lines of verse 1 of "The Only Way I Know."

5.8 Summary of Vocal Formant Usage in Case Studies

The three songs discussed above—Lil Nas X's "Old Town Road," Carrie Underwood's "Church Bells," and Jason Aldean's "The Only Way I Know"—contain perceived vocal backbeats. These songs feature similar mean frequencies of the formants for words on the backbeats more than those on beats 1 and 3 or the offbeats. Lil Nas X made the words on the backbeats more similar across all three formants, whereas Carrie Underwood made the F2 values the most consistent, and Jason Aldean sings so that F1 is the most consistent among the backbeat words compared to the words on odd-numbered beats and offbeats. The consistency among formant values in "Old Town Road" is largely due to the use of rhyming words, assonances, or recurring consonants on the backbeats, but for Underwood's and Aldean's songs, words that do not seem to share sounds have similar formant values as well, indicating a transformation of the vocal tract or position of the lips and tongue to make these words sound more similar. The consistency among mean frequencies of the formants results in a perceived emphasis of words in

the vocal melody in these mainstream country songs, all of which were released in the last ten years.

5.9 Vocal Backbeat and the Narrative Structure of Country Songs

Neal (2007) has suggested that aspects of musical structure, including form and harmonic structure, contribute to a country song's story. Here I suggest that the use of vocal backbeat or lack thereof also contributes to a song's narrative in recent mainstream country songs. Figure 5.19 summarizes the uses of vocal backbeat for narrative purposes. In sections with vocal backbeat, there is a tension between the phenomenal accents in the vocals and the conventional metric accents in the guitars and keyboards.[92] This often accompanies a conflict in the song's lyrics. Once the conflict resolves in the lyrics, the patterns of accent in the vocals, guitars and keyboards reconcile as well.

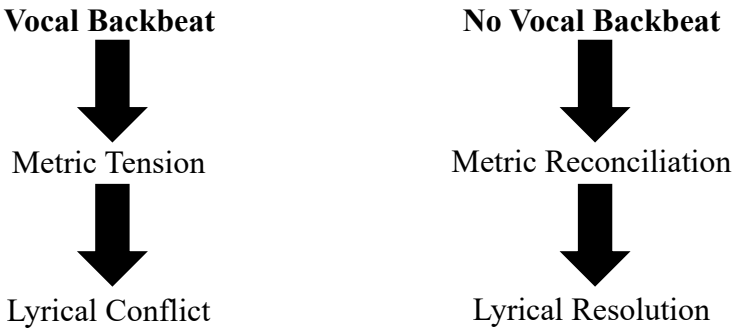


Figure 5.19 Summary of uses of vocal backbeat in recent mainstream country music.

"Church Bells" tells the story of a poor young woman who married a rich oil man but soon discovers that he is abusive. The verses feature vocal backbeat and recount the tumultuous story of how the man would get drunk and physically abuse his young wife, which leads to her

[92] William Rothstein (2022) describes a similar effect with "accentual equilibrium" in late-eighteenth century Italian opera.

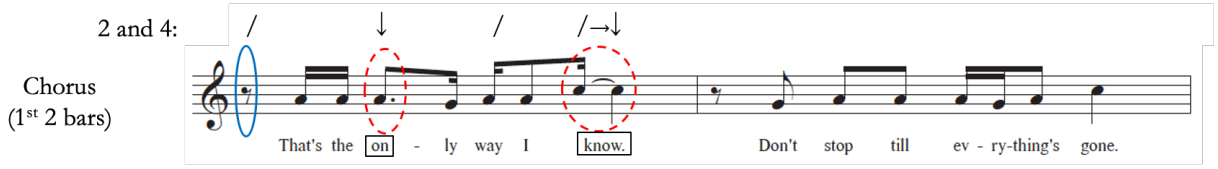
poisoning his drink. The divorce between the metrical accents and the phenomenal accents in the vocals and drums emphasizes the darkness of the story. The text setting and phenomenal accents in the chorus, on the other hand, align with the conventional metric accents. At this moment, Underwood sings of the young woman listening to church bells ring and a choir singing, which provides reassurance that everything will be all right.

Underwood's "Church Bells" follows a model for popular songs that David Temperley (2007) refers to as "loose verse/tight chorus." Temperley's definition of "loose verse/tight chorus" refers specifically to the unification or lack thereof of pitch organization between the melody and harmonic progression. We might extend this concept further to the use of vocal backbeat. In the verses of Underwood's "Church Bells," the vocals and the instrumental parts are stratified with a different pattern of phenomenal accentuation, but in the prechorus and chorus the vocals are unified with the instrumental parts with the phenomenal accents joining the conventional metric accents. While this is certainly a model used in some recent mainstream country songs, it is not used in all.

Jason Aldean's "The Only Way I Know," for instance, uses vocal backbeat throughout the song with one notable exception at the end of the chorus, a partial transcription of which is given in Figure 5.20. Aldean begins and ends the chorus with the line "That's the only way I know," but the placement of the line within the measure is different. At the beginning of the chorus, Aldean places "only" and "know" on beats 2 and 4. They create durational accents, and both feature the "oʊ" vowel sound. At the end of the chorus, these same words are positioned on beats 3 and 1 respectively, finally placing stressed syllables on metrically accented beats. This song is about the struggles of living and working in a small farming community. Aldean sings lines such as "you get tired and you don't show it," "we didn't know the odds were against us,"

and “maybe there’s another path that will get you there a little bit faster.” As the song’s protagonist questions his current way of life, the vocal melody consistently contradicts the conventional metric accents. Only when he realizes that his life is how he wants it are the two finally reconciled.

a) Shifted anacrusis backbeat placement



b) Metric 1 and 3 placement

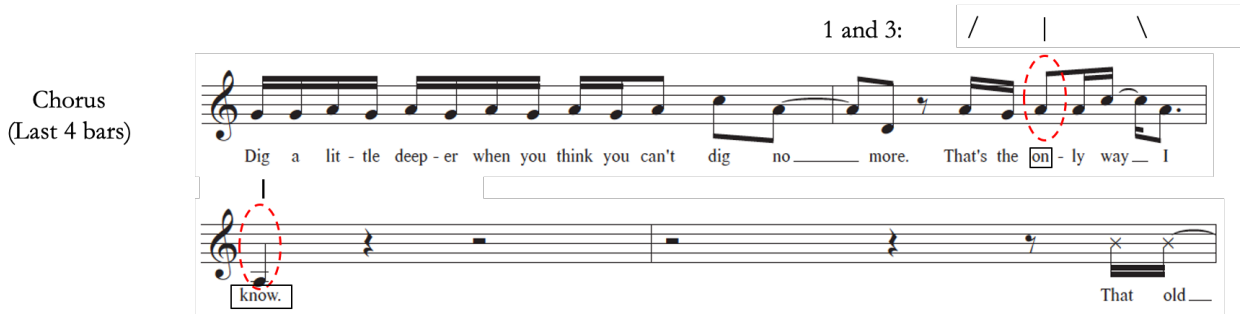


Figure 5.20 The shifting placement of the title line, “That’s the only way I know,” in the chorus.

Similarly, Lil Nas X’s “Old Town Road” features vocal backbeat throughout most of the song, though it still aids in telling the narrative of the lyrics. The road of which Lil Nas X speaks represents the path to success. The horse he is riding represents the fact that, despite having little in the way of material belongings, he believes he has the talent and determination to make it in the music industry. In the chorus and bridge, Lil Nas X declares that he is determined to become a success in the music industry despite the adversity he might face and all the people in his life telling him he will not succeed. The second verse is set in the future, where the protagonist has upgraded from a horse to a tractor and wears designer clothes, but he is also committing adultery

and doing drugs, suggesting that there are downsides to fame. While Lil Nas X sings of being confident that he will become successful in the music industry and wanting some of the excesses that come with it, the constant use of vocal backbeat suggests a tension and uncertainty in this belief. Even in the bridge, when we hear more of a focus on beat 3 with "nothing," the phenomenal accents in the vocals do not fully reconcile with the metric accents, as the word "tell" is still emphasized on beat 2. The tension between the conventional metric accents and the vocal backbeat could also represent his relationship with his detractors and their expectations of him to focus on a career path that is more conventional and stable. Lil Nas X rejects these lifestyle and metric conventions, opting for a vocal backbeat instead.

5.10 Conclusion

Returning to the mashup by Sir Mashalot, it is indeed true that many recent mainstream country songs sound similar. They use similar drum patterns and instrumentation, similar chord progressions, and have a comfortable walking tempo. On the other hand, as illustrated by the case studies in this chapter, these songs sound similar for a very specific reason—the vocal backbeat. In this chapter, I defined and identified instances of vocal backbeat in recent mainstream country music. The vocal backbeat is a recurring heavy emphasis in the vocals on beats 2 and 4 through the use of phonetic accents, rhyming words, and/or repetitions of vowel or consonant sounds. These similarities of vowels and consonants affect the formant frequencies. Formant frequencies are directly related to sound quality in vocals. Once this similarity of vowel quality is identified, anyone can make a similar mashup. A mashup that I created featuring three of the songs discussed in this chapter can be found at the following link: https://youtu.be/XQNIFaWuT0M.

I also presented some possible reasons as to why vocal backbeat has become a feature of this style of music. The instrumental parts support the story that is being told, as does the text setting. As any country music fan will likely tell you, storytelling is a mainstay of country music and one of the features most loved by fans. Songwriters of recent mainstream country music employ the text-setting device of vocal backbeat, a recurring heavy emphasis in the vocals on beats 2 and 4 throughout a section of their song, to support the story they are telling. The resolution of vocal backbeat with the coincidence of metric and phenomenal accents occurs at the moment of resolution in the song's narrative.

This chapter's main goal was to consider the ways in which the sensation of backbeat is created in the vocals of recent mainstream country music, but ethics also played a role in my analytical choices. I grounded my analyses in the perspectives of those who create the music, including songwriters (Devin Dawson and Michael Hardy), producers (Gregg Todd), the performers (the movements of Carrie Underwood and Jason Aldean during live performances), and listeners' movements while attending a concert, similar to how I tried to ground my analyses in Chapter 4 in Tagaq's perspectives on her own music and in Inuit epistemology. I also considered the context in which vocal backbeat is used in recent country songs, rather than only identifying instances of it. With this, I broadened my analyses to include the story that the songwriters tell and the emotion that they try to convey through it. Lastly, I purposely foregrounded Lil Nas X's "Old Town Road" in this chapter, a song that was deemed "not country enough" by *Billboard* and many in the country music industry. The thought that Lil Nas X's "Old Town Road" did not deserve to be on country radio or on the *Billboard* country charts shines a light on the rampant racism associated with country music. This chapter showed that, based on other country music released in the last decade, Lil Nas X followed a very similar

formula as many country songwriters and producers have in order to make a country hit. With these three choices, I have continued to place importance on taking an ethical stance in analysis, even when it might not be obvious that there is a need to do so.

Chapter 6: Sound Qualities and Form in Selected Songs of Tanya Tagaq

> "There's just basic human instincts, like fear, laughter, anger, lust. Like, if somebody goes, 'oh, that was really sexual.' In my mind, I'm like, 'oh, you're pretty sexually oppressed.' Or, they go, 'that was really angry'…It's like, 'oh, like maybe you've got anger." It's, like, whatever they want to see is…what they need to see. So, that's why I typically don't use words. I want to touch as many people as possible." – Tanya Tagaq (CBC Music 2016)

Tanya Tagaq is a storyteller, sometimes with words, such as in her novel *Split Tooth* and some of her more recent musical releases, and sometimes in wordless songs. In the above quote, Tagaq explains that she wants the listener to connect to and interpret her songs in the way that they need to. From this quote, I interpret that Tagaq hopes that listeners will experience her music in different ways, and because people listen differently, it is likely that different listeners will interpret her music in different ways. In Chapter 4, based on the data examined, it seemed that Tagaq and I often experienced beat in "Ilunikavi" (2009) in different ways. The previous chapters have also shown that the voice is malleable, and sound qualities can change rapidly if the vocalist changes their vocal tract in some way. With the aid of acoustic measurements to map continuous data for sound qualities, I am able to examine these changes over time through an entire piece. In the current chapter, I return to the same video performance of "Ilunikavi" that was discussed in Chapter 4; however, this time, I consider the effects of changes in sound qualities on the form of the song. To honor the above quote in this chapter, I present multiple analyses of the form of "Ilunikavi" and a story as an interpretation of Tagaq's "Sivulivinivut" from her 2016 album *Retribution*. The reader may interpret the song's meaning and the analysis's connection to the song for themselves.

I interpret Tagaq's description of throat singing as related to sound qualities. I also have assumed that sound qualities not only play a role in the classification of her throat-singing

sounds (discussed in detail in Chapter 4), but that they also contribute to the larger structure of her songs. In this chapter, I revisit the acoustic measurements used in analyzing sound qualities—RMS energy (loudness), periodicity (noisiness), and spectral centroid (brightness)—to analyze the form of the video performance of "Ilunikavi" that was initially discussed in Chapter 4. Later, I try an alternative method of presenting a musical analysis through a fictional story and consider the implications of such analyses. These methods of analyzing the form of Tagaq's songs break from the traditional Western reliance on pitch and harmony to determine the form of pieces and allow for multiple analyses of an individual song, an aspect of Tagaq's music that she relishes (CBC Music 2016; see the quote at the beginning of this chapter).

6.1 Multiple Analyses of Form in "Ilunikavi"

In this section, I analyze the structure of "Ilunikavi" (meaning "We are blood") with a focus on how the sound qualities of the music—analyzed using RMS energy, periodicity, and spectral centroid—reveal a process.[93] According to Tagaq (2004), she wrote "Ilunikavi" for her cousin, describing how "close they are and in tune with each other."[94] I analyze a performance of "Ilunikavi" that comes from a video series entitled *Inuit Throat Singing – for iPod/iPhone* from the Open University (2009a).[95] I use this version of "Ilunikavi," as opposed to the original recording, because Tagaq performs solo in the video version, allowing for more precise data regarding individual sounds.

[93] For other studies in how non-pitch/non-rhythmic elements in music may be viewed as form-shaping devices, see McAdams (1999) and Scotto (2017).
[94] In the original recording, to show this bond, Tagaq sings with herself in the original recording, imitating the traditional throat-singing style of singing with another woman. Two tracks of Tagaq's singing were overlaid to produce the final version of "Ilunikavi." To hear the version from *Sinaa*, consult Tagaq's YouTube channel: https://www.youtube.com/watch?v=2pzW4ITPxg0
[95] To hear the version of "Ilunikavi" from her video series with the Open University consult Apple Podcasts or the following link: https://www.youtube.com/watch?v=vJlrZNrbMfc

I have observed that Tagaq has certain performance tendencies that shape the form of her pieces. As the interviewer in a 2019 *60 Minutes* special notes regarding a performance he attended, "her voice flickers, then builds to a rhythmic panting, then the tension builds, and she unleashes a sonic storm." This description is somewhat reductive but captures an overall arc that is common in Tagaq's performances. Often, she will begin with a sweet-sounding melody with a very pure sound, gradually introducing the deep growls, and then at some point she switches entirely to a rhythmic throat-singing–inspired improvisation, which slowly builds in dynamic and tension. She may also add screams, cries, and moans into her performance. This is fairly similar to the formal structure of "Ilunikavi," which has a large-scale ABA' form. In the A sections, Tagaq sings a smooth, pure melody with lyrics using Exhaling High sounds. In contrast, the B section features fast-paced, rhythmic throat singing. Tagaq uses all four of the throat-singing sounds discussed in Chapter 4 during the B section of "Ilunikavi."[96] After a tension-filled B section, Tagaq returns to a relaxed, soothing vocal melody that is similar to the beginning in the A' section. The B section is much noisier, due to the use of the deep growls. In terms of brightness, the B section has an overall brighter aesthetic than the A sections. The A section has less variation in acoustic measurements throughout the section compared to the B section. In what follows, I discuss the changes in acoustic measurements over time within the A section and the B section.

6.1.1 The A section, featuring a melody with lyrics

Taking a holistic approach to the analysis of sound qualities over a greater span of time in the individual sections, I first examine the moving average trendlines of the data for the A

[96] As described in Chapter 4, Tagaq uses four primary sounds when throat singing: Exhaling High (EH), Inhaling High (IH), Exhaling Deep (ED), and Inhaling Deep (ID).

section in Figure 6.1.[97] Tagaq uses only Exhaling High sounds throughout this section with lyrics.[98] As shown at the top of Figure 6.1, the A section begins with high energy spikes with large dips down to silences (between moments of articulating a new syllable) or audible Breaths In. The energy in the A section begins at a relatively high level, with its peak (at approximately 0.5) arriving about a third of the way through the section. After that, Tagaq's singing becomes softer and softer until the next section begins. In terms of noisiness, the first section of "Ilunikavi" features mostly pure sounds, with most of the sounds having a periodicity at or near 1, as shown in the middle graph of Figure 6.1, but sometimes the graph shows rapid spikes down to the 0.2–0.4 range. These are due to moments of silence just before Tagaq articulates a new syllable or when she draws in breath. These fleeting moments break up the pure melody tones, adding a certain level of "noisiness" (in the sense that it is not a pure, soothing melodic pitch at those moments). In the bottom graph of Figure 6.1, the spectral centroid data show that Tagaq's singing is duller throughout this section, with most of the data points hovering between 500 and 1500 Hz. There is just one instance of a slightly higher average spectral centroid at 29.74s with data points between 2000 and 2500 Hz. There are also some spikes in brightness outside of this region; however, as in the graphs for energy and periodicity, these outliers happen where there are moments of articulation (right before Tagaq starts a new syllable) or audible Breaths In, when the higher frequencies are lingering in the room in which this performance was recorded. Ultimately, the most substantial changes happen only in RMS energy for the A section. In contrast, the B section, to be discussed next, shows substantial changes in all three acoustic measurements.

[97] The window size of all moving average trendlines in this chapter is 100.
[98] The lyrics for "Ilunikavi" are in Inuktitut; a translation is provided in the liner notes of *Sinaa* (2004).

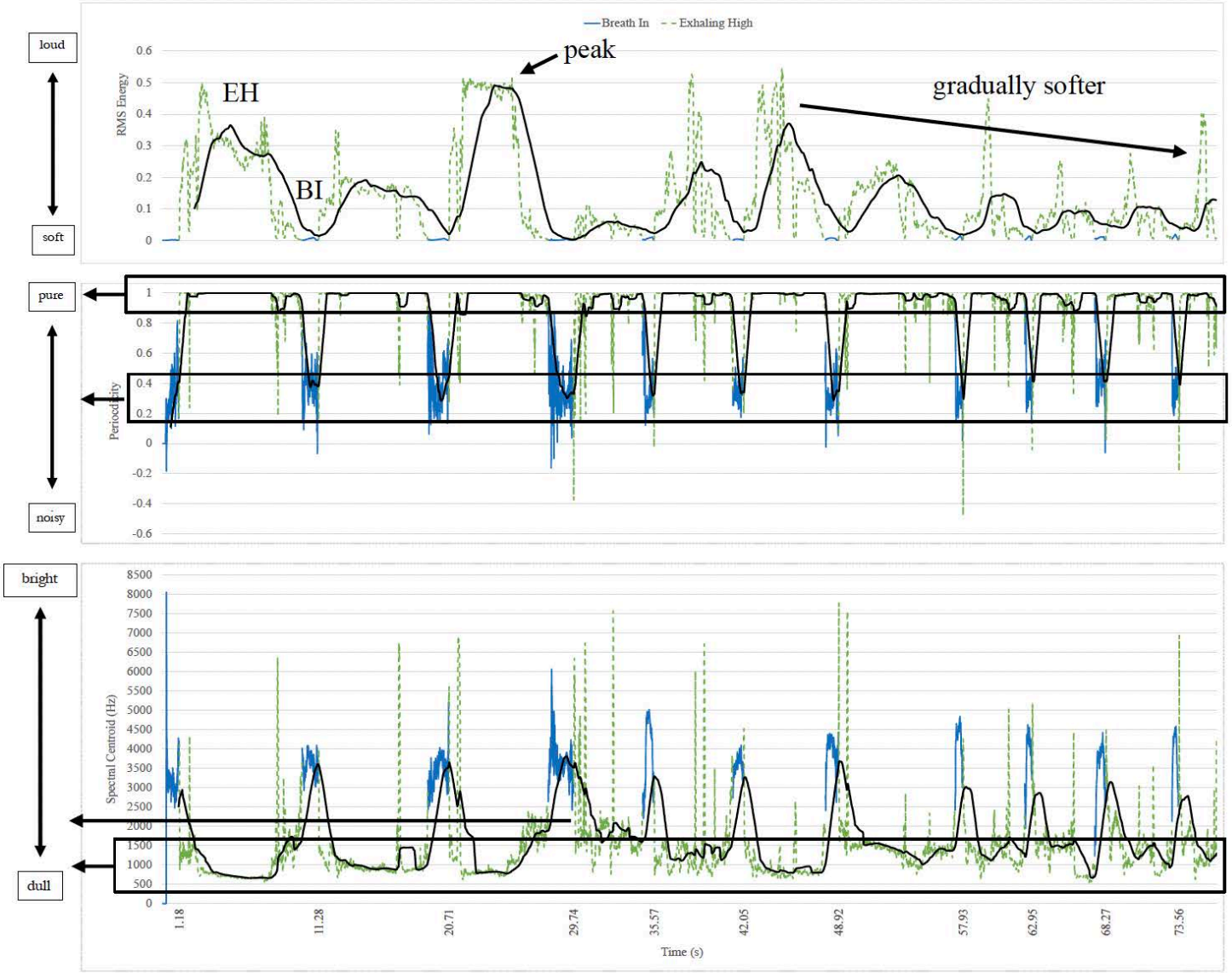


Figure 6.1 RMS energy (top), periodicity (middle), and spectral centroid (bottom) values and moving average trendline (black) for the beginning (A) of "Ilunikavi." Tagaq uses Exhaling High sounds exclusively throughout this excerpt.

6.1.2 The B section, featuring throat singing

The sounds of throat singing are typically performed cyclically in different combinations, which I refer to as *sound patterns*. The four throat-singing sound patterns used in the B section of the video performance of "Ilunikavi" are 1) EH followed by ED and a quick audible Breath In, 2) ED followed by a quick audible breath, 3) ED followed by IH, and 4) ED followed by ID.[99] The top of Figure 6.2 summarizes the four sound patterns with brackets and numbers.[100] A key to the right of the RMS energy graph summarizes the meaning of the numbers. The colors of the data points in the graphs indicate the individual sounds used, similar to the graphs found in Chapter 4, and can help to connect the graphs to the sound patterns. The first half of the B section is primarily a cyclic repetition of Sound Pattern 1 (EH followed by ED, then an audible Breath In).[101] In fact, this sound pattern comprises most of the B section. The second half of the B section features more variety in the sound patterns used, especially near the end. At approximately 127s, Tagaq briefly switches to Sound Pattern 3 (ED–IH) before returning to Sound Pattern 1. Toward the end of the B section, Tagaq regularly alternates between Sound Pattern 2 (ED–B) and Sound Pattern 1. She ends this section with a pattern not yet heard in this song, Sound Pattern 4 (ED–ID), which is followed by the first silence heard since the A section. This is the structure of the form for the B section if only pitches or the sounds used are considered.

If changes in RMS energy, periodicity, and spectral centroid are considered, three more interpretations of the form result. As shown at the top of Figure 6.2, Tagaq begins the middle

[99] Tagaq discusses these sound patterns in "The Sounds of Throat Singing" (2009).
[100] The sound patterns have been categorized based on Tagaq's descriptions of sound patterns that she demonstrated in "The Sounds of Throat Singing" (2009).
[101] In katajjait (plural of katajjaq), sound patterns are repeated cyclically until a new one is introduced. This is the way that Tagaq uses the patterns in the B section of "Ilunikavi."

section softly with low energy, then rises to a mid-level amount of energy around 92s. Next, she brings the energy level back down before building to the loudest part of the middle section approximately two-thirds of the way through, at 127s. This loudest moment in the B section could be considered the most tension-filled in terms of energy. Tagaq ends this section by returning to a softer dynamic with the least amount of overall energy heard in the B section.

In contrast to the double-arch shape of the energy values in which they twice begin low, rise, and fall again, the changes in periodicity (middle graph of Figure 6.2) suggest a growing tension toward the end of the section. Tagaq shapes the sounds with differing periodicities as the B section progresses. She begins singing with a mix of pure and noisy sounds as she moves out of the A section, which features only pure sounds. There is a dip to a mid-level around 85s before Tagaq returns to a purer sound quality at 93s. These changes in sound qualities happen while Tagaq sings only Sound Pattern 1. When Sound Pattern 2 is introduced at 115s, Tagaq's singing becomes noisier for a brief time, but her singing soon returns to a purer sound once she revisits Sound Pattern 1 for an extended length of time. Near the end of the section, when she begins regularly changing sound patterns, the overall sound quality becomes noisier, giving the song a tension it did not have earlier from a periodicity perspective. This noisiness at the end of the B section may connect to the song's lyrics, which are presented in the A sections. As mentioned above, Tagaq wrote this song for her cousin, and it discusses their close bond. The lyrics of "Ilunikavi" end with "We are blood/We are animals" (Tagaq 2004). While I tend to think of a noisy sound quality as tension-filled, perhaps the focus on deep growls that resemble the sound of animals and result in the noisy sound quality near the end of the B section is symbolic of this sweet sentiment. The deep growls could be interpreted as a representation of how tightly connected she and her cousin are.

The spectral centroid data, shown in a graph at the bottom of Figure 6.2, suggest a form just as complex as that of RMS energy and periodicity, though one with different moments of tension. Tagaq begins with a variable bright/dull sound quality. Soon after the beginning of this section, her singing gradually becomes brighter, reaching the brightest point in this section at 85s. Much of the B section features a mid-level brightness with a moving average between 1500 and 2000 Hz. There is a brief moment of duller sound quality between 127 and 129s. Toward the end of the section, Tagaq's singing gradually becomes brighter, though it does not quite reach the height of the peak brightness just before 93s. Bright sounds might be heard as tense or unpleasant sounding (Lerdahl 1987). In that case, there are two moments of tension in terms of brightness in the B section, one just before 93s and another at the end of the section. Importantly, this reading is significantly different from those based on energy and noisiness discussed earlier, in which the greatest tension in terms of loudness occurs at 127s (the loudest moment in the section) and in terms of noisiness occurs at 115s and again near the end of the section (the noisiest moments). This implies that listeners, depending on which sound qualities they attend to, could have markedly different interpretations of the same song.

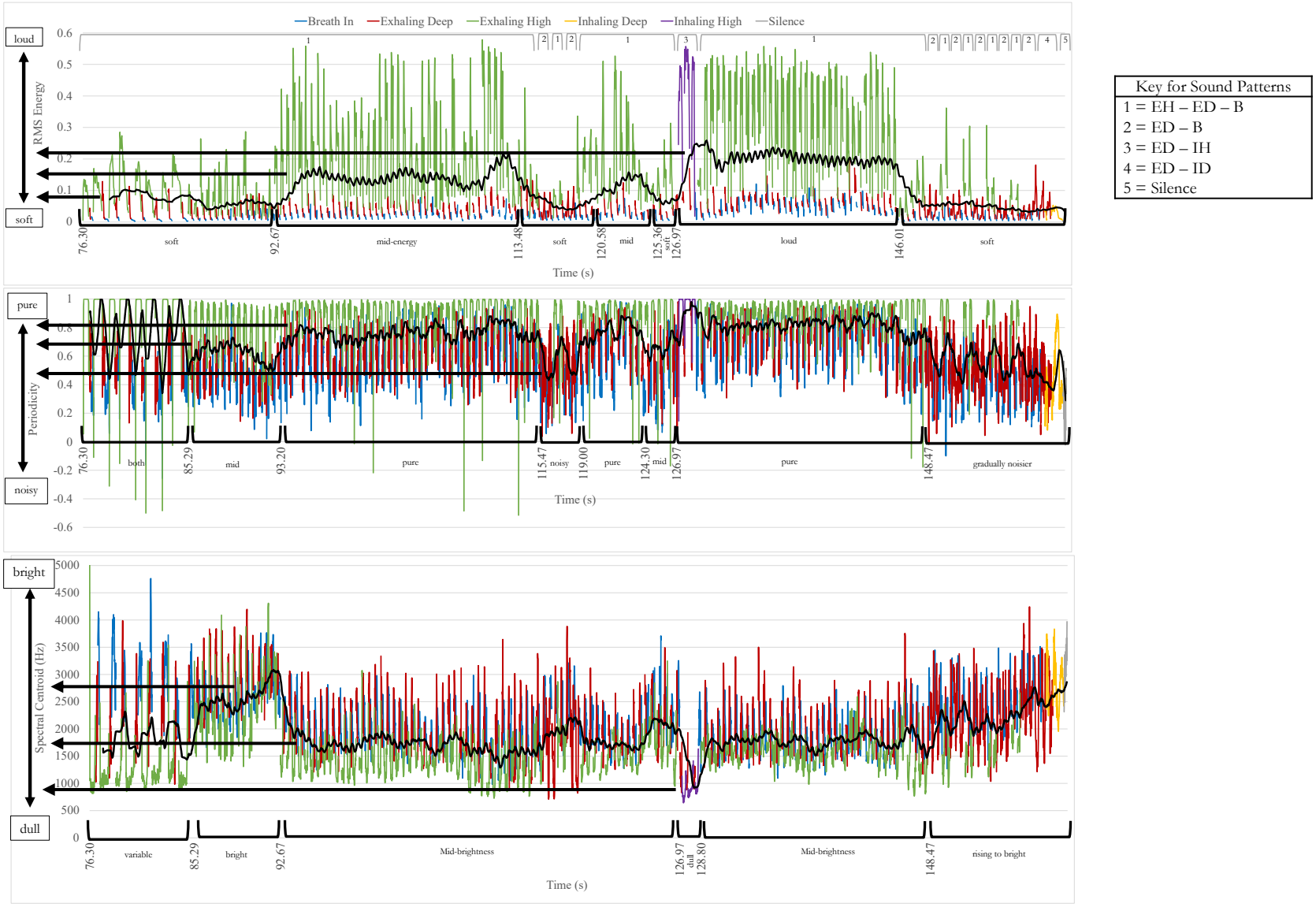


Figure 6.2 RMS energy (top), periodicity (middle), and spectral centroid (bottom) values and moving average trendline (black) for the middle section (B) of "Ilunikavi." Sound patterns are indicated at the top; a key is provided on the right.

Tracking the changes in loudness, noisiness, and brightness through acoustic measurements allows for multiple analyses of the form of the middle section of "Ilunikavi." In Figure 6.2, if one scans vertically where the brackets begin and end for pitch, loudness, noisiness, and brightness, there are only a couple of places where a substantial audible change happens in *all* four analyses. At approximately 127s and 148s, the sound patterns change, and each of the three sound qualities changes too. Otherwise, the analyses of pitch, loudness, noisiness, and brightness in relation to the form of the B section of "Ilunikavi" differ from each other rather substantially. I find it difficult to attend to all three sound qualities at once. When I listen to this song, I try to focus on one sound quality per listening experience: loudness, noisiness, or brightness. Attending to these different facets of sound allows me, as a listener, to hear a different form each time I listen.

While I have attempted to push the boundaries of what constitutes an analysis of form, the foregoing studies remain very traditional in a Western music theory sense by focusing on structural aspects of the music and by using visualizations (in this case, graphs rather than musical notation) to illustrate the music. The prose I use is also very scholarly and describes precisely what the visualizations represent. This is in an attempt to present the analysis as clearly as possible. In the following section, I turn to Tagaq's "Sivulivinivut" to practice a different way of presenting an analytical narrative of a musical piece, one that is ambiguous and requires the reader to add further interpretation to connect the narrative to the piece of music.

6.2 Toward a Different Way of Forming Analytical Narratives: A Story of "Sivulivinivut"

"Sivulivinivut," from Tagaq's 2016 *Retribution* album, begins quietly and calmly with whispers. The dynamic eventually increases slightly, but the aesthetic remains soft, gentle, and pure with Tagaq using mostly high, pure pitches throughout the track. "Sivulivinivut," meaning

ancestors, is one of the few songs on the album that features lyrics. The overall sound corresponds to the meaning of the song and its lyrics. Tagaq (2016b) describes "Sivulivinivut" as a "very heartfelt track" in which she "asks for help from those who came before or who's going to come after." In this song, the instruments blend so much with Tagaq's vocals that they create a unified texture and sound quality, making a discussion of the sound qualities in the song as a whole more appropriate than discussing individual instruments.

Music theory, as a Western academic discipline, is synonymous with "hungry listening." As Dylan Robinson explains, "hungry listening privileges a recognition of palatable narratives of difference, while in a more specific Western art music context, hungry listening takes part in content-locating practices that orient the ear toward identifying standardized features and types" (2020, 50). In analyzing a piece of music, settler music theorists, such as myself, tend to orient listening habits toward recognition. We want to recognize formal structures and categorize features of the music. There is a pleasure that we gain in feeling that we understand a piece of music or understand someone's analysis of that piece. We take pleasure in recognizing a feature of the piece of music as familiar and certain, as fitting in a framework already established in the discourse of music theory.

To decenter Western ways of analyzing music, one needs to forgo the hunger to know and understand with certainty as the receiver of an analysis, as well as the hunger to give others the ability to know and understand with certainty as the creator of an analysis. One way to do this is through performative writing. Performative writing in music theory has long been practiced and advocated for by feminist (McClary 1991; Guck 1994a, 2006; Luong 2017; McClary 1991), Indigenous (Avery 2012, 2014; Robinson 2020), and queer (Cusick 1994; Maus 1993) scholars. In addition to performative writing, Dylan Robinson also adopts apposite

methodologies—"processes for conveying experience alongside subjectivity and alterity" (Robinson 2020, 81)—in his book *Hungry Listening*, by providing poems throughout and "event scores" in between chapters. These event scores are themselves poetic and a form of art; they offer an analysis of a piece of music that is left open to interpretation. They express Robinson's understanding of a piece in a way that retains the affective qualities of the music. Readers of the event scores do not obtain a clear explanation of the musical work; instead, they need to further interpret the meaning of the performative writing in relation to the musical work. By adopting performative writing for music analysis, we decenter Western music-theoretical traditions. As feminist music theorists, such as Susan McClary (1991) and Marion Guck (1994a), suggest, we make our analyses deeply felt and open to interpretation when adopting this writing style. With this in mind, I have deliberately left open to interpretation the following analytical narrative for "Sivulivinivut" in an attempt to quell hungry listening further.

This style of analysis has connections to the narrative interpretations of Western art music that were common in program notes during the nineteenth and early twentieth centuries and the long tradition of hermeneutic/narrative analysis in music theory—beginning with the work of Momigny (1806) with his hermeneutic interpretation of a Mozart string quartet, Kretschmar (1902), and Tovey (1935), and continuing more recently with work by Cone (1977), Lochhead (2016), Klein (2004), Maus (1991), McClary (1991), Guck (1994b), and Okina (2020). There are some differences, in my opinion, between Robinson's work and these scholars that practice hermeneutics. While both approach analysis in an interpretive manner that goes beyond the purely structural or syntactic, the style of writing a narrative is largely (though not wholly) different. Often writers of narrative analyses in music theory make the music a character of the story or reference specific notes, harmonies, or instruments, which enact the action in the

narrative (Guck 1994b; Almén 2003). Using a fictional story-telling as the primary form of communicating of an analyst's interpretation of a musical work remains rare in music theory scholarship, especially without specific reference to moments in the music that connect to events in the story. My goal with this fictional story is to evoke the emotion brought on by changes in sound qualities through the narrative events. In this way, this narrative interpretation echoes work in affect theory (Luong 2019), not just the broader study of hermeneutics. My narrative is a method of communicating my analytical interpretation of this song, but it is also a creative piece of writing that could stand on its own, apart from the musical work, like the event scores in Robinson (2020).

What follows is a fictional story to accompany Tanya Tagaq's "Sivulivinivut," my initial attempt at a performative mode of writing for music analysis. As a suggestion for engaging with this analytical narrative, I hope that, at this moment, you will listen to "Sivulivinivut" in its entirety, with a focus on how the piece changes in loudness, noisiness, and brightness throughout the song, then come back to this document and read the following story as an analytical interpretation of the song. The recording with music video can be located at the following link: https://www.youtube.com/watch?v=qtAUoIbsTAQ. After reading the story, please return to the song or watch the music video with this story in mind.[102]

> *A young woman steps out of her house in northern Manitoba, Canada. She has been feeling lost, not knowing how to manage the situation presented before her. The world is feeling so large and overwhelming. It is dark. Surrounding her are tall but spindly pine trees lit by a*

[102] I first wrote this story with inserts describing what is happening in the music in terms of sound qualities in an effort to help the reader connect the story to the music and to make this interpretation clearer for the reader; however, this would promote hungry listening. Then I thought I might insert time markers to indicate moments of alignment between my interpretation and the recording to aid the reader in connecting the story to the music; however, this too would promote hungry listening. Instead, I consciously chose to remove any reference to the music in my interpretation. If, after reading my interpretation and listening to the song, you would still like to give in to hungry listening and view a graph detailing the RMS energy, periodicity, and spectral centroid data for "Sivulivinivut," it has been provided in the appendix at the end of this dissertation, though I implore you to sit with the uncertainty.

single light near the door of the home. Suddenly, the sky lights up and dances above with shades of green, yellow, and turquoise. She is greeted by ancestors.

The sight of the ancestors brings comfort to the young woman. She is no longer alone. But, her joy at their sight is subsumed by awkwardness. While she wants to visit with them, she also does not quite know how to approach the ancestors nor how to speak with them. What she really wants is to ask for help and guidance. The young woman first greets them by paying respects, telling them how much she loves them and misses them. After talking for a while, she finally builds up the courage to ask the ancestors for help in dealing with her troubles. The ancestors become more active, changing color more frequently and dancing about more fervently. It becomes clear that the young woman is not just wanting to visit but to seek something from the ancestors. She wants to be told what she should do; she wants a fix. Despite being irritated by the young woman's ulterior motives for welcoming them, the ancestors tell the young woman stories of their own…stories, not answers. Eventually, after the young woman has been comforted, the ancestors slowly fade away, as they must. The young woman reaches out for the ancestors, not wanting them to go. But, accepting that they must, the young woman thanks the ancestors for visiting and goes back inside.

This story is told from my perspective, but I also grounded my interpretation in the meaning of the title of the song and Tagaq's descriptions of the song. In an accompanying commentary for the album *Retribution*, Tagaq mentions that "Sivulivinivut" means ancestors. This song has lyrics in Inuktitut, but as someone who does not know Inuktitut, the meaning of the lyrics is inaccessible to me. Additionally, Tagaq (2016b) has purposefully not released a translation of the song, as she says she "didn't want it to be objectified in any way." To respect her wishes, I have not sought out a translation of the song's lyrics and instead have treated this song as though it has none. Instead of a translation, Tagaq (2016b) offers a description of the song, stating that it is about "asking for help from those who came before or who's going to come after...the idea of us being a total extension [of others]." She also says that "the translated lyrics are very beseeching and gentle" (Tagaq 2016b).

Additionally, I took inspiration for my narrative from the imagery in the music video for "Sivulivinivut," which features the Northern Lights. Some Inuit view the Northern Lights as ancestors, a concept I used in my interpretation. Though I used the meaning of the title and the

imagery of the Northern Lights as a jumping-off point for my fictional story, I also brought to the interpretation my own understanding of what having a connection with ancestors can be like—my experiences with going to cemeteries and talking to gravestones where my grandparents lie, and how I feel at the beginning of that, in the middle, and at the end.

This, I hope, will be understood as a subjective interpretation. It is a fictional story corresponding to my experience and thoughts while listening to "Sivulivinivut." But, really, *all* analyses are subjective, including the earlier studies of "Ilunikavi" that might have appeared objective due to the use of graphs with data compiled from a computer program.[103] Even the most seemingly objective analyses require interpretation of the music we are examining, which brings our biases and listening positionalities into play. In fact, the analyses of "Ilunikavi" and "Sivulivinivut" were done in much the same way. As I did when creating the analyses for "Ilunikavi," I viewed the graphs of RMS energy, periodicity, and spectral centroid data computed from the recording of "Sivulivinivut" while listening to the song to craft this narrative. The only differences between the analyses are that I used performative writing rather than scholarly writing for this one, and I did not reference my data or show the graphs with my interpretation.

This interpretation is also an attempt to push myself out of my comfort zone and to bring into my analysis another element of Indigenous research methodology: storytelling, though storytelling from an admittedly Western perspective.[104] There are no notated musical examples and no graphs to visualize the music, one aspect regularly found in Western music theory

[103] The acoustic measurements, descriptors for timbre, and the plug-ins used are biased. They were developed with a particular type of music in mind, namely Western art and popular music.
[104] Storytelling in Inuit culture is an oral tradition. Only certain people are given permission to be storytellers. Since storytellers are knowledge keepers, they are usually elders. Stories emphasize a close relationship with the land and animals, as well as teachings. Story, as Dylan Robinson notes, allows Indigenous peoples to "uphold Indigenous epistemic values (and refuse epistemic violence of other forms)" (2020, 101).

scholarship. My education in music theory has also taught me to make my writing "scholarly," not a story, but by writing in this way—as a story—this interpretation was more enjoyable for me as the writer and hopefully is also more enjoyable for you as the reader. The difficulties I had in writing the analysis of "Sivulivinivut" as a story stem from the expectations we have of academic writers, which center Western approaches that are often antithetical to Indigenous research methodologies. As Dylan Robinson (2019, 2020) has suggested, the modes of delivery of analysis are just one thing that needs to change in the move toward uncolonized music disciplines, including music analysis. If we continue to expand what we accept as musical analysis, the possibilities become far more interesting and vast.

6.3 Conclusion

This chapter builds on the analytical approach used in Chapter 4 to study the ways in which changes in sound qualities could be interpreted as having a role in the larger structuring of Tagaq's music. Tagaq's throat-singing style features contrasts of sound qualities that are illustrated by acoustic measurements, sometimes between songs on the same album or among the sections of a single song, as with "Ilunikavi." Changes in sound qualities can also affect our experience and understanding of the form of a song. In this chapter, I have suggested that a listener could interpret the form of a piece of music differently depending on the sound qualities to which they attend. The moments of perceived tension can change depending on the sound quality on which the listener focuses. In "Ilunikavi," is the greatest tension where the energy level is the greatest or where the most noise occurs? It all depends on what sound quality catches the attention of the listener and what they perceive as a tense sound. One could interpret the end of the B section as having the most tension due to its noisy aesthetic, or the louder part about two-thirds through the section as having the most tension if focusing on dynamic levels. It

depends on whether you gravitate toward the idea that loudness gives tension or feel more tension in a soft but noisier passage.

As the opening quote suggests, Tagaq herself seems to celebrate this multiplicity in her music with regard to her use, or lack thereof, of lyrics. She rarely uses lyrics in her songs because she wants the listener to develop their own meanings, and she revels in the idea that "whatever [the listeners] want to see is what they need to see" (CBC Music 2016). Wherever one feels the tension in Tagaq's music is where that person needs to feel the tension. This chapter has attempted to demonstrate that changes in sound qualities in Tagaq's music allow the listener to attend to different aspects of the music with every hearing. The method presented breaks from the Western reliance on pitch in analyzing form. It also allows for multiple readings and shows the complexities at work in her music. By focusing on sound qualities through acoustic measurements, rather than pitch, as a potential determinant of form, I also opened myself up to different modes of attention while listening to Tagaq's music and other possibilities for interpretation, including an alternative method of presenting a musical analysis as a story with characters and without reference to the musical work.

Chapter 7: Conclusions and Future Directions

This dissertation aimed to answer the question: how do sound qualities affect our experiences of beat, meter, and form in popular music vocals? The focus on interpreting songs from these three different angles, which progress from small-scale to large-scale structures, and reflecting on the choices that I make as an analyst have been threads that tied the chapters of this dissertation together, even while looking at wildly different repertoires and singing styles. I developed analytical approaches using acoustic measurements, specifically RMS energy, periodicity, spectral centroid, and formant frequencies, in conjunction with data visualization techniques and statistics, which allowed for this question to be considered while analyzing songs by Tanya Tagaq and recent mainstream country artists. The acoustic measurements of RMS energy, periodicity, and spectral centroid allowed me to visualize the sound qualities of loudness, noisiness, and brightness over time through a section of a song (Chapter 4) and over an entire piece (Chapter 6). Because these acoustic measurements provided continuous data for the audio file, I could track these changes visually in graphs and connect the differences in sound qualities to my experience of beat and form. In Chapter 4, I specifically focused on the relationship of sound qualities to beat structures, comparing my experience of the beat with Tanya Tagaq's through video recordings. As I conducted preliminary research for the chapters, I quickly realized that RMS energy, periodicity, and spectral centroid are effective when the voice is used percussively, as in Tagaq's throat singing, but formants are more effective when analyzing individual sung words. For that reason, in Chapter 5 I turned to formants to describe how vowel qualities affect our experience of meter in recent mainstream country songs. The acoustic measurement of formant frequencies confirmed what I had heard in country songs: similar vowel qualities are regularly placed on even-numbered beats in this style of music to create a feeling of

vocal backbeat. In Chapter 6, I returned to Tagaq's "Ilunikavi" to see how sound qualities can be considered a form-bearing element of the song. By analyzing the song through the lens of pitch and the acoustic measurements of RMS energy, periodicity, and spectral centroid, I could interpret the form of "Ilunikavi" in four different ways. I also experimented with presenting a form analysis as a story in Chapter 6 in an attempt to broaden the scope of how I communicate my analytical observations.

Not only can the approach developed and demonstrated here be used to relate changes in sound qualities to other aspects of musical structure, they are easily reproducible by others seeking to analyze sound qualities in pieces of music. In each chapter, I outline the specific steps that I use to create the graphs and the parameter settings used for each acoustic measurement. While the graphs are easily reproducible and can be adopted for other styles of music, each analysis may of course still differ slightly because of the subjective process of interpreting the graphs of data. This is where the analyst's individual positionalities plays a role in the analysis of sound qualities, even while using acoustic measurements.

To conclude, I critique some of my approach's technical and analytical limitations. I also suggest some future research avenues, which group into two broader categories, sound qualities research and repertoire. Finally, I revisit that which was addressed in Chapter 3, the chapter regarding the ethics of analysis and research and listening positionalities, to reflect on how my goal of approaching the study of Tagaq's music in a culturally appropriate and contextually informed way transpired throughout this dissertation. In this final section, I consider some of the challenges to incorporating ethics in the process of analysis and the ways forward from this dissertation.

7.1 Technical and Analytical Limitations of My Approach

While I have addressed some limitations in timbre theory by developing an easily reproducible approach to analyzing timbre and loudness that uses acoustic measurements, data visualization techniques, and statistical analysis, this approach is also not without limitations. Acoustic measurements are related to aspects of a sound that we can hear, but they are not exactly analogous to what we hear. Because of the reliance on acoustic measurements in this dissertation, some aspects of what we can hear might have been negated. In other words, the acoustic measurements might show more than what we can hear with our ears alone, or they might contradict what we hear.

Furthermore, acoustic measurements are not perfect. Errors can arise, leading to interpretations of a song that do not match aural experience. Sometimes the parameters are part of the reason measurements can have an error. For example, when retrieving mean formant values from Praat for Chapter 5, I found that if I changed the settings, I would get drastically different values for the same word. For this reason, once I decided on parameters that seemed to work for all songs, I maintained them while working on the chapter.[105]

Additionally, the researcher can sometimes affect the data that the programs provide in the interpretations they make before obtaining the measurements. For Chapter 5, I took a bit of a shortcut when identifying the formant frequencies of the country lyrics. Rather than breaking the words into their phonetic parts and obtaining formant frequencies for each phoneme, I used the mean formant frequencies of the entire word. While I do not see this as a problem for this study because the vowel is the longest part of a sung word, and because Praat visualized the formants for only sounded vowels and consonants, future research should probably take into consideration

[105] The settings used are summarized in footnotes in Chapters 4 and 5.

the phonetic parts of words to get a more thorough impression of formant structures in country (or other) lyrics.

Even though I have been careful about my choices while developing the analytical approaches used in this dissertation, it is important to note that there remain colonial undertones to this work. The descriptors I used for the sound qualities are ones established by (primarily white) scholars, and they are not necessarily ones Tagaq, or even the country artists or songwriters, would use to describe their music. Additionally, the primary tools I use to obtain the data from acoustic measurements, Sonic Visualiser and Praat, were developed and tested primarily using Western music and languages. As Tom Faber (2021) has explained, "unassuming as they may seem, these technologies are far from neutral. Like social media platforms, dating apps, and all data-driven algorithms, music production tools have the unconscious biases of their creators baked into their architecture." Music production tools assume certain things about music, including that it should be in a key and a regular metric structure with a steady tempo. Quantitative music research tools have been built with similar assumptions, which can affect how music that does not conform to these assumptions is analyzed. Aside from the tools used, I, as the analyst, must interpret the findings, and my interpretations are biased because of my positionality. Even when diversifying analytical approaches, it is still likely that a new system remains connected to (at least the effects of) colonialism, which should constantly be critiqued.

7.2 Future Research Avenues

7.2.1 Relating to Sound Qualities

This work has just scratched the surface in terms of analyzing sound qualities using acoustic measurements. Here I focused on RMS energy, but intensity could also be examined. It might be fruitful to compare data for the two measurements to see which is a better measure of

loudness. Other spectral descriptors, such as spectral flux, spectral flatness, spectral roll-off, and spectral spread, might also be useful for musical analysis. Spectral flatness can be used for analyzing noisiness, simar to periodicity. Again, it would be interesting to compare data for spectral flatness and periodicity of the same passage of music to see how these measurements differ. Mel-frequency cepstral coefficients (MFCCs) might also prove fruitful for the analysis of timbre in music, though I personally found the data from MFCCs to be extremely complicated to interpret. Beyond the use of measurements for analyzing sound qualities, there are measurements for analyzing pitch and time as well, and experimenting with these vamp plugins in Sonic Visualiser might lead to interesting analyses that would not be possible when analyzing a piece of music from staff notation. Estimated f0, for instance, is helpful for analyzing continuous pitch in melodies. These computer-assisted methods of analyzing music are particularly helpful when analyzing studio-produced music or performance practice, but have yet to be promulgated in music theory discourse.

In addition to experimenting with other acoustic measurements, there is room to analyze sound qualities in Tagaq's music and country songs further. The exploratory, computer-assisted method of analyzing sound qualities that I applied to Tanya Tagaq's throat-singing sounds helped show the differences between these sounds. I also showed how changes in sound qualities affect our experience of beat and form in her music. Nevertheless, I have not considered her entire collection of sounds, instead only examining her throat-singing passages in recorded performances. Thus, there is still room to consider how her screams and moans factor into these continua. Another research avenue that would likely be fruitful would be to analyze Tagaq's improvised live performances, which are significantly different in some cases from her studio recordings. In terms of country music (or any Anglophone popular music), it would be

interesting to analyze covers of songs to see how formant structures change with different singers and different pronunciations of words. This may lead to a different interpretation and understanding of meter in the vocals performed by the original singer compared to the cover artist.

The approaches established in this dissertation can be easily adapted for analyzing a variety of music genres. This approach could be used to analyze sound qualities in other repertoires as long as changes in timbre are elemental to the style. Any music that includes "extended" techniques will likely yield rich analyses. Some that come to mind are the music of Björk (who is often credited with giving Tagaq her big break in the music industry), metal singing, beatboxing, Tuvan throat singing, Sámi joik, as well as Western art music pieces (for voice or instrument), such as Luciano Berio's *Sequenza* series, John Cage's *Aria* and *Song Books,* Kaija Saariaho's *Sept Papillons*, Helmut Lachenmann's *Guero*, Liza Lim's *Invisibility*, and many others. The flexibility afforded by these analytical approaches, as demonstrated with the slight modifications made from chapter to chapter within this dissertation, will bring out subtle differences in the sounds used in these kinds of music and provide a way of describing them that could lead to other fruitful analyses.

Another avenue ripe for research is the examination of the sound qualities among layers in a mix. Here I have discussed vocals in popular music; other theorists have mainly analyzed sound qualities in popular songs within the mix as a whole (Peres 2016; Zagorski-Thomas 2018). A question that should be considered further is what role does each layer play in the experience of a song? How do instruments compare in terms of sound qualities for different parts of a song? How are instruments treated in terms of production effects, and how does that affect the sound qualities? In music production, sound qualities are manipulated through specific effects,

including slope, OER (odd/even ratio), pluck, inharmonicity, and distortion. This likely affects how we experience the song as listeners. In future research, I plan to examine how the vocals and other parts of a mix are manipulated in the studio with effects, the resulting timbral characteristics, and the result of that on our experience of music.

7.2.2 Relating to the Repertoire

In this dissertation, I focused on Tanya Tagaq's music and a few mainstream country artists from the last decade. Still, there are so many other artists that I could have written chapters about. While Tanya Tagaq is likely the best-known Inuit popular music artist, she is part of a burgeoning Inuit popular music scene. If interested in Inuit popular music, one should show support for artists such as Rise and Silla, PIQSIQ, Riit, The Jerry Cans, and Elisapie, as well as Tagaq.

Finally, I would be remiss if I did not mention the whiteness, homophobia, and racism prevalent in country music. It is part of the reason that two of the three songs I analyzed in Chapter 5 were by white, cis-gendered singers. While this aspect of country music was well known for many decades, it became extremely apparent with Lil Nas X's "Old Town Road" success and its subsequent removal from *Billboard*'s "Hot Country Songs" chart in March 2019. This event was one of the impetuses for me to write Chapter 5. Contrary to the reasoning given by *Billboard* at the time, "Old Town Road" features many elements of today's country music, as shown through the analyses in that chapter. If interested in country music, one should show support for the artists who rarely get played on country radio or make it onto the country music charts, including Black artists such as Rhiannon Giddens, the Carolina Chocolate Drops, Our Native Daughters, Mickey Guyton, Allison Russel, Amythyst Kiah, Yola, Brittney Spencer, Chapel Hart, Jimmie Allen, Breland, Shy Carter, The War and Treaty, Willie Jones; Indigenous

artists such as Shane Yellowbird, Don Amero, William Prince, Crystal Shawanda, Jayli Wolf, and Morgan Toney; and LGBTQ+ artists such as Orville Peck, D'orjay The Singing Shaman, Lily Rose, Brooke Eden, Brandy Clark, Chris Housman, Fancy Hagood, Ty Herndon, and T.J. Osborne. These lists are by no means exhaustive. There is diversity in country music, just not in the mainstream country industry or on country radio. Many of the artists listed above blend country music elements with elements from other genres in their music and label their music as country soul, country rap, Americana, or folk; these labels better describe the music that these artists make and open them up to a wider set of listeners.

7.3 Toward Ethical Approaches to Analysis: A Personal Reflection

In my efforts to expand the canon and my focus on Tanya Tagaq's music, examining the ethics associated with researching and analyzing music outside of my culture and background have become a primary concern for me. Since beginning to contemplate the ways in which I approach analysis, I have implemented several practices that I will carry with me as I continue to do research. I actively cite and base my interpretations on composers'/artists' discussions of their own music. To analyze Tagaq's music respectfully, I focus on aspects of her music that I interpret as important to her, based on what she has said in published interviews. I also make space for Tagaq's perspective by quoting her words in this dissertation rather than paraphrasing them. In my analyses, I break from the traditional reliance on pitch and harmony embedded in the style of Western art music and focus instead on sound qualities related to timbre and loudness. I have begun preferring the term "interpretation" to "analysis" since interpretations are open to multiple meanings, whereas analysis connotes a singular, correct understanding of a musical work (though you have probably noticed that "analysis" still creeps into my vocabulary and I have used it many times throughout this document). I have experimented with performative

writing as a way of presenting my interpretations in an effort to maintain the affective aspects of the music about which I write. Additionally, I sometimes presented multiple analytical narratives, rather than the "one correct analysis," to show that there are many ways of understanding a piece of music. I located myself in terms of my background, knowledge, and biases, all of which affect my analytical thinking and aural experience of music. that I bring to this analysis. Working toward an ethical approach to analysis has not been easy, and it will continue to be a journey for me. While my research has made significant steps in the right direction, there is much room for growth.

The greatest difficulty for me in writing about Tagaq's music came with Chapter 4, where I discuss beat structure in Tagaq's "Ilunikavi." First, I could not find any discussion by Tagaq about rhythm, beat, meter, or associated activities such as dance, in her published interviews. That necessitated looking for a way to discuss beat structure without feeling like I was speaking for Tagaq. I chose to study a video of Tagaq performing "Ilunikavi" and record myself conducting a beat while listening to the same performance. I then compared moments in which we each indicated an ictus with our movements, putting us on level ground, as both of us were subjects of the analysis. I also focused more on my movements than on Tagaq's, as I can speak more about the possible reasons I moved the way I did while listening to the song. I made some speculative arguments about potential reasons for the differences between our movements that I ground in Indigenous histories and epistemology regarding time by Indigenous scholars, such as Keyes (2009) and Robinson (2020). Nevertheless, this was the most challenging chapter to write as I was (and remain) uncomfortable with creating a dichotomy between myself and Tagaq and between Inuit ecological and Western mechanical conceptions of time, a dichotomy that is largely a colonial construct.

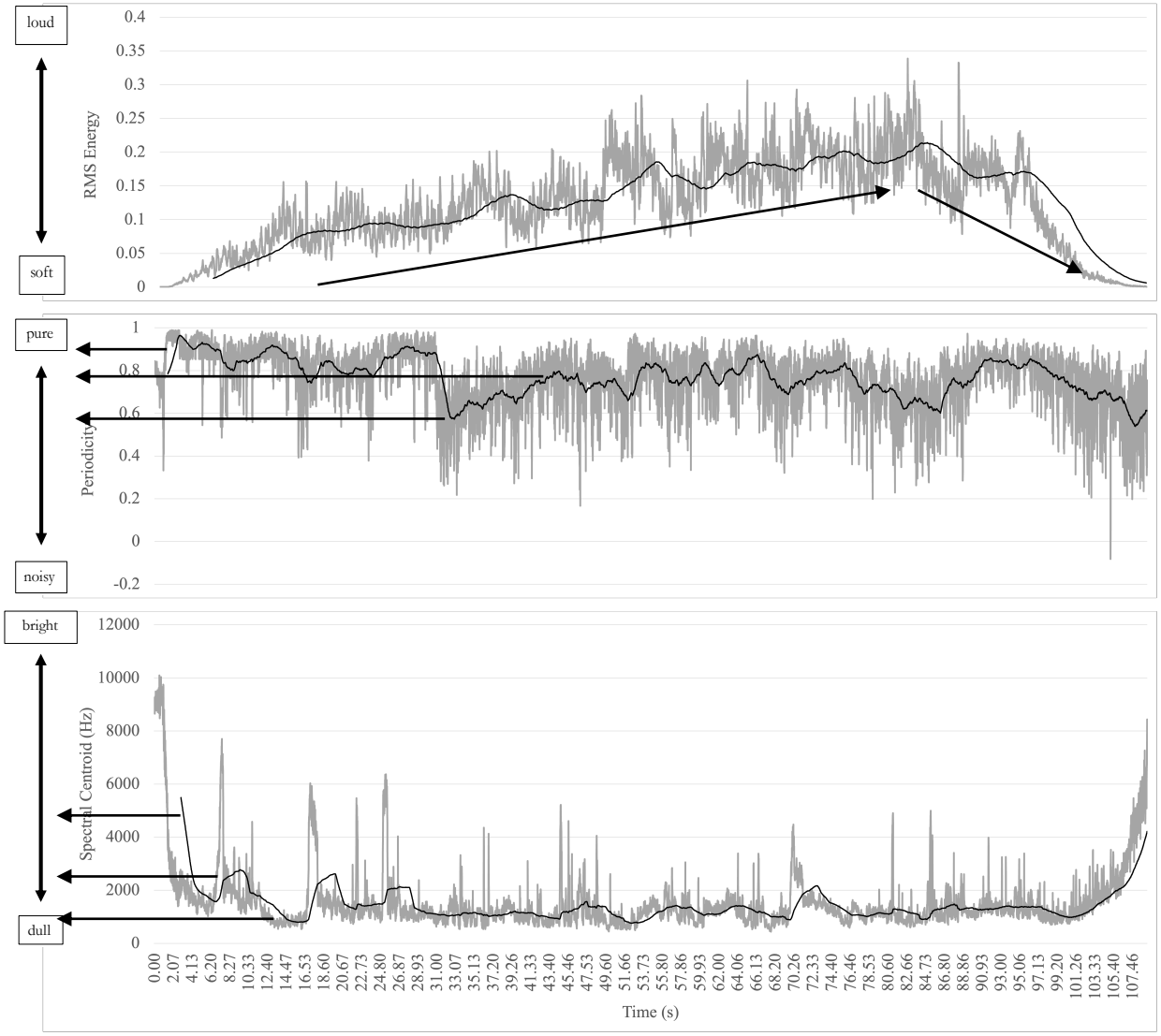
Much has been discussed regarding the ethics of analysis while working on music from a culture that is not your own. But what ethical position might be required when working on music for which you are an insider? In Chapter 5, I write about contemporary mainstream country music, a style drastically different from Tagaq's music and a type of music for which I feel I am an insider. How does analyzing country music, a historically (and even still currently) predominantly white genre, tie in to a discussion of ethical approaches to analysis? Regardless of the type of music analyzed and your position, there is always space to consider the ethical implications of the research. Because I have been listening to and playing country music since I was a teenager, I consider myself an insider of this style of music, which makes me biased in how I write about it. In my scholarship, I must also consider the whiteness associated with the genre and the racism prevalent in the country music industry. For this reason, I frame Chapter 5 around Lil Nas X, a Black, gay, trap country artist whose song was removed from *Billboard*'s country charts for supposedly not being country enough. Lil Nas X is an outsider to the country music industry, and they made it known when radio stations did not play his song. In Chapter 5, I argued that Lil Nas X's "Old Town Road" features lyrics and a melody that are highly similar to those featured in most recent mainstream country songs, contradicting the argument that it is not country enough.

This dissertation is written from a place of curiosity, which could very easily be interpreted as "hungry listening," to borrow Dylan Robinson's terminology. While I took steps to broaden my theoretical approach by attempting to center Tagaq's and other musicians' thoughts on their music, the analyses in this dissertation are largely my own interpretations of songs by Tagaq and country musicians. They are my, and only my, opinions and ways of understanding the music that I analyze. As Robinson pithily suggests, "Unlike iPhone photo filters, one cannot simply

select and add noncolonial, feminist, queer, or black listening filters in order to listen otherwise." I am a white settler/colonialist, and I hear as a white settler/colonialist; I cannot listen as Tanya Tagaq listens, even if I had met her and spoken with her about her music, nor will I likely listen as the country artists listen to their own songs. I listen as I listen; however, I can try new ways of listening, focusing on timbre or texture rather than pitch and harmony, or experimenting with different ways of conveying my understanding of a piece of music. In future research, I hope to continue expanding how I listen to and analyze music. I also plan to continue critiquing my work and my approach to analysis, as it is the only way to make them more inclusive.

Reflection and critique are essential in the continual process of decentering Western music theory and actively trying to center the voices that compose, create, perform, or own the music being analyzed. In future research, in addition to what I described above, I plan to continue this self-reflexive scholarship. I want to consider further ethical, culturally appropriate, and contextually informed ways of analyzing and teaching Indigenous popular music while working collaboratively with the musicians and developing new methods of interpreting music. It is my hope that I will be able to collaborate on analyses with musicians, but there are also benefits to maintaining my position as an armchair analyst that should not be ignored. As Kofi Agawu notes, fieldwork can be more harmful to cultures and musicians as it often involves intervention and sometimes deception (2003, 178). To uncolonize academia is a journey, and we are all in different stages of that journey. Pushing ourselves to change how we approach music analysis in order to interpret the music in more meaningful and culturally sensitive ways might eventually impact larger issues both within and outside of music theory discourses.

Appendix: Reference Graphs for Chapter 6 Analysis of "Sivulivinivut"



Bibliography

Abbate, Carolyn. 2004. “Music—Drastic or Gnostic?” *Critical Inquiry* 30: 505–36.

Abrahams, Rosa. 2019. “Mimicry as Movement Analysis.” *Analytical Approaches to World Music* 7/2.

Adams, Kyle. 2009. “On the Metrical Techniques of Flow in Rap Music.” *Music Theory Online* 15/5.

Agawu, Kofi. 2003. *Representing African Music: Postcolonial Notes, Queries, Positions*. New York and London: Routledge.

———. 2007. “To Cite or not Cite?: Confronting the Legacy of (European) Writing on African Music.” *Fontes Artis Musicae* 54/3: 254–62.

Alías, Francesc, Joan Claudi Socoró, and Xavier Sevillano. 2016. “A Review of Physical and Perceptual Feature Extraction Techniques for Speech, Music and Environmental Sounds.” *Applied Sciences* 6/143: 1–44.

Almén, Byron. 2003. “Narrative Archetypes: A Critique, Theory, and Method of Narrative Analysis.” *Journal of Music Theory* 47/1: 1–39.

American National Standards Institute. [1960] 1994. *Psychoacoustic Terminology: Timbre*. New York: American National Standards Institute.

Ammirante, Paolo, and Fran Copelli. 2019. “Vowel Formant Structure Predicts Metric Position in Hip-Hop Lyrics.” *Music Perception* 36/5: 480–87.

Attas, Robin. 2015. “Form as Process: The Buildup Introduction in Popular Music.” *Music Theory Spectrum* 37/2: 275–96.

———. 2019. “Strategies for Settler Decolonization: Decolonial Pedagogies in a Popular Music Analysis Course.” *Canadian Journal of Higher Education* 49/1: 125–39.

Aveling, Nado. 2013. “‘Don’t talk about what you don’t know’: on (not) conducting research with/in Indigenous contexts.” *Critical Studies in Education* 52/2: 203–14.

Avery, Dawn. 2012. “Tekeni-Two Worlds, Many Borders: A Look at Classical Native Music through Indigenous Eyes. *MUSICultures* 39(1): 129–68.

———. 2014. “Native Classical: Musical Modernities, Indigenous Research Methodologies, and A Kanienkéha (Mohawk) Concept of Non:wa (Now).” Ph.D. dissertation, University of Maryland.

Bader, Rolf. 2013. *Nonlinearities and Synchronization in Musical Acoustics and Music Psychology*. Berlin: Springer.

Bauer, Christine, and Johanna Devaney. 2022. "Constructing Gender in Audio: Exploring How the Curation of the Voice in Music and Speech Influences our Conception of Gender Identity." *Mediale Stimmentwurfe: Perspectives of Media Voice Designs*. https://www.grimme-institut.de/fileadmin/Grimme_Nutzer_Dateien/Medienbildung/Dokumente/SR_DG_NRW_07-Mediale_Stimmentwuerfe.pdf

Beaudry, Nicole. 1978. "Toward Transcription and Analysis of Inuit Throat-Games: Macro-Structure." *Ethnomusicology* 22/2: 261–73.

Biamonte, Nicole. 2014. "Formal Functions of Metric Dissonance in Rock Music." *Music Theory Online* 20/2.

Blake, David. 2012. "Timbre as Differentiation in Indie Music." *Music Theory Online* 18/2.

———. 2020. "Timbre." In *The Oxford Handbook of Critical Concepts in Music Theory*, edited by Alexander Rehding and Steven Rings, 136–59. New York: Oxford University Press.

Boerchers, Meredith. 2019. "'To Bring a Little Bit of the Land': Tanya Tagaq Performing at the Intersection of Decolonization and Ecocriticism." MA thesis, Carleton University.

Browner, Tara (ed.). 2000. "Making and Singing Pow-Wow Songs: Text, Form, and the Significance of Culture-Based Analysis." *Ethnomusicology* 44/2: 214–33.

———. 2002. *Heartbeat of the People: Music and Dance of the Northern Pow-wow*. Urbana: University of Illinois Press.

———. 2009. *Music of the First Nations: Tradition and Innovation in Native North America.* Urbana: University of Illinois Press.

Butler, Mark. 2006. *Unlocking the Groove: Rhythm, Meter, and Musical Design in Electronic Dance Music*. Bloomington: Indiana University Press.

Butterfield, Matthew W. 2006. "The Power of Anacrusis: Engendered Feeling in Groove-Based Musics." *Music Theory Online* 12/4.

———. 2010. "Variant Timekeeping Patterns and Their Effects in Jazz Drumming." *Music Theory Online* 16/4.

Cannam, Chris, Christian Landone, Mark Sandler, and Juan Pablo Bello. 2006. "The Sonic Visualiser: A Visualisation Platform for Semantic Descriptors from Musical Signals." In *Proceedings of International Society of Music Information Retrieval (ISMIR) conference*: 324–27.

CBC Music. 2016. “Tanya Tagaq | CBC Music’s Backstage Pass.” Online video clip. *YouTube,* 20 October 2016. https://www.youtube.com/watch?v=5hLQZ-Y1BkY.

Clarke, David. 2017. “North Indian Classical Music and Lerdahl and Jackendoff’s Generative Theory—A Mutual Regard.” *Music Theory Online* 23/3.

Clayton, Martin, Rebecca Sage, and Udo Will. 2005. “In Time with the Music: The Concept of Entrainment and its Significance for Ethnomusicology.” *European Meetings in Ethnomusicology* 11: 1–82.

Cleveland, Thomas, Johan Sundberg, and R. E. Stone. 2001. “Long-Term-Average Spectrum Characteristics of Country Singers During Speaking and Singing.” *Journal of Voice* 15/1: 54–60.

Cogan, Robert. 1984. *New Images of Musical Sound*. Cambridge, MA: Harvard University Press.

———, and Pozzi Escot. 1976. *Sonic Design: The Nature of Sound and Music*. Englewood Cliffs, NJ: Prentice Hall.

Cone, Edward T. 1977. “Three Ways of Reading a Detective Story or a Brahms Intermezzo.” *The Georgia Review* 31/3: 554–74.

Cooper, Grosvenor W., and Leonard B. Meyer. 1960. *The Rhythmic Structure of Music*. Chicago: University of Chicago Press.

Covach, John. 2005. “Form in Rock Music: A Primer.” In *Engaging Music: Essays in Musical Analysis*, edited by Deborah Stein, 65–76. New York: Oxford University Press.

———. 2020. “The Performer’s Experience: Positional Listening and Positional Analysis.” In *Investigating Music Performance: Theoretical Models and Intersections*, edited by G. Borio, G. Gioriani, A. Cecchi, and M. Lutzu, 56–68. New York: Routledge.

de Clerq, Trevor. 2017. “Embracing Ambiguity in the Analysis of Form in Pop/Rock Music, 1982–1991.” *Music Theory Online* 23/3.

Devaney, Johanna. 2020. “An Empirical Evaluation of Note Segmentation and Automatic Pitch-Extraction Methods for the Singing Voice.” In *The Routledge Companion to Interdisciplinary Studies in Singing*, vol. 1, edited by Frank A. Russo, Beatriz Ilari, and Annabel J. Cohen, 136–48. New York: Routledge.

———. In press. “Audio Tools for Music Corpus Studies.” In *The Oxford Handbook of Music and Corpus Studies*, edited by Daniel Shanahan, Ashley Burgoyne, and Ian Quinn. New York: Oxford University Press.

Diamond, Beverley. 2013. “Native American Ways of (Music) History.” In *The Cambridge History of World Music*, edited by Philip V. Bohlman, 155–80. New York: Cambridge University Press.

Dickie, Mary. 2014. “Tanya Tagaq Grabs the World by the Throat.” *Musicworks* 118. https://www.musicworks.ca/featured-article/tanya-tagaq-grabs-world-throat.

Dobrowohl, Felix A., Andrew J. Milne, and Roger T. Dean. 2019. “Timbre Preferences in the Context of Mixing Music.” *Applied Sciences* 9/1695: 1–17.

Dolan, Emily. 2013. *The Orchestral Revolution: Haydn and the Technologies of Timbre*. New York: Cambridge University Press.

———, and Alexander Rehding (eds.). 2018. *The Oxford Handbook of Timbre*. New York: Oxford University Press.

Dueck, Byron. 2012. “‘No Heartaches in Heaven’: A Response to Aboriginal Suicide.” In *Aboriginal Music in Contemporary Canada: Echoes and Exchanges*, edited by Anna Hoefnagels and Beverley Diamond, 300–22. Montreal and Kingston: McGill–Queen’s University Press.

———. 2013. “Rhythm and Role Recruitment in Manitoban Aboriginal Music.” In *Experience and Meaning in Music Performance*, edited by Martin Clayton, Byron Dueck, and Laura Leante, 135–58. New York: Oxford University Press.

Dukes, Billy. 2019. “Blake Shelton’s ‘God’s Country’ Lyrics are a Perfect Mix of Art and Practicality.” *Taste of Country*. https://tasteofcountry.com/blake-shelton-gods-country-lyrics-songwriters/.

Eidsheim, Nina Sun. 2019. *The Race of Sound: Listening, Timbre, and Vocality in African American Music.* Durham, NC: Duke University Press.

Everett, Walter. 1999. *The Beatles as Musicians:* Revolver *through* Anthology. New York: Oxford University Press.

———. 2009. *The Foundations of Rock: From “Blue Suede Shoes” to “Suite: Judy Blue Eyes*.” New York: Oxford University Press.

Ewell, Philip. 2020. “Music Theory and the White Racial Frame.” *Music Theory Online* 26/2.

Faber, Tom. 2021. “Decolonizing Electronic Music Starts with its Software.” *Pitchfork*. https://pitchfork.com/thepitch/decolonizing-electronic-music-starts-with-its-software/.

Fales, Cornelia. 2002. “The Paradox of Timbre.” *Ethnomusicology* 46/1: 56–95.

Feagin, Joe. [2009] 2013. *The White Racial Frame: Centuries of Racial Framing and Counter-framing*. 2nd ed. New York: Routledge.

Filice, Michelle. 2016. "Numbered Treaties." In *The Canadian Encyclopedia*. https://www.thecanadianencyclopedia.ca/en/article/numbered-treaties

Fink, Robert. 1990. "Rhythm and Text Setting in 'The Mikado'." *19th-Century Music* 14/1: 31–47.

———, Melinda Latour, and Zachary Wallmark (eds.). 2018. *The Relentless Pursuit of Tone: Timbre in Popular Music.* New York: Oxford University Press.

Fox, Aaron. 2004. *Real Country: Music and Language in Working-Class Culture*. Durham, NC: Duke University Press.

Freeman, Minnie Aodla. 2020. "Inuit." In *The Canadian Encyclopedia*. https://www.thecanadianencyclopedia.ca/en/article/inuit.

Galloway, Kate. 2020. "The Aurality of Pipeline Politics and Listening for Nacreous Clouds: Voicing Indigenous Ecological Knowledge in Tanya Tagaq's Animism and Retribution." *Popular Music* 39/1: 121–44.

Gombay, Nicole. 2009. "Today is today and tomorrow is tomorrow: Reflections on Inuit understanding of time and place." *INALCO Proceedings of the 15th Inuit Studies Conference, Orality*, Paris, France. Retrieved from http://www.inuitoralityconference.com/art/Gombay.pdf

Grant, Roger Matthew. 2014. *Beating Time & Measuring Music in the Early Modern Era*. New York: Oxford University Press.

Grey, John M. 1978. "Timbre Discrimination in Musical Patterns." *The Journal of the Acoustical Society of America* 64/2: 467–72.

———, and John W. Gordon. 1978. "Perceptual Effects of Spectral Modifications on Musical Timbres." *The Journal of the Acoustical Society of America* 63/5: 1493–1500.

Guck, Marion. 1994a. "A Woman's (Theoretical) Work" *Perspectives of New Music* 32/1: 28–43.

———. 1994b. "Analytical Fictions." *Music Theory Spectrum* 16/2: 217–230.

———. 2006. "Analysis as Interpretation: Interaction, Intentionality, Invention." *Music Theory Spectrum* 28/2: 191–209.

Hardman, Kristi. 2022. "The Continua of Sound Qualities for Tanya Tagaq's *Katajjaq* Sounds." In *Trends in World Music Analysis: New Directions in World Music Analysis*, edited by

Lawrence Beaumont Shuster, Somangshu Mukherji, and Noé Dinnerstein, 85–99. New York: Routledge.

Hasty, Christopher. 1997. *Meter as Rhythm*. New York: Oxford University Press.

Heidemann, Kate. 2016. "A System for Describing Vocal Timbre in Popular Song." *Music Theory Online* 22/1.

hooks, bell. 1992. *Black Looks: Race and Representation*. Boston: South End Press.

Hughes, Bryn, and Dominique T. Vuvan. Forthcoming. "Backbeat Placement Affects Tempo Judgments." *Psychomusicology*.

Ito, John Paul. 2020. *Focal Impulse Theory: Musical Expression, Meter, and the Body.* Bloomington, IN: Indiana University Press.

Jakobson, Roman, and Linda Waugh. 1979. *The Sound Shape of Language*. Bloomington, IN: Indiana University Press.

Keyes, Sarah. 2009. "'Like a Roaring Lion': The Overland Trail as a Sonic Conquest." *The Journal of American History* 96/1: 19–43.

Klein, Michael. 2004. "Chopin's Fourth Ballade as Musical Narrative." *Music Theory Spectrum* 26: 23–55.

Koestenbaum, Wayne. 1993. *The Queen's Throat: Opera, Homosexuality, and the Mystery of Desire*. New York: Poseidon Press.

Kovach, Margaret. [2010] 2021. *Indigenous Methodologies: Characteristics, Conversations, and Contexts*. 2nd ed. Toronto: University of Toronto Press.

Krahmer, Emiel, and Marc Swerts. 2007. "The Effects of Visual Beats on Prosodic Prominence: Acoustic Analyses, Auditory Perception and Visual Perception." *Journal of Memory and Language* 57/3: 396–414.

Krebs, Harald. 1999. *Fantasy Pieces: Metrical Dissonance in the Music of Robert Schumann*. New York: Oxford University Press.

Lavengood, Megan L. 2017. "A New Approach to the Analysis of Timbre." Ph.D. dissertation, City University of New York.

———. 2020. "The Cultural Significance of Timbre Analysis: A Case Study in 1980s Pop Music, Texture and Narrative." *Music Theory Online* 26/3.

Leight, Elias. 2019. “Lil Nas X’s ‘Old Town Road’ Was a Country Hit. Then Country Changed Its Mind.” *Rolling Stone*. https://www.rollingstone.com/music/music-features/lil-nas-x-old-town-road-810844/.

Lerdahl, Fred. 2006. “The Sounds of Poetry Viewed as Music.” *Annals of the New York Academy of Sciences* 930/1: 337–54.

———, and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Levin, Theodore, and Valentina Süzükei. 2018. “Timbre-Centered Listening in the Soundscape of Tuva.” In *The Oxford Handbook of Timbre*, edited by Emily I. Dolan and Alexander Rehding. New York: Oxford University Press.

Lewin, David. 2006. *Studies in Music with Text*. New York: Oxford University Press.

Library of Congress. 2021. “An Oral History with the Musicians of PIQSIQ, Tiffany Kuliktana Ayalik and Kayley Inuksuk Mackay.” Online video clip. *YouTube,* 1 March 2021. https://www.youtube.com/watch?v=o4DxxrFFbpI&t=1419s.

Light Echo Productions. 2014. “Echo Presents…Tanya Tagaq: Passion for Food.” Online video clip. *Vimeo,* 24 January 2014. https://vimeo.com/84979351.

Lochhead, Judith. 2016. *Reconceiving Structure in Contemporary Music: New Tools in Music Theory and Analysis*. New York and London: Routledge.

London, Justin. 2004. *Hearing in Time: Psychological Aspects of Musical Meter*. 2nd ed. New York: Oxford University Press.

Lorde, Audre. 1981. “The Uses of Anger: Women Responding to Racism.” *Women’s Studies Quarterly* 9/3: 7–10.

Luong, Vivian. 2017. “Rethinking Music Loving.” *Music Theory Online* 23/2.

———. 2019. “Analysis as Ethics: Experiments with Music Loving.” Ph.D. dissertation, University of Michigan.

MacLean, Taylor. “6 Common Myths about Treaties in Canada.” *University of Toronto Centre for Indigenous Studies*. Retrieved July 10, 2022. https://indigenousstudies.utoronto.ca/news/treaty-myths/

Malawey, Victoria. 2020. *A Blaze of Light in Every Word: Analyzing the Popular Singing Voice*. New York: Oxford University Press.

Malin, Yonatan. 2010. *Songs in Motion: Rhythm and Meter in the German Lied*. New York: Oxford University Press.

Manabe, Noriko. 2021. “Abe Road: Kuwata Keisuke’s Beatles Parody.” *SMT-V* 8.1: http://www.smt-v.org/archives/volume8.html#abe-road-kuwata-keisukes-beatles-parody.

Mather, Olivia Carter. 2017. “Race in Country Music Scholarship.” In *The Oxford Handbook of Country Music*, 327–45. New York: Oxford University Press.

Mauch, Matthias, and Simon Dixon. 2014. “pYIN: A Fundamental Frequency Estimator Using Probabilistic Threshold Distributions.” *IEEE International Conference on Acoustics, Speech, and Signal Processing*: 659–63.

Maus, Fred Everett. 1991. “Music as Narrative.” *Indiana Theory Review* 12: 1–34.

———. 1993. “Masculine Discourse in Music Theory.” *Perspectives of New Music* 31/2: 264–93.

McAdams, Stephen. 1989. “Psychological Constraints on Form-Bearing Dimensions in Music.” *Contemporary Music Review* 4/1: 181–98.

———. 1999. “Perspectives on the Contribution of Timbre to Musical Structure.” *Computer Music Journal* 23/3: 85–102.

———. 2013. “Timbre as a Structuring Force in Music.” *Proceedings of Meetings on Acoustics* 19/1: 1–6.

——— and Bruno L. Giordano. 2008. “The Perception of Musical Timbre.” In *The Oxford Handbook of Music Psychology*, edited by Susan Hallam, Ian Cross, and Michael Thaut, 72–80. New York: Oxford University Press.

McClary, Susan. 1991. *Feminine Endings: Music, Gender, and Sexuality*. Minneapolis: University of Minnesota Press.

Merriam, Alan P. 1964. *The Anthropology of Music*. Evanston: Northwestern University Press.

Meyer, Leonard B. 1989. *Style and Music: Theory, History, and Ideology*. Philadelphia: University of Pennsylvania Press.

Monelle, Raymond. 1989. “Music Notation and the Poetic Foot.” *Comparative Literature* 41/3: 252–69.

Morin, Peter. 2016. “This is What Happens When We Perform the Memory of the Land.” In *Arts of Engagement: Taking Aesthetic Action in and beyond the Truth and Reconciliation Commission of Canada*, edited by Dylan Robinson and Keavy Martin, 67–92. Waterloo, ON: Wilfred Laurier University Press.

Murphy, Nancy. 2022. "'The Times They are A-Changing': Metric Flexibility and Text Expression in 1960s and 1970s Singer-Songwriter Music." *Music Theory Spectrum* 44/1: 17–40.

Nattiez, Jean-Jacques. 1983. "Some Aspects of Inuit Vocal Games." *Ethnomusicology* 27/3: 457–75.

———. 1999. "Inuit Throat-Games and Siberian Throat Singing: A Comparative, Historical, and Semiological Approach." *Ethnomusicology* 43/3: 399–418.

Neal, Jocelyn. 2000. "Songwriter's Signature, Artist's Imprint: The Metric Structure of a Country Song." In *Country Music Annual 2000*, edited by Charles K. Wolfe and James E. Akenson, 112–40. Lexington, KY: University Press of Kentucky.

———. 2007. "Narrative Paradigms, Musical Signifiers, and Form as Function in Country Music." *Music Theory Spectrum* 29/1: 41–72.

———. 2018. "The Twang Factor in Country Music." In *The Relentless Pursuit of Tone: Timbre in Popular Musi*c, edited by Robert Fink, Melinda Latour, and Zachary Wallmark, 43–64. New York: Oxford University Press.

Okina, Risa. 2020. "Brahms and the Uncanny." Ph.D. dissertation, Temple University.

Osborn, Brad. 2013. "Subverting the Verse–Chorus Paradigm: Terminally Climactic Forms in Recent Rock Music." *Music Theory Spectrum* 35/1: 23–47.

Panagiotakis, Costas, and George Tziritas. 2005. "A Speech/Music Discriminator Based on RMS and Zero-Crossings." *IEEE Transactions on Multimedia* 7/1: 155–66.

Pecknold, Diane, ed. 2013. *Hidden in the Mix: the African American Presence in Country Music*. Durham, NC: Duke University Press.

Peeters, Geoffroy, Bruno L. Giordano, Patrick Susini, Nicolas Misdariis, and Stephen McAdams. 2011. "The Timbre Toolbox: Extracting Audio Descriptors from Musical Signals." *The Journal of the Acoustical Society of America* 130/5: 2902–16.

Peres, Asaf. 2016. "The Sonic Dimension as Dramatic Driver in 21st-Century Pop Music." Ph.D. dissertation, University of Michigan.

Phelan, Peggy. 1997. *Mourning Sex: Performing Public Memories*. New York: Routledge.

Red Bull Music Academy. 2016. "Tanya Tagaq Lecture (Montréal 2016) | Red Bull Music Academy." Online video clip. *YouTube,* 7 October 2016. https://www.youtube.com/watch? v=YkD00mGqBxE

Rifkin, Mark. 2017. *Beyond Settler Time: Temporal Sovereignty and Indigenous Self-Determination*. Durham, NC: Duke University Press.

Rigney, Lester-Irabinna. 1999. "Internationalization of an Indigenous Anticolonial Cultural Critique of Research Methodologies: A Guide to Indigenist Research Methodology and Its Principles." *Wicazo Sa Review* 14/2: 109–21.

Robinson, Dylan. 2019. "To All Who Should Be Concerned." *Intersections: Canadian Journal of Music/revue canadienne de musique* 39/1: 137–44.

———. 2020. *Hungry Listening: Resonant Theory for Indigenous Sound Studies*. Minneapolis: University of Minnesota Press.

Rockwell, Joti. 2011. "Time on the Crooked Road: Isochrony, Meter, and Disruption in Old-Time Country and Bluegrass Music." *Ethnomusicology* 55/1: 55–76.

Rosen, Jody. 2013. "Jody Rosen on the Rise of Bro-Country." *Vulture*. Retrieved from https://www.vulture.com/2013/08/rise-of-bro-country-florida-georgia-line.html

Rothstein, William. 1989. *Phrase Rhythm in Tonal Music*. New York: Schirmer Books.

———. 2011. "Metrical Theory and Verdi's Midcentury Operas." *Dutch Journal of Music Theory* 16/2: 93–111.

———. 2022. *The Musical Language of Italian Opera, 1813–1859*. New York: Oxford University Press.

Saariaho, Kaija. 1987. "Timbre and Harmony: Interpolations of Timbral Structures." *Contemporary Music Review* 2/1: 93–133.

Salamon, Justin, and Emilia Gómez. 2014. "MIR.EDU: An Open-Source Library for Teaching Sound and Music Description." *15th International Society for Music Information Retrieval Conference*.

Salley, Keith. 2011. "On the Interaction of Alliteration with Rhythm and Metre in Popular Music." *Popular Music* 30/3: 409–32.

Samarotto, Frank. 1999. "Strange Dimensions: Regularity and Irregularity in Deep Levels of Rhythmic Reduction." In *Schenker Studies 2*, edited by Carl Schachter and Hedi Siegel, 222–38. Cambridge: Cambridge University Press.

Scherzinger, Martin. 2019. "Temporalities." In *The Oxford Handbook of Critical Concepts in Music Theory,* edited by Alexander Rehding and Steven Rings, 234–70. New York: Oxford University Press.

Scotto, Ciro. 2017. “The Structural Role of Distortion in Hard Rock and Heavy Metal.” *Music Theory Spectrum* 38/2: 178–99.

Siedenburg, Kai, Charalampos Saitis, and Stephen McAdams. 2019. “The Present, Past, and Future of Timbre Research.” In *Timbre: Acoustic, Perception, and Cognition*, edited by Kai Siedenburg, Charalampos Saitis, Stephen McAdams, Arthur Popper, and Richard Fay Cham, 1–22. Cham, Switzerland: Springer.

Slawson, Wayne. 1981. “The Color of Sound: A Theoretical Study in Musical Timbre.” *Music Theory Spectrum* 3/1: 132–41.

Sluijter, Agaath, and Vincent van Heuven. 1996. “Spectral Balance as an Acoustic Correlate of Linguistic Stress.” *The Journal of the Acoustical Society of America* 100/4: 2471–85.

Smalley, Denis. 1997. “Spectromorphology: Explaining Sound-Shapes.” *Organised Sound* 2/2: 107–26.

Smith, Linda Tuhiwai. 1999. *Decolonizing Methodologies: Research and Indigenous Peoples*. New York: Zed Books.

Smith, Mark M. 2007. *Sensing the Past: Seeing, Hearing, Smelling, Tasting, and Touching in History.* Berkeley: University of California Press.

Spicer, Mark. 2004. “(Ac)cumulative Form in Pop-Rock Music.” *Twentieth-Century Music* 1/1: 29–64.

Sridhar, Rajeswari, and T.V. Geetha. 2008. “Music Information Retrieval of Carnatic Songs Based on Carnatic Music Singer Identification.” *International Conference on Computer and Electrical Engineering*: 407–11.

Sterbenz, Maeve. 2017. “Movement, Music, Feminism: An Analysis of Movement-Music Interactions and the Articulation of Masculinity in Tyler, the Creator’s ‘Yonkers’ Music Video.” *Music Theory Online* 23/2.

Stévance, Sophie. 2010. “Analysis of the Inuit Katajjaq in Popular Culture: The Canadian Throat-Singer Superstar Tanya Tagaq.” *Itamar. Revista de investigación musical: Territorios para el arte* 3: 82–89.

———. 2016. “Une Approche Contextualiste De Relations Voix/Gestes Dans Les Improvisations De Tanya Tagaq.” *Intersections* 36/2: 57–66.

———. 2017. “Tanya Tagaq’s Katajjaq Musical Signature.” In *The Routledge Research Companion to Popular Music and Gender*, edited by Stan Hawkins, 48–62. New York: Routledge.

Streefkerk, Barbertje. 1997. “Acoustical Correlates of Prominence: A Design for Research.” *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21: 131–42.

Strohm, Reinhard. 1986. “Zur Metrik in Haydn und Anfossis ‘La Vera Constanza.” In *Proceedings of the International Joseph Haydn Congress (1982)*, edited by Eva Badura-Skoda, 279–95. Munich: Henle.

Summach, Jay. 2011. “The Structure, Function, and Genesis of the Prechorus.” *Music Theory Online* 17/3.

Sundberg, Johan. 1987. *The Science of the Singing Voice*. DeKalb, IL: Northern Illinois University Press.

———. 1991. *The Science of Musical Sounds*. San Diego: Academic Press.

———. 1999. “The Perception of Singing.” In *The Psychology of Music*, edited by Diana Deutsch, 2nd ed., 171–214. San Diego: Academic Press.

Tagaq, Tanya. 2004. *Sinaa*. Lege Gordalua, CD (1), SS-137/05.

———. 2009a. “Illunikavi Performance.” Podcast audio. *Inuit Throat Singing – for iPod/iPhone*. The Open University. iTunes.

———. 2009b. “The Sounds of Throat Singing.” Podcast audio. *Inuit Throat Singing – for iPod/iPhone*. The Open University. iTunes.

———. 2016a. *Retribution*. Six Shooter Records, CD (1), SIX102.

———. 2016b. *Retribution Commentary*. Six Shooter Records, Spotify.

Tamburini, Fabio. 2003. “Prosodic Prominence Detection in Speech.” *Proceedings of the Seventh International Symposium on Signal Processing and Its Applications* 1: 385–88.

Tamlyn, Garry N. 1998. “*The Big Beat*: Origins and Development of Snare Backbeat and other Accompanimental Rhythms in Rock ‘n’ Roll.” Ph.D. dissertation, University of Liverpool.

Taylor-Neu, Robyn. 2018. “‘All There Is’: The Reconciliatory Poetics of a Singing Voice.” *American Anthropologist* 120/1: 113–25.

Temperley, David. 1999. “Syncopation in Rock: A Perceptual Perspective.” *Popular Music* 18/1: 19–40.

———. 2000. “Meter and Grouping in African Music: A View from Music Theory.” *Ethnomusicology* 44/1: 65–96.

———. 2007. “The Melodic-Harmonic ‘Divorce’ in Rock.” *Popular Music* 26/2: 323–42.

Tovey, Donald Francis. 1935. *Essays in Musical Analysis*, Vol. 1, *Symphonies*. London: Oxford University Press.

Tuck, Eve, and K. Wayne Yang. 2012. “Decolonization is Not a Metaphor.” *Decolonization: Indigeneity, Education & Society* 1/1: 1–40.

Wallmark, Zachary. 2014. “Appraising Timbre: Embodiment and Affect at the Threshold of Music and Noise.” Ph.D. dissertation, University of California, Los Angeles.

———, and Roger Kendall. 2018. “The Sound of Evil: Timbre, Body, and Sacred Violence in Death Metal.” In *The Relentless Pursuit of Tone: Timbre in Popular Music*, edited by Robert Fink, Melinda Latour, and Zachary Wallmark, 43–64. New York: Oxford University Press.

Wilson, Shawn. 2008. *Research is Ceremony: Indigenous Research Methods*. Black Point, NS: Fernwood.

Woloshyn, Alexa. 2017. “‘Welcome to the Tundra’: Tanya Tagaq’s Creative and Communicative Agency as Political Strategy.” *Journal of Popular Music Studies* 29/4: 1–14.

Zagorski-Thomas, Simon. 2018. “The Spectromorphology of Recorded Popular Music.” In *The Relentless Pursuit of Tone: Timbre in Popular Music*, edited by Robert Fink, Melinda Latour, and Zachary Wallmark, 345–66. New York: Oxford University Press.