Timing Is Everything . . . Or Is It?
Investigating Timing and Sound Interactions in the Performance of Groove-Based Microrhythm

Guilherme Schmidt Câmara

UNIVERSITAS OSLOENSIS
MDCCCXI

Thesis submitted for the degree of Philosophiae Doctor

Department of Musicology

Faculty of Humanities

RITMO Center for Interdisciplinary Studies in Rhythm, Time, and Motion

Faculty of Humanities, University of Oslo

Abstract

This thesis investigates the expressive means through which musicians well versed in groove-based music shape the timing of a rhythmic event, with a focus on the interaction between produced timing and sound features. In three performance experiments with guitarists, bassists, and drummers, I tested whether musicians systematically manipulate acoustic factors such as duration, intensity, and volume when they want to play with a specific microrhythmic style (pushed, on-the-beat, or laid-back).

The results show that all three groups of instrumentalists indeed played pushed, on-the-beat, or laid-back relative to the reference pulse and in line with the instructed microrhythmic styles, and that there were systematic and consequential sound differences. Guitarists played backbeats with a longer duration and darker sound in relation to pushed and laid-back strokes. Bassists played pushed beats with higher intensity than on-the-beat and laid-back strokes. For the drummers, we uncovered different timing–sound combinations, including the use of longer duration (snare drum) and higher intensity (snare drum and hi-hat), to distinguish both laid-back and pushed from on-the-beat strokes. The metronome as a reference pulse led to less marked timing profiles than the use of instruments as a reference, and it led in general to earlier onset positions as well, which can perhaps be related to the phenomenon of “negative mean asynchrony.” We also conducted an in-depth study of the individual drummers’ onset and intensity profiles using hierarchical cluster analyses and phylogenetic tree visualizations and uncovered a diverse range of strategies.

The results support the research hypothesis that both temporal and sound-related properties contribute to how we perceive the location of a rhythmic event in time. I discuss these results in light of theories and findings from other studies of the perception and performance of groove, as well as research into rhythm and microrhythmic phenomena such as perceptual centers and onset asynchrony/anisochrony.

Acknowledgments

This thesis was borne of blood, sweat, and tears, as well as a whole lotta love. (Musical puns intended, though mainly for the in crowd [Oops, I did it again].)

First of all, I would like to thank my supervisors, Anne Danielsen and Kristian Nymoen, for their invaluable guidance and support throughout this journey. Anne, thank you first and foremost for providing the opportunity to research groove music in a scholarly, scientific manner for all these years —the combination of praxis and theory have bolstered its awesome power for me many times over. Your shrewd mentorship, unceasing kindness and patience, and limitless passion for knowledge have been a constant source of learning and inspiration. It is an honor to work alongside such a juggernaut scholar (and fellow funk head!), and I hope to continue unraveling the mysteries of groove together with you. Kristian, thank you for painlessly leading me through a new and wonderful technological path, one that has opened up so many analytical possibilities to which I would otherwise not be privy save for your masterful guidance and boundless patience. Thank you for helping me navigate the murky waters of Matlab, and for the wonderful scripts you tailored to my endless requests. You and Anne have taught me so much, and I could not have hoped for better instruction in interdisciplinary and collaborative academic research.

I would also like to thank my fellow co-authors of this thesis's papers, Olivier Lartillot and George Sioros. Olivier, thank you not only for your excellent contributions but also for sharing your great expertise and your remarkable eagerness to help out with any problem, such as casually creating a brand-new onset-detection function tailored specifically to my needs over a weekend. George, thank you for your brilliant, visionary contributions, as well as for your solid, sober guidance and advice. I have thoroughly enjoyed, and learned a great deal from, our countless thought-provoking discussions on groove, rhythm, and life over the last few years (very nice!).

Thanks also to my fellow TIME project team members. Thank you, Mari Romarheim Haugen—fellow "Quants" group and Blue Room (+band!) member—for your guidance and advice on various Ph.D.-related matters, and for the many insightful discussions on rhythm and groove (one day we will figure out what the hell the "beat" really is). I look forward to collaborating with you on upcoming follow-up articles regarding the performance experiment data where your mocap and rhythmic-movement expertise will be invaluable. Thanks also to Sofia Dahl for her help with designing the mocap setup for the experiments, and with movement performance matters in general (I am also looking forward to our upcoming collaborations), and to Sabine Leske for her helpful advice and conversations on statistical methods in psychological research. Thanks also to the members of the TIME "Quals" group—Ragnhild Brøvig-Hanssen, Mats Johansson, Eirik Jacobsen, Bjørnar Sandvik, and Kjell Andreas Oddekalv—for their fruitful commentary and discussions during our shared workshop sessions. Special thanks to Kjetil Klette Bøhler for all the encouragement and advice on Ph.D. life, and the numerous intense discussions of groove and politics as well as the joys and frustrations of academic research.

Thank you to Carl Haakon Wadelaand for his help early on with designing the performance experiments, and with the recruitment of so many excellent musicians. Thanks to Dag-Erik Eilertsen for his advice on statistical methods, and for the extremely useful, custom-tailored SPSS syntax scripts, which saved me millions of clicks. Thanks to Martin Torvik Langerød and Victor Gonzales Sanchez for their help in designing and testing the audio and mocap recording setups, respectively.

Thanks also to Justin London for his encouragement and support, for his commentary and feedback on early manuscripts of papers I and II, and for being an outstanding opponent during my midway Ph.D. evaluation. Thanks to Werner Goebl for his feedback on approaches to the analysis of the performance data. Thanks to Rainer Polak for the many thoughtful and informative discussions during his visits to IMV and at various conferences around the globe. Thanks to Juan Diego Diaz for the support and advice on Ph.D. matters, and for the groovy conversations. A massive shout-out to Olivier Senn and the entire Swiss Groove Crew at Hochschule Luzern: Toni Bechtold, Florian Hoesel, Lorenz Kilchenmann, Rose Dawn, Rafael Jerjen, Alessandri Elena, and Antonio Baldassarre. Thanks so much for graciously hosting me during my exchange semester—your inspiring, groundbreaking groove research, amazing hospitality, and glorious mountain ranges and funky cheeses left me yearning to return. Thanks also to my terrific copywriter, Nils Nadeau, for adroitly streamlining my dense and idiosyncratic prose.

Thanks to all of my colleagues at RITMO and the Department of Musicology (IMV) for creating a friendly and stimulating work environment. Thanks to Peter Edwards, Målfrid Hoaas, Alexander Refsum Jensenius, and Anne Cathrine Wesnes for their excellent leadership and adept administrative support, as well as to Ingrid Stange Yttersad, Victoria Tømmeraas Berg, Ellen Filmberg, and Rolf Inge Godøy. Thanks to my fellow Ph.D., postdoc, and professor colleagues, who not only let me pedantically rave on about groove at every possible opportunity but also freely shared their time, skills, and knowledge: Agata Zelechowska, Tejaswinee Kelkar, Merve Acka, Emil Kraugerud, Dana Swarbrik, Çağrı Erdem, Julian Fuhrer, Ulf Holbrook, Tore Størvold, Maja Dyhre Foldal, Alejandro Omar Blenkmann, Áine Mangaoang, Simon Høffding, Alan Hui, Chris Stover, Kyle Devine, Jonna Vuoskoski, Erling Guldbrandsen, Hans T. Zeiner-Henriksen, Roger Arntzen, Asbjørn Lerheim, and Anders Tangen. Special shout-out to Connor Spiech for all the much-needed, homebrew-fuelled coronavirus sanity checks.

Thanks so much to all of the amazing instrumentalists for lending their groovy gifts to science in the experiments (who unfortunately cannot be named, due to research privacy laws, but you know who you are). A big thanks to my fellow band members from the Professors of Funk, and Chakras, as well: Hugo Pereira (aka Baba Soul), Stian Nordviste Skog, Øystein Bendos Aune (aka Police), Andreas Løvold, Audun Kjeldahl Berntsen (aka Vincent Velur), Michael Strütt, Lius Baruch Machado, Tiago Mendes, Ingvild Andersgaard, Audun Øksendal, and Carl Helsvig. Without all of you, I would never have begun to understand the many pleasures and meanings that groove music can hold nor have developed my skills as a performer and composer/arranger—all of which have inspired and sustained my interest in groove research. A huge thanks to the late great Alfredo Bessa (aka "*O rei do Arpoador*"), one of the finest Afro-Samba percussionists to have ever graced this Earth. You were my *mestre*, *amigo*, and *companheiro velho de guerra*, and without you I might never have had the fortitude to walk down the path of music. Thank you showing me the kind of dedication and discipline it takes to be a professional groove musician, and for passing on just a little bit of that magical *suínge* and *balanço* of yours, which has irreversibly shaped all of my playing and thinking about music. *Saravá*!

Last, but absolutely not least, thank you to my family: my best friend and wife, Hedda Lingaas Fossum, not only did you turn me on to jazz music years ago, opening up a whole new world of groove, but also you are just pretty much the grooviest person I've ever met (despite your dubious predilection toward an Elaine-esque finger-pointing style of dancing). Forgive the cliché, but I really couldn't have done this without you: your moral and emotional support, intellectual prowess, tenacious skepticism (mighty handy for keeping me on my toes), and tremendous love and

patience have sustained and vitalized me throughout this gargantuan trial. To my mother, Carmen Schmidt Câmara, thank you for instilling a deep sense of discipline and ethics in me, something you likely acquired from your own trials and tribulations as a medical doctor and professor, as well as an iron-fisted yet fair and loving matriarch. You have always been my unwavering number-one fan and supporter (a remarkable, yet surely misguided, character trait of truly wonderful mothers, I suppose) and an immutable source of parental affection, guidance, and motivation. To my dear father, Eduardo Gomes Câmara, thank you for your love and support, not to mention your "*paitrocínio*." Your encouragement of my intellectual endeavors has been paramount in my personal and professional journey so far and always amplified by the example and high standard that you set as a cosmopolitan, jet-setting naval architect. Thanks to my truly archetypical big sister, Leticia Câmara Roinesdal—fun and loving, yet also wise, strong, and pragmatic, and always aware of what to do and say (because, as the older sibling, you've naturally been there and done that). Like our parents, you too have set the bar high in terms of both career achievement—far outshining us all in terms of mileage points acquired in the line of duty—and dedication to family. Finally, thanks to my brother from another mother, (soon to be Dr.) Vincent Møystad, for all the love, wisdom, support, and encouragement, and for constantly introducing me to endless swaths of sophisticated, groovy music over the years (scoopty woop!).

Guilherme Schmidt Câmara (aka Dr. Gui)
December 2020

Abbreviations

Some of the abbreviations regularly used throughout thesis are as follows:

IOI	Inter-Onset-Interval
Mocap	Motion Capture
NMA	Negative Mean Asynchrony
P-center	Perceptual Center
RMANOVA	Repeated Measures Analysis of Variance
RMS	Root mean square
SC	Spectral Centroid
SD	Standard Deviation
SPL	Sound Pressure Level
TIME	Timing and Sound in Musical Microrhythm (project)

List of Publications

Papers included in this thesis

I Guilherme Schmidt Câmara, Kristian Nymoen, Olivier Lartillot, and Anne Danielsen, Effects of instructed timing on electric guitar and bass sound in groove performance, in *Journal of the Acoustical Society of America* 147(2), 1028–1041 (2020).

II Guilherme Schmidt Câmara, Kristian Nymoen, Olivier Lartillot, and Anne Danielsen, Timing Is Everything . . . or Is It? Effects of Instructed Timing Style, Reference, and Pattern on Drum Kit Sound in Groove-Based Performance, in *Music Perception* 38(1), 1–26 (2020).

III Guilherme Schmidt Câmara, Georgios Sioros, and Anne Danielsen, Mapping Timing and Intensity Strategies in Drum-Kit Performance of a Simple Back-Beat Pattern, submitted to *Journal of New Music Research*.

Other papers

- Câmara, G. S., & Danielsen, A. (2019). Groove. In A. Rehding & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory* (pp. 271–294). Oxford University Press.
- Sioros, G., Câmara, G. S., & Danielsen, A. (2019). Mapping timing strategies in drum performance. 20th International Society for Music Information Retrieval Conference, 776–783.
- Danielsen, A., Nymoen, K., Anderson, E., Câmara, G. S., Langerød, M. T., Thompson, M. R., & London, J. (2019). Where is the beat in that note? Effects of attack, duration, and frequency on the perceived timing of musical and quasi-musical sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 45(3), 402–418.
- Lartillot, O., Nymoen, K., Câmara, G. S., & Danielsen, A. (2021). Computational localization of attack regions through a direct observation of the audio waveform. *Journal of the Acoustical Society of America*, 149(1), 723–736.
- Senn, O., Bechtold, T., Rose, D., Câmara, G. S., Düvel, N., Jerjen, R., Kilchenmann, L., Hoesl, F., Baldassarre, A., & Alessandri, E. (2020). Experience of groove questionnaire: Instrument development and initial validation. *Music Perception*, 38(1), 46–65.

Contents

Abstract .... ***i***

Acknowledgments .... ***iii***

Abbreviations .... ***vii***

List of Publications .... ***ix***

Table of Contents .... ***xi***

1 Introduction **1**

1.1 Background 1

1.2 Motivation 2

1.3 Research Questions and Hypotheses 3

1.4 Approach 4

1.5 Thesis Outline 5

2 Groove, Rhythm, and Microrhythm **7**

2.1 The Multiple Understandings of Groove 7

2.1.1 Cultural and Historical Origins of “Groove” 7

2.1.2 Groove as a Performance Approach 9

2.1.3 Groove as a Psychological Construct (“the Pleasurable Urge to Move”) 10

2.2 Rhythm Production and Perception 13

2.2.1 Meter 14

2.2.2 Pulse 15

2.2.3 Subdivision 17

2.3 Microrhythm 18

2.3.1 Boundaries and “Norms” of Perceptual Categories 19

2.3.2 Swing and Non-Isochronous Subdivisions 21

2.3.3 Beat Delay/Anticipation (Laid-back versus Pushed) 22

2.3.4 Microrhythm as More Than (Onset) Timing 25

3 Timing and Sound Interaction **29**

3.1 P-Center Studies 29

3.1.1 Effects of Sound Features 30

3.1.2 P-Center “Shapes” as Indicative of “Beat-Bins” in Musical Contexts 33

3.2 Anisochrony and Asynchrony Studies 35

3.2.1 Effects of Sound Features 35

3.2.2 Just-Noticeable Differences in Relation to IOI Rates 39

4 Methods 43

4.1 Experiment Data Sets 43

4.2 Audio Recording 45

4.2.1 Guitar and Bass 45

4.2.2 Drums 46

4.3 Audio Processing and Analysis 51

4.3.1 Importing Data and Extracting Single Strokes 51

4.3.2 Calculating Stroke Features 53

4.4 Statistics 60

4.4.1 Finding Group Trends: Within-Subjects Tests of Difference between Averages 60

4.4.2 Exploring Diversity of Individual Strategies: Hierarchical Clustering and Phylogenetic Tree Visualizations 62

5 Research Summary 65

5.1 Papers 65

5.1.1 Paper I 65

5.1.2 Paper II 67

5.1.3 Paper III 70

5.2 Contributions 73

6 Discussion and Conclusions 77

6.1 General Discussion 77

6.1.1 Combined Timing–Sound Strategies as a Means of Conveying Microrhythmic Feel in Groove Performance 77

6.1.2 JNDs: Do Timing–Sound Differences between Instructed Microrhythmic Feels Matter Perceptually? 82

6.1.3 Influence of Timing Reference 86

6.1.4 The Role of Microrhythm in “Groove” 89

6.1.5 Limitations 91

6.2 Summary and Conclusions 93

6.3 Future Work 95

Bibliography 97

Papers 107

Paper I 109

Paper II 125

Paper III 153

Appendix 189

Chapter 1

Introduction

It's right on time,
It's right on time,
It's right on time, it's right on time, it's right on time, it's right on time,
Looking fine, you're looking fine, and get on 1999
—Red Hot Chili Peppers

Timing is everything
—Everybody

1.1 Background

When thinking about good "timing" or "time" in groove-based music, the most likely thing that springs to the mind of a performing musician or enthusiastic listener will likely be how the rhythms of the different instruments are positioned within the given metrical context of the groove. However, what most might take for granted rather than formalize explicitly is that the nature and shape of the sounds that comprise those rhythms—their timbre, duration, and intensity—also play an important role. Imagine, for example, a typical funk groove as played by James Brown and his band in the late 1960s or early 1970s. Most musicians and fans would agree that the grooves they played sound quite "tight," in the sense that the core rhythm section instruments of guitar, bass, and drums are played with great consistency and accuracy relative to a common metrical framework—that is, each is played in its own right time and place throughout the performance, if you will. Now, imagine that we take the sounds of the guitar—its typical fast, sharp attack, bright twangy tone, and short staccato phrasing—and replace them with a wobbly synth sound that has a slow rising attack profile (like a softly bowed note on a violin) and a dull and dark timbre, wherein each note is played in a very stretched-out, legato fashion. It is clear that not only would the overall groove sound highly unidiomatic but also that the perceived "timing" of the exquisitely groovy and intricate counter-rhythmic patterns that the guitar typically plays in a funk groove would no longer seem as "tight" but would instead become quite "loose" and ambiguous. That is, they would no longer sound in sync with the bass and drums in the right, or expected, way, and they might even sound muddy or messy. Such a thought experiment shows that, even if it is tacit knowledge, the "timing" of rhythmic events in groove-based music and the qualitative feels they might engender are quite dependent on their "sound."

In other words, timing and sound are inextricably connected, and in this thesis I seek to expand the common view of microrhythm beyond just the aspect of micro*timing* to the articulation of rhythm at the micro level, including all aspects of a rhythmic event—both temporal and sonic features as well as their interactions. In the act of listening, we generally do not separate between these different aspects (a change in a tone's loudness is not heard separately from, say, its timing location or timbre) but rather perceive them more holistically, and such an approach is also likely in performance: when one changes one aspect of a produced sound, one likely changes another aspect, either directly or indirectly. Precisely which aspects are interrelated and how they change in tandem likely vary depending on musical context (instrument, pattern, timing reference, and so on), which is precisely what I aim to investigate further here.

1.2 Motivation

For over a decade, I have been fascinated with the micro- and macro-rhythmic structures of groove-based musics. As a practicing instrumentalist, having played guitar and percussion in various soul/funk, rock, reggae, bossa nova, and samba ensembles over the years, I have constantly striven to understand more about what constitutes "good timing" in different groove-based performance traditions. The widespread notion that "timing" is paramount to good groove practice is one that I, like many musicians, have particular taken to heart. In any given style, my fellow band members and I often stop and ask ourselves these questions: In order to groove best here, should I play a little ahead, behind, or right on the "beat"? And in relation to whose beat, exactly (drummer, bassist, guitarist, etc.)? Also, should I do so consistently throughout the course of the basic groove pattern or only at key metrical locations, pushing and pulling to different degrees at different times? As such, in my previous academic research on groove, I have largely focused on empirically investigating onset-timing relationships between rhythmic elements in styles such as funk, jazz-funk (Câmara, 2016), and soul and r&b (Câmara and Danielsen, 2019). At the same time, while I fully recognize that "timing" is absolutely vital to groove, it is surely not everything, for achieving a good "tone" on one's instrument, which contributes to a cohesive overall "sound" in the band, is also a key contributing factor to the aesthetic success of good groove music. That is, it is simply not sufficient to play a groove with the right "timing" (that is, by placing events in the correct metrical and micro-temporal positions) if the types of sounds used, and the ways in which they are shaped, are not also right for any given instrument and style (for example, with the correct dynamic and durational accentuation or the appropriate bright or dark tone). However, while empirical groove research has proliferated over the past twenty years due to the greater accessibility and availability of digital audio-analysis tools, it is still not especially common to encounter investigations of groove-based music that focus on both the "timing" and the "sound" aspects of performances.

The end of my master's thesis (Câmara, 2016) coincided with the early stages of the establishment of the TIME project (Timing and Sound in Musical Microrhythm)[1] at the University of Oslo's RITMO Centre for Interdisciplinary Studies in Time, Rhythm, and Motion. The project's objective was precisely to explore the relationship between timing and sound features in various rhythmic musical contexts. One prime motivator driving this research impetus was the fact that various psychological studies of rhythm show that the moment in which we subjectively perceive an objective, physical rhythmic sound event as occurring depends not only on its onset location but

---

[1] See https://www.uio.no/ritmo/english/projects/time/.

also on various other acoustic properties, such as duration, intensity, and timbre. In my personal ventures as an engineer and producer in the recording studio, I had frequently noticed during the post-recording editing process that shifting the onset location of a performed rhythmic event to the zero millisecond mark of a given beat on the isochronous grid did not always make it sound squarely "on-the-beat" relative to a click-like metronome that externalized that grid, and, conversely, that events that were in fact performed consistently on the zero mark could sound rather behind or ahead of the beat, depending on the type of instrument in question and its duration, attack/decay profile, and frequency/timbre. As such, the TIME project provided an excellent opportunity to expand upon the work pioneered by Anne Danielsen and colleagues (Danielsen, Waadeland et al., 2015) and contribute to the further expansion of the concept of "timing" in groove-based microrhythm by incorporating under-investigated acoustic features. If timing is indeed everything in music, then "sound" is a tacitly understood part of that everything—and the goal of this thesis, and the TIME project, is to make explicit what seems to be taken for granted in the performance and perception of groove.

1.3 Research Questions and Hypotheses

The primary research objective of this thesis is to further understand the role of timing and sound features in the performance of groove-based microrhythm. From this, the following research questions emerge:

Main Research Question:

> How do temporal and sound-related features interact in the performance of simple groove-based patterns by drums, electric bass, and electric guitar, respectively?

Sub-questions:

> A: What are the characteristic (timing and sound-feature) interactions in these different instrumental settings?
>
> B: To what extent do sound–timing interactions used in performance correspond to perceptual sound–timing interactions?

Main hypothesis:

> Musicians **systematically manipulate** sound features such as **intensity, duration, and frequency** of rhythmic events **in addition to onset** timing location in order to communicate different desired microrhythmic styles/feels ("laid-back" [behind-the-beat], "on-the-beat,"[2] and "pushed" [ahead-of-the-beat]).

---

[2] Although in this thesis's papers I used "on-beat" in a shorthand manner to refer to this microrhythmic feel, in this summary document I will use the extended term "on-the-beat" to more clearly distinguish the microrhythmic feel from singular beats/events that are "on beat" relative to a metrical scheme. That is, the first, third, fifth, and seventh eighth

Sub-hypotheses:

A:

- While there is no hypothesis regarding particular instruments as such, we expected timing–sound strategies to be governed by the constraints of the instrument (that is, the range of sound features that are manipulable).
- In line with findings by Danielsen, Waadeland, and colleagues (2015), we expected drummers to implemented greater intensity on laid-back strokes. An additional/alternative hypothesis was that intensity would be used to make both laid-back and pushed strokes stand out as offbeat. We further hypothesized that bassists and guitarists would likewise incorporate strategies of systematic sound manipulation in order to communicate various desired microrhythmic feels.
- For drummers, who can produce several simultaneous rhythmic layers with different parts of the drum kit, we expected to uncover a rich diversity of inter-instrument onset timing and sound manipulation strategies.

B:

- Perceptual studies suggest that, in rhythmic contexts, the perceived timing of a sound event is affected by its shape/acoustic envelope. Based on findings from those studies, we hypothesized that a laid-back feel would be characterized by slower/longer (maybe softer/darker) strokes and a pushed feel by faster/shorter (maybe louder/brighter) strokes.

1.4 Approach

At its core, the research conducted in this thesis is highly interdisciplinary. Its main object of investigation—the performance of microrhythm in groove-based music—and related core musical concepts of groove and microrhythm are primarily informed by research from various fields of musicology, including mainly ethnomusicology (Keil and Feld, 1994; Monson, 2006), popular music studies (Abel, 2014; Danielsen, 2006; Iyer, 2002), and music theory and structural performance analysis (Butterfield, 2006; Pressing, 2002). Therefore, despite the many meanings attributed to "groove" in recent times (see discussion in section 2.1), it is here primarily understood through a musicological lens as an "approach to performance"—a diverse set of structural elements common to a wide gamut of musical styles historically and originally designated as groove-based within musical communities themselves, with a focus on macro- and microrhythmic patterns, as well as the manner in which they are performed, particularly via various timing "feels" (Câmara and Danielsen, 2019). Key concepts and theories from rhythm production and perception studies are also pertinent to this thesis, since, at the very least, the experience of musical rhythm presumes an interaction between sounding rhythms in the physical environment and subjective reference

---

notes in a 4/4 meter are commonly referred to as "on-beat" notes, because they coincide with the main quarter-note beats of the meter, whereas the second, fourth, sixth, and eighth notes are referred to as "offbeat" notes for the opposite reason.

structures in the perceptual/cognitive apparatus of performers/listeners that condition and guide the experience of the musical rhythm's meaning in any given context (Clarke, 1987; Danielsen, 2006; London, 2012). Furthermore, both psychophysical and analytical performance studies show that the nuances of both sound and timing features of an acoustic rhythmic event, and their interaction, play a role in how they are perceived, whether in terms of rhythmic P-center or onset (Goebl and Parncutt, 2002; Gordon, 1987; Hove et al., 2007; Tekman, 2002; Villing, 2010; Wright, 2008). Heuristic just-noticeable difference (JND) thresholds derived from either exploratory or experimental studies conducted in a wide range of rhythmic contexts can also help us to gauge the extent to which physical feature differences present in this thesis's collected performances may have a perceptual effect upon both performers and listeners, thus further engendering a range of qualitative microrhythmic timing "feels" (Butterfield, 2006; Frane and Shams, 2017; Friberg & Sundberg, 1994; 1995; Friberg and Sundström, 2002).

At the same time, the methods used here to investigate how microrhythm is expressed in groove-based performances derive from those typically used in experimental psychology and psychoacoustics, as well as statistics. Computational approaches from music information retrieval (MIR) are also vital to the analysis of the performances, where large data sets of various acoustic features are calculated mathematically from recorded audio signals. Overall, then, the work in this thesis perhaps most closely aligns with that of empirical musicology (E. F. Clarke & Cook, 2004), which places greater focus on observation and experimentation. This is because the thesis is largely based on the behavioral observation of musicians' performance strategies via a controlled experimental paradigm, the audio data of which are recorded and analyzed via MIR and statistical tools, and the results of which are interpreted via concepts of rhythm performance and perception.

1.5 Thesis Outline

This thesis consists of two parts, whereby the first part is an introduction to the theoretical framework and methods used in the investigation of the experiments conducted, and the second part is a collection of papers published or submitted to peer-reviewed journals. In the summary section of the thesis, I introduce these papers and elaborate upon their key findings.

Chapters 2 and 3 provide an overview of the theoretical motivations and key concepts related to the main themes explored in this thesis: groove, rhythm, and microrhythm (chapter 2), and the interaction of timing and sound features in the production and perception of rhythm (chapter 3).

Chapter 4 describes the experiments investigated in the papers and the methods utilized to analyze the audio data and conduct the statistical tests.

Chapter 5 provides a summary of the main results in each of the thesis's papers, discussing them in relation to the research questions raised in this introduction and highlighting their contributions to the fields of groove and rhythm research.

Chapter 6 offers a more in-depth discussion of the papers' results in relation to several key themes related to groove, microrhythm, and timing–sound interactions, as well as a concluding summary and avenues for future work.

All of the papers published in this thesis are either currently openly available or will be available in the near future on the University of Oslo's website. The data sets will also be released and made freely publicly available upon completion of the TIME project.

Chapter 2

Groove, Rhythm, and Microrhythm

In this chapter I give an overview of the theoretical framework related to the concepts of "groove," "rhythm," and "microrhythm." In section 2.1, I discuss certain understandings associated with the term "groove." In section 2.2, I present an important theoretical assumption for research into rhythm, namely how production and perception of musical rhythm derives from the interaction between physical sounding events and subjective reference structures of temporal organization, focusing in particular on rhythmic elements assumed to be important to groove-based music. In section 2.3, I engage with the concept of microrhythm and its specific production and perception in groove-based musical styles.

2.1 The Multiple Understandings of Groove

In the following sections,[3] I will explore some of the ways in which "groove" has been conceptualized in both popular and academic spheres. Ultimately, for the purposes of this thesis's papers, I subscribe to an understanding of groove that is at once broad *and* narrow. It is broad in that groove here encompasses a wide variety of music-theoretical rhythmic structural features as well as the manner in which they are applied in performance (this is "a performance approach"; see section 2.1.2)—that is, groove is not limited to its clear-cut operational definition in recent music psychology theory ("the pleasurable urge to move"; see section 2.1.3). It is narrow in that I associate groove specifically with musical performance and composition traditions that have some direct connection to musical styles originally described as "groove-based" (such as African American styles of jazz, soul, and funk), rather than with all forms of rhythmic music that may also feature some of the same rhythmic aesthetic elements (such as Western classical or European folk styles; see section 2.1.1). While other conceptual approaches to groove exist, including important phenomenological ones (Danielsen, 2006; Roholt, 2014; see also Câmara & Danielsen, 2018), I will not discuss them in depth here, as they are not directly relevant to this thesis's papers.

2.1.1 Cultural and Historical Origins of "Groove"

Most modern-day musicians associate "groove music" with styles derived from African American performance traditions, such as jazz, R&B, soul, funk, disco, and hip-hop, and especially with their rhythmic qualities (Abel, 2014; Câmara & Danielsen, 2019). However, references to the term "groove" were already characterizing jazz styles in the first half of the twentieth century. According to R. J. Gold (quoted in Kernfeld, 2003), "groove" was used by swing-era and bop musician

---

[3] Some of the text in subsections 2.1 and 2.2 has been adapted from my parts of a book chapter reviewing the various concepts of "groove" that I wrote in collaboration with Anne Danielsen while preparing this Ph.D. thesis ("*Groove*" in Câmara and Danielsen, 2018).

communities to convey an aesthetic judgment about performances that encompassed the degree of "excellence" and/or "sophistication" they displayed. The direct implication of such a usage was that those artists or bands who were thought to "groove" or to produce performances that were "groovy"/"in the groove" were superior to those who did not. This desirable quality is also reflected in song titles of the era, such as "In the groove" (1937, Decca 1621) by Andy Kirk's big band or "Groovin' high" (1945, Guild 1001) by Dizzie Gillespie. By the 1960s and 1970s, the term was more closely associated with soul and funk artists, due in part to their frequent use of the term in song titles and lyrics, including James Brown's "Ain't that a groove" (1966, King 6025) and Parliament/Funkadelic's "One nation under a groove" (1978, WB 56-539). Kernfeld (2003) points out that, since the 1960s or so, the term "groove" has most commonly operated in the popular sphere in reference to musical styles that "utilize characteristic accompanimental ostinatos drawn from African-derived dance music."

In musicology and music theory in recent decades, the denominator "groove-based music" has also come to be associated primarily with elements of the rhythmic aesthetic practices found in African American musical styles, such as the presence of a strong and regular beat and the generous use of syncopation and cross/counter-rhythmic patterns (Burnim & Maultsby, 2015; Bowman, 1995; Câmara & Danielsen, 2018; Danielsen, 2006; Iyer, 2002; Wilson, 1974; see also section 2.1.2). While groove, understood as a rhythmic aesthetic practice, may very well be featured in many other Afro Diasporic–derived styles as well, such as Afro-Cuban (e.g., salsa), Afro-Brazilian (e.g., samba), and other forms of Latin American or Caribbean dance music, these traditions tend to use other terms to describe those aspects associated with groove in the African American tradition (examples include "balanço"/"suingue" in samba or "sabor"/"bomba" in salsa; see Bøhler, 2013; Gerischer, 2006). The term groove is also not typically used within West African drumming traditions, which likewise display many of the formal aspects of groove-based music (see, for example, Locke, 1982; London, Polak, & Jacoby, 2017; Nketia, 1974; Polak, London, & Jacoby, 2016). Groove thus seems to be used principally but not exclusively to describe the foundation and aesthetic qualities of African American rhythmic music.

While the term groove has nevertheless been operationalized of late in ways that allow it to be applied to any rhythmic musical style that elicits the "pleasurable sensation of wanting to move" (see section 2.1.3)—even styles seemingly very far apart in their cultural origin and aesthetic approach to rhythm (such as those from the Western classical or Scandinavian folk traditions)—it may be prudent to reserve its use for musical styles that have some overt connection to those styles originally designated as such. Abel (2014, p. 23), for example, rejects conceptions of groove that allow for the inclusion of styles from Western classical and European folk traditions in particular, arguing that doing so "dilutes the concept of groove" to such an extent that it is of limited use in explaining why the term only arises in connection with the popular musics of the twentieth century. Instead, he proposes that groove should be understood as an aesthetic approach to rhythm, though one that he claims to be unique not only to African American styles originally described as "groove-based" but also to much of the twentieth-century Western popular music that derives from it in great part and likewise displays such structural rhythmic qualities. Pressing (2002) also extends the use of the term to musical styles of the cultural and geographical "Black Atlantic" regions ("African and African Diaspora") that display a range of similar structural features. Despite such debates about the scope of the styles to be included in culturally oriented definitions of groove-based music, most researchers and stakeholders seem to agree on a core set of common rhythmic features displayed by all relevant styles, such as aforementioned presence of a strong and regular beat and the generous use of syncopation and/or cross-rhythmic devices (see section 2.1.2)

2.1.2 Groove as a Performance Approach

Music-theoretical definitions of "groove" have been derived from attempts to delineate a set of rhythmic patterns and devices that is common to various styles designated as "groove-based," as well as the ways in which musicians apply such features in performance contexts (Abel, 2014; Câmara & Danielsen, 2019; Danielsen, 2006; Iyer, 2002; Pressing, 2002). What is or is not considered "groove-based" thus depends on the degree to which a given musical work demonstrates certain structural and performative features. This approach often stems from one of the simpler applications of the term "groove"—to denote a particular rhythmic pattern typical of a given musical style (such as "a swing groove," "a funk groove," and so on). Accordingly, groove has been defined as "a persistently repeated pattern" (Kernfeld, 2003) or as the "rhythm matrix" of a particular style (Monson, 1996, pp. 67–68). It has been widely stressed that the types of patterns typically found in groove-based styles readily serve to establish the sense of a regular pulse. For example, Iyer (2002, p. 397) describes groove-based music as featuring an "interlocking composite of rhythmic entities" that gives rise to "a steady, virtually isochronous pulse," while Pressing (2002, p. 288) describes groove as "a cognitive temporal phenomenon emerging from one or more carefully aligned concurrent rhythmic patterns" that engenders the "perception of recurring pulses, and [a] subdivision structure to such pulses." The mere ability to establish a steady pulse, however, is not generally considered sufficient to distinguish groove-based music from a myriad of other rhythmic musical styles that may do the same. In addition, most scholars agree that groove-based patterns display the generous use of rhythmic devices such as "syncopation" and "cross-rhythm" that serve to simultaneously establish and challenge the pulse, thus introducing various degrees of rhythmic complexity in aesthetically successful ways that engage the listener's attention both mentally and physically (Abel, 2014; Câmara & Danielsen, 2018; Iyer, 2002; Pressing, 2002).

In music-theoretical terms, these patterns become equivalent to the basic rhythmic "structure" that characterizes a given style, and this structure then becomes what one would commonly transcribe as the groove. When used as such to denote typical stylistic rhythmic patterns, however, groove invariably also encompasses the particular prescribed manner in which those patterns are implemented in relation to both the timing and the sound and shape of the rhythmic events, which is the particular topic of this thesis. Groove has further been associated with patterns of "microtiming"—variations from normative metrical reference structures on an order ranging from tens to hundreds of milliseconds and encompassing sounds that are often felt more than heard (Butterfield, 2006; Danielsen, 2010; Iyer, 2002; Senn et al., 2017). Relatedly, groove has been conceptualized as an aesthetic quality or "feel" that emerges from the various performed microrhythmic patterns both within or between the instrumental parts in an ensemble. This concept is often related to the "PD theory" of groove that is attributed primarily to the research of Keil (Keil, 1987, p. 275), who suggested that it is the "participatory discrepancies" (PDs) produced between performers' rhythms at the microrhythmic level that create various "process"-related feels of "'beat,' 'drive,' 'groove,' 'swing,' 'push,' etc." Despite the breadth and richness of Keil's concept of groove, which encompassed "textural" discrepancies ("timbre," "sound," "tone qualities") as well as discrepancies related to the dynamic shaping of the acoustic envelope of sounds ("inflection," "articulation"; see 1987, p. 275), most studies investigating microrhythm in groove-based music have instead focused on the purported claim of PD theory that the "groove feel" "emerges . . . from musicians' use of expressive microtiming at the sub-syntactical level" (Butterfield, 2006, para. 2).[4]

As Bengtsson, Gabrielsson, and Thorsén (1969) observed, however, such PDs or, in their terminology, "systematic variations" might act as constituents of the norm or as the groove pattern itself. Thus, rather than asserting the traditional divide between structure (notation) and expression (performance) of notation-based music, metrical grids in groove analyses should be seen to supply a pragmatic means of measurement of actual locations of rhythmic events—one that leaves open to discussion what the pattern actually is. In any case, what constitutes "structure" in groove-based styles involves both micro- and macro-rhythmic features of performed stylistic patterns, as well as the subjective reference structures that perceivers and performers alike rely upon to organize these patterns in time.

2.1.3 Groove as a Psychological Construct ("the Pleasurable Urge to Move")

Groove-based music is commonly described as music that imparts a feeling of "motion," "vital drive," or "rhythmic propulsion" and that is well suited to the synchronization and entrainment of bodily actions such as dancing (Iyer, 2002; Keil & Feld, 2005; Pressing, 2002). This particular aspect of groove is the foundation of recent psychological attempts to establish an operational definition of groove as "the sensation of wanting to move some part of the body in relation to some aspect of the sound pattern" (Madison, 2006, p. 201). Janata and colleagues (2012) also stress the coupling of groove with dance and motion but relate this urge to move to groove's pleasure aspect. Based on their survey of 153 American undergraduate students' ratings of a variety of preformed descriptive phrases believed to be "associated with the concept of the groove to varying degrees" (and informed by structural music-theoretical definitions of groove and the authors' own intuitions), these researchers concluded that groove is "that aspect of the music that induces a pleasant sense of wanting to move along with the music" (Janata et al., 2012, p. 54). Accordingly, they theorize groove as a "pleasurable drive toward action" that results from sensorimotor coupling—that is, from an "engagement of the brain's motor action systems while listening to music"—and induces a "positive affective state" (2012, p. 54). Groove as a psychological construct, then, focuses on the degree to which a given rhythmic music elicits the experience of wanting to move and/or some form of pleasure or enjoyment, among other experiential features. This, in turn, allows "groove" to be associated with any musical style that involves such experiences, regardless of cultural origin (Senn et al., 2019; see also Senn et al., 2020

---

[4] In his original work on PDs, Keil claimed that "[m]usic, to be personally involving and socially valuable, must be 'out of time' and 'out of tune,'" (1987, p. 275), and that "it is the little [processual] discrepancies within a jazz drummer's beat between bass and drums, between rhythm section and soloists, that create 'swing' and invite us to participate" (1987, p. 277). Various scholars have objected to such a sweeping claim, questioning whether the structure of "syntactical" patterns themselves, rather than simply those microtiming patterns that emerge in performance, should not also constitute a unique "groove-based" feature and contribute to different types of groove "feels" that emerge from those syntactical patterns' application (Butterfield, 2006; 2010; 2011; Câmara, 2016; Danielsen, 2006; Iyer, 2002). Butterfield, for example, maintains that groove "must draw some of its affective power from the specific nature of a given groove pattern, and nothing is gained by insisting upon the irrelevance of the syntactical level for its production" (2006, para. 4).

for an overview of typical questionnaire items utilized in “grove rating” studies). However, comparisons of genre groove ratings have found that African American styles of soul, funk, r&b, and jazz as well as West African and Afro-Latin styles in particular tend to received higher ratings for movement or pleasure than other styles, in line with many musicological definitions of groove-based music (Janata et al., 2012; Madison et al., 2011).

Interestingly, musicians themselves often use the term groove to refer to a “state of mind” wherein the creation of music becomes apparently effortless, engendering an intense, almost euphoric feeling (Berliner, 1994; Monson, 1996; Pressing, 2002). Such a state is often referred to as “being in the groove” and likewise related to groove’s pleasurable aspects (Danielsen, 2006, pp. 11–12, 215; Janata et al., 2012; Roholt, 2014, p. 108). Unlike the pleasurable states of more teleological musical forms, groove’s pleasurable state derives first and foremost from the musical process itself. That is, a groove mode of listening or dancing (Danielsen, 2006, pp. 177–179) is not directed toward a goal (such as tonic closure); instead, it demands one’s presence in the groove’s “here and now.” Witek (2017) applies a similar perspective to explain the widespread appearance of syncopated structures in groove-based musical styles: “When synchronizing our bodies to the beat, we enact parts of the musical structure by filling in the gaps; as long as the syncopations are repeated, we continue to participate, and processual pleasure is prolonged” (Witek, 2017, p. 151). Groove may therefore be characterized by structural tension at the level of the stylistic pattern that requires active interpretive participation, such as filling in beats at structurally salient positions that are not explicitly articulated in the sound. As Witek points out, this means that pleasure in groove is not necessarily caused by some cognitive-physical stimulation but emerges through one’s enactment of aspects of the musical structure and, thus, one’s constant engagement, almost as a part of the groove itself.

Various studies have investigated which structural properties correlate to listeners’ self-reported ratings of “want to move” and/or elicit some form of “pleasure,” “enjoyment,” or “liking” (simply referred to as “ratings” in the following discussion, unless stated otherwise). Madison and colleagues (2011) found that “beat salience” (“the degree of repetitive rhythmical patterning around comfortable movement rate, on the time scale up to a few seconds”) and event density (“the density of sound events between beats”) were generally positively correlated to higher ratings. Moderate syncopation, understood as the “violation” of a normative metrical expectation scheme, has widely been associated with the ratings of movement induction and/or pleasure (Madison and Sioros (2014); Sioros et al., 2014; Witek et al., 2014). Senn and colleagues (2018) also found that both syncopation and event density correlated highly with ratings of various drum patterns (though beat salience did not). Overall, findings from these studies are in accord with musicological conceptualizations, in the sense that they also observe that typical groove-based patterns tend to display syncopated events that add complexity for aesthetic purposes, but not to such an extent that they disrupt the sensation of a regular pulse (Danielsen 2006).

Regarding microtiming, various studies have investigated the effect of onset asynchronies within or between instrumental layers of an ensemble and listeners’ ratings, partly inspired by claims from strong versions of PD theory (see section 2.1.2) regarding the affective motional qualities that participatory discrepancies engender in groove-based music. Findings have been, as of yet, inconclusive. Early studies involving presentation of real musical excerpts from various genres found no significant effect (Madison et al., 2011). Controlled listening experiments that artificially introduced systematic and fixed intervals of onset displacements to synthesized stimuli generally found lower ratings for stimuli with higher onset asynchrony magnitudes compared to quantized stimuli (Frühauf et al., 2013). However, some notable exceptions to this general trend have also arisen—Davies and colleagues (2013) found that conditions of a “jazz” pattern with higher magnitudes of cyclic onset displacements (i.e., swung note pairs) were in fact rated higher than quantized conditions with no displacement, and also that both “samba” and “funk” patterns received ratings just as high as quantized stimuli when they were matched using the same type of onset manipulation method. Matsushita and Nomura (2016) tested fixed onset shifts of bass guitar rhythm (steady eighth notes) against an unchanging isochronous drum beat (“rock pattern”) and found that the condition where the bass onsets anticipated the drums (i.e. “pushed”) by about 30 ms milliseconds received ratings similar to the quantized condition. They also found that “the center of the distribution of asynchronies with higher scores was biased toward the bass-precedence side . . . indicat[ing] that complete synchrony is not always the best condition in terms of groove [ratings]” (2016, p. 127). Skaansar and colleagues (2019) also tested fixed onset shifts of an acoustic bass against isochronous drum rhythms and found that in the “high” complexity pattern condition (which most resembled a real neo-soul performance), participants rated the microtiming condition where the bass onsets were delayed (i.e. “laid-back”) relative to the drums by 40 milliseconds as high as a quantized version. They interpreted this result as demonstrating that ratings of microtiming depend on stylistic context as well as potential interactions of onset microtiming profiles and other structural features of the pattern. The difference in results from those of Matsushita and Nomura (bass anticipation also obtained high rating) and Skaansar and colleagues (bass delay also obtained high rating) may be partially explained by differences in pattern complexity and in participants’ cultural backgrounds (Japanese vs. Norwegian).

Further experiments involving stimuli with microtiming onset patterns derived directly from original performances, on the other hand, have found that expert participants produce more movement to original performances than to quantized conditions (Kilchenmann & Senn, 2015), and both experts and non-experts subjectively rated original performances just as high as those of quantized conditions; if the microtiming patterns were artificially exaggerated in magnitude (via percentage scaled, rather than fixed, intervals), on the other hand, the ratings went down (Senn et al., 2016, 2017). Hofmann and colleagues (2017) found that listeners did not prefer fully quantized versions of original performances of jazz trios (drums, bass, saxophone) but rather versions that displayed “tight” rather than exaggerated onset asynchronies. They also observed that greater magnitudes of asynchrony were preferred when participants were listening to the rhythm section together with the saxophone than when they were listening only to the rhythm section, providing evidence that, in more ecological listening contexts, listeners are “more tolerant to ensemble asynchronies” (Hofmann et al., 2017, p. 338). Finally, while not a “groove rating” study per se, Neuhoff and colleagues (2017) presented Malian participants with phrases from a manjanin rhythm with various onset timing manipulations and found that listeners rated the conditions with appropriate typical onset microtiming patterns as “a better example of the characteristic rhythm” than quantized, isochronous versions.

Overall, findings in these studies show that onset asynchronies as such may not be strictly mandatory for promoting a higher or lower sense of wanting to move and/or pleasure. Pertaining to this thesis’ topic, it should also be noted that to simply manipulate onset timing in listening experiments without altering any other acoustic-envelope features of sounds (attack, duration, intensity, or timbre), as virtually all of the aforementioned studies did, may not be entirely ecological given the evidence from perception studies regarding how acoustic features interact with perceived timing (see Chapter 3). Also, Danielsen, Waadeland and colleagues (2015) have shown that musicians may also systematically manipulate such sound features when intending to play with different microrhythmic feels—a finding that has been further tested in the present thesis (see also section 2.3.4 and this thesis's papers).

To summarize, most psychological studies of groove have operationalized it as a transcultural perceptual/experiential phenomenon focused extensively on the aspect of music that elicits "pleasure" and/or "the desire to move" and have identified certain structural rhythmic features that correlate either more (e.g., event density, syncopation) or less (e.g., beat salience, onset microtiming) consistently to such experiential aspects depending on various contextual factors. While such functional approaches capture valid and importance experiential effects that groove-based styles may engender, it must be acknowledged that, like more structural music-theoretical approaches reviewed in section 2.1.2, they may only reveal a limited set of aspects associated with the multifarious concept of groove.

2.2 Rhythm Production and Perception

As we saw in the previous section (2.1), despite the multiple meanings and definitions ascribed to groove, one central component found in all of them is rhythm. This leads us to ask the following: How do we perceive and produce rhythms? What is a rhythm to begin with? When a series of sound events (whether produced by human, machine, or natural process) presents regular time intervals between each successive event, it is often said to be "rhythmic." In a musical context, however, a "rhythm" tends to refer more specifically to a pattern of durations formed by intervals between the perceived time points of events (London, 2001). Several studies have described the experience of musical rhythms by humans as involving an interactive process between exogenous events demonstrating regular temporal patterns and endogenous cognitive frameworks with which humans categorize and organize those events in time. Various terms have also been proposed in the fields of musicology and music psychology to delineate this distinction, such as "rhythm" versus "meter" (Clarke, 1987; Gjerdingen, 1989; Lerdahl & Jackendoff, 1983; London, 2012) or "actual/sonic events" versus "virtual/underlying reference structures" (Danielsen, 2006; Haugen, 2016). While the two respective terms may vary, the general consensus is that their interaction is twofold—that is, externalized musical rhythms always generate some form of internal reference structure within us, while, in order to correctly externalize a musical rhythm in the first place (whether by performing it on an instrument or simply moving a body part along to it), we must rely upon an internal reference structure to guide our actions. In other words, reference structures exist not only to allow us to cognitively represent external sounding rhythms within our minds but also to anticipate, predict, and organize our actions. In addition, the nature of these structures may vary depending on one's cultural background and degree of musical expertise, and the stylistic context (Jacoby & McDermott, 2016; Polak et al., 2017; see also section 2.3.1).

In this thesis, I will use the term (*subjective) reference structures* to label the various types of organizational frameworks we use to perceive and produce musical rhythm, whether they consist of pulses, subdivisions, specific stylistic patterns/figures, or total metrical systems. Here, the occasional use of the prefix "subjective" simply announces that these structures are essentially cognitive constructs that stem from within us as human "subjects." Regarding musical rhythms, I shall also broadly distinguish *physical rhythms*, or those that exist as physical sound phenomena in the environment and can be objectively measured via recorded audio signals, from *perceived rhythms*, or those that are related to how we subjectively interpret and categorize physical rhythmic sounds via various reference structures. While physical rhythms may or may not directly manifest reference structures—consider, for example, highly syncopated groove-based patterns that do not express a clear-cut pulse—the assumption is that, in order to produce those patterns, they require a perceptual temporal organizational framework in the first place. In the following sections, I will examine theoretical concepts of reference structures assumed to be important to groove-based rhythms, such as meter (2.2.1), pulse (2.2.2), and subdivision (2.2.3).

2.2.1 Meter

In music theory, "meter" is commonly held to denote the total temporal framework with which we organize rhythmic events in time, and it is typically described as a hierarchically nested system of pulsations occurring at different periodicities (London, 2012; Poudrier & Repp, 2013). While inter-onset intervals (IOIs) between rhythmic events in music can be said to occur at various time spans ranging roughly from 100 milliseconds to 6 seconds, London (2012, p. 33) points out that we do not "perform or perceive all durations and durational patterns within this range in the same way." At the heart of most theories of meter resides the concept of the "pulse" or "tactus" as the primary metrical level. Citing evidence from several perceptual experiments, London (2012, pp. 30–33) theorizes that the lower and upper IOI thresholds for events comprising the pulse in a meter are approximately 250 milliseconds and 2 seconds, respectively. Rhythmic pulses in this periodicity range are strongly associated with the human sensorimotor system, and in music often correspond to comfortable rates to which one readily synchronizes via body movements such as hand claps or foot taps (Jones, 1976; Large & Jones, 1999; London, 2012; Merker et al., 2009)

Conventional music-theoretical conceptions of meter based on notational approaches posit that a sense of pulse is established via the subjective accentuation of alternating strong and weak beats (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1984). Here, the greater the overlap between metrical levels, the greater the subjective accentuation or weight ascribed to a beat location within a pulse. In common 4/4 meter, for example, the first and third (quarter note) beats of the pulse within a measure tend to be described as "stronger," since they overlap with beat locations of faster (both eighth and sixteenth notes) and slower (half note) metrical levels, whereas the second and fourth quarter note beats are described as "weaker," since they overlap only with beat locations of faster metrical levels. In recent times, scholars have questioned the universality of metrical accentuation schemes derived from traditional Western notation systems, based on observations of rhythms from various musical styles that do not reflect such schemes (Abel, 2014; Danielsen, 2006; London, 2012; Polak, 2010). For example, many groove styles in 4/4 meter feature a "back-beat" pattern that involves sounding events occurring only on beats 2 and 4. If dynamic accentuation patterns of physical rhythms are assumed to reflect subjective metrical accents as closely as possible (Lerdahl & Jackendoff, 1983), then the resultant metrical accent scheme might be (a) inversely felt as "weak–strong–weak–strong" or maintained as "strong–weak–strong-weak," and where the back-beats are either (b) heard as salient syncopated events that contradict the metrical accentual scheme or (c) incorrectly identified as events occurring on the expected strong metrically accented beats of 1 and 3. However, if the link between dynamic accent and metrical accent is broken, there is no need to assume that the rhythmic events on beats 2 and 4 are felt as either accentual contradictions to the meter or a potential confusion of the downbeat of the pattern as displaced/shifted by a quarter note in the pulse scheme. Zuckerkandl (1956), for example, proposes that meter should not be viewed as produced from a pattern of strong and weak accents but rather likened to virtual cyclical motion "waves" comprised of "to-from" or "away-back" movements. Abel (2014, p. 50) embraces such a conception to explain how events on beats 2 and 4 of the back-beat pattern can be emphasized without usurping the "strong" beats of the meter. He suggests that 4/4 meter can be conceived more simply as a 2/4 meter with duple divisions at every level comprising oscillating patterns of "one–two–one–two," where events falling on "two" are not considered weak as such but rather as "away-from-one." Consequently, "the status of beat 'one' depends not on its relative strength but on its character as an 'on' or 'away' beat and on its role in marking the start of the cycle," and "[c]onversely, beats 'two' and 'four' have quite a different character, that of a 'return' or preparation for the next on- or away-beat" (2014, p. 50). The broader implication of such a view is that any event coinciding with the beats of a given metrical level may simply be considered "on-" or "off-beat" relative to faster or slower levels without requiring that accentuation patterns present in physical rhythms reflect any particular assumed metrical accent scheme.

2.2.2 Pulse

Groove-based styles are overwhelmingly described as music for dancing, or, more generally, as music eliciting a strong urge to move one's body in some form or fashion. As such, the perception of a regular, steady pulse—most commonly referred to as "the beat" in popular parlance within musician communities—would appear to be vital for groove-based styles, for without a fundamental reference structure such as a steady beat to guide dancers' feet or musicians' fingers, there is arguably no experience of perceived rhythm or groove (in any of its definitions). When a pulse level is not explicitly manifested by any particular instrument in a groove-based style, London, Polak, and Jacoby (2017, p. 479) argue that cyclically recurring patterns are what gives rise to a sense of regular beat, as well as "the higher and lower levels of metric organization." However, in many African American–derived groove styles, the pulse level is in fact often systematically externalized by one or more rhythm section instrumentalists. Whether it be funk, R&B, or rock, one very often finds an instrument playing sounds regularly spaced out by even intervals.

Both Abel (2014) and Iyer (2002) claim that adherence to a near- or virtually isochronous pulse is the key feature that distinguishes the groove rhythm aesthetic from those of other stylistic traditions, such as Western art music, especially since the groove aesthetic tends to eschew the use of global tempo variation for expressive purposes. Abel (2014, pp. 26–27) notes that, in nineteenth-century Romantic music, disruptions to the sense of a steady pulse via tempo manipulation by an entire ensemble are commonplace, either in the form of slowing down and speeding up (ritardandos and acellerandos) or long indeterminate pauses (fermata), but in groove-based styles such operations are extremely rare (with the exception of occasional "novelty effect" slow-downs or speed-ups at the ends of songs). Groove-based popular music, that is, "puts great emphasis on avoiding involuntary" tempo changes, which are considered "'destructive' or 'anathema' to a good sense of groove" (2014, pp. 26–27).

The rate at which the pulse level occurs in such rhythmic music often corresponds to the rate at which one most readily entrains to a musical rhythm via body movements such as hand claps or foot taps (Jones 1976; Large & Jones, 1999; London, 2012; Merker, Madison, & Eckerdal, 2009).

Recently, dynamic-systems approaches to pulse and meter have posited that meter is a musically specific instance of the general human capacity of "entrainment," or as a "sympathetic resonance of our attention and motor behavior to temporal regularities in the environment" (London, 2012, p. 190). These approaches are largely based on the perceptual theory of *dynamic attending* (Jones, 1976; Large & Jones, 1999) which assumes that humans possess internal biological oscillators (referred to as "attending rhythms") that can synchronize with "external rhythms" that are present in the environment. When confronted with external rhythms, these internal oscillators are set into motion, generating self-sustaining "periodic activity" that tunes into and synchronizes with the period and phase of salient features in the external rhythms via a process termed "abstraction" (Jones, 1990). This synchronized activity in turn leads to "generation" (Jones, 1990), whereby the oscillators are reinforced and maintained, creating peaks of "attentional energy" or temporal ranges where salient features of the external rhythm, such as events coinciding with the beats of a pulse, are expected to occur (Large & Jones, 1999). Unlike the static clock models of rhythm perception (Povel & Essens, 1985), self-sustaining oscillators are able to regularly adapt their period and phase to perturbations and changes in the external rhythms and are therefore said to be capable of dynamic "entrainment" rather than simple synchronization.

Danielsen (2018) and Danielsen, Haugen, and colleagues (2015) argue that onset asynchronies between near-simultaneous events at the pulse level may encourage a widening of the "attentional focus," thus allowing multiple events to be positioned within the "beat bin" of a given metrical level such as the pulse. Attentional focus is described by Jones (1976) as the result of a process whereby attentional energy is allocated over time as attending rhythms synchronize their phase and period with those of external rhythms. Large and Jones (1999) posit that, as synchronization is increased, the width of the attentional peak is narrowed, and, conversely when synchronization degrades, the peak widens. Danielsen, Haugen and colleagues (2015, p. 136) suggest that this may explain why a timing deviation from an expected beat location is "more likely to be noticed when attention is highly focused (has a narrow pulse) than when it is broadly focused (has a wide pulse)." They further suggest, however, that in a musical context, listeners are able to direct their attention toward either single or multiple instrumental rhythmic layers at any given time, which may lead to alternative widening or narrowing of attentional focus, thus affording a certain degree of flexibility to the perceived beat bin.

Dynamic attending models of meter also have parallels with conventional metrical accentuation schemes, in that "the sense of accent that accrues to the downbeat is the result of the mutual reinforcement of the component oscillators" (London, 2012, p. 21). That is, when the abstracted periodicities, or metrical levels, of external rhythms overlap, they are said to strengthen the attentional peaks and thus increase the expectancy of an event at those locations. However, when one is met with external rhythms with which one is not entirely familiar, London suggests that one's inference of a metrical framework "does not involve extracting invariant information such as component periodicities" via a process of abstraction "but rather matching the musical figure against a repertoire of well-known rhythmic/metric templates" (2012, p. 28). He also uses the example of the back-beat pattern, arguing that the reason why enculturated listeners readily correctly identify it as belonging to a 4/4 metrical framework, where the salient rhythmic events on beats 2 and 4 serve to strengthen their sense of pulse and meter rather than subvert it, is because they "have a bevy of metrically familiar templates at their disposal, and in recognizing these commonplace gestures they are readily able to establish [correct] metric entrainment" (2012, p. 28).

2.2.3 Subdivision

Most rhythmic music displays events occurring at faster periodicities than those of the supposed pulse level. Few would claim, for example, that the ticking of an isochronous metronome stream alone would constitute a groove, let alone an instance of music. Events occurring at faster levels are ubiquitous in groove-based music, and certain styles can even be partly identified based on their "density referent" (Nketia, 1974) or "metric floor" (London, 2012)—that is, the metrical level comprising the shortest practical subdivision unit of a given meter. For example, in rock styles, eighth notes often tend to comprise the shortest frequently utilized unit, whereas, in funk and disco styles, sixteenth notes are highly commonplace (Abel, 2014; Câmara & Danielsen, 2019; Danielsen, 2006, 2018; Stewart, 2000). Groove styles are also further identifiable based on whether subdivision levels are organized in duple/quadruple or triplet grouping patterns (Abel, 2014)—for example, blues styles are regularly organized in 12/8 meter, where the eighth notes are subdivided into triplets, while funk, soul, and rock styles tend to feature duple or quadruple subdivisions in 4/4 meter.

However, the density referent in groove is not always so straightforward, as Danielsen (2006) has noted, since it depends not only on the onset locations of events but also on their durations. When events corresponding to the onset locations of a given metrical level are played with shorter durations than the categorical note span of said level, they could be perceived to imply a faster metrical level. She points to the example of the horn stabs in James Brown's "Papa's got a brand new bag," (1965, King 5999), the onsets of which coincide with the eighth-note level but are played in an extremely snappy and staccato fashion, leading to an ambiguity in the "emergent non-fit of subdivision [phrasing] and density referent" and thus implying the presence of a faster sixteenth-note level (Danielsen, 2006, p. 75). Stewart (2000, p. 309) also notes that, in various styles of funk, "[e]ven when it is not played . . . the sixteenth note is always felt or implied." As such, perhaps even the most basic of backbeat drum patterns with event onsets occurring only on eighth-note beat locations could be said to imply faster subdivision levels due to their short and impulsive character, especially since every drum sound event never lasts a full quarter note duration (how odd, dull, and, I dare say, "ungroovy" such an experience would be!). Whether explicit or implicit, though, events occurring at faster subdivision levels of the main pulse are often described as imparting qualitative sensations of rhythmic "drive" (Danielsen, 2006) or "motional energy" to grooves (Butterfield, 2011). Psychological studies have also shown that faster metrical levels tend to facilitate greater entrainment to the pulse level by providing extra temporal cues (Madison, 2014).

In addition, the degree to which subdivisions are "swung" can be a further identifying aspect of groove-based styles. "Swing" as a rhythmic device is often associated with the use of unequal long–short eighth-note duration patterns in jazz styles (Benadon, 2006; Butterfield, 2006, 2011; Friberg & Sundström, 2002), though various groove-based styles outside of jazz have also been shown to demonstrate different duration ratios ("swing ratios") at the sixteenth-note level, such as funk (Câmara, 2016) and hip-hop (Frane, 2017). In both jazz and funk, as well, swing ratios can vary both globally across a performance and locally between beats within the same performance, and they have been shown to depend on tempo, sub-style, instrument, and/or individual player preference (Câmara, 2016; Frane, 2017; Friberg & Sundström, 2002; Haas, 2007; Honing & de Haas, 2008). In an expanded sense, then, swing can be considered one type of systematic microrhythmic relationship that may be present in groove styles, alongside beat delay/anticipation and the presence of early/late onset timing asynchronies between instruments. Accordingly, when events at the pulse subdivision level are systematically swung by performers to such a consistent and stable degree across an entire stylistic tradition, it may be more accurate not to speak of microrhythmic variations (or "deviations") from a supposed (near-)isochronous subdivision reference as such but rather to find that unequal swung note patterns may constitute non-isochronous subjective reference structures in and of themselves (see sections 2.3.1 and 2.3.2).

2.3 Microrhythm

As touched upon in the previous section, in recent decades, several empirical studies have found that, when they are measuring the timing of physical rhythmic events in performances, musicians reveal different degrees of variability around assumed subjective reference structure locations. This phenomenon has been variously referred to as "participatory discrepancies" (Keil, 1987; Keil & Feld, 2005; Prögler, 1995), "expressive timing" (E. F. Clarke, 1989), "expressive variation" (London, 2012), "systematic variation" (Bengtsson et al., 1969), "microtiming" (M. W. Butterfield, 2006; Frane, 2017; Iyer, 2002; Madison et al., 2011; Naveda et al., 2011; Senn et al., 2017), and "microrhythm" (Benadon, 2006; Câmara, 2016; Câmara & Danielsen, 2019; Danielsen, 2006, 2018; Gerischer, 2006). Though I also use the term "microtiming" in papers I–III, in keeping with groove-based research literature, here I shall mainly use the term *microrhythm* to denote the phenomenon. First of all, I do so because the suffix "-timing" is heavily connoted with event onset locations or inter-onset intervals (IOIs), to the point that they are somewhat synonymous. However, the perceived timing of a rhythmic event does not always correspond to the onset of its physical signal and instead has been shown to be influenced by other sound parameters, such as intensity, duration, and frequency/pitch (Danielsen et al., 2019; London et al., 2019; Gordon, 1987; Wright, 2008; Villing, 2010; see also Chapter 3). The suffix "-rhythm," therefore, is a broader term that more readily encompasses the acoustic shape of rhythmic events beyond their simple onset timing. Second, I find the prefix "micro-" apt in its highlighting of the fact that the magnitude of variation of either the physical or perceived timing of rhythmic events occurs at a time resolution below presumed categorical levels of rhythmic perception, commonly termed the "macro" or "structural/syntactical" levels (Butterfield, 2006, 2011; Câmara, 2016; Clarke, 1987; Danielsen, 2006; Keil, 1987; Keil & Feld, 2005; Kvifte, 2004; Polak, 2010).

In section 2.3.1, I will first examine what types of perceptual categories may be at work in our subjective reference structures at the microrhythmic level, with a particular focus on how we perceive rhythmic events at the boundaries of such categories. I will also consider the meaning of microrhythm in terms of structural "norm" versus "variation/expression." Then, I will review some of the literature that has provided evidence of the presence of microrhythmic events in groove-based performances, conceptualized as "swing" or non-isochronous subdivision structures (section 2.3.2) and instances of beat delay/anticipation (section 2.3.3). Finally, in section 2.3.4, I will address why the production and perception of groove-based microrhythmic patterns must encompass both the onset timing profiles of events and the manner in which they are sonically shaped.

2.3.1 Boundaries and “Norms” of Perceptual Categories

It is commonly assumed that perception of musical rhythm has a “categorical” aspect (Clarke, 1987; Desain & Honing, 2003; Schulze, 1989). That is, while physical rhythm and time are “continuous” in nature, our perception of them is not necessarily so—instead of operating with an infinite number of rhythmic categories within our arsenal of subjective reference structures, we distinguish among rhythmic events using a limited and specific set of familiar perceptual categories of varying timespans, such as metrical pulse beats or subdivisions. London (2012, p. 122) proposes that the “hallmark of perceptual studies” is that “we are poorer at making discriminations within a category . . . than at category boundaries.” Therefore, while rhythmic events occurring within the timespan of the beat of an expected metrical level are more readily considered to be manifestations of the beat itself, events occurring just before or after those boundaries may not necessarily be heard as belonging to a faster metrical level but instead heard as “subsumed” within the nearest slower category. Exactly how many or what types of perceptual rhythmic categories we are able to distinguish, as well as what constitutes the boundaries of such categories, are questions various scholars have posed.

Traditional notation-based conceptions of meter tend to either explicitly or implicitly assume that our rhythmic categories are isochronous in nature at the level of both categorical pulse and subdivision levels (Lerdahl & Jackendoff, 1985), and this assumption is reflected in the grid representations of metrical notation. However, based on empirical evidence of performance timing in both Western and non-Western stylistic traditions, more recent conceptions of meter have acknowledged that both pulse and subdivision categories may also be non-isochronous in nature (Danielsen, 2010; Danielsen, Haugen et al., 2015; Gerischer, 2006; Haugen, 2016; London, 2012; Polak, 2010; Polak & London, 2014). Experimental studies of rhythm production and perception have also assumed that motor-synchronization production tasks can serve as probes into the nature of subjectively perceived categories of rhythm. One such common experimental paradigm involves finger tapping, whereby participants either tap along to a rhythmic “target” stimulus or continue tapping after an initial synchronization period (see Repp, 2005; Repp & Su, 2013). Many studies show that, when asked to tap along to rhythms comprised of two IOI intervals with more complex integer relationships than 1:1 (isochronously spaced events) or 1:2 (such as an eighth note followed by a quarter note), participants regularly tend to “soften” or “sharpen” their tapping ratio toward the simpler ratio (Jacoby & McDermott, 2017; Polak et al., 2018). In these studies, the reliable reproduction of a target stimulus ratio is generally considered to be evidence of a valid a priori perceptual rhythmic category. When a performance consistently differs from the stimulus ratio, however, it is not considered to reflect additional, more complex perceptual rhythmic categories as such but instead considered to be subsumed by other simpler rhythmic categories that “distort” the rhythmic production toward themselves. As such, simple 1:1 and 1:2 ratios have been described as “attractor ratios” (Repp et al., 2012), and some studies claim them to be universally valid across cultures (E. F. Clarke, 1987; Drake, 1993; Fraisse, 1982)

Polak and colleagues (2018), however, claim that most of these studies involve selective samples of participants mainly with Western cultural backgrounds and argue that rhythmic categories of perception (“rhythmic prototypes”) are not solely determined by innate biological factors (and thus “universal” as such) but also influenced by cultural background and cognitive mechanisms of statistical learning (and thus “relativistic” instead). To test this hypothesis, they devised a cross-cultural tapping study with participants from different performance traditions and cultural backgrounds (Mali, Bulgaria, and Germany) and demonstrated that participants were indeed able to faithfully reproduce rhythms featuring more complex relationships such as 2:3 at the pulse level (Bulgarians) and 3:4 at the subdivision level (Malians), which they explained in part by the presence of such rhythmic prototype ratios in those respective cultures and musical traditions. These researchers also found that successful production of certain complex rhythmic prototypes only emerged at slower or faster tempos, demonstrating that the combination of durational ratio and tempo range determines in part whether a rhythmic category will be perceptually valid or not. In any case, it would appear that, as London (2012, p. 123) suggests, "categorical determinations are not simply 'stimulus driven' but a product of the interactions between stimulus and listener, a listener who has learned to categorize certain durations in a certain context in a particular way."

Further evidence regarding the extent to which physical rhythms are perceived as isochronous or not is provided by P-center (perceptual center) and PAT (perceptual attack time) studies. In general, these studies show that, in a series of rhythmic events with onsets spaced isochronously (thus displaying "physical isochrony"), those events are not necessarily perceived as such. Rather, the perceived timing of the events depends in part on the shape of their acoustic envelope, where, in particular, both the duration of the attack segment (or "rise time") and the total extent of the given sound affects whether or not a steady stream of events will be heard as "perceptually isochronous" (Danielsen et al., 2019; Gordon, 1987; London et al., 2019; Villing, 2010; Wright, 2008; see also section 3.1). While no P-center study has specifically investigated thresholds of isochronicity in music per se, Danielsen and colleagues (2019) suggest the variance (standard deviation) of derived average P-center locations can provide an indication of the expected temporal width, or beat bin (Danielsen, 2010), of sounds in perceptually isochronous contexts. Related to their interpretation of dynamic attending in pulse perception, Danielsen and colleagues (2015) also theorized that, because expected pulse locations can be considered distributions of attentional energy that may vary in temporal extent (see also Danielsen, 2018), events that occur within the bounds of an attentional peak coinciding with an expected pulse location will be heard as falling "on the beat." Several music scholars have also observed that the extent to which something is considered on-the-beat or not may in part be determined by the musical context as well as the rhythmic aesthetic norms present in a given musical style (Abel, 2014; Haugen, 2016; Johansson, 2010; Snyder, 2000; Stover, 2009). Based on evidence from various music perception and performance studies, then, it would appear that strict physical isochrony is not a prerequisite feature of subjective reference structures in either the perception or the performance of groove-based rhythms at the pulse or subdivision levels.

In the context of rhythmic events said to occur at the microlevel of music (that is, at the boundaries of subjective perceptual rhythmic categories), the often associated modifiers of "expressive," "variation," and "systematic" all essentially hinge upon the concept that there exists a referential "norm" from which events differ or "deviate." However, as Bengtsson, Gabrielsson, and Thorsén (1969) note, what constitutes "the norm" for a pattern or any subjective reference structure is an open question that might, in fact, be impossible to answer due to the "systematic variations" present in so many styles of music. Polak (2010, para. 149) argues that, in fact, "the common understanding of expressive timing as deviation from nominally (or categorically) isochronous pulsation is useful primarily in the context of musical styles that employ roughly isochronous meters," and that it may be "relevant mainly for those musical traditions that rely on the writing and reading of music." While I agree with this statement, I would add that such an approach may be valid, as well, for many groove-based styles, such as funk, that rely on combinations of notation (particularly horn-section instrumentalists) and oral learning (rhythm-section instrumentalists). However, as Polak rightly suggests, in musical traditions where "deviations" from an isochronous norm are particularly systematic and regular throughout the course of a performance at a given metrical level, what is often referred to as timing or rhythmic "feel" in music may alternatively be described as "metrical feel," or "an instance of departure for expressive variation . . . rather [than] the point of departure for expressive variation" (2010, para. 149).

2.3.2 Swing and Non-Isochronous Subdivisions

When events perceived to occur at a faster metrical level than the pulse are performed with unequal duration patterns to a consistent degree and in a stable fashion throughout a performance, this is commonly referred to as "swing" and described in numerical relationships ("swing ratios"). In jazz performances studies, for example, swing ratios of drummer's rhythms have tended to focus mainly on the degree of "global swing"—that is, average swing ratio across entire excerpts (Câmara 2016)—in steady streams of eighth-note patterns, such the standard jazz swing pattern played by drummers on the ride cymbal. Investigations of commercial jazz recordings (Benadon, 2006; Friberg & Sundström, 2002; Prögler, 1995; Rose, 1989) as well as performance experiments (Haas, 2007; Honing & Haas, 2008; Prögler, 1995) have found that drummers apply a wide range of ratios from approximately 1:1 (isochronous, or "even/straight" eighth notes) to 4:1 (a dotted eighth note followed by a sixteenth), depending on subgenre, tempo, and/or individual player preference. When examining solo-instrumentalist rhythms exclusively, studies tend to shift the investigative focus toward "local swing" ratios instead—that is, swing ratios of all successive eighth-note pairs within the measure (Câmara, 2016). Several in-depth analyses of highly regarded solo performers have found that local swing tends to fluctuate quite significantly throughout a performance in extended solo passages (Benadon, 2006; M. W. Butterfield, 2011; Collier & Collier, 2002; Ellis, 1991). The use of different degrees of swing ratio in such passages has been theorized to convey various qualities of "motional energy," or "the force of momentum with which some musical events are directed toward others" (Butterfield, 2011, p. 4).

Swing ratios higher than 1:1 have also been found at the sixteenth-note level in jazz-funk, funk, and hip-hop styles, both at the global and the local levels (Butterfield, 2006; Câmara, 2016; Danielsen, 2006; Frane, 2017), albeit to a subtler degree overall. Based on the findings of a psychoacoustic swing discrimination study by Frane and Shams (2017), it would seem that events swung at a ratio of about 1.2:1 or higher in the tempo range of the musical excerpts in these studies would more likely be perceived as categorically swung (or different from "straight/even" events) by expert drummers, at least, but also by a particularly discriminating population of listeners, and more so in rhythms with high rather than low swing density. In general, these swing ratio findings provide evidence that an ambivalent mix of both straight and swung events at the subdivision level may be a characteristic aesthetic-stylistic trait of various funk-derived styles.

Overall, in both eighth note–based jazz and sixteenth note–based funk, jazz-funk, and hip-hop groove-based styles, different instruments may display conflicting degrees of either global or local swing, even within the same performance. As such, the notion of a single normative non-isochronous subjective reference structure shared by all instrumentalists in an ensemble becomes complicated. In any given style, pattern, or tempo, then, which musician's swing ratio should be considered the “norm” from which all other musicians systematically deviate or express variation? And which performances by which performers in which recordings should be considered prototypical? While it is impossible to answer these questions, we might instead consider that, in any given context, an individual player operates with their own degree of swing at the subdivision level—one that may or may not match that of the other players—thus affording multiple possible combinations of synchronicity or ambiguity in the total musical texture and engendering various potential degrees of swing feel from the subtle to the stark.

2.3.3 Beat Delay/Anticipation (Laid-back versus Pushed)

Another form of systematic microrhythm thought to contribute to the overall rhythmic “feel” in groove-based music involves the interaction between and within the perceived timing of various instruments in an ensemble. On the one hand, it has been argued that “tightness,” or the close synchronization between the instruments within an ensemble, is the primary timing approach in groove-based performance (Abel, 2014). On the other hand, it has also been considered a hallmark of technical proficiency (in African American–derived groove performance practice, at least) to be able to play flexibly around a given timing reference in a controlled fashion while maintaining a steady tempo—either “behind the beat”/“laid-back” or “ahead of the beat”/“pushed” (Danielsen, Waadeland et al., 2015; Iyer, 2002; Kilchenmann & Senn, 2011). In fact, a myriad of references to musicians displaying laid-back or pushed approaches in performance can be found in ethnographic accounts and empirical investigations of multiple groove-based genres, including jazz (Berliner, 1994; Friberg & Sundstrom, 2002; Keil, 1987; Monson, 2006), funk (Butterfield, 2006; Câmara, 2016; Danielsen, 2006, 2012), classic soul (Bowman, 1995), neo-soul/hip-hop (Bjerke, 2010; Carlsen & Witek, 2010; Danielsen, 2010; Danielsen, Haugen et al., 2015; Peterson, 2008; Stadnicki, 2017), and African American music in general (Iyer, 2002).

Regarding groove-based drumming practices, certain timing feels are commonly associated with a given rhythmic pattern, such as the presence of a slight delay (laid-back timing) in the snare strokes of the ubiquitous “back-beat” pattern. Iyer (2002) observed that, in drum-kit performance of many African American styles, if one considers the kick drum as establishing the primary reference of an assumed isochronous pulse comprised of four beats, then the back-beats of the snare on beats 2 and 4 seem to arrive slightly later than an ideally expected beat location. He notes that despite the magnitude of effect being “much subtler than the salient rhythmic categorization of the long and short durations of swing,” it nonetheless contributes to different qualitative feels: a back-beat with delay is regularly described as relatively more “relaxed” compared to a “stiff” or “on top” rendition without delay (2002, p. 406). Iyer further speculated that back-beat delay may result from the time delay caused by the different distances that nerve signals must travel between the brain and the hand and foot, thus creating a micro-asymmetry when one is producing a steady stream of events at the pulse level. Butterfield (2006), however, rejected this suggestion based on evidence from tapping synchronization studies by Wohlschläger and Koch (2002) that showed that tempo, practice, degree of expertise, and stimuli type can reduce synchronization error in production, as well as theoretical speculation that performers intentionally utilize microrhythmic deviations from isochrony in order to exploit various qualitative effects of “anacrusis” (see section 2.3.2).

Frane (2017, p. 299) also found that, in a sample of thirty classic hip-hop “drum breaks,” the snare backbeats “were often (albeit not always) slightly delayed at beat 2, but not typically delayed at beat 4.” In such cases, he notes, one might well hear the IOI between beats 4 and 1 as shorter than the IOI between beats 1 and 2, meaning that events on beat 1 would appear to arrive slightly early (as pushing). This alternative model has been termed the "downbeat in anticipation" by Danielsen (2006), who found a similar microrhythmically early tendency in the timing of events on beat 1 in several of the funk grooves of James Brown and Parliament. Câmara (2016) also recognized that sixteenth-note syncopations or pick-ups that are swung to a high ratio in funk and jazz-funk could be perceived as early instances of an ensuing subjective downbeat location. However, the distinction between back-beat delay and a downbeat in anticipation amounts to a matter of perspective that depends on which instrument layer, or subjective reference structure, one uses to establish a timing reference—either may be considered "correct."

Back-beat delay has been called a prominent feature of the soul recordings of the Stax label. Bowman (1995) describes how, when the Stax recording studio location moved to an old movie theatre in the mid-1960s, musicians were spaced far apart in the room during recording sessions, which led to a slight delay between the rhythms of house band guitarist Steve Cropper and drummer Al Johnson in recordings. Cropper recalls that at first this delay was unintentional, but because it was received as aesthetically desirable, he and Johnson eventually refined a strategy of establishing tight synchronicity on beats 1 and 3 while laying back on beats 2 and 4 ("which became an actual physical thing, not room delay at that point"), as well as synchronizing more tightly on the back beats at strategic transitional points in a groove, such as when a drum fill would mark the end of a section (Bowman, 1995, 309). Such statements by expert musicians provide evidence, albeit anecdotal, that microrhythmic expression may very well be intentional rather than an inevitable consequence of human sensorimotor system constraints.

Several studies have examined the "elasticity" or "controlled flexibility" (Berliner, 1994, chapter 13) of jazz musicians' performances via systematic laying back or pushing one instrument's timing against another. Friberg and Sundstrom (2002) found that, in a sample of five commercial jazz recordings, bassists and solo horn/keyboard musicians tended to play late relative to the drummers' ride cymbal at on-beat quarter-note locations but synchronized more closely on the eighth-note off-beat locations. Ellis (1991) instructed three saxophone players to play on a wind MIDI controller against a computer-generated quarter-note walking bass line and found that all three delayed their onsets relative to the bass. In a swing-jazz recording from a play-along record, Rose (1989) found that the onset timing of the bass and piano tended to be behind that of the drums. From these studies, we might also conclude that drummers strive to play ahead of the rest of the rhythm section and solo instrumentalists. But, once again, what is considered laid-back or pushed is simply a matter of perspective—it depends on which instrument constitutes the timing reference layer, as well as what the "beat" is considered to represent. Hoffman and colleagues (2017), for example, found that in performance of three jazz standards by six ensemble combinations (three drummers on a MIDI kit [hi-hat pedal, ride cymbal], two bassists [electric], and one saxophonist [acoustic]), relative to an idealized pulse (calculated as the average onset location of all the instruments), the drummers tended to play the ride cymbal late but the hi-hat early. Also, while the bassists generally played behind the idealized pulse in terms of the absolute onset asynchrony among the instruments, they played early relative to the drummers' ride cymbal but late relative to the hi-hat. Therefore, the straightforward assumption that playing behind or ahead of the "beat" (thus, laid-back or pushed) might be the "preferred approach" for any given instrumentalist in any given style is yet again challenged when one takes into account the multiple rhythmic reference layers (both physical and idealized) that are present in polyphonic performances.

Other examples of intentionally produced microrhythmic variations in groove-based music can be found in the dramatically "lilting" grooves of neo-soul styles from the 2000s influenced by hip-hop producer J-Dilla, drummer Ahmir "Questlove" Thompson, and artist Michael "D'Angelo" Archer. In these styles, considerable onset asynchronies often appear among the different instrumental layers, either intentionally produced in performance by musicians or programmed and manipulated in digital audio editing software post-performance, or a combination of both (Danielsen, 2010; Stadnicki, 2017). In analytical investigations of the track "Left and right" (D'Angelo, *Voodoo*) Danielsen (2010) measured onset asynchronies of up to 80 milliseconds between the rhythmic layers of guitar and drums around beats 2 and 4 of an idealized isochronous pulse. The track tempo was stable and constant, interestingly, at around 92 beats per minute, situating the pulse in an isochronous context, at least at the measure level (idealized pulse beat IOI = about 652 milliseconds). Although these onset asynchrony values are well above heuristic values of just noticeable differences derived from quasi-musical psychophysical studies of both onset asynchrony detection (at ca. 5 percent of IOI = ca. 31 milliseconds; "cyclic displacement", Friberg and Sundberg, 1995) and the perception of a steady, regular pulse (at ca. 8.6 percent of IOI = ca. 54 milliseconds; "pulse attribution", Madison & Merker, 2002), listeners are still able to maintain the perception of a steady, regular pulse despite the "seasick" or "lilting" feel that the track can engender. Danielsen and colleagues (2015) suggest that this is because, based on a dynamic attending model of pulse entrainment, the temporal width of one's subjective beat bin can increase to allow the multiple events falling around idealized isochronous pulse locations as part of the beat. Therefore, it would appear to be possible to maintain a sense of a regular, virtually isochronous pulse even when faced with rhythms replete with onset asynchronies around idealized pulse locations. While a main purported feature of groove-based music involves rhythms that "give rise to the perception of a steady pulse," often via "strict adherence" (Abel, 2014, pp. 26–29), as Standnicki (2017, p. 256) correctly observes, the grooves of neo-soul styles such as D'Angelo's "Left and right," become "endowed with extra expressiveness because of their isochronous context" by "blurring the pulse location and tapping into a particularly wobbly rhythmic swagger."

While instrumentalists often have a "preconceived" feel they wish to communicate prior to playing, Prögler (1995, p. 35) cautions that "an intention to vary the qualitative elements of a performance can indeed result in measurable, sub syntactical quantitative differences, even though they may not match the stated intentions of the performer." A few studies, however, have provided empirical evidence of the extent to which musicians are able to intentionally play laid-back or pushed (as opposed to on-the-beat) in controlled experimental contexts along to fixed prerecorded metronome stimuli. The earliest such study seems to have been by Reinholdsson (1987), who instructed a bassist to play a jazz pattern along to a metronome with "pushed" and "laid-back" feels, respectively, and found that the bassists successfully placed their strokes ahead and behind the metronome, demonstrating conscious control of the feel. Kilchenmann and Senn (2011) instructed two drummers to play a jazz-rock rhythmic pattern with laid-back, on-the-beat, and pushed timing. While both drummers displayed distinctive onset microtiming patterns for each feel, one showed a tendency to play all the timing conditions early in relation to the metronome (a "pushy" player), while the other played the pushed and on-the-beat feels early but the laid-back feel late (a more "laid-back" player). The tendency to anticipate an external timing reference in synchronization tasks along to click-like metronome stimuli (termed "negative mean asynchrony" [NMA]) has also been found in numerous tapping studies, though it tends to be dampened when more musical stimuli are used (Repp, 2005; Repp & Su, 2013; Wohlschläger & Koch, 2000). In a similar instructed timing performance study with ten drummers playing a standard “back-beat” rock pattern at three different tempi (64, 96, and 148 beats per minute), Danielsen, Waadeland and colleagues (2015) found that all were able to produce onsets in the laid-back and pushed conditions early and late relative to their own average on-the-beat timing, though most showed an anticipatory tendency toward the metronome in the on-the-beat condition as well. They also hypothesized that the drummers would systematically manipulate additional sound features such as intensity (sound pressure level [SPL]) and timbre (spectral centroid [SC]) in order to distinguish the different feels, and they confirmed that they showed a tendency to play laid-back strokes louder (higher SPL) relative to on-the-beat strokes at the medium tempo (96 beats per minute), and with a “darker” timbre (lower SC) as well. These tendencies were explained as strategies to further communicate the intentionally produced early or late onset asynchrony discrepancies. The researchers also found that the overall microrhythmic feel partly depended upon tempo, since onset and sound-feature differences between timing style conditions were greatly diminished at the faster tempo (148 beats per minute), likely due to motoric limitations.

In summary, in groove styles where “close-to-isochronicity” at the subjective reference levels of pulse and subdivision is assumed to be the norm, producing synchrony between instrumental rhythmic layers in a strict “on-the-beat” fashion is only one of several possible approaches to creating a microrhythmic “feel” in the music. Various configurations of either relatively more “laid-back” or “pushed” approaches are just as valid and aesthetically appropriate depending on several contextual factors.

2.3.4 Microrhythm as More Than (Onset) Timing

As we have seen in various sections of this chapter, most studies of microrhythmic variations in groove-based music tend to focus on the onset timing relationships among rhythmic events. In analytical investigations of polyphonic music from commercial recordings, this focus stems in part from the fact that it is easier to measure and compare instruments’ onsets (which can be readily identified in signal waveforms and spectrogram representations with a decent degree of accuracy) than aspects related to timbre, intensity, or duration (which are harder to isolate and quantify with confidence without filtering or distorting signals). However, even in experimental performance studies of groove-based music where better signal sources enable such investigations, scholars have continued to be concerned with onset timing relationships, including the degree to which musicians can closely synchronize onsets to, or systematically anticipate or delay against, a timing reference (Fujii et al., 2011; Kilchenmann & Senn, 2011) or the degree to which musicians displace onsets in swung note pairs (Ellis, 1991; Haas, 2007; Honing & Haas, 2008; Prögler, 1995; see also sections 2.3.2 and 2.3.4). Studies involving perceptual listening tasks with groove-based stimuli have also primarily focused on onsets—an inclination that stems from a preoccupation with testing PD theory’s insistence that it is mainly the onset microtiming profiles of performances that influence the qualitative “feel” of their grooves (Butterfield, 2010; Davies et al., 2013; Frühauf et al., 2013; Kilchenmann & Senn, 2015; Matsushira & Nomura, 2016; Senn et al., 2016, 2017; Skaansar et al., 2019; see also section 2.1.3).

We must remember, however, that Charles Keil, the progenitor of PD theory, in fact considered the notion of participatory discrepancies to encompass not only “processual” aspects related to timing onset asynchronies in performance but also “textural” aspects related to “‘timbre,’ 'sound,' [and] 'tone qualities'" (1987, p. 275). In fact, Keil even explicitly states that instead of the term participatory discrepancies, "one could substitute 'inflection,' 'articulation,' 'creative tensions,' 'relaxed dynamisms,' [or] 'semiconscious or unconscious slightly out of syncnesses [*sic*]'" (1987, p. 275)—that is, features related to both discrepancies in timing location ("syncnesses") and the dynamic shaping of the acoustic envelope of sounds ("inflection," "articulation"). He also quotes from musician Chester "Hoot" Filipiak's description of the rhythm section of a well-known Polish American polka ensemble:

> the articulation of the bass line has also been a crucial variable in defining "beat" and style and sense of tempo. The amplified bass from a concertina dictates a "slower," looser pace and feeling than the bass from the accordion or the still "quicker" bass from a chordovox. The old combination of string bass and piano working together creates still another "time feel." (Hoot in Keil, 1987, p. 277)

Statements like these indicate that both Keil and musicians themselves are aware of the impact that the combination of different instruments' rhythms and differing attack qualities may have upon how their timing is perceived in relation to the rest of the ensemble (see also Brøvig-Hanssen et al., forthcoming, and Johansson et al., in preparation). In short, both "textural" and "processual" features can affect the overall perceived timing "feel" of a groove.

Various other scholars undertaking analytical investigations of groove-based performances have touched upon how the shape of performed rhythmic events may affect our perception of their timing in relation to either other physical rhythms or subjective metrical structures. Danielsen (2006, p. 75) notes that events occurring at a slower metrical level can imply the presence of a faster metrical subdivision level when they are phrased with fast attacks and short durations. Abel (2014, p. 31) observes that the typically sharp attacks of plucked and percussive groove-based rhythm-section instruments such as drums, guitar, and piano tend to "expose the slightest timing discrepancy" unless they are played with "tight" synchrony relative to one another. However, both "tighter" and "looser" timing may be equally appropriate here, depending on stylistic context and individual player preference (see section 2.2.3). Regarding timing and intensity, Iyer (2002, p. 409) demonstrates that the perception of a swing timing feel in jazz not only arises from delaying the onset of the second note in note pairings at a given swing ratio but also "arises from complex variations in timing, intensity, or articulation." Butterfield (2011) further theorizes that jazz musicians may produce various qualitative effects of "anacrusis" by emphasizing the beats of a quarter-note pulse level (the "downbeats") by either delaying offbeat eighth notes at different ratios ("durational inequality") or "slurring [notes] from offbeats to downbeats" and "varying the intensity of the downbeats and offbeats" (either alone or in combination with one another). Iyer (2002) speculates that back-beat–based drum patterns tend to combine with both a slight timing delay and a greater dynamic (intensity) accent. Danielsen, Waadeland and colleagues (2015) found evidence for this this in their instructed timing performance of a "basic rock pattern" in noting that drummers tended to play intentionally laid-back snare strokes louder than on-the-beat strokes. However, Hofmann and colleagues (2017) found that drummers instructed to play three jazz standard tunes with no specific timing feel displayed a variety of onset timings and accentuation-pattern combinations on the ride cymbal. As such, it would seem that, just as with onset asynchrony or swing ratio, the degree to which rhythmic events are dynamically accentuated in groove-based performance depends on stylistic context and personal preference (at the very least).

As we have seen from just these few examples, variations in both onset timing and other sound features are often involved in the production of groove-based music at the microrhythmic level, and the different perceptual effects that these interactions may engender are often not a simple matter but the complex product of several contextual factors. In Chapter 3, I will review further experimental research from both perceptual and performance studies that have investigated perception of timing in various rhythmic contexts more generally, which may serve as a guide to how timing and sound features interact in the experience of groove-based microrhythm.

Chapter 3

Timing and Sound Interaction

Perceived timing involves more than simple acoustic cues for physical onset locations, because the way in which sound events are shaped in terms of duration, intensity, or frequency can modulate how we hear them in time. Furthermore, all of these constituent sound features can interact with one another to various extents in our auditory system. Thus, the distinction between “timing” and “sound” is not absolute, and when we speak of the perceived “timing” of a given microrhythmic pattern (e.g., laid-back, on-the-beat, pushed), we must acknowledge the full range of sound features that are involved in musicians’ articulation, or shaping, of the performed rhythmic events.

To try to unpack the ways in which interactions of timing and sound affect our subjective experience of microrhythm, we can turn to experimental studies that have systematically investigated the role of acoustic features in the perceived timing of sounds. In section 3.1, I will review selected experimental perceptual center (P-center) studies that have shown that various acoustic features influence the perceived rhythmic “moment of occurrence” of sounds. Then, in section 3.2, I will review selected studies that have focused on how different sound factors may affect perception of anisochrony and asynchrony.

3.1 P-Center Studies

Various studies of speech, quasi-musical, and musical rhythms have found that the perceived rhythmic “moment of occurrence” (Morton et al., 1976) or “moment of rhythmic placement” (Wright, 2008) of a sound—commonly referred to as its “perceptual center” (P-center) or “perceived attack time” (PAT)—does not always equal its onset. That is, there are salient differences between a sound’s “physical onset” (the moment when its energy rises above a certain threshold, which can be measured from audio signals; see section 4.3.2), its “perceptual onset” (the subliminal moment when it can be recognized as “beginning”), and its P-center, which is, again, the subjective “reference point for where a sound is placed relative to other sounds in a rhythmic sequence” (London et al., 2019, p. 2088). Most P-center studies find that P-centers tend to be heard at some point between physical onset and maximum energy peak. Therefore, when two or more different sounds are vertically aligned based on their physical onset locations in an isochronous sequence (that is, they are *physically* synchronous) or horizontally spaced out with equal inter-onset-intervals (that is, they are *physically* isochronous), they may not be perceived as *perceptually* synchronous or isochronous, respectively.

It would appear, then, that in rhythmic contexts, the auditory system imposes various degrees of temporal delays when subjective “event detector units” are engaged in the perceptual processing of sound signals (Collins, 2006). The magnitude of these delays, or “P-center shifts,” relative to a sound’s onset location has been shown through experiments to be contingent upon various acoustic features such as attack and total duration in particular, as well as intensity, frequency, and various interaction effects between these features that I will review in section 3.1.1. Then, in section 3.1.2, I will consider the possibility that, rather than being discrete time points, P-centers have a shape or spread that may give further clues as to the width of the “beat-bins” of subjective reference structures (such as metrical pulse).

3.1.1 Effects of Sound Features

P-centers are percepts that can be measured using different psychophysical methods, the most common of which are “adjustment” tasks, where participants are asked to align a probe sound with a target sound in a repetitive isochronous loop, either (a) in phase relative to each other, to the point of perceptual synchrony, or (b) in anti-phase relative to each other, to the point of perceptually isochrony, and “production” tasks, where participants are asked to reproduce a target sound via tapping in ways akin to tasks used in sensorimotor synchronization studies (Repp, 2005; Repp & Su, 2013; for a comprehensive overview of P-center methods, see London et al., 2019; Villing, 2010). Each type of task has its drawbacks.[5] Despite this, both in-phase and anti-phase adjustment tasks tend to produce highly similar P-center location and variability results, indicating that they measure the same percept (London et al., 2019).

Production tasks such as tapping, on the other hand, while displaying directional P-center location shifts for various sound features similar to those of adjustment tasks, often lead to the problem of onset negative mean asynchrony (NMA), where participants tend to anticipate tap onsets relative to the external pacing stimuli (Danielsen et al., 2019; London et al., 2019). Overall, in musical performance, it can be argued that musicians are faced with both of the types of “tasks” used in experimental studies. They need to infer the P-centers of sounds comprising external rhythmic sequences, enacting some sort of in-phase “mental alignment” in order to create a sense of a pulse, and they need to motorically produce their own sounds relative to their subjective pulse as well as the physical external sounds themselves, in synchrony or asynchrony, depending on the microrhythmic feel they intend to create.

In the following subsections, I will review findings from various experiments that used these types of tasks to gauge the effect of various sound features on P-centers.

Attack and Total Duration

The time interval required for a sound’s amplitude energy to reach its maximum value (“peak”) from signal onset (variously referred to as “attack time,” “attack duration,” “rise time,” or simply “attack”) has a particularly strong delaying effect on the P-center. Findings from studies involving musical sounds (whether artificial/synthesized or acoustic instrumental) show a strong consensus that the shorter or “faster” a sound’s attack, the earlier its average P-center tends to be heard relative

---

[5] In-phase alignment tasks, whose purpose is to align two sounds to the point of perceptual synchrony, can create problems related to masking and sonic blend. Anti-phase alignment tasks, on the other hand, which avoid masking problems to some extent, represent relatively less ecological situations for musicians, since “in most instances auditory cues for synchronization are heard non-dichotically” (2019, p. 2089). That is, although musicians may occasionally play rhythms strictly in tandem with one another, they most often encounter situations where they are presented with an external rhythm as a timing reference that marks out a steady stream of events (such as the ticks of a metronome or a linear drum pattern that manifests all the subdivision locations of a meter) and must consequently perceive and synchronize their own produced sounds along with such a reference in an in-phase manner.

to its onset, and, conversely, the longer or slower the attack, the later the P-center (Bechtold & Senn, 2018; Danielsen et al., 2019; Gordon, 1987; Villing, 2010; J. Vos & Rasch, 1981; P. G. Vos et al., 1995; Wright, 2008). A similar, yet weaker, effect has been found in relation to the total duration of sounds, or the time interval between signal onset and offset. Similar to attack, longer durations lead to later P-centers, and shorter durations lead to earlier P-centers (Danielsen et al., 2019; Scott, 1998; Seton, 1989; P. G. Vos et al., 1995). In a recent study by Danielsen and colleagues (2019), an interaction effect was found between attack and total duration: "positively correlated" combinations of attack and total duration (fast attack/short duration, slow attack/long duration) shifted P-centers either earlier or later in time, respectively (causing a "redundancy gain"), whereas in "negatively correlated" combinations (fast attack/long duration, slow attack/short duration), the cumulative shifting effect was either attenuated or canceled out because one factor tends to shift the P-center earlier while the other one shifts it later (causing a "redundancy loss").[6] While these interactions reveal a potential sensory-perceptual interference between the two factors, it still appears that attack duration has a stronger effect on P-center, such that sounds with fast attacks but either short or long total durations do not tend to present very different magnitudes of P-center shift. In other words, longer durations seem to have a stronger P-center delaying effect in rhythmic sounds with slower attacks.

Intensity

Limited systematic research has been conducted regarding effects of P-centers on intensity (loudness) in musical and quasi-musical sounds. However, a few studies seem to indicate that louder sounds tend to lead to earlier P-centers. In one early study, Vos and Rasch (1981) noticed that musical instruments played with greater intensity produced earlier "perceptual onsets"[7] than those played with lesser intensity. Seton (1989) examined the effect of loudness presentation on the P-center of a sawtooth tone and found no effect, though results for mixed-level presentation compared to blocked-level presentation of sounds revealed a dependency on relative loudness that resulted in smaller P-center shifts (that is, shifts that were earlier in time) for greater intensity levels. In a study of various instrumental sounds with differing acoustic profiles, Gordon (1987) found that a saxophone sound played with greater intensity led to an earlier P-center than one played with lesser intensity. Bechtold and Senn (2018) found further evidence that, for saxophone sounds with complex attack profiles, the combination of more articulation (a stronger "tongue attack" technique) and higher dynamics (a louder sound) led to an earlier P-center on average in comparison to that of sounds with weaker articulation and lower dynamics. Villing (2010, p. 46) suggests that P-centers may depend in part on the relative intensity level of preceding and succeeding events in a sequence, and that if the "auditory system continuously adapts to the short term average sound level, then it is easy to imagine that the onset of a quiet sound following a loud sound may be more difficult to detect (or alternatively that it will be detected only at a higher level relative to the sound's peak)."

---

[6] Note that the terms "redundancy gain(s)" and "redundancy loss(es)" as used by Danielsen and colleagues (2019) in a P-center context have meanings different from their original Garnerian context, where they relate to the relative ease/difficulty of performance in classification tasks within a classic Garner interference paradigm (see my discussion of Tekman [2002] in section 3.2.1).

[7] As Danielsen and colleagues (2019, p. 403) note: "Vos and Rasch (1981) defined perceptual onset as 'the moment at which the temporal envelope during the rise portion passes a certain relative threshold amplitude' (p. 325). However, their method of adjustment was to produce perceptually isochronous sequences of sound, which implies that they measured a percept very similar to the P-center of the sounds."

Frequency

A few studies have also focused on the effects of frequency/pitch. On one hand, an early study by Seton (1989) found that high-frequency synthetic tones generally led to later P-centers than did middle-frequency tones. On the other hand, at least two studies have found that musical tones with lower central frequencies lead to later P-centers. Hove and colleagues (2007) found that, in sine tone dyads with onset asynchronies, when a lower-frequency tone followed a high-frequency tone, the P-centers of the compound tones were later. Danielsen and colleagues (2019) found that, in a musical stimulus set, sounds that were darker in timbre generally led to later P-centers, but only when those sounds had slow attacks and were of longer duration. Gordon (1987) also found that P-center was influenced by spectral features, but only for sounds with slower attacks. The latter studies point to further interaction effects between the factors of duration and frequency, where further effects of frequency on P-center may only occur in sounds with slower attack durations. The reason why sounds with lower frequencies lead to later P-centers may be related to bio-acoustic constraints in the auditory pathway, where, at earlier stages of processing, when sounds reach the inner ear (cochlea), a greater "frequency-dependent" delay occurs for lower-frequency sounds than for higher-frequency sounds, because they take more time to complete a wave cycle (Wojtczak et al., 2012, 2017). However, Wojtczak and colleagues (2017, p. 1204) point out that, at higher levels of auditory processing (perceptual and neural), the brain may compensate for frequency-dependent delays to some extent, given that lower-frequency sounds aligned in physical synchrony with higher-frequency sounds of equal duration are often heard as perceptually simultaneous despite "different [physiological] delays in traveling through the auditory periphery." This, then, may explain why Danielsen and colleagues (2019) found that lower-frequency musical sounds only led to later P-centers when the sounds also had slower attack durations, since they would, in effect, further delay P-centers and attenuate the effects of frequency-dependent delay compensation for lower-frequency tones.

Compound sounds with onset asynchronies

While virtually all of the abovementioned P-center studies involved experiments with single sounds, very few studies have been undertaken on the P-centers of compound sounds, or sounds involving combinations of two or more event onsets. Seton (1989) tested a noise sound combined with a delayed sine tone at various degrees of onset asynchrony and found that, when participants were asked to attend to the tone component (rather than the noise), P-centers were delayed as onset magnitude increased. This was partially interpreted as evidence for streaming effects on P-center (Bregman, 1990)—that is, the extreme difference in frequency and spectrum between noise and tone sounds allowed participants to parse the compound sounds into separate streams. In a study involving sine tones, Hove and colleagues (2007) found that, in a tapping task, participants' taps were drawn toward the lagging (second) tone's onsets in both "low late" and "high early" asynchronous conditions. In a follow-up anti-phase adjustment task, P-center locations were also estimated to be closer to the later onset, regardless of frequency and tone order (though both taps and P-centers were even later when the low tone followed the high tone). The researchers interpreted the tapping task's findings as evidence that participants synchronized their taps to some point other than simply the onset of either the first or second tone—namely, the P-centers of their compound sound—because "when the [asynchronous] tone P-centers occur later than the P-centers of chords with simultaneous tone onsets, taps must be delayed so that their P-centers are aligned with the [compound] tone P-centers" (2007, p. 706). From this limited research, then, it would appear that P-centers of compound sounds with multiple onset asynchronies may be delayed toward the onsets of later constituent events in general, and that the degree of P-center shift is further affected by frequency differences between leading and following events.

3.1.2 P-Center "Shapes" as Indicative of "Beat-Bins" in Musical Contexts

P-centers were originally conceived of as discrete time points occurring sometime after the physical onset of a sound, thought to be best represented by the central tendency of the mean of P-center measurements (Villing, 2010). However, various scholars have noticed that participants often display a range of values that seem "equally correct" when they are aligning sounds, indicating that mean P-center locations fluctuate to a certain degree depending on the characteristics of the specific sounds (Danielsen et al., 2019; Gordon, 1987; London et al., 2019; Wright, 2008). As such, P-centers have also been conceived as having a temporal duration or width that can be represented as a probability distribution (Wright, 2008), where sharp sounds with fast attacks and short durations tend to produce narrow P-center shapes, and sounds with slower attacks and longer durations tend to produce wider shapes. Danielsen and colleagues (2019) further proposed that the extent of the variability (or standard deviation) around estimated mean P-center locations in adjustment tasks for any given sound may provide a clue as to the width of a sound's "beat-bin" in a rhythmic musical context, or the perceived temporal window in which sounds are perceived as occurring within their categorical boundaries. However, these researchers also note that P-centers are psychophysical phenomena, and that "while a P-center is produced by a sonic event, the event is not a point or a bin in itself"—instead, P-centers and beat-bins may be viewed as a product of sounds' "affordances for action" in musical contexts (p. 416).

As such, in a musical listening context, we assume that when we are presented with a regular rhythmic stream of sounds that establishes a subjective pulse within us, the beat-bin of this pulse can be either wide or narrow, depending on the shape of the sounds of the pulse-carrying rhythmic layer(s) of the music. If the sounds have a fast attack and a short duration, then they afford a narrower beat-bin against which additional external rhythmic events will be heard as asynchronous if their own P-centers are heard as clearly occurring outside of the reference's bounds. On the other hand, if the sounds have a slower attack and a longer duration, then they would afford the possibility of a pulse with a wider beat-bin. Accordingly, the same additional external events could then be perceived as occurring "on the beat" (or, rather, "within the beat"). A similar process may apply in a production context (either musical performance or a tapping task) as well, since before musicians can produce sounds against an external rhythm, they too must first establish a sense of pulse (London, 2012; Repp & Su, 2013). According to dynamic attending interpretations of the beat-bin (Danielsen, Waadeland et al., 2015; Danielsen et al., 2019; see also section 2.3.1), we can broaden our attentional focus at any given stage during the process of entraining to a pulse, adjusting the beat-bin width to allow multiple events to fall within it and still be considered part of the beat—which might, in turn, further enhance a steady sense of pulse—or we can engender a more ambiguous situation where vertically asynchronous or horizontally non-isochroous events chafe against the beat more conspicuously, creating a tension in terms of the event being either slightly early or late.

While the perceptual process that might allow for the dynamic or flexible adjustment of subjective beat-bins in a repetitive/rhythmic context is relatively unproblematic in more “passive” listening contexts (see Danielsen, 2010), “active” production contexts such as musical ensemble performance present a more complex case, because they involve not only perceptual judgments but also physical/bodily interaction with external rhythms. For example, if musicians want to play in a laid-back or pushed fashion—that is, to synchronize in a systematically delayed/anticipated “phase-displaced” manner relative to the sounds of a regular external timing reference as in the experiments of this thesis’ papers—they must maintain a sense of pulse with a more or less stable beat-bin width based on the external timing reference in order to consistently coordinate their actions so that their own produced sounds will be heard as falling just outside or on the edge of the beat-bins afforded by that reference. If they do not do so, and the beat-bins are constantly adapted based on the resultant compound sounds of the musicians’ rhythms and those of the external timing reference, this would eventually create a recursive or feedback situation where the musicians would have to constantly increase the magnitude of their anticipation/delay so that events would remain just beyond the bounds of the ever-widening beat-bins at every subsequent expected beat location. In situations where the timing reference is flexible, one could theoretically imagine that if one performer kept constantly trying to anticipate/delay events and another attempted to compensate for the increasing shifts, then adaptive beat-bins based on the compound sounds of the entire ensemble would eventually generate constant increases and decreases in tempo, which is not aesthetically appropriate in most groove-based contexts (however, it is not uncommon for less practiced ensembles to speed up when attempting to play with a “pushed” feel or slow down when attempting a “laid-back” feel, unless one instrumentalist is able to maintain a steady tempo).

In those groove-based contexts, though, the tempo of a timing reference rhythm usually remains more or less isochronous, by way of a metronome, a prerecorded instrumental backing track, or a highly steady musician as the timekeeper. It would appear, then, that musicians are able to selectively attend to the sounds of specific rhythmic layers in an ensemble as their primary reference (the drummer’s hi-hat cymbal, for example, or the combination of drums and bass), “locking into” the P-centers and beat-bins these layers afford while maintaining the production of their own rhythms via motor-synchronization processes that are somewhat independent from the perceptual timekeeping mechanisms. Relatedly, while P-center variability has been shown to be sensitive to sound features in adjustment tasks, tapping IOI variability tends to be highly stable and less sensitive to sound stimuli of differing shapes in production tasks (Danielsen et al., 2019; London et al., 2019). Therefore, in performance, it seems that once musicians are locked into the beat-bins afforded by the external rhythms, they are able to produce repetitive rhythmic events either on-the-beat or slightly off-the-beat relative to those locations without being overly influenced by the P-center shifting effects of their own performed sounds. However, just as more automatic (i.e., non-conscious) mechanisms for period and phase correction are engaged that allow musicians or tappers to maintain relatively stable IOI timing when random timing variations are introduced to rhythms in sensorimotor synchronization tasks (Repp, 2005; Repp & Su, 2013), similar adaptive “beat-bin correction” mechanisms may be engaged when musicians find themselves pushing or laying back too much against a timing reference. And just as training and practice allow musicians to produce in-phase synchrony with more accuracy (less NMA) and stability (less IOI variability) (Repp, 2005; Repp & Su, 2013), those who are more practiced in controlled microrhythmic variation might also be capable of greater accuracy and stability when synchronizing in a slightly off-phase manner. This may potentially reflect their ability to maintain consistently narrow subjective beat-bins in the face of additional phase-displaced rhythms, but it may also be the case that a musician does not want to play consistently off the beat throughout a given groove-based pattern but instead seeks a looser style with a certain degree of ambiguity (see also section 2.3). As such, it might be that aesthetic-stylistic and individual preference factors also influence how narrow or wide a musician might "tune" their beat-bin.

3.2 Anisochrony and Asynchrony Studies

Several studies have investigated the effects of *anisochrony*[8] and *asynchrony* in experimental contexts. While both constitute types of "rhythmic irregularity," the former considers horizontal timing relationships and the latter considers vertical relationships between events. Many studies (though not all) have focused on finding perceptual thresholds regarding how large an onset timing "deviation" is required before an event is perceived as either horizontally non-isochronous or vertically asynchronous relative to other events.

In section 3.2.1, I will focus on studies that have investigated interaction effects between onset timing and other sound features, and more precisely the interaction between onset anisochrony and asynchrony and other sound features such as attack/total duration, intensity, and frequency. In section 3.2.2, I will turn to the relationship between sensitivities to different types of onset differences at various inter-onset-interval (IOI) rates.

3.2.1 Effects of Sound Features

Similar to findings from P-center studies (see section 3.1), anisochrony and asynchrony studies may provide additional clues as to the boundaries of subjective reference structures (such as metrical pulse and subdivisions) and the ways in which changes in duration, intensity, or frequency increase or decrease the perceptual salience of produced onset asynchronies or anisochronies in performance.

Methods regarding onset manipulation, stimuli, and task type[9] vary widely among these studies. Anisochrony perception studies tend to involve monophonic stimuli and tasks where participants are asked to subjectively determine the threshold at which they can hear a "difference" between physically isochronous baseline sequences and altered versions of such sequences. These typically aim to "identify psychophysical law[s] for time perception" (Ehrlé & Samson, 2005, p. 134) and discern just-noticeable difference (JND) thresholds across a range of tempi. Typical stimuli onset manipulations may perhaps be broadly categorized as either "local" or "global."[10]

---

[8] In microrhythmic studies of groove-based musics, the term "non-isochronous" is more commonly used for rhythms displaying systematic "deviations" from a normative isochronous grid (see section 2.3.2). However, perceptual studies often use the synonymous term "anisochronous" for such types of rhythms, as well as "anisochrony" to label the phenomenon in general. I shall use these various terms interchangeably, as they relate to the same concept—that of timing "deviations" from isochrony.

[9] In addition to tapping tasks and adjustment tasks, these studies also commonly apply various types of "discrimination" tasks. Here, participants tend to be instructed to make judgments between isochronous/synchronous stimuli and altered stimuli with either gradually increasing or decreasing magnitudes ("adaptive" methods) or fixed magnitudes of event onset shifts ("constant" methods), amongst others (Ehrlé & Samson, 2005).

[10] In local anisochrony methods, the onset of a single event in a rhythmic sequence is shifted early or late to varying degrees whereas the onsets of all other events either remain fixed to an isochronous grid ("displacement") or arrive correspondingly early or late after the displaced events occur ("lengthening/shortening"; see Ehrlé & Samson, 2005; Friberg & Sundberg, 1995). Global methods, on the other hand, involve shifts of multiple event onsets in a sequence, either: displacement of every even event while odd events remain fixed to an isochronous grid ("duple rhythms" or

Asynchrony perception studies, on the other hand, necessarily involve polyphonic stimuli, where at least two sound events are presented as "pairs" (or "dyads") with varying degrees of onset magnitude difference between them. They often involve tasks where participants are asked to either subjectively determine whether pairs are perceived as synchronous or not (from which average "asynchrony detection" JND thresholds are often calculated) or identify which sound is heard as arriving first/second in the pair (from which "temporal order" thresholds are calculated).[11]

In the following, I will review select findings from anisochrony and asynchrony studies that I have found to be relevant to the experimental groove-based performance context explored in this thesis's papers.

Attack and Total Duration

It would appear that no studies investigating asynchrony/anisochrony thresholds have tested the effects of attack duration. In most anisochrony (and asynchrony) studies, in fact, the sounds utilized have identical or very similar attacks and total durations. Regarding total duration, however, it would appear that at least Erhlé and Samson (2005, experiment 1c and 1d) systematically tested the effects of varying total durations (between 10 and 300 milliseconds) on JNDs of anisochrony in sequences with IOIs between 300 and 900 milliseconds using a local displacement method and a discrimination task. They found no significant effects of duration in any of the IOI ranges. However, since only one sound was utilized rather than two or more different sounds with varying acoustic profiles, this result is perhaps not surprising, since they would all have the same P-center (see section 3.1). Regarding JNDs of duration between two different sounds irrespective of onset anisochrony/asynchrony, Johnson and colleagues (1987) found that, at least for simple psychoacoustic stimuli (sine tones), relative JNDs can range between 9 and 16 percent of stimulus duration depending on the discrimination methods used (average of 12.5 percent across various studies).

Intensity

Regarding asynchrony, analyses of solo classical piano performance have found that pianists often emphasize a melody tone (that is, increase its perceptual salience) by playing its onset earlier relative to accompaniment chord tones (Palmer, 1996). This has been partially attributed to a "velocity artifact," where, in order to achieve earlier timing, pianists depress the key of the melody tone more rapidly than the key of the lower tone, resulting in a louder sound (Goebl, 2001). Using

---

"cyclic displacement", e.g. Frane & Shams, 2017; ten Hoopen et al., 1994); or lengthening/shortening of IOIs of all events in a sequence to varying degrees ("systematic" lengthening/shortening; e.g., Merker & Madison, 2002). Studies that specifically investigate musical rhythms tend to utilize melodic stimuli and a more performance-stylistic-oriented version of the systematic onset manipulation method, where certain specific events deemed important to subjective metrical and musical structures are manipulated based on prototypical onset timing profiles of real performances ("expressive variation"; e.g., E. F. Clarke, 1989; Repp, 1998). An altogether different type of method involves gradual "tempo" changes, where all event IOIs in a sequence are gradually increased/decreased after a certain number of events are presented (e.g., Drake & Botte, 1993).

[11] Synchronous or asynchronous pairs can be presented as (1) separate stimuli ("lone asynchrony"; e.g., Goebl & Parncutt, 2002), (2) rhythmic sequences where a single asynchronous pair is introduced in a sequence otherwise comprised of synchronous pairs ("local asynchrony"; e.g., Wojtczak et al., 2017), or (3) separate isochronous sequences exclusively comprised of either synchronous or asynchronous pairs ("global asynchrony"; e.g., Hove et al., 2007). Relatedly, a few sensorimotor synchronization studies (SMS) have also investigated how asynchronous tone sequences affect the production of tapping (Hove et al., 2007, 2014). Here, participants are asked to tap in synchrony to various altered tone pairs using mainly global onset manipulation methods, where the locations of participants' produced taps relative to asynchronous sequences are interpreted as providing insights into their perceived timing.

a lone asynchrony method and a discrimination task, Goebl and Parncutt (2002, Experiment III, "piano" conditions) tested the effect of relative intensity (± 0, 10, and 20 MIDI velocity units) on onset asynchrony detection in various piano tone pairs with different pitches (high vs. low range: 494 to 942 Hz) and onset asynchrony magnitudes of 27 and 54 milliseconds. They found that, regardless of pitch, at 27 milliseconds of asynchrony, it was generally more difficult to detect an asynchrony when the leading tone in the pair was louder ("melody lead" conditions, i.e., early + loud pairs: between ca. 20 to 30 percent detection rate) than when the lagging tone was louder (late + loud pairs: between ca. 90 to 100 percent). When the magnitude of asynchrony was increased to ±54 milliseconds, detectability was enhanced in both earlier and later directions, though it remained more difficult to detect an asynchrony between tones in early + loud pairs (ca. 40 percent) than in late + loud pairs (ca. 100 percent). The overall greater difficulty of detecting early + loud pairs was explained as the result of either a psychoacoustic forward-masking effect, where leading louder tones decreased the salience of lagging softer tones for a certain period of time, or a decreased sensitivity to asynchrony due to a greater familiarity with the early and loud tone combinations typically found in classical piano performance.

As for anisochrony, Tekman (1997, 2001) initially found that IOIs that precede higher-intensity tones in tone sequences tend to be perceived as being longer (or, alternatively, as arriving later than expected relative to previous tones). In a subsequent study, Tekman (2002) systematically tested the interaction effects of timing and intensity using a lengthening method and a classification task within a so-called "Garner interference" paradigm[12] (see Algom & Fitousi, 2016; Melara & Marks, 1990). Participants were asked to selectively attend—that is, focus their attention exclusively—to either timing or intensity variations of locally manipulated tones in various different isochronous sequences (0 or +25 ms lengthening, 0 or +3 dB). His findings showed that, while it was easiest in general to detect intensity variations, it was easier to correctly detect timing variations in "positively correlated" sequences where manipulated tones both arrived later and were louder than preceding tones. The results were more generally regarded as evidence that timing and intensity variations in tone sequences are not processed independently in the auditory system but rather interact at some perceptual level. This was also interpreted as offering a perceptual explanation for why delaying the timing of a note (an "agogic accent") is often perceived as creating a similar effect to that of increasing its intensity (a "dynamic accent"), where both devices are found to "accentuate" musical structures in classical performance. Also, Clarke (1988) and Palmer (1993) have found that accented beats tend to be lengthened in performance. Perceived duration might thus be altered as a consequence of a dynamic accent both prior to and following the accent.

Regarding thresholds of intensity irrespective of onset anisochrony/asynchrony, it has been found, for sinusoid tones, that intensity discrimination is generally dependent not on signal frequency but on sensation level or perceived loudness of stimuli presentation. The louder the presentation level, the lower the JND of intensity of a given tone (Jesteadt et al., 1977; Johnson et al., 1987). At louder presentation levels of 70–80 dB (which are not uncommon in live music performance contexts), typical reported JNDs tended to be around 0.4–0.8 dB.

---

[12] Tekman (2002) explains that in a Garner paradigm, if two dimensions of stimuli are processed in independent perceptual channels, then variations in one attended dimension (e.g., timing) should not affect the classifications of stimuli in the other, unattended dimension (e.g., intensity). Very simply put, if uncorrelated variations in the unattended dimension cause interference in a classification task (i.e., classification either becomes easier ["redundancy gain"], or more difficult ["redundancy loss"]), it indicates that there is "cross talk" between the two channels, and that they therefore must interact at some perceptual level.

Frequency

Regarding asynchrony, Hove and colleagues (2007) used a global method and tapping task involving sine tones with higher and lower frequencies (1400 Hz and 350 Hz) and three fixed magnitudes of onset displacement between the tones (0, 25, and 50 milliseconds) at an IOI rate of 500 milliseconds. They found that, compared to conditions with synchronous pairs, taps were more drawn toward the lagging tone when it was lower in frequency ("low late" condition) than when it was high. As mentioned in section 3.1.1, they interpreted these results as evidence that participants synchronized their taps to the P-centers of the tone combinations, which were found to be later in "low late" combinations. Repp and Su (2013), alternatively, suggest that if participants regarded the onset of the leading tones as targets for tapping synchronization (which was not specifically instructed, however), then the second late tones acted as "distractors," exerting an "attracting effect" that has been found in other tapping experiments by Repp (2003, 2004, 2006). In a follow-up study with a variant of the local displacement method and a tapping task involving piano tones (466.2 Hz and 196.0 Hz) and 50 millisecond onset displacements, Hove and colleagues (2014) reproduced their results in the sense that taps were drawn more toward the lower-frequency tones in "low early" pairs than toward the higher-frequency tones in "high early" pairs of the sequences. Results from a concomitant electroencephalography (EEG) experiment found that neural mismatched negativity (MMN)[13] responses were higher for "low early" compared to "high early" pairs, which led the researchers to hypothesize that the auditory system is more sensitive to the timing of low-pitched tones (that is, it presents a "superior neural encoding of lower-frequency sounds").

However, in a later similar study with synthesized harmonic tones (pitch: 196 Hz and 466 Hz, IOI: 500 milliseconds), Wojtczak and colleagues (2017, "tone pairs" experiment) more systematically tested all possible combinations of early/late and low/high tones. Using a discrimination task with a local asynchrony method, they found that listeners were more readily able to detect an asynchrony in sequences with low-leading pairs ("low early" and "high late") than in sequences with high-leading pairs ("high early" and "low late") by a small margin (mean onset delay thresholds ca. 3 vs. 5 milliseconds, respectively). Wojtczak and colleagues (2017, p. 1203) also found in an EEG experiment that low-leading pairs in general led to greater MMN responses compared to high-leading pairs, which led them to further suggest that there is no superior neural time encoding for low over high frequencies but rather that there exists a "perceptual and neural asymmetry in the processing of temporal asynchronies between tones of different frequencies" related to the greater physical frequency-dependent delays of lower-frequency tones in the cochlea (see section 3.1.1). As such, they suggested that the reason why Hove and colleagues (2007; 2012) found that taps tend to be more drawn toward lower tones in asynchronous dyads (whether they led or lagged) may be because low-lead tone asynchronies are perceptually more salient than high-lead tone asynchronies.

Regarding anisochrony, Wojtczak and colleagues (2017, "single tones" experiment) also used a displacement method and a discrimination task with the same stimuli and tempo mentioned above. They found no effect of either tone frequency or order: an average of approximately 20 milliseconds of onset displacement was required for participants to detect the presence of a rhythmic irregularity (anisochrony), whether the single tone arrived either early or late and regardless of

---

[13] Hove and colleagues (2014, p. 10383) state: "MMN is generated primarily in the auditory cortex and represents the brain's automatic detection of unexpected changes in a stream of sounds."

whether it was high or low in frequency. The threshold was almost an order of magnitude higher compared to their asynchrony experiment above, indicating that at least at an IOI rate of 500 milliseconds (i.e., 120 beats per minute), it is more difficult to detect anisochrony in monophonic sequences than it is to detect a local asynchrony between two tones in a polyphonic sequence.

3.2.2 Just-Noticeable Differences in Relation to IOI Rates

Several studies involving monotonic rhythms in non-musical contexts have sought to determine whether sensitivity to onset anisochrony in terms of JND thresholds increases proportionally with IOI, and whether it conforms to Weber's law[14] at all ranges. Though methodological factors (in addition to onset manipulation types) vary widely, making direct comparisons slightly tenuous, literature reviews by Friberg and Sundberg (1995) and Ehrlé and Samson (2005) seem to indicate that, at medium to slow IOI rates (ca. 200–1400 milliseconds), relative JND onset thresholds tend to be constant, in agreement with Weber's law (maximum values: local = ca. 7 percent; global = ca. 6 percent; tempo = ca. 3 percent). At faster IOIs (below ca. 200 milliseconds), however, relative JND thresholds tend to increase steeply as IOI decreases, mostly in disagreement with Weber's law (maximum values: local = ca. 17 percent; global = ca. 12 percent; tempo = ca. 7 percent). The change from constant relative JNDs in the mid-slow IOI range to a different JND relationship for faster IOIs has been interpreted by various researchers as evidence of the potential existence of separate perceptual strategies for detecting differences—either more "memory-based" mechanisms for faster IOIs, where participants detect onset differences in altered non-isochronous stimuli based on holistic comparisons with previously memorized versions of isochronous stimuli, or "beat-based" timing mechanisms for slower IOIs, where an internal subjective beat scheme serves as a reference against which onset differences are judged (for an overview, see Ehrlé & Samson, 2005). Regardless of which mechanism might apply in either range, the general trend found in these studies seems to indicate that, in order to detect an anisochrony in simple rhythmic sequences, the slower the IOI rate, the greater the absolute onset difference needed and the lower the relative JND thresholds, and the faster the IOI rate, the lower the absolute onset difference needed and the higher the relative JND thresholds. Also, it would appear to be easier to detect anisochrony if multiple event onsets are shifted in a global fashion, rather than if only single events are shifted in a local fashion, and easiest of all to detect it when the whole rhythm gradually speeds up or slows down in tempo. The lower thresholds to both tempo and global compared to local onset manipulations may be because, in the former case, they affect "a greater number of intervals and consequently more information is available to the perception" (Friberg & Sundberg, 1995, p. 2527).

As for sensitivity to anisochrony in musical contexts, Clarke (1989) found that in simple piano melodies based on classical repertoires with baseline IOIs of 400 milliseconds, when compared to a "metronomic" baseline condition where all notes were positioned isochronously, if key structural notes were lengthened, participants required around 20 milliseconds of lengthening to detect the presence of a rhythmic irregularity above 50 percent chance level. If one manipulated baseline sequences with typical expressive onset profiles based on real performances where note IOIs varied systematically ("rubato" conditions), however, participants required at least 50 milliseconds of lengthening of the same altered tones to detect irregularity. This is because, in many classical music performance traditions, in which isochrony at subjective metrical levels of both

---

[14] Weber's law "supposes that the just-noticeable difference between two stimulus durations is a constant proportion of the shorter of the two values" (Ehrlé & Samson, 2005, p. 131).

pulse and subdivision is not the norm, familiar listeners tend to already "tolerate" or expect larger onset deviations from isochrony, thus decreasing their sensitivity to further onset variations. Such "obligatory" expectations of anisochrony are, in fact, considered aesthetically appropriate and can range up to 30 percent of IOI in classical piano performances (Repp, 1998a, 1998b), where they are generally viewed as expressive devices that "accent," "highlight," or "bring out" (i.e., increase the perceptual salience of) key structural elements of the music, such as meter, melody, and harmony, among others.

Very few studies have investigated sensitivity to anisochrony in groove-based musical contexts. Regarding IOI rates resembling those of faster metrical subdivisions, which are often non-isochronous to subtle degrees in certain groove styles (see section 2.3.2), investigations have mostly focused on "swing" (cyclical displacement) detection thresholds. Friberg and Sundberg (1994) found that in a synthesized jazz melody with an eighth-note IOI of 176 milliseconds (170 bpm), participants required about a 35 millisecond difference between on-beat and off-beat note pairs to detect swing (1.2:1 ratio). Frane and Shams (2017) found that for sixteenth-note hi-hat rhythms set against a basic kick and snare drum back-beat pattern with IOIs between 107 and 188 milliseconds (80–140 bpm), the magnitude of onset displacement needed to detect swing was 8–14 milliseconds for expert drummer participants (ca. 1.2:1–1.8:1 ratio) and 18–31 milliseconds for non-drummers (ca. 1.1:1–1.3:1 ratio), where the higher values were obtained for rhythms with more swung events (or a "high swing density"). As for sensitivity to anisochrony at pulse IOI levels, it would appear that no studies have specifically focused on groove-based contexts, since they tend to be quite isochronous at this level (see section 2.2.2). Even so, Madison and Merker (2002) investigated thresholds in a simple sequence with percussive clave-like sounds using a systematic lengthening/shortening onset-manipulation method that evoked natural pulse-level IOI fluctuations in non-metronomic groove-based contexts to some extent. Focusing on IOI rates of 630–570 milliseconds (95–104 bpm), the researchers asked participants whether they heard a "rhythmic irregularity" and whether there was "a pulse in [the] sound sequence, such that [they] would be able to beat along with it" (Madison & Merker, 2002, p. 203). They found that the average threshold for pulse attribution was much higher (8.6 percent relative JND = ca. 49–54 milliseconds) than that for simply detecting a horizontal rhythmic irregularity (3.5 percent relative JND = ca. 20–22 milliseconds). As such, their findings suggest that even when a simple rhythm contains no musical "structural" information about meter or pattern, one may subjectively maintain a steady sense of pulse that is very flexible about or tolerant of physical anisochronies. In groove-based contexts that often supply clear and stable metrical information (for example, by drummers marking out all pulse and subdivision beats with different instruments), it stands to reason that even greater IOI variability might be possible without disturbing one's sense of a pulse.

Note that several of the abovementioned studies have found that the degree of musical training or instrument specialization can affect sensitivity to onset anisochrony. Ehrlé and Samson (2005) found that classical percussionists presented lower JNDs of single displacement in monotonic tone sequences over a wide IOI range of 80–1400 milliseconds compared to non-percussionist musicians and non-musicians alike. Frane and Shams (2017) found that drummers displayed lower overall JNDs of swing detection compared to non-drummers. Merker and Madison (2002) found that participants with more musical training were more likely to detect a "rhythmic irregularity" (but the threshold for disturbing the "sense of regular pulse" was the same for more vs. less musically trained participants). Overall, most explanations of why participants with musical training tend to perceive anisochrony more readily are based on the assumption that they are more attuned to a wider range of rhythmic categories afforded by subjective metrical structures (whether pulse or subdivision) and, as such, more sensitive to onset irregularities when they break with subjective expectations of normative onset timing profiles, regardless of whether they are relatively more (groove-based) or less (classical) isochronous (Frane, 2017; Madison & Merker, 2002; Yee et al., 1994).

Regarding sensitivity to onset asynchrony, on the other hand, it would seem that thresholds are quite dependent on sound stimuli type and context (non-musical vs. musical). For simpler psychoacoustic stimuli in non-musical contexts, very small onset asynchronies of around 1–2 milliseconds are easily detected (Hirsh, 1959; Zera & Green, 1993).[15] However, in classical performance contexts involving more complex musical sounds with overlapping spectra such as piano (Goebl & Parncutt, 2003), string/reed/recorder instruments (Rasch, 1979), or vocal ensemble (D'Amario et al., 2019), greater onset magnitudes of 30–50 milliseconds or more appear to be required to detect the presence of asynchrony. In groove-based performance contexts, once again, no systematic studies have been conducted. However, although their participants were not specifically tasked with discerning onset asynchronies in groove patterns with exaggerated onset manipulations, "groove rating" studies (see section 2.1.3) suggest that onset asynchronies between rhythm section instruments with sharper defined attacks, such as drums and bass, can be heard with magnitudes as low as 10–20 milliseconds, as they often lead to significant differences in ratings of "urge to move" and/or "pleasure" compared to fully quantized patterns (Hofmann et. al, 2017; Frühauf et al., 2012; Skaansar et al., 2017).

Summing up, findings from P-center and anisochrony/asynchrony studies suggest that the perceived timing of a rhythmic event might be influenced by sound features such as duration, intensity, and frequency. In compound sounds, this impact is further complicated by the way in which such features create different degrees of perceived vertical synchrony/asynchrony and/or isochrony/anisochrony when heard against other physical events or subjective reference structures. In performance, then, it is likely that musicians align their own physical rhythms in such a fashion that the P-centers of the resultant compounds between their own sounds and those of external timing reference rhythms correspond to the overall intended microrhythmic feel they wish to convey. While it is clear that this can be effectively achieved through the application of greater magnitudes of earlier/later onset asynchrony relative to a timing reference's sounds, if the instrumental sounds are further shaped via faster/slower attack, shorter/longer duration, greater/lesser intensity, or higher/lower frequency, the perceived earliness or lateness of the resultant compound sounds can be further diminished or enhanced.

---

[15] Onset asynchrony sensitivity also concerns "temporal order" thresholds, which is to correctly judge which event comes first. Early studies by Hirsh (1959) found that around 20 milliseconds was needed to correctly detect temporal order in combinations of simple non-musical stimuli (clicks and tones), largely independent of their frequency. Pastore and colleagues (1982) further found that the shorter the duration of asynchronous sinusoidal sounds, the smaller the threshold—as small as 5 milliseconds if their durations were 10 milliseconds. In groove-based contexts, Butterfield (2010) more recently found that participants were not able to detect whether bass or drums led or lagged in relation to one another when onset asynchronies were below 30 milliseconds, both in simple synthesized jazz patterns and in commercial recordings.

Chapter 4

Methods

This chapter provides an overview of the methods I used in the papers included in this thesis. In section 4.1, I describe the TIME project performance experiments that yielded the data analyzed in all the papers. In 4.2, I give details about the recording of the audio data. In 4.3, I present an overview of the audio processing and analysis pathway. In 4.4, I describe the statistical methods used to analyze the data.

4.1 Experiment Data Sets

As mentioned in chapter 1, the work presented in this thesis was conducted as a part of the TIME project at the RITMO Centre (University of Oslo) during the first three years of the project's five-year time span. In collaboration with several TIME project members, I designed a series of performance experiments for guitarists, bassists, and drummers to investigate whether expert instrumentalists systematically manipulate acoustic features other than stroke onset in order to express different intended timing styles (or "microrhythmic feels") in groove-based performance. The basic design of each experiment contained three instruction categories (independent variables) with two or three conditions (levels):

1. Pattern type
    a. Two conditions
        i. Pattern 1 ("simple")
            1. Guitar: alternating two-chord "back-beat" figure (beats 2 + 4)
            2. Bass: alternating single-note "down-beat" figure (beats 1 + 3)
            3. Drums: basic "back-beat" figure (hi-hat: steady eighth-note stream, kick drum: beats 1 + 3, snare drum: beats 2 + 4)
        ii. Pattern 2 ("complex")
            1. Guitar: three-chord syncopated version of "back-beat" figure (beats 2 + 2-and + 4)
            2. Bass: "down-beat" figure with preceding pick-up notes (beats 1+ 2-and + 3 + 4-and)
            3. Drums: "back-beat" figure with additional off-beat snare and kick drum strokes (hi-hat: steady eighth-note stream, kick drum: beats 1 + 3-and + 4-and, snare drum: beats 2 + 2-and + 4)
2. Timing reference
    a. Two conditions
        i. "Metronome": comprised of woodblock sounds
        ii. "Instrumental backing track": comprised of guitar/bass/drum sounds

3. Timing style
    a. Three conditions
        i. “Laid-back”: play behind the beat of the timing reference(s)
        ii. “Pushed”: play ahead of the beat of the timing reference(s)
        iii. “On-the-beat”: play as synchronized as possible with the timing reference(s)

The experiments had a “within-subjects” design (see section 4.4.1) whereby each participant performed all the instructed conditions, thus maximizing data collection while minimizing cost and time per experiment. (For further details of the experiment conditions and procedures, see papers I and II).

For the experiments, I recruited, recorded and collected audio data (section 4.2) from sixty-three musicians (twenty-one guitarists, twenty-one bassists and twenty-two drummers) with experience in groove-based performance. From that data, several temporal and sound features were extracted (section 4.3), then analyzed further via statistical methods (section 4.4) and reported on in papers I and II (in collaboration with Kristian Nymoen, Olivier Lartillot, and Anne Danielsen) and paper III (with George Sioros and Anne Danielsen), in chronological order.

In paper I, we reported on a portion of the guitar and bass experiment data that was limited to performances of the “simple” pattern (1) along with the “metronome” timing reference in all three timing style conditions (laid-back, on-the-beat, pushed). We did this because, within the scope of the journal article format, we wanted to establish from the outset the methodology for the audio data analysis, as well as present an in-depth description and discussion of the guitar and bass results—both pioneering aspects of performance-timing research. For these purposes, the analysis of a single pattern and timing reference condition was considered sufficient. An analysis of the performances from the remaining conditions (pattern 2: “complex,” and timing reference: “instrumental backing track”) will be reported in a future publication. I will, however, briefly comment on some predicted results of that further analysis in section 6.1.3.

In paper II, we focused on the drummers’ experiment data from performances in all instructed pattern, timing reference, and timing style conditions. As the core experiment design and methodology for the audio data analysis had already been developed and reported on in paper I, we were able to encompass all of the instructed conditions within the scope of the journal article format. In paper III, we further investigated the drummers’ experiment data by limiting the focus once again to performances of the “simple” pattern (1) along with the “metronome” timing reference. This was because we wanted to focus on the development of a novel hierarchical clustering and phylogenetic visualization methodology and had go into selected data in much more depth to demonstrate the method’s feasibility and validity.

In addition to capturing audio data, I also simultaneously collected motion-capture data from all of the participants using an optical tracking system (Qualysis AB, Gothenburg, Sweden). While these data will likely shed further light on the audio data, they will not be reported on here, for two main reasons: first, motion-capture methods and analyses are beyond the scope of this thesis, and second, the original purpose of such data collection was to inform later analyses in the TIME project timeline via collaborative publications led by other TIME project researchers (ideally, after the audio data analyses have been reported upon). At the present time, those researchers are developing novel motion-profile categorization and statistical methods that are appropriate and precise enough to analyze the data from this thesis’s experiments, as well as the motion-capture data of other TIME experiments. Both the audio and motion capture materials I collected during the course of this thesis will be published and made publically available upon completion of the TIME project.

4.2 Audio Recording

For the performance experiments reported on in papers I–III, we conducted audio recordings of the guitarists, bassists, and drummers in the spring of 2018 at the Motion Capture Laboratory, Department of Musicology, University of Oslo, Norway. We chose to provide the same instrument setup for all participants in order to ensure that any differences in produced temporal or sound features would be primarily the result of intentional participant strategies rather than the result of differences in instrument design and construction (certain instruments allow for greater range of feature manipulation than others). In addition, controlling for the instrument set-up facilitates the automatic detection and calculation of stroke features in the analysis stage, because the range of adjustable global parameters and thresholds is decreased overall—that is, the variation in sound signals produced from the same instrument played by twenty participants is smaller than the variation implied by twenty different signal types.

Instrument signals were sent into an RME BabyFace Pro sound card (Audio AG, Heimhausen, Germany) and recorded in the audio software Reaper 5.77 (Cockos, Inc., San Francisco, CA) at a sampling frequency of 44.1 kHz and 24-bit resolution. We supplied monitoring to the participants as follows. For the guitarists and bassists, amplified instrument signals were routed to an analog mixer (Yamaha MG10XU, Hamamatsu, Japan) and then fed back to the participants via semi-open headphones (Beyerdynamic DT-990 Pro, Heilbronn, Germany) at a level they found comfortable. The timing reference tracks were routed into the mixer from the recording software and then into the headphones, also at a self-determined comfortable level, which, combined with the direct monitoring of the instrument signals, ensured a “near zero” latency monitoring situation in which the delay between what performers heard in playback and what they played on their instruments was negligible (1–2 samples ≈ 0.02 ms). For the drummers, only the reference track audio was provided in their headphones, and they monitored their own drum sounds through the semi-open headphones, which is also a common practice in studio recording situations.

4.2.1 Guitar and Bass

All guitarists were provided with a 2002 Fender Telecaster (Scottsdale, AZ), and all bassists with a 2016 Fender Precision Bass (Ensenada, Mexico). Overall, participants reported that they were comfortable and highly familiar with the make and model of the instruments provided, which was unsurprising given the historical popularity of the Fender brand. Both instruments were “set up” to a high standard in a workshop by professionals prior to the experiments—that is, they underwent a series of adjustments to bring them to their optimum playability and sound. We received no subsequent comments from bassists or guitarists as to the quality of the instruments. Pick-up tone and gain output settings were pre-configured for both instruments and not alterable for the duration of the experiment, in order to control for the range of frequency, timbre, and intensity output for all participants and to facilitate automatic detection of features in the audio analysis stage.

The bassists were instructed to play with only their index and middle fingers (not their thumb), mainly to standardize the simultaneous collection of motion capture data. All participants reported being comfortable doing so within the experiment's pattern context. The guitarists, on the other hand, were allowed to use a plectrum of their choice, because they tend to be accustomed to a narrow range of plectrum thicknesses. Providing them with an unfamiliar plectrum might have adversely affected their "natural" motor dynamics, which we were interested in preserving as much as possible between the experiment conditions. (That is, a guitarist accustomed to a thinner plectrum might have had a hard time playing a softer stroke with a thick plectrum, and vice versa.) Furthermore, since the experiments had a within-subjects design primarily concerned with the produced differences among conditions within, rather than between, participants, this choice of plectrum option was deemed an acceptable compromise between experimental control and ecological validity.

While professional musicians tend to prefer monitoring their instrument sound from amplifier speaker cabinets, they are also generally familiar with monitoring via headphones, particularly in studio-recording situations. Since the latter eliminates any potential noise or signal leakage issues as well, we decided not to record signals captured by microphone signals from an external speaker cabinet; instead, we provided monitoring of participants' own sounds exclusively via direct-input (D.I.) amplified signals through headphones. In order to increase the overall realism and comfort, however, the D.I. signals from the instruments were first "colored" via analogue amplifier emulators (guitar: Tech 21 Blonde V2; bass: Tech 21 VT-Bass DI [New York, NY]) with preset configurations, in order to reproduce the more characteristic timbres to which guitarists and bassist are accustomed in live-performance settings.

4.2.2 Drums

Since our aim was to investigate performers in the most realistic setting possible, we opted against an electronic MIDI drumkit—while they readily provide clean onset and intensity data, most drummers are not accustomed to playing them. The kinesthetic feedback from the plastic drumhead membranes of electronic drumkits is very different from that of acoustic membranes, and the difference between hi-hat "cymbals" is particularly stark, as acoustic ones are made of metal. In addition, electronic kits do not directly produce drum sounds as such but rather trigger predetermined audio samples when struck, further straining the relationship between produced stroke timing and dynamics and expected sound output in comparison to an acoustic setup, where the relationship among these things is direct. In other words, electronic drum-kits tend to require drummers to play in an entirely different manner than they would with an acoustic kit, and since our experiments were designed to investigate differences in produced timing and sound that may be quite subtle and delicate at times, thus potentially difficult, it was best to not require drummers to utilize a set-up any more challenging than necessary. An acoustic drum-kit was provided to all drummers (see paper II for details). Though not as extensive a drum-kit as most drummers tend to use, the basic kick, snare, and hi-hat set-up is considered the minimum unit which most players are accustomed to.

In order to simultaneously record motion capture data for future investigations, the shiny reflective surfaces of the drum kit had to be covered with black fabric and/or masking tape, which resulted in a certain degree of sound dampening (particularly in the snare, as the fabric had to be wrapped around the top rim edge and was taped to the muffle ring to secure it in place). This resulted in the drum-kit producing 'drier' and 'snappier' sounds, rather than 'wetter' or 'boomy/more resonant' ones. However, most participants were not bothered by this, reporting that they were accustomed to tuning drums to both relatively 'drier' or 'boomier' configurations depending on the genre/style context of any given performance. Once again, though, due to the experiment having a within-subjects design, possible differences in dampening were not an issue as long as the instrument conditions were similar for each participant between conditions.

Contact versus Air Microphones

From the outset, a challenge presented itself in how best to capture as clean and most isolated audio signals possible from each acoustic drum instrument separately—a common problem for music scholars and audio engineers alike. As expected, test recordings revealed that conventional close-microphone techniques with midrange dynamic microphones (SM57, Shure, Niles, IL) led to audible leakage (or "bleed") between instruments in the hi-hat and snare microphones, regardless of positioning and placement. Waveform and spectrogram visualizations of test recordings of two strokes played in close proximity to one another (alternately leading and following) reveal the extent of the leakage between snare and hi-hat signals in dynamic microphones placed close to the hi-hat [figure 4.1(b) and (c)] and snare [figure 4.1(e) and (f)] in comparison to respective single instrument strokes [figure 4.1(a) and (d)]. For the kick drum, on the other hand, a dynamic microphone specifically designed for lower-frequency capture (Beta 52, Shure) placed a few centimeters into a hole in the front drumhead obtained a cleaner and more isolated signal, as it was partially shielded from the other drum kit instruments [figure 4.1(g-i)].

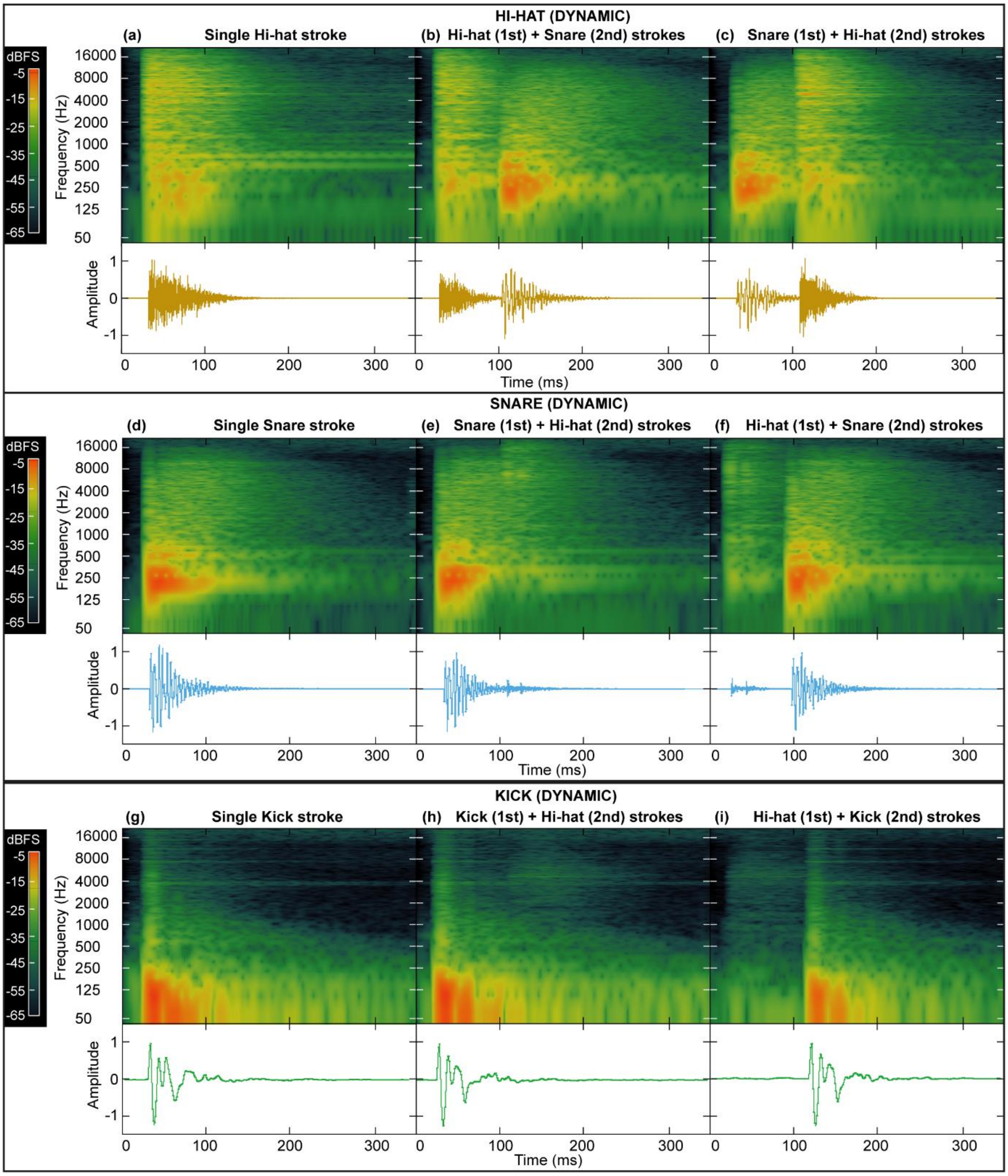

Figure 4.1. Dynamic microphone recordings: waveforms and spectrograms of target single (a) hi-hat, (d) snare, and (g) kick drum strokes and double drum strokes where the target drum was (b, e, h) played ahead of [“1st”] or (c, f, i) behind [“2nd”] a near simultaneous non-target drum stroke. Spectrograms produced in Sonic Visualizer [ver. 3.0.3] using hanning windows with length = 1024 samples and 93.75 percent overlap.

Signal leakage presents problems for stroke-feature detection in the early stages of audio analysis, and since the aim of our investigations in papers I, II, and III was to statistically compare features of individual drum-instrument strokes such as stroke onset/offset, duration, and intensity, obtaining good signal-source isolation was vital. While it is still possible to obtain the stroke-feature information of a single instrument from a mixed-source signal, it requires filter processing to separate the sources that invariably alters or distorts the original signals. While such a method may be viable for simple onset-detection purposes, it is less so for the calculation of individual drum-stroke intensity and duration, because the three drums overlap significantly in frequency range. Even simple filtering, then, would likely lead to relatively inferior or reduced signals compared to those recorded with piezoelectric, or contact, microphones. Unlike dynamic air microphones, contact microphones convert vibrations from solid objects to electrical signals and are highly insensitive to air vibrations, making them less susceptible to signal leakage between close sources.

Tests of all the drum instruments with C411 contact microphones (AKG, Vienna, Austria) placed on the front and top drumheads of the kick and snare, respectively, and the top hi-hat cymbal revealed considerably less signal leakage for all three instruments in relation to the dynamic microphones [see figures 4.2(a–i)]. Also, when a non-target instrument signal is early in relation to a target signal, any frequency overlap caused by the leaked signal tends to be masked to a greater degree by the contact microphones than by the dynamic microphones. For example, in the snare contract microphone, when the snare arrives later than the hi-hat [figure 4.2(f)], it is almost invisible in the waveform, though a small degree of hi-hat leakage can be seen in the 125–1000 Hz range in the spectrogram. However, when the hi-hat arrives later than the snare [figure 4.2(e)], its leaked signal becomes masked and virtually undetectable in both the waveform and spectrogram. This is not the case with the corresponding signals captured by the dynamic snare microphone [figure 4.1(e) and (f)], where the hi-hat leakage is significant in both the waveform and spectrogram, and masking is less effective. The only real concern with the contact microphone signals, then, seemed to be minor pre-stroke leakage that is easy to distinguish from actual target strokes via threshold settings in the analysis stage (see section 4.3.2), so we opted for the contact microphone signals for the snare and the hi-hat in the analysis. For the kick drum, both dynamic and contact microphones seemed suitable and were used in the analyses in papers I and III, respectively.

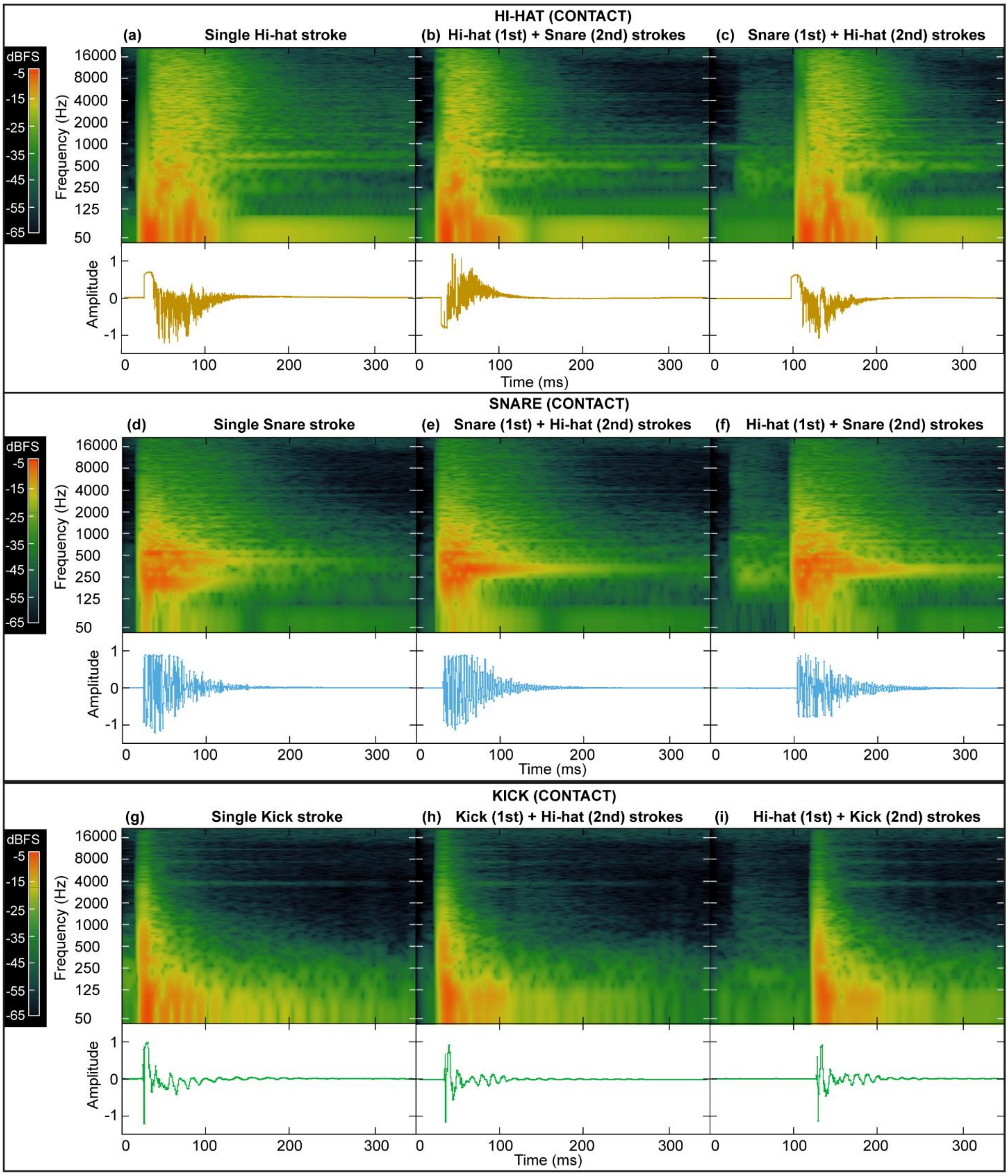

Figure 4.2. Contact microphone recordings: waveforms and spectrograms of target single (a) hi-hat, (d) snare, and (g) kick drum strokes and double drum strokes where the target drum was (b, e, h) played ahead of [“1st”] or (c, f, i) behind [“2nd”] a near simultaneous non-target drum stroke. Spectrograms produced in Sonic Visualizer [ver. 3.0.3] using hanning windows with length = 1024 samples and 93.75 percent overlap.

4.3 Audio Processing and Analysis

An overview of the audio processing and analysis pathway is illustrated in figure 4.3. All processing was done via custom scripts made in Matlab [version 2018a][16] (MathWorks, Natick, MA).

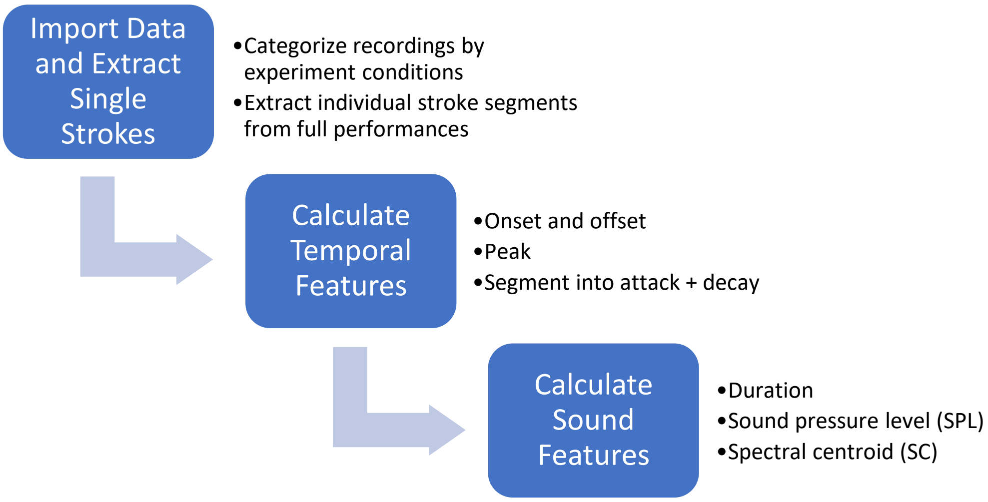

Figure 4.3. Audio processing and analysis pathway.

4.3.1 Importing Data and Extracting Single Strokes

Before we could calculate precise temporal and sound features from individual strokes, we had to isolate and clean the areas around each target stroke in order to facilitate the next stage of the automated extraction process (section 4.3.2).

We first categorized all recordings according to participant number ID, microphone type used, tempo, timing-reference condition, pattern condition, and timing-style condition. Then, to isolate each stroke, we set up a semi-automated process to reduce the time required by manual annotation and to ensure some degree of consistency across the entire dataset. We began by extracting a *pattern grid*, which determined the ideal onset location of each target note/stroke in a given specific instructed pattern relative to a quantized grid comprised of equally spaced quarter-note beat intervals. Each audio recording was cropped into rough individual segments starting 400 milliseconds before, and ending 900 milliseconds after, each target pattern grid location—a relatively large time window that encompassed both the target stroke signal and any pre- and post-stroke noise or potential minor leakage (in the drum recordings).

Next, we produced an estimate of the individual stroke by calculating a rough envelope with two cascaded Butterworth filters, then searching for the fastest rising point closest to the pattern grid point on the envelope curve. The envelope was weighted such that it would prioritize peaks slightly ahead of the pattern-grid point in the pushed condition and slightly behind the pattern-grid point in the laid-back condition. Once this point was detected, a prior local minimum provided a rough estimate of the beginning of the stroke. We roughly estimated the offset as the first local minimum of the envelope after it dropped below a set threshold, then silenced the signals before the beginning and after the end of the rough stroke estimate. At this stage, we plotted all of the individual stroke-segment extractions, then manually inspected and validated them. We also

---

[16] In collaboration with Kristian Nymoen.

manually adjusted certain parameters for envelopes and thresholds for individual recordings to ensure that the entire signal of interest was sufficiently encompassed and passed on to the precise feature calculation step (section 4.3.2).

An example of the rough segmentation of a target guitar stroke appears in Figurefigure 4.4. The green dashed line indicates the pattern-grid time point; the black solid curve indicates the rough envelope calculated; and the red solid line indicates the rough beginning, peak, and end points of the target guitar stroke signal. All of the portions of the signal before and after the red line, such as the non-target muted or "ghost" strokes before and after the target stroke at ca. -300 and +650 milliseconds in this example, were cropped out by the rough individual stroke-segmentation process for all of the instruments.

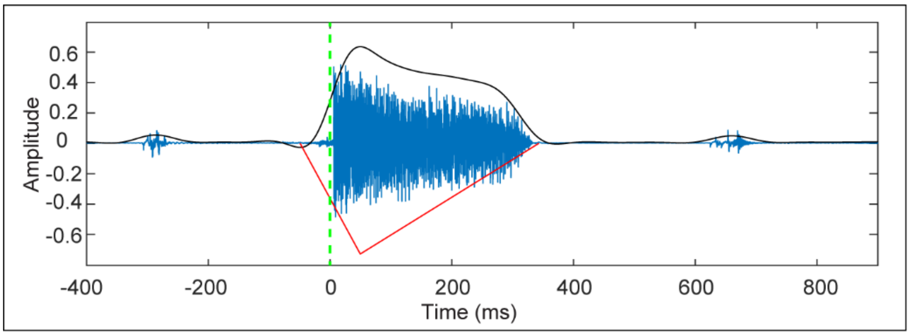

Figure 4.4. Rough individual stroke segmentation of a guitar stroke played with an on-the-beat instruction. Time window is -400 to + 900 ms to either side of the target pattern grid point (at 0 ms, green dashed line); the figure reflects detection of the rough envelope (black solid curve) and beginning, peak, and end points (red solid line) of the signal.

Segments that contained no signal around the grid point above a root-mean-squared (rms) rms energy threshold (individually tailored to each instrument) were automatically marked as *missing strokes* and removed. These were usually the result of small performance errors (participants failing to produce an instructed stroke at a given pattern-grid location), though we also found that, at the start of each recording, participants tended to wait one or two measures before beginning to play in order to entrain to a given timing reference track, resulting in two to four missing strokes per recording on average.

Individual stroke-segment beginnings detected more than 200 milliseconds early or late in relation to a respective location on the pattern grid were also automatically marked as *timing mistakes* and removed. London (2012) argues that, in a rhythmic context, two sequential events spaced by an interval of approximately 100 milliseconds or less tend to be perceived as a single fused event, though Polak (2010) and Haugen and Danielsen (2020) argue that this threshold may be even lower, citing the smaller intervals produced in in Brazilian samba and Malian jembe performances, respectively. Therefore, at the medium tempo of 96 bpm tempo instructed in our performance experiments, any stroke produced later or earlier than this conservative threshold would already be slightly beyond the location of an ideally inferred sixteenth note (± 156 milliseconds from each target pattern-grid note) and thus likely well outside the time window of an expected target quarter-note or eighth-note metrical location. Rather than signal instances of microrhythmically early or late quarter or eighth notes, such timing mistakes would be heard as outside of the relevant "beat bin" (Danielsen 2010) for this musical context—that is, they would be categorically perceived as events belonging to an alternative (and faster) metrical level, such as a preceding or ensuing sixteenth note, in turn producing the qualitatively different sensation of a syncopation.

4.3.2 Calculating Stroke Features

Once each stroke was roughly separated and isolated, the next stage of the automated process involved determining more precise temporal features—that is, onset, peak, and offset points that would serve as reference points for the further separation of each stroke into attack and decay segments. Finally, based on these temporal points, we calculated sound features (duration, intensity [SPL], and brightness [SC]) for each stroke segment.

Though useful for the purposes of this thesis, the conceptual distinction between *temporal* and *sound* features is arbitrary in nature, in that all duration, intensity, or brightness values relate to measurements of a sound signal at either a specific time point or over the course of a timespan, and, conversely, all onset, offset, and peak values describe moments in time when an event displays a certain sound-feature value, such as amplitude energy. For analytical clarity, however, we use the term *temporal features* to refer to the time points of onset, peak, and offset that were extracted from the stroke signals, and the term *sound features* to refer to duration, SPL, and SC measurements over the timespans of attack, decay, and total segments of strokes.

In our choice of quantitative audio descriptors, we strove to use the most precise measurement techniques possible, and those that best described the temporal and sound features that were present in the musicians' performances. While different parameter and threshold settings might have provided even more precise results, and better methods might have been available, we feel confident that we produced a fair comparative analysis of performance data across a large sample, principally because we applied a consistent methodology to the measurement of quantitative features across all participants. It is also true that some amount of subjective interpretation is always present as one defines the so-called objective features of musical performance, because, as perceptual experimental studies of various auditory features have shown, listeners perceive those features in different ways. In our analyses, then, we chose from the outset to study features that have been shown to be perceptually relevant to listeners, then interpreted how these features could be qualitatively perceived, based on both heuristics derived from experimental studies and our own qualitative analyses.

Temporal Features

The temporal features we calculated are those derived from physically recorded signals (as opposed to perceptually derived measurements from perceptual experiments). Thus, with onset, offset, and peak, we refer to the *physical* signal measures thereof, calculated via reliable audio descriptors. As we saw in Chapter 3, research into the P-centers of musical sounds shows that the average moment in time when rhythmic events are perceived to occur tends to be located somewhere between onset and maximum amplitude peak point. While measured P-centers are often very similar in value to physical onset and peak locations in transient-rich sounds with fast rise times, they tend to occur relatively later than those physical locations in slower-rising sounds. As of yet, however, no existing computational models are able to accurately identify P-center location for a wide range of instrumental sounds (Nymoen et al., 2017; Villing, 2010). Consequently, we contented ourselves with onset and peak as valid, albeit approximate, indicators of the temporal location of strokes for the comparative purposes of this thesis's analyses.

Onset/Offset

Before calculating any sound feature of a given stroke, one must first define the total extent of the individual stroke iteration, which involves delimiting the beginning and ending time points (onset and offset) as precisely as possible. Onsets and offsets are typically measured as points in time when the amplitude energy of a signal rises and falls above or below a certain threshold, which is usually either zero (Nymoen et al., 2017) or, somewhat less commonly, a predetermined baseline noise level. The common basis of most algorithms for determining the time points of the beginning and end of a sonic event is a function that first describes the energy envelope of a sound signal, then calculates onset and offset based on various thresholds applied to this envelope and its derivatives, with the method and parameters specified in the envelope calculation strongly influencing the estimated features (Nymoen et al., 2017).

At an early stage of the thesis, we tested the detection of onsets from pilot recordings of single instrument strokes using state-of-the-art algorithms available at the time, including *mirevents* from the MIR Toolbox [version 1.7] (Lartillot et al., 2008). *Mirevents* first calculates the envelope curve of a signal with the default envelope extraction method ("*Envelope*" option), computing a spectrogram whereby the sum of the amplitudes in each frame of a certain specified length and hop factor determines the shape of the envelope. The first time derivative of the envelope energy (e') is used as the basis for estimating the beginnings and endings of the attack and decay phases, (default: "*Derivate*" method) whereby onset and offset points (termed "attack start" and "decay end" in the MIR Toolbox) can be extracted, alongside other temporal features. In general, shorter frame lengths and smaller hop sizes produce more jagged envelopes but retain more temporal information from the waveform, whereas longer frame lengths and larger hop sizes produce smoother curves but retain less temporal information from the waveform.

According to Nymoen and colleagues (2017), the default envelope parameters of *mirevents* prior to version 1.7 were not oriented towards the precise capture of sounds with faster rise times (time span between onset and energy peak), but rather those with relatively slower energy evolution. Based on an optimization of envelope parameters against a set of sounds from a P-center study, these researchers recommended using shorter frame length and hop size as well as a lower e' percentage[17] for faster rising sounds, such as those used in this thesis's experiments. Applying similar parameter values with minor adjustments on selected sounds of all the instruments from our performance experiments, we can see how the detected onsets from the *mirevents* envelopes generally tend to anticipate the points in the waveforms where the signals begin to rise, which were manually annotated for comparison (figure 4.5). While the discrepancies are relatively small, ranging between ca. 10 to 30 milliseconds, as we saw in chapters 2 and 3, microtiming differences between different instructed timing styles in performance are often within this subtle range. Since one aim of this thesis was to provide the most detailed and accurate information possible regarding the timing and sound features of instrumentalists, such values displayed a loss of temporal information that was far too great.

---

[17] Consequently, starting with MIR Toolbox versions 1.7 and onwards, Nymoen and colleagues (2017) suggestions of frame length = 30 ms and hop factor = 2 percent were incorporated as the new default parameters for *mirevents* ("*Attacks*"), while the e' thresholds for attack/decay start defaults were set to 10 percent and the attack/decay end thresholds to 20 percent as a default. (However, since version 1.7.1, the threshold for attack end has been set to 7.5 percent).

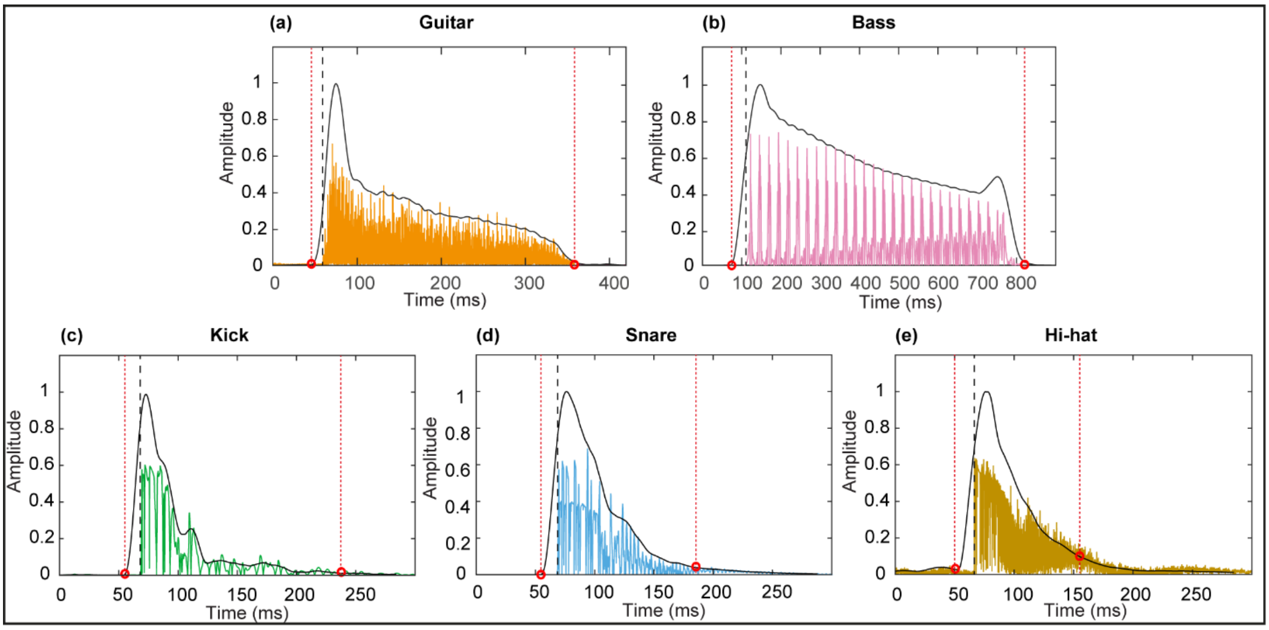

Figure 4.5. Examples of calculated envelope (solid black lines), onset [“Attack Start”] and offset [“Decay End”] (red circles + dotted vertical red lines) points using *mirevents* (“*Attacks*”, “*Decays*”, “*Single*”) [version 1.8] against rectified waveforms of single (a) guitar and (b) bass strokes recorded via amplified direct input signals, and (c) kick, (d) snare, and (e) hi-hat strokes recorded with contact microphones. Manually annotated onset points indicated by dashed vertical black lines. Frame length = 30 ms and hop factor = 2 percent for all sounds except bass (frame length = 100 ms), in order to ensure a smooth enough envelope that encompassed the entire waveform. Onset (e' “attack start”) and offset (e' “decay end”) threshold values = 3 percent for all instruments except hi-hat (10 percent), in order to account for baseline noise level in the recordings.

Despite attempts to further fine-tune the envelope parameters, we failed to achieve better automated onset detection results compared to those we obtained manually, as it would appear that attack/decay phase feature detection based on envelope extraction via either low-pass filtering or frame-decomposed energy summation is problematic, due to the “inevitable loss of time resolution” (Lartillot et al., 2021). Because the data sets to be analyzed were far too large to depend upon manual annotation, we approached Olivier Lartillot (main creator of the MIR Toolbox) to inquire whether there might be a way to adjust the *mirevents* algorithm to detect onset and offset features directly from the audio waveform, rather than from the envelope, and we provided several sound examples from our data set for his investigation. Lartillot then proceeded to develop a new option for the *mirevents* algorithm called “*Waveform*”[18] that defines:

1. *Onset* as the starting point of the “attack phase,” where the “attack phase” is a line fitted on the audio waveform through an optimization of both the slope and the maximum amplitude of the envelope (given by *mirevents* with the “*Attacks*” option set to “*Waveform*”).
2. *Offset* as the ending point of the “decay phase,” computed using the same approach as the onset (given by *mirevents* with the “*Decays*” option set to “*Waveform*”).

A “*WaveformThreshold*” parameter further enables the adjustment of the percentage of the envelope’s maximum amplitude, from which the onset and offset thresholds are then determined. In the experiment, different instruments’ signals required different onset/offset threshold parameters, depending chiefly on the level of baseline noise produced by the microphones. In figure 4.6, we can see that, with the new “*Waveform*” option, *mirevents* provides much more precise onset detection and more consistent offset detection for all instruments in comparison to the purely envelope-based approach (figure 4.5). We therefore decided to utilize the former approach to

---

[18] Now included in the most recent MIR Toolbox, version 1.8.

calculate onsets and offsets in our analyses, with manually adjusted *WaveformThreshold* values for each of the instruments as our starting points (certain recordings were further manually adjusted where necessary).

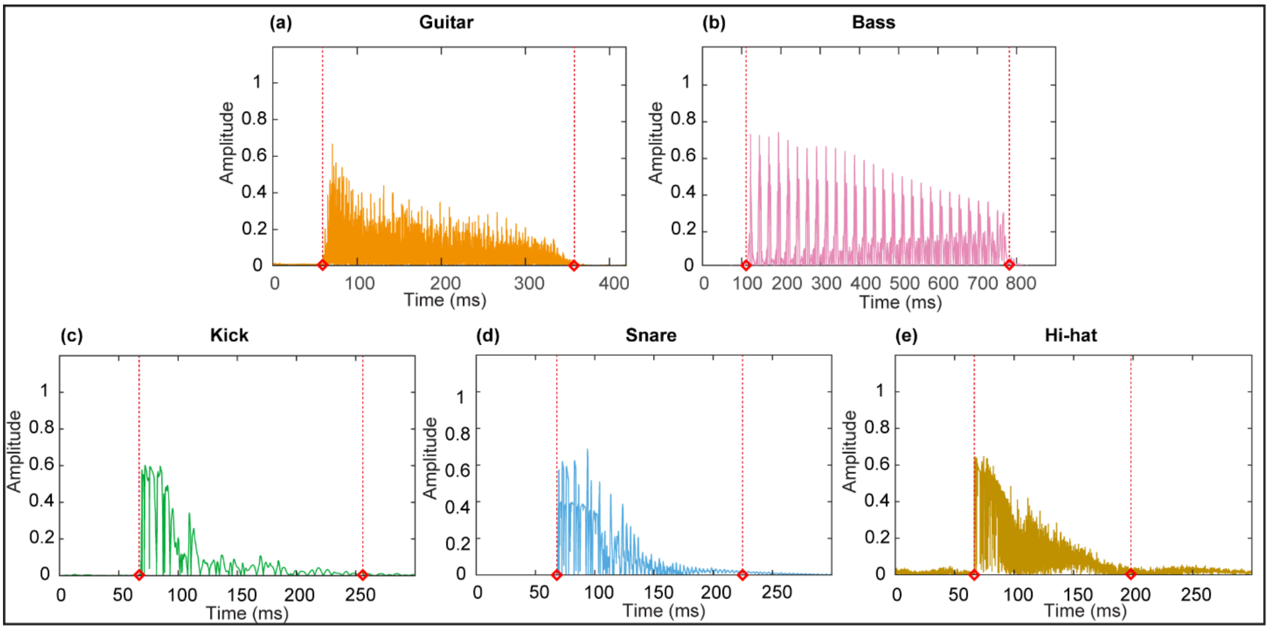

Figure 4.6. Examples of calculated onset and offset (red circles + dotted vertical red lines) points using *mirevents* (“*Waveform*”, “*Attack*”, “*Decays*”, “*WaveformThreshold*”) [version 1.8] against rectified waveforms of single (a) guitar and (b) bass strokes recorded via amplified direct input signals, and (c) kick, (d) snare, and (e) hi-hat strokes recorded with contact microphones. *WaveformThreshold* values (both onset and offset) = 3 percent for all instruments except hi-hat (10 percent), to account for baseline noise level in the recordings.

Peak and Stroke Attack/Decay Segmentation

Strokes from both the string (guitar and bass) and percussive (kick, snare, hi-hat) instruments tended to display an acoustic energy shape over time that can be roughly described in two stages: an initial impulsive, transient stage caused by the direct striking of strings or drumheads (the “attack” segment), followed by a resonating stage, where strings or drumheads are allowed to “ring out” until they naturally gradually decay over time or are actively stopped by a performer action (the “decay” segment). Because we wanted to analyze how these different segments may be manipulated in specific ways by musicians in order to express different timing styles, we had to determine a consistent third time point between onset and offset to separate the strokes into “attack” and “decay” segments.

A common approach in attack/decay segmentation is to use the maximum amplitude energy, or “peak,” of the signal as the separating time point. Like onset and offset detection, the signal peak can be extracted either directly from a waveform or indirectly via a smoothed envelope. Through a comparison of peaks calculated from waveforms and envelopes (using *mirevents*), we found that our instrument signals often displayed several transients in their early stages, which led to waveform peaks being detected somewhat inconsistently relative to signal onsets across strokes. On the other hand, the peaks of the envelopes tended to be detected more consistently relative to onsets, regardless of whether the waveforms displayed one or several transients in the early signal stages. Figure 4.7 illustrates two guitar and bass strokes from two different participants’ recordings. In guitar example 1 (4.7a), the singular transient in the early stage of the signal leads to a very similar peak detection for both the waveform (indicated by a red diamond and red dotted line) and the envelope (red circle and black dashed line) relative to the waveform onset (green diamond and green line). In guitar example 2 (4.7b), however, the signal shows several transients throughout the stroke, and the maximum waveform peak is subsequently detected much later than the envelope peak in comparison to guitar example 1. The bass strokes show a similar pattern: bass example 2 (4.7d) displays a strong singular transient early in the signal, leading to a marginal difference between detected waveform and envelope peaks, while the absence of a strong transient in the early signal stage in bass example 1 (4.7c) leads to the waveform peak being detected much later in relation to the onset than the envelope peak in comparison to bass example 2.

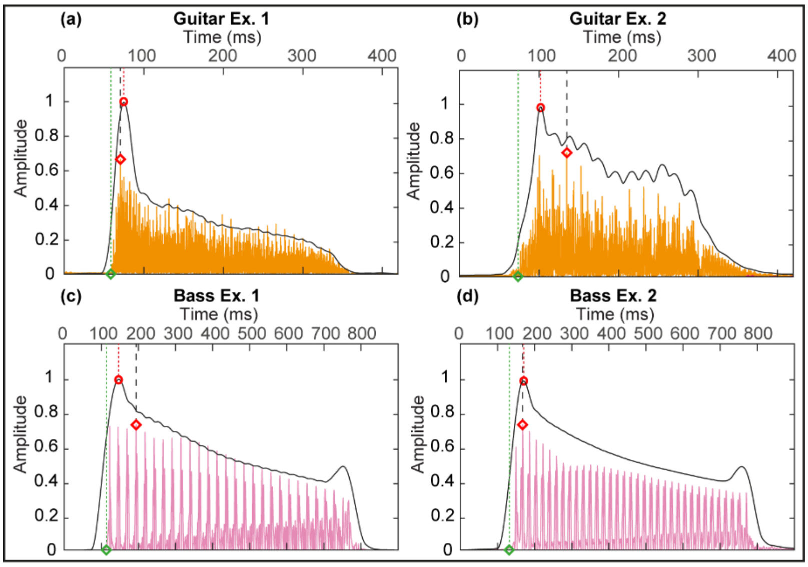

Figure 4.7. Examples of the maximum envelope amplitude peak (indicated by red circles + horizontal solid black lines), maximum waveform peak (red diamonds + vertical red dotted lines), and waveform onsets (green diamond + vertical green dotted lines) in rectified waveforms of two single (a–b) guitar and (c–d) bass strokes. Envelopes calculated using *mirevents* (“*Attacks*”, “*Decays*”, “*Single*”) with frame lengths = 30 ms (for guitar) and 100 ms (for bass) and hop factor = 2 percent (for both instruments), and onsets calculated using *mirevents* (“*Waveform*”, “*Attack*”, “*Decays*”, “*WaveformThreshold*”) with *WaveformThreshold* values = 3 percent (for both instruments) [MIR Toolbox version 1.8]. Waveform peak calculated simply as the maximum amplitude time point of the signals.

The drums also showed the same kind of inconsistency in waveform peak detection. Figure 4.8 illustrates three typical kick, snare, and hi-hat strokes (recorded with contact microphones) where the waveform peaks frequently tend to be detected either earlier than (hi-hat [figure 4.8 c]), later than (snare [figure 4.8 b]), or very close to (kick [figure 4.8 a]) the envelope peak relative to signal onset.

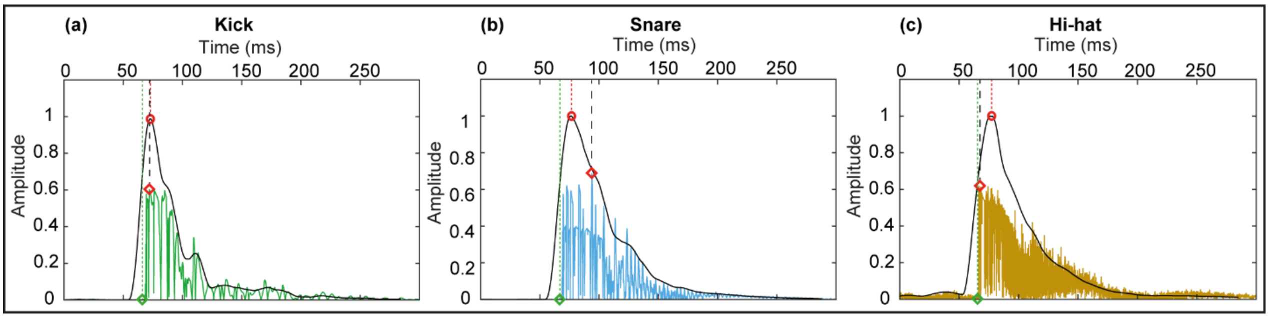

Figure 4.8. Examples of the maximum envelope amplitude peak (indicated by red circles + horizontal solid black lines), maximum waveform peak (red diamonds + vertical red dotted lines) and waveform onsets (green diamond + vertical green dotted lines) in rectified waveforms of a single (a) kick drum (b) snare drum and (c) hi-hat cymbal, recorded with contact microphones. Envelopes calculated using *mirevents* (“*Attacks*”, “*Decays*”, “*Single*”) with frame lengths = 30 ms and hop factor = 2 percent (for all instruments), and onsets calculated using *mirevents* (“*Waveform*”, “*Attack*”, “*Decays*”, “*WaveformThreshold*”) with *WaveformThreshold* values = 3 percent (for kick and snare) and 10 percent (for

hi-hat) [MIR Toolbox version 1.8]. Waveform peak calculated simply as the maximum amplitude time point of the signals.

In summary, since envelope peaks tended to be less sensitive to the amplitude fluctuations of transients in the early stages of stroke signals and were detected as more consistently distant relative to signal onsets, we decided to use the maximum-energy peak time of the stroke envelope to separate strokes into attack and decay segments. More specifically, we defined the peak for the analyses as the global maximum amplitude of the envelope of a given stroke, where the envelope is calculated as the sum of bin amplitudes across the columns of a spectrogram (using *mirevents* with the “Attacks” option and default parameter settings; frame length = 30 milliseconds, frame hop factor = 2 percent).

As a result, all strokes were subsequently separated into the following segments:

a) Attack: interval between onset (waveform) and maximum peak (envelope);
b) Decay: interval between maximum peak (envelope) and offset (waveform);
c) Total: interval between onset (waveform) and offset (waveform).

Sound Features

Once the temporal locations of onset, peak, and offset were determined, we proceeded to calculate the duration, intensity (sound pressure level [SPL]), and brightness (spectral centroid [SC]) of each stroke’s attack, decay, and total segment. We chose these three sound features in particular because each has been found to affect the perceived timing of sounds (see Chapter 2 and Chapter 3). Numerous perceptual studies report that the total elapsed time of a sound, as well as the duration of its attack segment in particular (its “rise time”), has a significant effect on the P-center, whereby longer total and attack durations tend to shift the experienced P-center later in time. A few studies also suggest that sounds with increased intensity (higher SPL) and pitch/frequency (higher SC) tend to lead to earlier P-centers, and vice versa.

Duration

Duration was simply calculated as the elapsed time interval of the stroke segments (attack/decay/total), measured in milliseconds. Above, we saw how different onset/offset and peak measurement methods based on waveform or envelope can significantly impact time-point detection, thereby also affecting the subsequent attack, decay, and total duration measurements.

During a performance, instrumentalists are able to manipulate the total duration of strokes to some extent by controlling how long strings or drumheads are allowed to resonate, using techniques such as the muting of strings via fretting or the strumming/plucking hand for guitarist and bassists, and the regulating of stick/pedal “bounce” versus “choke” for drummers. The analyses in papers I and II show that, for bassists and drummers, the shortening or lengthening of strokes tends to occur mainly in the decay segment, as it is apparently difficult to manipulate the attack duration of individual bass or drum strokes pre-impulse. For the guitarists, on the other hand, who were instructed to play polytonal chords that require the striking of multiple strings simultaneously, it was possible to shorten or lengthen both the attack of the strokes (by decreasing or increasing the time interval between each successive string struck in the chord) and the decay of the strokes (by letting the chord ring out post-impulse). Figure 4.9 illustrates two vastly different guitar attack-duration approaches. Example (a) shows a "swept" stroke, where the chord is arpeggiated rather slowly. Here, the spectrogram in particular clearly reflects the individual strings being struck, as indicated by the impulsive transient sections with higher degrees of noise (energy spread across a wide frequency range); the onset is detected when the first string/note is struck (with the lowest energy), while the peak is detected when the last string is struck (with the highest energy). The resultant calculated attack duration (indicated by the green dashed line) is quite long, at around 180 milliseconds. Example (b), conversely, shows a "swift" stroke, where all four strings/notes are played in fast succession. Here, both graphs show far more dramatic separation of the string transient impulses, and the onset and peak are detected as closer to each other, resulting in a much shorter calculated attack duration of around 30 milliseconds. In both cases, the decay durations (or distance between peak and offset, indicated by the red dotted line) are rather similar, at around 200 to 250 milliseconds, and the abrupt decrease in energy in the spectrograms indicates that both strokes were actively stopped (or "muted") by the guitarists rather than allowed to decay gradually.

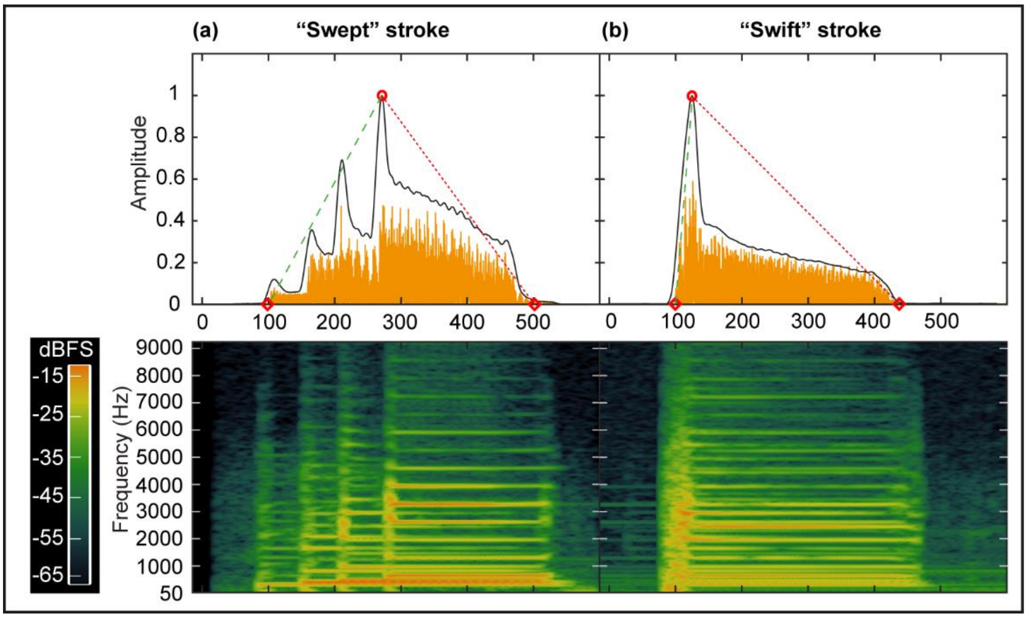

Figure 4.9. Comparison of a single guitar stroke played with (a) "swept" and (b) "swift" attack techniques. Spectrograms produced in Sonic Visualizer [ver. 3.0.3] using hanning windows with length = 1024 samples and 93.75 percent overlap.

SPL

To calculate stroke intensity, we began by calculating the global energy (unweighted root-mean-square [rms] amplitude) of each stroke segment using *mirrms* (MIR Toolbox, version 1.8). Then, in order to be able to compare relative stroke intensities based on a more familiar perceptual scale later in the statistical analyses, we converted the rms energy values into sound-pressure-level units (dB SPL), with the 0 dB reference defined as the average rms amplitude of all strokes in each individual instrument's dataset.[19] SPL is a relative intensity measure that is highly correlated to perceived loudness and remains a widely used metric in perceptual studies (Rossing et al., 2002). The resultant attack, decay, and total SPL measurements produced the average (mean) SPL values of each stroke segment.

---

[19] The conversion formula used was $20\log_{10}$(individual stroke RMS/mean $\text{RMS}_{\text{all strokes}}$).

SC

To calculate stroke brightness, we calculated the spectral centroid (SC) of each stroke segment using *mircentroid* (MIR Toolbox, version 1.8), which gives the weighted mean of the frequencies present in the signal via a Fourier transform, with their magnitudes as the weights, measured in Hz. SC is a frequency descriptor that purportedly accounts for the perceived brightness or sharpness of sounds (Donnadieu, 2007; Schubert & Wolfe, 2006). The resultant attack, decay and total SC measurements produced the average (mean) SC values of each stroke segment.

We investigated the SC features in the analyses of the guitar and bass datasets (reported on in paper I) but not the analyses of the drums (paper II). This is because, as mentioned in section 4.2.2, we opted to use contact microphones to capture the snare and hi-hat signals in the recording process. Unlike the dynamic microphone, the contact microphone provided recordings from which superior temporal feature (onset, peak, offset) and loudness (SPL) information could be extracted due to better signal-source isolation. When they are reproduced, however, contact microphone signals sound very “tinny” or “metallic” and relatively “distorted” (perceptually, not in the signal) relative to the signals of dynamic microphones, partly because of the former microphone’s higher impedance. As such, we did not deem the contact microphone recordings to faithfully reproduce the timbre of real drum sounds as actually heard by listeners, and thus any eventual comparisons of their brightness (SC) would have had very low validity.

4.4 Statistics

Once all of the strokes from each instrument dataset had been isolated, the temporal features (onset, peak, offset) had been extracted and separated into attack/decay/total segments, and the sound features had been calculated (duration, SPL, SC), we organized the information into data arrays and exported them for further analysis.

In papers I and II, the aim of the statistical analyses was to discern group trends across all participants in the guitar and bass (paper I) and drum (paper II) experiments. Here, within-subjects tests of differences of means/medians (RMANOVA/Friedman) indicated the extent to which the musicians produced strokes with statistically significant differences in temporal and sound features among the various instructed timing conditions and patterns (see section 4.4.1). In paper III, the focus shifted away from group average trends and towards the mapping of the range of individual strategies present within the group, here limited to the analysis of the drummers’ experiment data. Using a hierarchical clustering method combined with phylogenetic tree visualizations, we gained further insight into how drummers were able to distinguish laid-back and pushed from on-the-beat timing styles in a variety of ways (see section 4.4.2).

4.4.1 Finding Group Trends: Within-Subjects Tests of Difference between Averages

The three performance experiments reported on in papers I (guitar and bass) and II (drums) were specifically designed using a within-subjects structure. Such a design applies when several repeated quantitative measurements of a dependent variable (e.g., duration of strokes) are taken under different qualitative conditions of one or more controlled independent variables (e.g., instructed timing style: laid-back, on-the-beat, pushed). Within-subjects tests then compute the main effects of each independent variable for all subjects individually, which together represent how much a given dependent variable tends to vary among the different instructed conditions for the average participant in the sample. By analyzing the data via such tests, we were therefore able to gauge the extent to which the average guitarist, bassist, or drummer in our sample changed the way they produced strokes in terms of temporal or sound features when playing under different instructed timing-style, timing-reference, and/or pattern conditions.

Since the experiments had a within-subjects design, I found it appropriate to within-subjects standardize the data before checking for normality (Fischer & Milfont, 2010). I did this by averaging all of the strokes across timing-style conditions for a given participant, then subtracting the average from each individual timing style for the same participant, which essentially centers each participant's data around the mean of their own performances.[20] Then, I computed standardized residuals for each data series and manually screened them for normality via histograms and Q-Q plots for each dependent variable and instrument. For the series that showed normal distribution of standardized residuals, I conducted repeated-measures analyses of variances (RMANOVAs) using the means; for the series that departed from normality, I conducted corresponding non-parametric Friedman tests using the medians (both cases, using raw, unstandardized data). I used post-hoc pairwise comparison tests (paired samples t-test for RMANOVA, or Wilcoxon test for Friedman) to gauge which conditions differed significantly from one another more specifically as appropriate, with Bonferroni correction of alpha values to adjust for multiple comparisons. I corrected violations of the sphericity assumption (Mauchly's test) using Greenhouse-Geisser.

In paper I, I ran several univariate one-way RMANOVAs or Friedman tests on the guitar and bass data, depending on whether normality assumptions were satisfied for each dependent variable. The investigated independent variables here involved the instructed timing style (three levels: laid-back, pushed, on-the-beat), and the dependent variables involved two temporal (onset, peak) and nine sound-related stroke features (attack/decay/total × duration, SPL, and SC). In paper II, because all of the drummers' data series satisfied normality assumptions, I ran three-way RMANOVAS on each individual percussion instrument (kick, snare, hi-hat), where the three investigated independent variables involved instructed timing style (three levels: laid-back, pushed, on-the-beat), timing reference (two levels: metronome and instrumental backing track), and pattern (two levels: simple and complex), and the dependent variables involved one temporal (onset) and six sound stroke features (attack/decay/total × duration and SPL). Here, I tested not only the main effects but also the interaction effects between the different independent variables.

All statistical analyses were performed using SPSS [ver. 25] (IBM, Inc., New York).

---

[20] It should be noted, though, that results of RMANOVAs, which assume normality of residuals, are not affected by such a transformation—they are identical regardless of whether raw or within-subjects standardized values are computed.

4.4.2 Exploring Diversity of Individual Strategies: Hierarchical Clustering and Phylogenetic Tree Visualizations

While within-subjects investigations can inform us about the average changes in stroke feature manipulation between different instructed timing styles, average performances do not represent real performances as such but rather indicate general trends in a given sample of musicians. In paper III, then, we sought to further explore and map the diversity of individual strategies actually used by all twenty drummers in the sample to distinguish laid-back and pushed from on-the-beat performances. Here, we also wanted to explore the range of possible *combinations* of stroke relations produced between all three drums at the same time (rather than separately, as in paper II), as well as how they were produced at different metrical locations in the pattern. We limited the scope of this investigation to the exploration of stroke onset and intensity (total SPL) based on group average findings from paper II, which indicated that differences between these features were likely the most salient (see discussion in section 5.1.3). For the purpose of this investigation, we developed a novel methodology (Sioros, Câmara & Danielsen 2019). Based on a methodology developed by Sioros and colleagues (2019), we used an agglomerative, hierarchical cluster analysis of stroke onset and intensity feature combinations with phylogenetic tree visualizations of the clusters, which allowed to us to provide a more in-depth overview of the main, or “archetypical,” strategies used by the drummers to express laid-back and pushed timing styles.

Hierarchical clustering algorithms group sets of data points according to a measure of their similarity. In our case, the data points were profiles of selected onset and intensity relations between drum instruments extracted from each drummer’s performances, treated as arrays of variables, and the we calculated the similarity between two profiles as the averaged Euclidean distance between them (Euclidean distance divided by the number of dimensions). The distance between groups of points is then a function of their pairwise distances. In this study, we used the “unweighted group average” linkage criterion (UPGMA) for this purpose. The end point of the hierarchical algorithm is a dendrogram or tree. Clusters of similar performance profiles were created using the similarity profiles algorithm (SIMPROF) (K. R. Clarke et al., 2008), as implemented in the Fathom Toolbox for Matlab (D. L. Jones, 2017). We then applied the SIMPROF method, which was initially developed in the field of bioinformatics, to the pre-calculated dendrogram. This method performs a series of permutations that tests the statistical significance of the dendrogram branches’ internal structure. We set the number of iterations of the permutation test at 10,000, and the significance level alpha at 0.025—the probability value at which the hypothesis of an internal structure is rejected. In addition, we used the Bonferroni correction (Legendre & Legendre, 2012, p. 745) to progressively adjust the probability values for multiple simultaneous tests. In the end, the visualization of the clustering results as phylogenetic trees provided a simple and accessible overview of the onset and intensity strategies and the ways in which they were related.

Although it would have been possible in principle to analyze any combination of performance stroke features using more conventional multivariate statistical approaches, it is often difficult to formalize hypotheses to be tested given the variety of strategies that musicians seem to exhibit. Hierarchical clustering methods, on the other hand, group and sort recordings according to their similarity and reveal their relations without the need for a priori hypotheses about the existence of specific strategies—they require, in fact, only a similarity metric between the data points. Also, unlike many multivariate approaches, they do not impose restrictions on the distribution of the data. However, in future studies, the two approaches could be used in a complementary fashion, where hierarchical clustering could initially assist in formalizing concrete hypotheses that could be subsequently tested with conventional analyses, thus providing more robust statistical results.

Regarding “tolerance” thresholds, or the parameters which essentially define whether performances are categorized as “different” in terms of stroke timing and sound features (relative to either a timing reference grid or between instruments), while the thresholds in papers I and II were fixed (determined by conventional statistical parameters), we chose to derive various “production” thresholds in paper III based on each individual drummer’s own on-the-beat performances. As such, we chose to use a “relative” approach, where the mean variability (2 x SD) of a drummer’s own stroke onset and intensity reference values in the on-the-beat condition was used as a tolerance threshold. That is, for laid-back and pushed strokes to be categorized as early/late or louder/softer at any given metrical beat, they had to occur 95 percent outside of the mean distribution of the on-the-beat reference values. The rationale for this was that we assumed that, in their on-the-beat performances, drummers played the various onset and intensity features with a certain degree of intentional control and accuracy. While other approaches, such as more “absolute” ones relying on fixed tolerance thresholds determined by average psychoacoustically derived JND values (section 3.2), may be just as valid, our approach, while admittedly somewhat arbitrary, is nonetheless a conservative one that cautiously ensures that such strokes played with onset and intensity differences were done purposefully, rather than occurring due to random error or chance. Note that, because these relative tolerance thresholds were tailored to each drummer individually, the lesser the variability of produced onset and intensity features in the on-the-beat condition, the greater the chance that laid-back and pushed strokes with lesser magnitudes of onset asynchrony and intensity would be categorized as earlier/later or louder/softer, and vice versa. As such, the analyses in paper III are best suited to gauging what each individual drummer did to specifically distinguish laid-back and pushed from on-the-beat performances, and how consistently they did so, rather than to directly comparing onset and intensity difference magnitudes between different drummers. In addition, as mentioned in paper III (note 15), had we chosen less conservative threshold values for the classification of strokes determined as louder/softer and early/later (for example, 1 x SD), an even greater number of individual performances would have been subsumed by archetypes characterized by systematic intensity and/or onset differences rather than by non-difference archetypes (though the latter still remained in the minority).

Note also that, in paper III, we chose to base the metrical “grid” reference on the onset timing of drummers’ own hi-hat strokes in the on-the-beat condition rather than on metronome or backing-track timing references. Thus, the “grid” was conceptualized here as a subjective reference structure representing where drummers might have perceived the location of the main quarter-note beats of the 4/4 meter. The hi-hat was specifically chosen as a proxy for the average location of the drummers’ perceived metrical grids, since findings from paper II as well as other studies (Fujii et al., 2011) reveal that the hi-hat often exhibits the least NMA or is most closely synchronized to external timing references during in-phase (or on-the-beat) tasks, and thus likely serves as the main “timekeeper” for drummers.[21] We could have alternatively chosen to use a superimposed idealized grid based on local (measure level) or global (entire performance) average onset timing of a single

---

[21] Note further, though, that the magnitude of onset differences between drums and timing references naturally differs from drummer to drummer and is oftentimes negligible or even positive, especially when the stimuli are more ecological (paper II; Danielsen, Waadeland et al. 2015; Kilchenmann and Senn 2011). In addition, it likely depends on how “accurately” drummers are able to synchronize in an on-the-beat fashion based on their skill level. Personal aesthetic preference may also play a role, since what is considered to be “on the beat” in the context of the pattern for any given drummer might entail relatively “tighter” or “looser” produced onset relations relative to the timing reference.

drum instrument (such as that used by Câmara [2016] and Bilmes [1993], or of all drum instruments (Hoffman et al., 2017). However, as Haugen (2016) points out, in general performers' own produced movements provide suitable clues as to how they themselves construe subjective metrical reference structures in performance.

Chapter 5

Research Summary

The three papers that form the core of this thesis aimed to shed light on the various ways in which groove-based rhythm-section musicians (paper I: guitarists and bassists; papers II and III: drummers) distinguish between different microrhythmic performance approaches (“laid-back”, “pushed” and “on-the-beat”) through the systematic manipulation of both timing and sound features. In papers I and II, our interpretation of the perceptual effects is based on the statistical average timing–sound combinations across the entire group of musicians. As such, they only represent one possible “archetypal” timing–sound combination—the actual combinations of stroke onset/duration/intensity in any given individual participant’s performance may or may not conform to this archetype and may even produce different perceptual effects. In paper III, we therefore sought to map out the diverse range of individual strategies used by the drummers to achieve the different microrhythmic feels.

In section 5.1, I present the main findings of these three papers. In section 5.2, I elaborate on the papers’ findings in relation to the research questions outlined in chapter 1.

5.1 Papers

5.1.1 Paper I

Guilherme Schmidt Câmara, Kristian Nymoen, Olivier Lartillot, and Anne Danielsen, “Effects of instructed timing on electric guitar and bass sound in groove performance,” in *Journal of the Acoustical Society of America* 147(2), 1028–1041 (2020).

Abstract. This paper reports on two experiments that investigated the expressive means through which musicians well versed in groove-based music signal the intended timing of a rhythmic event. Data were collected from 21 expert electric guitarists and 21 bassists who were instructed to perform a simple rhythmic pattern in three different timing styles—“laid-back,” “on-the-beat,” and “pushed”—in tandem with a metronome. As expected, onset and peak timing locations corresponded to the instructed timing styles for both instruments. Regarding sound, results for guitarists revealed systematic differences across participants in the duration and brightness [spectral centroid (SC)] of the guitar strokes played using these different timing styles. In general, laid-back strokes were played with a longer duration and a lower SC relative to on-the-beat and pushed strokes. Results for the bassists indicated systematic differences in intensity [sound-pressure level (SPL)]: pushed strokes were played with higher intensity than on-the-beat and laid-back strokes. These results lend further credence to the hypothesis that both temporal and sound-related features are important

> indications of the intended timing of a rhythmic event, and together these features offer deeper insight into the ways in which musicians communicate at the microrhythmic level in groove-based music.

In this paper, we focused on the data from the guitarists and bassists in the performance experiments outlined in section 4.1, limiting the analysis to one pattern per instrument (guitar: “backbeats,” bass: “downbeats”) and one external timing reference rhythm (a wood-block metronome). Based on findings from perceptual studies of P-centers and anisochrony/asynchrony (see chapter 3), we expected that musicians would shape the sounds of their own instruments in ways that might enhance the perceptual salience of intentionally late or early strokes. In particular, we hypothesized that they would produce strokes with sound features that would shift the P-centers of their own sounds relative to those of the metronome in a later direction (via slower attack/longer duration/lower SC/lesser intensity) to convey a more laid-back feel or an earlier direction (via faster attack/shorter duration/higher SC/greater intensity) to convey a more pushed feel. Conversely, in the on-the-beat performances, we expected that the musicians would seek to align their onsets with those of the metronome as closely as possible to create perceptually vertical synchronous relationships (and thus also horizontal isochronous relationships).

We found that, on the one hand, the average onsets of on-the-beat strokes for both instruments tended to anticipate those of the metronome, a common effect often found in tapping synchronization studies that is termed “negative mean asynchrony” (NMA) (Repp, 2005; Repp and Su, 2013). On the other hand, we found that the average peak of on-the-beat strokes was closer to the metronome onsets and not anticipatory, which suggests that guitarists and bassists might rely upon a moment closer to their own sounds’ maximum peaks when synchronizing their strokes to the metronome reference sounds. That is, they may be aligning the P-centers of their strokes to those of the metronome in such a fashion that they would be heard as falling within the subjective “beat-bin” afforded by the timing reference sounds (Danielsen 2010; Danielsen et al., 2019; see also section 3.1.2).

In the laid-back and pushed conditions, as expected, we found that, compared to the on-the-beat condition, participants produced both the onset and the peak of their strokes between 35 - 46 ms earlier or later in time on average, respectively. The variability of average laid-back and pushed onset and peak locations was also found to be numerically greater than those of on-the-beat strokes, which could be attributed to the more ambiguous, and arguably difficult, task of deciding where to place strokes when aiming to be “slightly off-the-beat” rather than squarely on-the-beat. In the paper, we suggested that this greater variability reflected the fact that musicians’ subjective beat-bins widened slightly when they were thinking in a laid-back or pushed manner.

Regarding sound differences between the conditions, we found that, for the guitarists, laid-back strokes had slower attacks, longer durations, and lower SC frequency compared to both pushed and on-the-beat strokes, potentially achieved by a variety of techniques (“sweeping/sliding out” the strokes). The results are in accord with a simple P-center interpretation, in that these directional changes in sound features might either increase or decrease the distance between the P-centers of the guitar sound relative to the timing reference or the musicians’ subjective sense of the beats (that is, the quarter note pulse of the meter), thus enhancing the perceived early/late asynchronous character of strokes. To achieve a laid-back feel, playing strokes with late onset/peak timing in tandem with a slower attack/longer duration/darker timbre would move the P-center later in time and enhance the perceived asynchrony. Conversely, playing strokes with early onset/peak timing in tandem with a faster attack/shorter duration/brighter timbre would move the P-center earlier in time and therefore increase the perceived synchrony for on-the-beat strokes or early asynchrony for pushed strokes.

As for the bassists, we found that pushed strokes were played with a greater intensity than on-the-beat and laid-back strokes on average. While early strokes that are played louder would theoretically emphasize the asynchrony between the bass and the metronome/subjective beat in terms of P-center, louder leading tones in an asynchronous pair tend to mask ensuing lagging tones to a certain extent, thus decreasing the salience of the asynchrony (Goebl and Parncutt, 2002). We therefore suggest that, when playing with an intentional pushed feel, bassists might utilize louder strokes to ensure that they are heard as early in relation to the previous beat, emphasizing the microrhythmic earliness of a pushed stroke in relation to the metronome reference by increasing its intensity and thus pushing against the beat reference. For a laid-back feel, softer strokes would conversely prevent the bass from being perceived as taking the lead over the corresponding timing reference sound. However, the difference between on-the-beat and laid-back was not significant.

Overall, this paper's findings lend support to the hypothesis that both temporal and sound-related aspects are important to musicians in order to communicate an intended timing feel in performance, and that they may be related to perceptual effects whereby the lateness/earliness of sounds is enhanced/attenuated via strategic timing-sound combinations by performers.

5.1.2 Paper II

Guilherme Schmidt Câmara, Kristian Nymoen, Olivier Lartillot, and Anne Danielsen, "Timing Is Everything . . . or Is It? Effects of Instructed Timing Style, Reference, and Pattern on Drum Kit Sound in Groove-Based Performance," in *Music Perception* 38(1), 1–26, 2020.

Abstract. This study reports on an experiment which tested whether drummers systematically manipulated not only onset but also duration and/or intensity of strokes in order to achieve different timing styles. 22 professional drummers performed two patterns ([1] a simple "back-beat," and [2] a complex variation) on a drum kit (hi-hat, snare, kick) in three different timing styles (laid-back, pushed, on-the-beat), in tandem with two timing references (metronome and instrumental backing track). As expected, onset location corresponded to the instructed timing styles for all instruments. The instrumental reference lead to more pronounced timing profiles than the metronome (pushed strokes earlier, laid-back strokes later). Also, overall the metronome reference led to earlier mean onsets than the instrumental reference, possibly related to the "negative mean asynchrony" phenomenon. Regarding sound, results revealed systematic differences across participants in the duration (snare) and intensity (snare and hi-hat) of strokes played using the different timing styles. Pattern also had an impact: drummers generally played the rhythmically more complex pattern 2 louder than the simpler pattern 1 (snare and kick). Overall, our results lend further evidence to the hypothesis that both temporal and sound-related features contribute to the indication of the timing of a rhythmic event in groove-based performance.

In this paper, we focused on the data from the drummers (N=20) in the performance experiments outlined in section 4.1. The aim here was the same as in paper I—to discern whether drummers on average systematically manipulated not only the (onset) timing of strokes but also other sound features (duration and intensity) when producing the different instructed microrhythmic feels (or “timing styles”) of laid-back, on-the-beat, and pushed. Here too, we hypothesized that drummers might express the lateness/earliness of strokes relative to timing references via the manipulation of sound features that have been shown to influence the perceived timing of events in experimental P-center and asynchrony/anisochrony studies. We also included the influence of two musically related factors that might further affect performances: “pattern,” where drummers played both a “simple” backbeat pattern and a slightly more “complex” variation, and two external “timing references” comprised of either short percussive woodblock “metronome” sounds or a guitar and bass “(instrumental) backing track.” We only used onset data in our analysis for this paper, since onset and peak tend to be very close to one another for percussive drum sounds (Danielsen et al., 2019).

Compared to the laid-back and pushed conditions, the drummers on average played stroke onsets closer to the timing references in the on-the-beat condition, as expected. However, they tended to anticipate their on-the-beat stroke onsets relative to the timing references (displaying NMA), especially when playing along to the metronome. On-the-beat stroke onset NMA decreased overall when they were playing along to the sounds of the instrumental backing track. We saw this as related to the fact that guitar and bass sounds have slower attacks and longer durations than the metronome sounds, and thus they likely afford later P-centers. Drummers may therefore have placed their own strokes slightly later when playing along to the backing-track rhythms to achieve greater perceptual synchrony.

As for laid-back and pushed strokes, the average onset location for all drum kit instruments was significantly later (between 21 and 27 ms) and earlier (between -37 and -42 ms), respectively, compared to the on-the-beat strokes across all patterns and timing references. Interestingly, the average pushed and laid-back onset values presented a slight asymmetry in magnitude—that is, pushed stroke onsets tended to be slightly more anticipatory than laid-back strokes were delayed on average. We saw this as related to the relatively greater unfamiliarity and thus difficulty reported by drummers regarding pushed versus laid-back playing, which may have led them to exaggerate stroke onsets in the former. We also speculated that the difference could have been more aesthetic or perceptual nature—that is, greater degrees of onset asynchrony might be required to create the perception of a pushed (vs. laid-back) feel.

Timing reference also had a further influence on average onset location, whereby the backing-track sounds generally led to even more pronounced early and late onset timing for pushed and laid-back strokes, respectively. We suggested that the drummers might have done this to ensure that their intentionally early/late asynchronous sounds were not as readily masked, spectrally or dynamically, by the relatively wider instrumental bass and guitar sounds, which had slightly slower attacks and longer durations than the metronome sounds. Seen from a beat-bin perspective (see section 3.1.2), the instrumental sounds may have afforded a wider subjective beat-bin for drummers than the metronome sounds, thus requiring them to increase the magnitude of the onset asynchrony in order to remain just barely at the edges of the wider beat boundaries.

Regarding sound-feature differences, while we found main effects of intensity and duration on timing style across all pattern and reference conditions, not all instruments applied the same *direction* of sound-feature change to uniformly distinguish either laid-back or pushed strokes from on-the-beat strokes. Regarding duration, laid-back snare strokes were, as expected, played on average with slightly longer total durations (which generally led to later P-centers). Laid-back kick-drum strokes, on the other hand, were played with slightly shorter average total durations (which led to earlier P-centers). We linked this to the different musical functions of the beats played by the kick and the snare—that is, short kick strokes combined with long snare strokes may increase the perceived time interval between the end of the kick and the beginning of the snare and could be heard to enhance the perceived lateness of the snare strokes. We also offered explanations as to how drummers achieve these subtle duration differences in the kick and snare drum via their manipulation of the degree of stick or pedal bounce versus muffling (or “choke”).

Regarding intensity, while snare strokes were generally played with higher intensity in the laid-back condition on average relative to the pushed (significant difference) and on-the-beat conditions (numerical difference), the hi-hat strokes were played with higher intensity in both laid-back and pushed conditions relative to the on-the-beat conditions (both significant differences). The late and loud snare strokes reinforce previous findings by Danielsen, Waadeland, and colleagues (2015), who suggested that late plus loud snare combinations may more effectively convey the presence of an intentionally produced late asynchrony against a steady timing reference (such as a metronome). Indeed, findings from perceptual studies suggest that a late sound combined with greater intensity increases the detectability of a late event relative to preceding sound events due to a smaller forward-masking effect (Gobel and Parncutt, 2002).

Hi-hat strokes, on the other hand, were played louder in both pushed and laid-back conditions compared to the on-the-beat condition. We saw this as related to its assumed role as the main timekeeper of the drum kit in groove-based musical contexts in both asynchronous conditions. Greater intensity would generally increase its perceptual salience to the benefit of other ensemble instrumentalists. When pushing against a steady external rhythm, an early, loud hi-hat might clearly communicate a pushed feel to the rest of the ensemble, and a late, loud hi-hat might clearly communicate a laid-back feel.

The two patterns also led to differences in both sound and timing features. For example, snare and kick strokes were played louder, and hi-hat strokes earlier, in pattern 2 than in pattern 1. Since the main difference between the two patterns was structural in nature, it would appear that the differences can more generally be attributed to the greater complexity and note density of the extra snare pick-ups and kick syncopations in pattern 2.

Overall, this paper’s findings represent further evidence that performing a simple groove-based pattern with a particular microrhythmic feel seems to involve more than just creating stroke onset asynchronies between instruments and/or timing references. At the same time, in the case of drum-kit playing, onset timing manipulation does appear to be the most effective means of distinguishing laid-back and pushed from on-the-beat performances. Unlike the results for the guitarists and bassists in paper I, the results for the drummers are complex, in that (1) laid-back strokes with late onsets were not *only* combined with longer duration and lower intensity (which would theoretically enhance perceived asynchronies by increasing P-center distances between drum sounds and timing reference sounds/subjective beat locations), and (2) pushed strokes were not *only* combined with shorter duration and greater intensity (which would conversely decrease P-center distances). Other combinations of onset location and intensity were used, most likely to make off-the-beat rhythmic events and/or layers more salient. This complexity probably results from the fact that drum-kit performance entails creating more complex combinations of interacting sound events: in addition to the asynchronies created by each individual drum instrument with the external timing references/subjective reference structures, further inter-instrument asynchronies can be created when snare and kick coincide with hi-hat strokes at the beat locations of the meter. While this paper's findings of statistical average trends do show that durational and intensity manipulations are used in some way to differentiate laid-back and pushed from on-the-beat performances, exactly how they serve to enhance either "laid-backness" or "pushedness" warrants closer study. We therefore attempted to map out some of the more complex inter-drum instrument timing and sound relationships in paper III.

5.1.3 Paper III

Guilherme Schmidt Câmara, Georgios Sioros, and Anne Danielsen, "Mapping Timing and Intensity Strategies in Drum-Kit Performance of a Simple Back-Beat Pattern," submitted to the *Journal of New Music Research,* 2020.

Abstract: In this article, we identify various ways in which twenty-two drummers express a simple "backbeat" pattern with different intended timing styles (laid-back, on-the-beat, pushed) via manipulation of stroke onset and intensity features. Based on hierarchical clustering analyses and phylogenetic tree visualizations, we found three main strategies used to distinguish pushed/laid-back from on-the-beat performances: (1) strong "general earliness/lateness" strategies, where most instruments are consistently played earlier/later in time relative to a metrical grid; (2) subtler "early/late flam" strategies, where most instruments are played synchronously with the grid but at least one instrument is distinctively played as an early/late flam against both its simultaneous drum stroke and the grid; and (3) even subtler "ambiguously early/late compound sound" strategies, where two instruments are played synchronously in relation to each other as a compound sound, but one is played synchronous with the grid, while the other is played early/late. The majority of drummers utilized additional consistent intensity strategies, the most common being via application of greater hi-hat or snare intensity, however, intensity was not necessarily applied uniformly to exclusively distinguish laid-back/pushed from on-the-beat timing but instead as a potential means to enhance or diminish the effect of off-the-beat rhythmic elements more generally.

In this paper, we once again focused on the data from the drummers (N=20) in the performance experiments outlined in section 4.1. Our findings in paper II indicated average trends of onset timing and sound-feature changes between microrhythmic feels across the entire sample of drummers.[22] However, these trends do not necessarily represent how these changes were actually implemented in the individual performances per se. In paper III, then, we conducted an exploratory investigation where we attempted to map out the range of strategies actually used by the twenty drummers in order to distinguish laid-back and pushed from on-the-beat performances while confining the scope

---

[22] Additional results of within-subjects tests (RMANOVAs) also showed that drummers on average systematically manipulated acoustic features other than onset to distinguish performances of basic groove-based patterns played with intentionally "laid-back," "pushed," and "on-the-beat" microrhythmic feels (or "timing styles"). However, these results were not reported in the paper, as they fell outside its primary focus on trends across all drummers.

of the analysis to the simple backbeat pattern (pattern 1) and stroke onset and intensity (total stroke SPL) features.[23] Here, we wanted to specifically explore the different asynchronies between the drum kit and the beats of the meter, as well as between the drum instruments themselves, that the drummers utilized to convey different types of microrhythmic feel. In addition, we explored how they utilized stroke intensity changes between laid-back/pushed and on-the-beat performances to further modulate the qualitative effects of onset timing features in different ways.

Based on a methodology first developed by Sioros, Câmara, and Danielsen (2019), we used a hierarchical cluster analysis of the various onset and intensity features combined with phylogenetic tree visualizations to provide an overview of the strategies used by the drummers to distinguish laid-back and pushed from on-the-beat performances (see section 4.4.2). In the separate onset and intensity similarity matrices and trees, each drummer's individual feature profile allowed us to identify how consistently a single drummer performed strokes earlier/later relative to the grid or another instrument, or softer/louder relative to their own beat performance, at various beat locations of the meter. We further computed "archetypes" to summarize the general characteristics of each onset and intensity cluster by averaging the probabilities of the features of the drummers in each cluster. The archetypes were then symbolically visualized next to clusters in the trees and serve as a guide for roughly identifying a given drummer's strategy and tactics.

Note that, in this paper, the "grid" was conceptualized as a subjective reference structure representing where drummers might have perceived the location of the main quarter-note beats of the 4/4 meter. We based this reference on the onset timing of drummers' own hi-hat strokes in the on-the-beat condition rather than the timing of the external metronome or backing track timing references themselves. This is because we were interested in discerning how drummers distinguished on-the-beat performances from laid-back ("behind-the-beat") and pushed performances on an *individual* basis. We specifically chose to use the hi-hat onset timing in the on-the-beat condition as a proxy for the average location of a given drummer's perceived metrical grid, since findings from paper I as well as other studies (Fujii et al., 2011) show that the hi-hat is often the drum kit instrument most closely synchronized to external timing references (that is, showing the smallest NMA)—therefore, it likely serves as the main "timekeeper" for drummers. Furthermore, we used the mean variability (2 x SD) of a given drummer's own stroke onset and intensity reference values in the on-the-beat condition as the "tolerance" threshold. That is, for laid-back and pushed strokes to be categorized as early/late or louder/softer on any given metrical beat, they had to occur 95 percent outside of the mean distribution of the on-the-beat reference values. This was a conservative approach that assumed that such strokes were played with purposeful onset and intensity differences rather than occurring due to error or chance.

Regarding the results of the onset feature analyses, laid-back and pushed performances were generally characterized by some form of consistent late or early timing in relation to the drummers' perceived metrical grids, respectively. This was anticipated based on group average onset findings from paper II, as well as previous drum performance studies (Danielsen, Waadeland et al. 2015; Kilchemann and Senn 2011). However, we found that drummers utilized a variety of distinct strategies involving the combination of instrument-grid and inter-instrument onset asynchrony

---

[23] While average changes in stroke duration were also found in paper II, we chose to focus on onset and intensity features in this paper as they presented magnitude differences between timing styles that are both more perceptible and potentially more readily controllable for the drummers. Since this was an exploratory investigation, limiting the number of features included in the cluster analysis also helped to mitigate the overall complexity of the interpretation of the results, allowing us to try and identify general trends more easily.

relationships and identified three overarching categories of them: (a) a “general earliness/lateness” approach, where drummers played most, if not all, of the instruments earlier/later in time relative to the grid and in synchrony with each other (that is, presenting no clear inter-instrument asynchrony); (b) an “early/late flam” approach, where drummers played most of the instruments in sync with the grid and each other but also played at least one instrument early or late relative to the grid (that is, presenting a clearer inter-instrument asynchrony); and (c) an “ambiguously early/late compound sounds” approach, where, within a simultaneous kick/snare + hi-hat stroke, the onset distance between the two instruments was not large enough to be considered asynchronous relative to the individual drummers’ tolerance thresholds but one instrument was played in closer synchrony with the grid while the other was played early/late.

We also observed variations of these main onset strategies. One particularly popular example was the “synchronized downbeats + late back-beats” variation, where the hi-hat/kick were played more on-the-beat relative to the grid on beats 1 and 3, but either the hi-hat and/or the snare was played late relative to the grid on beats 2 and 4. This strategy was used by half of the drummers (ten out of twenty) to distinguish laid-back from on-the-beat performances against the backing track, which recalls the observation from empirical groove performance studies that drummers often favor the use of delayed snare strokes on beats 2 and 4 (a “backbeat delay”) when they are conveying a laid-back feel in an ensemble performance context (Butterfield 2006; Frane 2017; Iyer 2002).

Regarding intensity strategies, our analyses showed that drummers did not use greater or lesser intensity *uniformly* in order to exclusively distinguish either laid-back or pushed from on-the-beat timing feels. As was the case with onset timing, a range of intensity strategies arose within the group. However, the most common strategies were those that consistently applied greater hi-hat or snare intensity, which is in line with the average group statistics results in paper II, in the sense that consistently louder hi-hat or snare strategies in pushed and laid-back performances were more common than softer ones. The role of intensity as a mediator for communicating intended timing feel in drum performances thus once again proved to much more than simple stroke-onset manipulation. This is because, even in the case of laid-back and pushed performances that displayed consistent early/late onset timing profiles, an increase in any given instrument’s intensity both increased its own perceptual salience in the sonic output of the performance and enhanced or diminished the salience of the other instruments’ earliness/lateness, depending on the type of onset strategy present as well as the context in which the performance was heard (that is, solo or against an external reference). One interesting observation was that, in the instances of pushed and laid-back performances where drummers did *not* utilize any consistent onset asynchrony strategy, most of them saw the consistent manipulation of intensity features to distinguish those conditions from their respective on-the-beat performances. In some of these performances, it would appear that increasing the intensity of onset features with borderline early/late tendencies can emphasize their slightly offbeat asynchronous character.

We also assessed the different qualitative effects that these various onset and intensity strategies engendered by listening to and comparing individuals’ laid-back and pushed performances with their on-the-beat performances. In most cases, the pushed and laid-back performances were clearly conveyed. Ultimately, however, how these performances would be perceived by listeners will vary depending on which given instrumental layer one chooses as a “primary” reference in establishing a subjective sense of pulse and meter—a process that, in turn, might be influenced by the listening context and experience of the listener.

Overall, the clustering and phylogenetic analyses showed that drummers utilized a range of different onset-timing strategies to convey laid-back and pushed microrhythmic feels via various combinations of onset asynchronies relative to a subjective metrical grid/external timing reference and relative to the interaction among the constituent instruments of the drum kit. Regarding intensity strategies, while no clear patterns emerged, it does appear that intensity manipulations can serve as a means of enhancing or diminishing the effect of intentionally produced asynchronous rhythmic relationships, though this is not always the case.

5.2 Contributions

In chapter 1, I established that the main concern of this thesis was to explore how musicians communicate different intended microrhythmic feels in groove-based performance. I will elaborate on the findings below in relation to the thesis's main and sub-research questions.

Main question. How do temporal and sound-related features interact in the performance of simple groove-based patterns?

Main hypothesis. Musicians systematically manipulate sound features such as intensity, duration, and frequency of rhythmic events, in addition to onset timing location, in order to communicate different desired microrhythmic styles/feels.

Findings: The interaction between timing and sound features has only been investigated in one previous study to date (solo snare drum onset, intensity, and SC: Danielsen, Waadeland et al., 2015). The performance experiments in this thesis show that sound features can be important to the overall feel resulting from groove-based microrhythm. They are the first of their kind to explore the relationship between (a) onset timing and duration, as well as the effect of timing reference, in drum performance; (b) onset timing and duration, intensity, and brightness in guitar and bass performance; and (c) effect of pattern in drum, guitar, and bass performance. Findings from this thesis's papers suggest that all three instrumentalist groups—guitarists, bassists, and drummers—do in fact systemically manipulate both the physical timing (onset) and at least one other sound feature in order to distinguish between preconceived laid-back, on-the-beat, and pushed microrhythmic feels in groove-based performance. As such, they support the main hypothesis and contribute to the field of empirical groove performance and rhythm studies more generally.

Furthermore, while numerous analytical investigations of microrhythm in groove-based music have shown that musicians tend to display varying degrees of onset beat delay and anticipation (see section 2.3.3), this thesis's findings provide direct evidence of musicians' high degree of intentional control of onset manipulation in the performance of groove-based microrhythm. Our results are in accord with findings from a handful of previous experimental instructed-timing studies demonstrating this high degree of control in drummers (Danielsen, Waadeland et al., 2015; Kilchenmann and Senn, 2011) and bassists (Reinholdsson, 1987). The investigation in paper I, on the other hand, is the first controlled performance experiment, to our knowledge, demonstrating this characteristic for guitarists as well.

Sub-question A: What are the characteristic timing and sound-feature interactions in the different instrumental settings?

Sub-hypothesis A: While we had no particular hypothesis regarding any given instrument in particular, we expected that timing–sound strategies would be governed by the constraints of the instrument. For instrumentalists such as drummers, who can produce several simultaneous rhythmic layers with different parts of the drum kit, we expected to encounter a rich diversity of inter-instrument onset-timing and sound-manipulation strategies. In line with findings by Danielsen, Waadeland, and colleagues (2015), we also expected that drummers would introduce greater intensity to laid-back strokes. An additional/alternative hypothesis was that intensity would be used to make both laid-back and pushed strokes stand out as “off-the-beat.” We furthermore hypothesized that bassists and guitarists would similarly incorporate strategies for systematic sound manipulation in order to communicate various desired microrhythmic temporal beat profiles.

Findings: The average group timing–sound strategies for the different instrumentalist groups across all conditions (in terms of statistically significant differences) were as follows:

- Guitarists:
    - Laid-back = slower attack + longer duration + lower SC
    - On-the-beat/pushed = faster attack + shorter duration + higher SC
- Bassists:
    - Pushed = greater intensity
    - On-the-beat/laid-back = lower intensity
- Drummers:
    - Snare:
        - Laid-back = greater intensity + longer duration
        - On-the-beat = shorter duration
        - Pushed = lower intensity
    - Kick drum:
        - Laid-back = shorter duration
        - Pushed = longer duration
    - Hi-hat:
        - Laid-back/pushed = greater intensity
        - On-the-beat = lower intensity

As we expected, the ways in which the musicians specifically combined physical timing and sound features appeared to be dependent upon contextual factors such as instrumental constraints (number and range of manipulable sound features) and musical context (pattern, timing reference), as well as individual player preference (and perhaps basic motor-physical limitations/constraints). All of these factors influence the ways in which physical sound–timing interactions can be/are explored in performance by musicians, which in turn influence how they are perceived by both performers and musicians alike.

The results for the drummers were in accordance with those of a previous instructed-timing experiment by Danielsen, Waadeland, and colleagues (2015), in that intensity was systematically used on average to distinguish off-the-beat from on-the-beat microrhythmic feels. Paper III also demonstrated the diverse range of onset timing and intensity strategies among the drummers at the individual participant level—that is, we found multiple iterations of onset asynchrony among the drum-kit instruments as well as relative to the metrical grid, likely owing to both individual player preference and skill level. Though it was not directly investigated in this thesis, the guitarists and bassists also presented a range of individual timing–sound strategies beyond the average group trends discerned in paper I.

Overall, the findings regarding timing–sound strategies further contribute to the scholarly fields of groove and rhythm studies. They also have pedagogical value, in the sense that, by explicitly shedding light on typical timing–sound combinations used in performance, they reveal important facets of the tacit knowledge that skilled groove-based musicians accumulate through extended practice. As such, students of groove-based performance may capitalize upon this knowledge to improve, or expand upon, their microrhythmic performance skills. For example, guitarists attempting to achieve a laid-back feel might try to think in terms of not only late stroke-onset positioning but also longer durations and maybe a darker tone (and vice versa for a pushed feel).

Sub-question B. To what extent do sound–timing interactions used in performance correspond to perceptual sound–timing interactions?

Sub-hypothesis B. Perceptual studies suggest that, in rhythmic contexts, the perceived timing of a sound event is affected by its shape/acoustic envelope. Based on findings from P-center studies, we hypothesized that laid-back events would be characterized by slower/longer (maybe softer/darker) strokes and pushed events by faster/shorter (maybe louder/brighter) strokes.

Findings: Based on P-center findings of musical sounds (see section 3.1), one could theoretically expect musicians to combine laid-back strokes with relatively later onset timing and slower attack/longer duration/lower intensity/lower SC, and pushed strokes with earlier onset timing and faster attack/shorter duration/greater intensity/higher SC, compared to on-the-beat strokes. In the on-the-beat condition, we assumed that the musicians would aim their strokes so that they would be perceived as vertically synchronous relative to the sounds of the timing references (that is, with P-centers closely aligned). In the laid-back and pushed conditions, on the other hand, we expected that the musicians would seek to produce sounds that enhanced the perceived asynchrony between drum and timing-reference sounds (that is, with P-center distances increased/less aligned) and came across as relatively later or earlier than on-the-beat strokes, respectively.

The group trends from paper I for the guitarists and bassists agreed with this prediction, in that, on average, they utilized timing–sound combinations that would in theory enhance the perceptual asynchronous character of laid-back and pushed strokes (see research sub-question A above). However, the results for the drummers did not agree with a simple P-center interpretation, since the laid-back and pushed strokes were not uniformly combined with sound features that would lead to perceptual redundancy gains (Danielsen et al. 2019; see section 3.1). This might be because the drum instruments themselves constrain the palette of possible sound–timing interactions. Accordingly, more complex group and individual archetypical combinations of onset, duration, and intensity were utilized, which might more generally serve to distinguish both laid-back and pushed from on-the-beat strokes by making them “stand out” as perceptual accents (Danielsen, Waadeland et al. 2015; Palmer and Drake, 1993).

By shedding further light on the microrhythmic dynamics and nuances of timing and sound in groove-based performance, this thesis not only contributes to the scholarly research on groove but also benefits music makers and producers seeking to “humanize” computer-programmed grooves that may otherwise sound mechanical or deadpan. The results from these performance studies also shed light on the interaction between the systematic effects of acoustic factors on perceived timing found in controlled perceptual experiments and contextual factors such as instrumental constraints, musical context, and individual training and preference—that is, on the ways in which psychoacoustic mechanisms unfold under the constraints of actual musical performance.

Chapter 6

Discussion and Conclusions

The main concern in this thesis was to explore how temporal and sound-related features interact in the performance of simple groove-based patterns, focusing on three different microrhythmic feels: “laid-back,” “on-the-beat,” and “pushed.” I investigated this interaction for three different instruments that are typically used in a groove-based setting: drums, electric bass, and electric guitar. I was also interested in the extent to which timing–sound interaction in performance accords with findings from perceptual studies of the P-centers of rhythmic events, which suggest that both the physical timing (onset, peak) and the sound features (duration, intensity, frequency) of rhythmic events influence their perceived timing in relation to other physical rhythms and/or subjective reference structures (such as pulse and subdivisions). In section 6.1, I will discuss some of the central topics of this thesis more generally in relation to the papers’ findings. In section 6.2, I will present some general conclusions; finally, in section 6.3, I will offer some thoughts on avenues for future study.

6.1 General Discussion

In the following sections, I will discuss some of the central topics of this thesis more generally in relation to the papers’ findings.

6.1.1 Combined Timing–Sound Strategies as a Means of Conveying Microrhythmic Feel in Groove Performance

The overarching theme explored in this thesis’s papers (I–III) concerns what musicians do when they intend to play a rhythm with a particular preconceived “timing,” or microrhythmic feel—in either on-the-beat, pushed (“ahead of the beat”), or laid-back (“behind the beat”) fashion—as well as why they do what they do. These conditions immediately raise the question of what constitutes the “beat” against which said musicians are to play ahead of, on, or behind. If it is simply considered to be the onsets of a regular external timing reference, then we might assume that musicians will simply want to “vertically” align, or “synchronize,” the physical timing (or onsets) of their own produced rhythms relative to the beat of the external timing reference, depending on the overall intended microrhythmic feel. However, just as “rhythm” can be defined as an interaction between external physical rhythms and internal subjective reference structures (see Chapter 2), the beat can be defined by way of both a physical component—that is, sound signals—and a subjective, perceptual component conceptualized as an internal, cognitive “timekeeper” of some sort that affords a “beat-bin,” or regular time window in which sounding beats are expected to occur periodically. Regardless of the true nature of the ever-elusive beat, findings from P-center and asynchrony/anisochrony studies (see Chapter 3) suggest that the perceived timing of sounds that regularly occur at any given beat location in a rhythmic context is not determined solely by their physical onset locations but also by other acoustic features, such as duration, intensity, and frequency. In performance, then, it may be the case that musicians align their own produced rhythms in such a fashion that the P-centers of their sounds and those of any external timing reference rhythms and/or subjective reference structures correspond to a preconceived microrhythmic feel.

As I explored in section 3.1.2, when musicians seek to play in an “on-the-beat” fashion, we might expect that they want to position their produced sounds in such a way that their P-centers are more closely aligned with the P-centers of the external timing reference sounds presented to them, thus ensuring that their sounds fall squarely within the subjective “beat-bin” afforded by such reference sounds (though always governed, of course, by motor-physical constraints and acuity of perceptual timing judgment). However, when they play with preconceived “laid-back” or “pushed” feels, they instead likely position their sounds so that their P-centers are less directly aligned with those of the external reference sounds compared to an “on-the-beat” approach, thus coming across as relatively more perceptually asynchronous or anisochronous. However, they should not be so early or late as to fall too far outside of a target beat’s temporal boundaries. In such cases, they could risk falling within the beat-bins of preceding or following beats belonging to higher or lower metrical levels, thus categorically coming across as instances of syncopations rather than as slightly anticipated or delayed “beats.” While it is clear from various perceptual experiments that conveying the lateness or earliness of a given rhythm can be effectively achieved through the application of greater magnitudes of onset asynchrony relative to a timing reference’s sounds alone, if the target sounds are further shaped via faster/slower attack and shorter/longer duration, in particular, and potentially greater/lesser intensity and/or higher/lower frequency, their perceived earliness or lateness can be further diminished or enhanced without altering the onset position.

Therefore, in a manner reminiscent of experimental listening contexts, performing musicians can presumably capitalize upon perceptual “redundancy gains” and “losses” (Danielsen et al., 2019) across the dimensions of timing and sound (whether consciously or subconsciously) in order to further convey an intentional pushed, laid-back, or on-the-beat microrhythmic feel to a given performed rhythm. They may do this by enhancing or attenuating the cumulative delaying effects of P-centers. For example, redundancy gains may be achieved by increasing the attack duration of a stroke played with a late onset. This might enhance its laid-back character, as both factors tend to lead to later P-centers. Conversely, shortening a stroke played early might enhance its “pushedness” (in both cases, relative to either physical rhythms or subjective reference structures). On the other hand, redundancy losses may be obtained when late strokes are combined with short attacks, or early strokes with long attacks, since in both cases the two factors tend to lead to opposite directional shifts in P-center. Musicians will, however, always be constrained by the range of sound features that they are able to physically manipulate given the construction and design of their instruments, as well as the particular musical pattern to be performed.

At the outset, all three instrumentalists (guitarists, bassists, and drummers) were in theory able to dynamically accentuate their strokes with greater or lesser intensity by simply striking strings/drum membranes with greater or lesser velocity. Regarding attack duration, because it was analytically defined as the time interval between onset and peak of a single instrument’s sound across all of the thesis’s experiments for consistency’s sake, it would appear that the guitarists, in particular, would have had more ability to manipulate it, since they were instructed to play a pattern comprised of four-note chords and thus could easily lengthen or shorten chords by playing individual notes either more slowly or more quickly in succession. On the other hand, both bassists and drummers—the former when plucking single strings to produce the instructed monotonic (single-tone) rhythms, and the latter when striking individual drum instruments (hi-hat, snare, kick drum) with either drum sticks or foot pedal—would appear to have been limited in their manipulation of the attack duration of their sounds. This is because once a single string or drum membrane is struck, the produced sound will generally reach its acoustic peak rather quickly and proceed to decay naturally unless further analog/digital envelope-shaping effects processors are used. As for decay duration (the time interval between signal peak and offset), guitarists and bassists are easily able to control tone length either by allowing strings to ring out and naturally decay depending on how intensely they are struck (we found moderate positive correlations between SPL and duration for both guitarists and bassists in paper I) or by muting with the fretting or striking hand post impulse, or both. Drummers would seem to have less direct control of a stroke's length via striking intensity (paper II revealed only significant weak correlations between stroke duration and intensity). In turn, we discerned that drummers might lengthen their strokes by controlling the degree of stick/pedal rebound, allowing "bounce" or "choking" the membrane by pressing the stick/pedal firmly on it post-impulse. As for timbre manipulation specifically in terms of the perceived brightness of the stroke (with which SC measurements tend to correlate positively), for guitarists and bassists, at least, striking the strings at different positions on the bridge (further away from or closer to it) can subtly affect the tone "color" and brightness of the produced sounds, though this would need to be further investigated in future mocap analyses of the data. Due to limitations with microphone signal leakage (see sections 4.2 and 4.3), we were unable to gauge brightness differences among drum instruments. Nevertheless, in their performance practice, drummers are clearly aware that striking drums at different locations on the drum skin or cymbal (closer or further away from the center and toward outer rim) can produce different timbres. Future mocap analyses combined with correlated frequency-related audio features may reveal whether drummers utilized certain drum locations to distinguish between timing styles.

Regarding the average timing–sound strategies found for each instrumentalist group, then, the guitarists seemed to have more access to the further manipulation of all three sound parameters measured (duration, intensity, and frequency/brightness), in addition to the onset timing location, because they were instructed to perform a backbeat pattern comprised of polyphonic chords. In paper I, the results of the statistical tests revealed that the general trend across all guitarists was to utilize at least two of these sound parameters (duration and SC) in addition to onset/peak timing in order to achieve and distinguish between the three instructed microrhythmic feels. Laid-back strokes were played on average not only with later onset but also with slower attack, longer decay/total duration, and a darker SC; conversely, on-the-beat and pushed strokes were played with more synchronous and earlier onset timing, respectively, and combined with relatively faster attacks, shorter durations, and brighter SC compared to the laid-back strokes. The guitarists, that is, tended to utilize "positively" correlated average timing and sound strategies that exploit perceptual P-center redundancy gains and further enhance the perceived asynchrony of late and early strokes relative to the "beat" in the laid-back and pushed conditions compared to the on-the-beat condition. Why guitarists did not also further manipulate the intensity of strokes in accordance with findings from either P-center or asynchrony/anisochrony studies remains unclear. However, it may be that intensity changes do not necessarily affect either P-centers or synchrony thresholds as effectively, or that, within the context of the musical pattern, guitarists did not find it either aesthetically appropriate or mechanically intuitive to play strokes louder or softer in any of the given timing style conditions.

The bassists, as mentioned above, were limited in their ability to manipulate attack duration because they were instructed to play single-note patterns as opposed to chords (which are highly atypical for the bass in groove-based performance traditions). The statistical tests revealed that the average trend across all participants was chiefly to further manipulate the intensity of tones in addition to onset/peak location. More specifically, we found a tendency for bassist to play pushed and on-the-beat strokes with greater intensity than laid-back strokes. Though previous findings of the effects of intensity on the P-centers of musical instrumental sounds have been few and somewhat inconclusive (see section 3.1.1), most studies indicate that greater intensity tends to invite an earlier perception of P-centers. Therefore, combinations of early and loud, and late and soft, strokes may also be considered timing–sound strategies that exploit perceptual redundancy gains, since both would theoretically invite the perception of the P-centers of the bass strokes as earlier and later relative to physical or subjective beat reference structures, respectively. In addition, a few asynchrony/anisochrony threshold studies suggest that intensity levels affect how synchronous/isochronous two sounds are perceived to be (Goebel and Parncutt, 2001; 2003; Tekman 1997; 2002; see also section 3.2.1), whereby in simple listening contexts, early and loud tone-pair combinations tend to sound more synchronous than late and loud ones. In the more complex listening context of our experiments' performances, it is possible that by playing pushed strokes with greater intensity, the bassists' rhythms were perhaps more clearly accentuated and thus heard as leading the reference track rhythms, whereas in the laid-back performances, playing late strokes softer sets up the opposite—the bass is heard as following the reference rhythms.

As for the drummers' timing–sound strategies, the results of both the average group trends (paper II) and the individual strategies (paper III) were perhaps the most "ambiguous" in terms of predictions and findings from previous P-center and asynchrony/anisochrony studies. Though the manipulation of onset location was clearly the most effective means of conveying microrhythmic feel, stroke intensity as well as duration were also readily utilized by drummers, though not in a fashion that might consistently exploit perceptual redundancy gains and thus potentially enhance the lateness or earliness of strokes in the laid-back and pushed conditions, respectively. In paper II, we found significant differences in intensity between group *average* trends for both hi-hat (pushed louder) and snare (laid-back louder). The follow-up analyses of paper III revealed further nuances at the individual-drummer level. As such, it would appear that, when conveying a particular timing feel with a highly percussive instrument such as the drum kit, which is clearly more limited in terms of sound-parameter manipulation than string instruments, intensity might be instead utilized as a way to accentuate "offbeatness" more generally (whether laid-back or pushed) relative to an on-the-beat feel. Intensity is then primarily used to make a pushed or laid-back stroke into a perceptually salient event that "stands out and captures a listener's attention" (Drake & Palmer, 1993, p. 344).

One interesting finding regarding onset timing location was that, on average, drummers showed a tendency to play pushed strokes with larger onset magnitude differences (relative to on-the-beat strokes) compared to laid-back strokes. We suggested a few reasons why drummers produced greater asynchronies to achieve a pushed feel. On the one hand, pushing against the beat of the timing reference may simply be a comparatively more difficult or ambiguous task for drummers, which might be demonstrated by its greater variability (20 ms for pushed vs. 14 ms for laid-back and 9 ms for on-the-beat strokes). Pushed in particular was more often reported by drummers in the post-experiment interviews to be the most difficult or unfamiliar condition. On the other hand, the explanation may be more perceptual in nature—that is, greater degrees of onset asynchrony might be required to effectively communicate the perception of the drum instruments pushing against the reference in comparison to a laid-back feel, and we suggested that drummers may also display a decreased sensitivity to early as opposed to late asynchronously produced rhythmic events in the backbeat-based pattern context. To our knowledge, no studies of onset asynchrony/anisochrony have thoroughly compared JNDs of off-the-beat early versus late events in an isochronous rhythmic context. However, two studies involving single local anisochronous displacements of a monotonic sequence that have touched upon the matter show no consensus: Wojtczak and colleagues (2017) found no differences between early and late displacements at an IOI rate of 500 milliseconds, while Hibi (1983) found that, at an IOI rate of around 330 milliseconds, JNDs for early displacements were significantly higher than those for late displacements (around 2 percent higher relative JND).[24]

Relatedly, regarding asynchronous events, Repp (2003, 2004, 2006) conducted three sensorimotor synchronization (SMS) tapping experiments wherein participants were instructed to tap in synchrony along to an isochronous "target" sequence while ignoring a simultaneous "distractor" sequence with onsets aligned off-phase (either early or late at various magnitudes). He found that participants' taps tended to be shifted in the direction of the distractor tones' onsets. Interestingly, the attraction effect was asymmetrical, in the sense that the tapping shifts were far more attracted to leading rather than lagging distractor target tones. It remains unclear why this asymmetric effect occurred, and Repp (2003) simply suggests that it may be attributed to "temporal precedence"—that is, simply because leading tones arrive first, they are more difficult to ignore and thus have a stronger attracting effect, whether they are distractor or target tones. The distractor effect was also found in a wide range of IOIs, though it mainly occurred when tone onsets were within 0 to 120 milliseconds of one another, or within a temporal window where tones were likely "perceptually grouped" together and "jointly engage the [internal] phase correction process that helps maintain synchrony between the taps and the target sequence" (Repp, 2004, p. 410). (Note that this value is similar to the smallest categorically perceived unit subjective metrical subdivision suggested in music perception of around 80 to 100 milliseconds [London, 2012; Iyer 2002: Polak 2010; Haugen and Danielsen 2020].) On the other hand, the asymmetrical tendency to produce pushed strokes earlier than laid-back strokes late was not present for bassists or guitarists. This might indicate that guitarists and bassists are more accustomed to pushing against timing references than drummers (thus favoring the first explanation for the drummers), and that this "pushing" tendency is aesthetically more appropriate for guitar and bass in medium-tempo backbeat-based contexts.

In retrospect, of course, we can see limitations to our analytical definition of attack duration in particular, concerning the way in which we focused on the attack of separate individual instruments in the drum kit. Because P-center studies have shown that attack duration has a stronger P-center shifting effect than other sound features, instruments that do not readily afford the physical manipulation of attack duration (such as bass or drums) may not be able to capitalize upon delaying effects. However, if onset asynchronies between individual instruments' rhythms are viewed as *simultaneously* creating compound sounds with an extended attack, then, in actuality, drummers are also able to employ multiple event articulations (that is, "flams") among the different instruments of the drum kit. The highly limited research on the P-centers of two-sound compound events suggest

---

[24] Which, interestingly, translates to a metrical beat interval roughly equivalent to our experiments' eighth-note level (312.5 milliseconds). However, Hibi (1983) found no differences at IOIs close to the quarter-note main pulse level in our experiments (625 milliseconds).

that P-centers tend to be delayed in the direction of the second event's onset, and potentially even more so when the second event is lower in frequency (Seton, 1989; Hove et al. 2007). While we did not consider the effects of compound drum sounds in our investigation in paper II, focusing instead on whether the drummers manipulated features of each individual instrument separately, the descriptive statistics of the average group trends across all pattern and reference conditions (table 2) indicate that the onset asynchrony magnitudes between any two drum instruments in all the three timing style conditions were not greater than around 10 milliseconds on average. Therefore, if these *average* inter-instrument onset-asynchrony trends are perceived as extended drum attacks, their effects are likely very subtle, according to most heuristics derived from simple psychophysical experiments (section 3.2.1). However, the descriptive statistics of the *individual* drummers' onset timing (paper II, appendix) do indicate a range of inter-instrument asynchrony magnitudes. In paper III, we further found that drummers displayed a range of consistent inter-instrument onset asynchrony relationships between hi-hat and snare/kick in pushed and laid-back compared to on-the-beat performances. Future listening experiments involving archetypical stimuli based on the various systematic inter-instrument onset-asynchrony strategies found there may reveal whether greater onset-asynchrony magnitudes correlate with later P-centers, as predicted by the two aforementioned compound-sound P-center studies, or whether certain instruments in the drum kit attract P-centers more than others based on their frequency/timbre and intensity level, among other potential sound factors.

Finally, we must acknowledge that if the combination of guitar/bass/drum strokes and reference track sounds is heard more holistically as a compound sound, then, in addition to creating vertical asynchronies relative to the reference sounds, laid-back and pushed strokes may also engender a certain degree of horizontal anisochrony. *How* non-isochronous each early or late asynchronous sound pair would be perceived relative to preceding and ensuing events, however, would depend not only on the onset/peak timing of either the lagging or leading sound (as noticed by Hove et al., 2007) but also on the P-centers of their compound sound—which, in turn, would be affected by the magnitude of the physical onset asynchrony between the constituent sounds and these sounds' duration, intensity, and frequency in myriad ways.

6.1.2 JNDs: Do Timing–Sound Differences between Instructed Microrhythmic Feels Matter Perceptually?

Psychophysical experiments tend to be conducted under simple and highly specific controlled listening contexts that rarely resemble real musical performances. As such, just noticeable difference (JND) threshold findings and predictions derived from experiments involving monotonic or two-tone stimuli may only serve as rough heuristic guidelines, offering partial and speculative, but not exhaustive, explanations for *why* musicians play the way they do in a given groove-based context, and what perceptual effects they may engender. At the same time, we may speculate upon the degree to which the produced timing–sound differences between the microrhythmic feels may be heard or felt by listeners, since both musicians and scholars have historically ascribed a great deal of significance to rhythmic devices of "timing" in groove-based traditions (see section 2.3) While heuristic onset asynchrony/anisochrony and intensity JNDs exist in the current psychoacoustic literature and will be examined in light of our experiments' results in this section, we did not find any relevant JNDs regarding SCs (for example, the threshold at which two sounds will tend to be perceived as “darker” or “duller” in comparison), thus they were not explored in relation to our experiments’ SC results in this thesis summary.

Onset

If we consider the group-trend averages[25] from papers I (guitar/bass) and II (drums) to be somewhat representative of the degree of intentional beat delay and anticipation that skilled musicians control at the microrhythmic level, then, at least in terms of stroke-onset placement, our findings suggest a high level of such control—a level that has only been demonstrated in controlled experimental settings in a handful of previous investigations (Danielsen, Waadeland et al., 2015; Kilchenmann and Senn, 2015; Reinholdsson, 1987) and that is thus not part of the sparse empirical groove literature on performance timing. At the experiments’ tempo (96 beats per minute, IOI quarter note = 625 ms), the difference in onset-location magnitude between both offbeat conditions (laid-back and pushed) and the on-the-beat condition ranged between 20 and 50 milliseconds across all three instrumentalist groups (guitarists, LvO = 35 ms, PvO = 37 ms, LvP = 72 ms; bassists, LvO = 44 ms, PvO = 46 ms, LvP = 90 ms; drummers [all three drum instruments], LvO = 21–27 ms, PvO = 37–42 ms, LvP = 58–69 ms). Comparing these values with the oft-utilized JND threshold of onset anisochrony—that of local displacement of a monotonic rhythmic sequence (about 2.5 percent of IOI or about 16 ms; see Friberg and Sundberg, 1995)—and allowing that acoustic onset cues alone determine the perceived timing of instrumental strokes, average pushed and laid-back strokes would likely be heard as just slightly earlier and later, respectively, than either the physical (metronome and backing track) or subjective timing references in relation to the on-the-beat strokes for all the instruments. If we further compare with slightly more conservative JND thresholds involving regularly occurring displacements that perhaps more closely resemble the performance contexts of our experiments—that is, either that of global cyclic anisochronous displacement in a monotonic context (about 5 percent of IOI or about 31 ms; ten Hoopen, 1994; see section 3.2.2) or that of asynchrony between two instruments (about 30 ms; Butterfield, 2010 [jazz rhythm section]; about 27 ms; Goebl and Parncutt, 2001 [piano tone dyads])—then the difference between laid-back/pushed and on-the-beat strokes relative to the timing references either just skirts these thresholds (guitarist and bassists, both LvO and PvO; drummers, PvO) or falls just short of it (drummers, LvO).[26] This would suggest that musicians are able to play flexible, slightly off-the-beat events that reside at the very boundaries of perceptual timing categories. As for the differences between the two opposite offbeat conditions laid-back and pushed (LvP), they are substantially larger than all these onset JND threshold values for all instruments.

While the abovementioned JND thresholds mainly gauge the extent to which one is able to subjectively detect a “rhythmic irregularity” in simple rhythmic contexts, whether in terms of horizontal anisochrony or vertical asynchrony, they do not necessarily translate to the threshold at which one’s subjective sense of pulse is simultaneously disturbed in similar contexts. Though it appears to be much less investigated, Madison and Merker (2002) provide heuristics for the JND of “pulse attribution” (the extent to which listeners can feel “a pulse in [the] sound sequence, such that

---

[25] While individual average onset timing and variability (SD) values naturally vary between individual musicians, thus likely producing effects ranging from stark to subtle, the group average values nonetheless indicate that skilled musicians are generally able to produce intentionally pushed and laid-back strokes with systematically earlier and later onset timing, respectively, compared to on-the-beat strokes.

[26] As noted in paper II (p. 17), based on these more conservative heuristics, pushed strokes would likely be slightly more detectable than laid-back strokes relative to the timing reference sounds or an underlying regular subjective “beat.”

[they] would be able to beat along with it" [2002, p. 203]). Though they also derived their results from simple monotonic sequences, they found that the average threshold for pulse attribution was much higher (about 8.6 percent of IOI) than the threshold for simply detecting a horizontal rhythmic irregularity in the same sequences (about 3.5 percent of IOI). When we compare our performance experiments' results with these JNDs on average, the average laid-back and pushed strokes mostly reside above the threshold for the detection of a rhythmic irregularity (about 22 ms), but they simultaneously fall just short of the threshold for disturbing one's sense of pulse (about 54 ms). This suggests that, while one might hear or feel rhythmic irregularities at the main metrical pulse level in the guise of asynchrony/anisochrony between or within instrumental layers, one would, at the same time, likely retain a sufficiently strong sense of pulse. Furthermore, like the JNDs derived from simple experimental listening contexts, actual JNDs for pulse attribution in real groove-based musical contexts may be higher as well, especially since groove-based musical styles tend to be comprised of repetitive rhythms that engender a strong and steady sense of pulse that is virtually isochronous despite the presence of physically anisochronous or asynchronous relationships between events—relationships with which enculturated listeners are already familiar in any case (see chapter 2). In addition, while Madison and Merker's experiment involved simple rhythms with no explicit musical "structural" information regarding meter, groove-based music always provides clear and stable metrical information that likely decreases one's sensitivity to larger IOI magnitudes or variability without disturbing one's sense of the beat. In our experiments, that is, the external timing references themselves always provided musicians/listeners with a highly isochronous background, and both the timing references and the patterns performed by the musicians supplied structural cues as to the meter, in terms of either tonality or timbral/durational/dynamic accentuation. Ultimately, the fact that the psychophysical literature indicates that even large performance onset asynchronies/anisochronies of up to 50 milliseconds or beyond in groove-based contexts may therefore be heard or felt without, in fact, disturbing one's sense of pulse (though the comparison is incomplete and limited) corresponds nicely with the purpose of microrhythmic strategies such as beat delay and anticipation in groove performance traditions: to play flexibly around the beat without challenging its primacy in order to convey different "feels" to the groove.

Duration and Intensity

Regarding whether the produced differences in duration may be perceived by listeners and/or the musicians themselves, we may look at the two instrumentalist groups that displayed significant differences between timing styles on average across all participants: guitarists and drummers. Based on JNDs of total tone duration from the psychoacoustic literature (about 12.5 percent of stimulus duration based on the average of the reported values from various studies reviewed by Johnson and colleagues [1987]), as noted in paper I (p. 1037), it is plausible that, for the guitarists, the average laid-back stroke would be heard as longer in terms of total duration than the average on-the-beat stroke (diff. = 35 ms, JND = 12.5 percent of 286 ms [average on-the-beat stroke total duration] = 36 ms), though the values are borderline. As for attack duration (which the guitarists also tended to prolong in the laid-back vs. on-the-beat strokes by 7 ms), while we were unable to find literature dealing with JNDs of attack-tone durations specifically, it is very possible that laid-back strokes struck in a more clearly arpeggiated or "swept" manner—so that the individual strings of the four-note chord are played more slowly (see section 4.3.2)—are perceived not necessarily as explicitly longer but as "texturally" different in comparison to the "swifter" nature of both the on-the-beat and pushed strokes. That is, listeners may perceive these various strokes as holistically different from one another rather than different exclusively in terms of the independently differing aspects of longer attack, decay and total duration (and this goes for all the other timing and/or sound features as well).

As for the drums (snare and kick), while we did not attempt to discern in paper II whether the average absolute differences in duration between timing style conditions would be perceptually salient, it appears they would be below the threshold according to the same heuristic JND thresholds.[27] For example, though laid-back snare strokes were played slightly longer and kick strokes shorter, the average difference in total duration between laid-back and on-the-beat conditions was sub-threshold (diff. = 7 ms, JND = about 18 ms: 12.5 percent x 145 ms [average on-the-beat stroke total duration]), as was the average difference in kick stroke decay duration between the laid-back and pushed conditions (diff. = 5 ms, JND = about 15 ms: 12.5 percent x 117 ms [average on-the-beat stroke decay duration]). However, as was the case for the guitarists, while the duration differences are admittedly far subtler in terms of absolute magnitude, it may be that, in combination with simultaneous changes in other sound parameters related to intensity timbre or frequency during the different stroke segments (such as SC, which we intend to investigate for the drums in the future), these durational differences may contribute to more holistic perceptual differences that might bring attention to strokes played early or late.

Regarding intensity, however, it would seem that the group-trend significant differences between timing styles found for the drums (snare and hi-hat) and bass would likely be evident (that is, perceptually softer/louder) according to an absolute 0.4–0.8 dB JND range for a 70–80 dB stimuli presentation level (Johnson et al., 1987; Jesteadt et al., 1977)—a level not uncommon in live-performed groove-music contexts. For the drummers, the average laid-back snare strokes were played with greater intensity than pushed strokes, with values above the lower JND range (+0.41 dB), whereas average differences in hi-hat stroke total SPL between the laid-back and on-the-beat conditions (+0.76 dB), as well as the pushed and on-the-beat conditions (+0.79 dB), were just above the lower range value and just shy of the higher range value. As for the bassists, the intensity differences between pushed and on-the-beat strokes (+1.49 dB) and pushed and laid-back strokes (+2.24 dB) are both well above the JND thresholds values, and thus likely detectable.

Once again, we must note that, as was the case for onset JNDs, such theoretical comparisons of stroke duration or loudness are limited, in that they mainly provide clues as to whether these often-subtle sound differences may or may not be easily detected by an average listener in direct comparative experimental listening contexts—a mode of listening common in experimental paradigms, that is, but rare in real music-listening situations.[28]

---

[27] These sub-threshold values were one of the reasons we chose to omit duration differences in our investigation of individual drummers' strategies in paper III, along with reducing the complexity of the novel mapping/categorization method.

[28] With the exceptions, perhaps, of the astute, enthusiastic listener exploring differences between recordings for curiosity's sake or the musician/music producer seeking to learn timing–sound performance strategies to apply to their own playing or productions.

6.1.3 Influence of Timing Reference

Timing reference stimuli: Do they afford wider versus narrower beat-bins?

How narrow or wide the subjective beat-bin of an internal timing reference is in and of itself another aspect of how musicians use timing–sound strategies to communicate intended microrhythmic feels. According to the view of P-centers as shapes rather than discrete points in time (Danielsen, 2018; Danielsen et al., 2019), if the timing-reference rhythms with which musicians must synchronize in either an on-the-beat or off-the-beat fashion are comprised of sounds with faster attacks and shorter durations, they theoretically afford narrower subjective beat-bins; conversely, sounds with slower attacks and longer durations afford wider beat-bins. In other words, depending on the acoustic envelope of the reference rhythm's sounds, the physical temporal window in which additional produced sounds will tend to be heard as perceptually synchronous (or not) may be longer or shorter surrounding the mean P-center. If this is so, as well, "wider" timing reference sounds may require musicians to produce greater onset asynchrony magnitudes than "narrower" sounds in order for them to be heard as perceptually off-the-beat to some extent. Danielsen and colleagues (2019) suggest that the standard deviation of average P-centers (click-alignment [CA] task) may be indicative of the relative beat-bins of different sounds, and they found that sounds with fast attacks and short durations (such as a click-like metronome) tend to have relatively lower SDs than sounds with slower attacks and/or longer durations (such as piano or bass tones). In this thesis's three experiments, the metronome reference consisted of woodblock sounds, and we found the SDs of their mean P-centers (measured in a simple CA task) to range from two to 7 milliseconds (see paper II, appendix).[29] Regarding the backing-track references, due to time constraints and logistical challenges, we were unable to collect P-center data for the various different instrumental sounds, but, based on most of the existing P-center studies of musical sounds, it is safe to assume that not only would their mean P-centers be relatively later than those of the metronome sounds but also their spread (SD) would be greater. Thus, theoretically, the temporally more extensive guitar/bass and drum sounds would engender wider beat-bins than the shorter woodblock sounds.

In paper II (drummers), we found that using instruments as timing-reference stimuli had an amplifying effect on the magnitude of the asynchrony—that is, we found a significant interaction between timing style and reference where, overall, the instrumental reference led to larger magnitudes of early and late onset timing in the pushed and laid-back conditions, respectively. This could be because the drummers (and, eventually, the listeners) perceived the subjective beat-bin widths of the reference sounds as wider than those of the metronome reference. It would follow, then, that drummers would benefit from producing larger asynchronies in the instrumental reference condition than in the metronome one in order to more effectively signal that early/late strokes should be heard as off-the-beat (that is, relatively asynchronous/anisochronous). In paper I (guitarists/bassists), we did not include the instrumental reference condition (only metronome

---

[29] This is not to say that one standard deviation (2–7 ms) from the average P-center would be directly indicative of the temporal width of the beat-bin of these metronome sounds, because this would be analogous to an onset asynchrony/anisochrony JND of 0.3–1.1 percent at an IOI of 625 ms (quarter note), whereas the lowest onset asynchrony/anisochrony JNDs derived from the psychoacoustic literature tend to reside around 2.5 percent of IOI (that is, 16 ms; Friberg and Sundberg, 1995). Instead, the findings of Danielsen and colleagues (2019) show that, as is the case with average P-centers, the magnitude of SD between two different sounds can differ based on their acoustic properties (such as attack and total duration) and thus may indicate the relative width of the subjective beat-bins they may afford.

sounds were used), but future analyses will likely find a similar amplifying effect of wider reference stimuli for those instruments.

As for the variability of the actual produced strokes (rather than the potential subjective beat-bin widths), the average standard deviation of the laid-back and pushed onset locations for all instrumentalists was found to be numerically greater than that of the on-the-beat strokes. On the one hand, this could simply be attributed to the arguably more difficult task of deciding where to place strokes when aiming to be slightly offbeat rather than squarely on the beat, or to the fact that the degree of timing asynchrony considered aesthetically "appropriate" in a laid-back or pushed performance of a groove pattern is greater (though this also likely depends on the player's individual preferences). On the other hand, as suggested in papers I and II, this greater produced variability may suggest that musicians subjectively hear the beat-bins as wider when thinking in a laid-back or pushed manner as opposed to an on-the-beat manner (but not so wide as to risk strokes that come across as a syncopation of a higher categorical subdivision). However, as suggested in section 3.1.2, it may alternatively be the case that, during a performance, musicians might conversely perceive a somewhat stable, delineated beat width (either wide or narrow) afforded by the timing reference sounds at hand in order to be able to accurately produce strokes that would, in fact, be heard as falling at the very edges of the "beat" (the width of this reference beat-bin, however, could still be influenced by the reference sound's acoustic features, as suggested above). For listeners, on the other hand, the more variable average onset timing of the laid-back and pushed strokes relative to the reference tracks in general (whether metronome or instrumental) might alternatively afford wider or narrower beat-bins depending on which particular rhythmic layers they perceive to be establishing the "beat" reference at any given time. If so, then, in fact, multiple beat-bin widths become possible, engendered by a multitude of compound sound configurations between different instruments and rhythmic layers that produce various degrees of vertically synchronous and/or horizontally isochronous relationships.

Onset NMA in performance: P-centers rather than onsets as target points for on-the-beat synchronization

Timing reference was included as a factor in our experiments in order to test whether, and to what extent, the typical onset NMA would be obtained in a musical performance context by skilled musicians. Studies of sensorimotor synchronization have shown that NMA typically arises when pacing stimuli are comprised of sounds with short attacks and durations, such as metronome-like clicks or brief tones. Training effects of musical and instrumental performance do reduce NMA, however—while they vary from 20 to 80 milliseconds for participants with no musical training, they fall between 10 and 30 milliseconds for musicians (Repp, 2005; Repp & Su, 2013) and can be even lower for drummers in particular (Danielsen, Waadeland et al., 2015; Fujii et al., 2011; Kilchenmann and Senn, 2011). However, when the pacing stimuli are musical sounds with broader sound shapes (slower attacks/longer durations/multiple onsets), the onset NMA of produced target sounds tends to be reduced, pointing to the possibility that participants use P-centers of sounds (rather than onsets) as targets or cues for synchronization (Danielsen et al., 2019; Hove et al., 2007; London et al., 2019; P. G. Vos et al., 1995).

In our experiments, we found that the average stroke onset location in the on-the-beat timing condition anticipated the metronome timing reference by 16 and 17 milliseconds for guitarists and bassists, respectively, and by slightly less for drummers—between 9 and 14 milliseconds—in accordance with results from the NMA literature. Furthermore, in paper II, we noted that drummers displayed lesser NMA values in the instrumental compared to the metronome reference (approximately 8 to 9 milliseconds), and though we have yet to analyze the effects of timing reference for the guitarists and bassists, the results would likely be similar. One potential explanation for the instrumental reference leading to lower average onset NMA is that, as mentioned above, the instrumental sounds have longer attack and total durations than the faster and shorter metronome woodblock sounds. Therefore, as predicted by the findings of previous research into the perceptual centers of musical sounds, the synchronization target location for musicians would be slightly later in the instrumental reference condition. Thus, in seeking to produce a groove pattern as on the beat as possible (where the "beat" reference is supplied by the external reference track at hand), musicians likely intuitively align their strokes with those of the reference in such a fashion that they would be as perceptually synchronous as possible in relation to each other, with the produced instrument strokes falling firmly within the subjective "beat-bin" afforded by the timing reference sounds (Danielsen 2010; Danielsen et al., 2019; see also section 3.1.2). In other words, while onset asynchrony may be present to a small extent, the instrument and reference sounds may in fact not be heard as asynchronous, and NMA may be partially explained by P-centers as targets for synchronization, rather than onsets, as suggested by P. G. Vos and colleagues (1995).

Further evidence for the fact that musicians may be aligning with P-centers rather than onsets to produce on-the-beat performances was that, for both guitarists and bassists, the average peak of on-the-beat strokes was closer to the metronome onsets rather than anticipatory. While we did not include peak as a measure in paper II,[30] I did effectively define peak in this thesis as the end time point of the attack duration. Therefore, it is possible to gauge the average peak locations across all timing references by adding the average attack duration to the average onset location: snare, -10 + 7 = -3 ms; kick, -14 + 13 = -1 ms; hi-hat, -9 + 21 ms = 12 ms (see table II, paper II). As such, we can see that the average peaks for all the drum instruments are also in fact quite close to (kick, snare) or not anticipatory/negative relative to (hi-hat) the timing references, which demonstrates that the decreased presence or absence of NMA produced by drummers when intentionally playing as synchronously as possible may in part be attributed to P-center alignment.

Note, however, that NMA in the production of in-phase synchrony cannot be fully accounted for by a P-center interpretation. As Danielsen and colleagues (2019) observe, onset NMA is often obtained in tapping tasks where the target rhythms are comprised of sounds with fast attacks and short durations (such as metronome clicks and short instrument tones), wherein physical onset and P-center are virtually the same. Because taps also tend to be short and percussive, negative mean asynchrony will tend to be present regardless of whether onset or P-center is the reference point. An alternative hypothesis regarding what exactly leads to the typical NMA in production of synchrony, then, is that people's subjective timekeeping mechanism tends to perceptually underestimate event IOIs, thus leading to anticipatory motor behavior in synchronization tasks. Wohlschläger and Koch (2002) found that, when one introduces subdivision events at a faster metrical category between the IOIs of the pacing stimuli (and therefore offers more temporal information regarding its metrical structure), NMA tends to decrease. According to Repp and Su

---

[30] This was done primarily in order to limit the number of dependent variables included in the already extensive analysis. In addition, preliminary comparisons of RMANOVAs using peak as a timing feature showed very similar results to those of onset. This is because the distance between onset and attack tends to be quite small for the drum instruments. Even so, Danielsen and colleagues (2019) found that the average P-centers snare and kick sounds with fast attacks (2 and 5 milliseconds, respectively), at least, still tend to be heard as slightly later than their onset locations (about 2 and 6 milliseconds later, respectively).

(2013), however, findings from other studies testing this hypothesis have been inconclusive, and indeed, in paper II, though the instrumental reference track in pattern 2 (complex) provided drummers with additional information regarding the metrical subdivision structure (off-beat eighth-note guitar and bass events), we found no significant effect of pattern on average onset location and no indication that NMA was further reduced in relation to pattern 1 simple (involving only quarter-note main-beat metrical information).

6.1.4 The Role of Microrhythm in “Groove”

Though it is often fruitful for empirical musicologists (who deal with quantitative analyses of musical properties) to borrow methods from psychology, occasional conflicts arise between the disciplines regarding how to define musical concepts of mutual interest. One such concept is groove itself, and a related debate concerns the purported role that microrhythm/microtiming plays in the production and perception of groove. Senn and colleagues (2017, p. 18) note that the high preference for quantized stimuli found in “groove rating” studies is often taken to undermine the PD theory claim that microtiming variations are “essential to groove” (see sections 2.1.2 and 2.1.3) and, in fact, often leads such studies “to the opposite conclusion that music should be played with as little microtiming as possible in order to have high groove.” Such counterclaims are both intriguing and provocative, especially given that many musicians in groove-based styles are very concerned with timing and sonic nuances in performance, and that both controlled experimental and analytical performance studies have revealed a wide range of systematic microtiming patterns (even within a single musical style) rather than any single prototypical pattern (see sections 2.3.2 and 2.3.3). These counterclaims also imply certain paradoxes—for example, if highly quantized music is the “grooviest,” why do records from certain groove-based styles rife with pronounced systematic onset asynchronies such as contemporary neo-soul and hip-hop with “highly lilting” beats, to mention two extreme examples (Danielsen, 2010; 2018; Stadnicki, 2017), sell in the millions and are both enjoyed and moved to by countless enthusiasts? It would appear that the psychological counterclaims regarding the detrimental role of microtiming in groove dismiss an otherwise important aspect of certain groove-based performance-practice traditions.

Fortunately, this conflict stems in part from divergent definitions of groove itself, which in turn generate semantic formulations in music psychology studies such as “microtiming is negatively associated with groove.” On the one hand, experimental psychological studies are predicated on the assumption that it is, in fact, possible to measure and quantify “groove” as an experience or set of behaviors via methods such as self-reported questionnaires that can be subsequently tested against various musical stimuli, thus revealing which structural features present “higher” or “lower” amounts of “groove” experience. They also tend to interpret groove ratings from a more functional rather than aesthetic perspective (though this is not always the case). As such, the presence of microtiming (read: onset asynchronies/anisochronies) has been thought to *hamper* the ability of listeners to predict and anticipate temporal events against a subjective timekeeping mechanism, keeping them from synchronizing successfully with the rhythms and in turn leading to a lesser sense of “groove” (read: lower ratings of wanting to move and of pleasure) (Frühauf et al., 2013; Janata et al., 2012; Madison & Sioros, 2014; Merker, 2014). However, studies that conversely find that ratings related to the urge to move and/or pleasure do not necessarily decrease for stimuli with onset asynchronies, especially when their patterns resemble those of real performances (Davies et al., 2013; Matsushita & Nomura, 2016; Senn et al., 2016; Skaansar et al., 2019), challenge the notion that music "with microtiming" cannot generate stable synchronization. In fact, dynamic entrainment models of pulse perception may very well be able to account for the possibility that even substantial onset asynchronies above heuristic perceptual thresholds derived from psychoacoustic studies may allow for a regular sense of pulse (Danielsen, 2018; Large & Jones, 1999; London, 2012; Polak & London, 2014; see also sections 2.2.2 and 2.2.3). Madison and Merker (2002), for example, specifically show that that thresholds of "pulse attribution" (the subjective quality of being able to feel a "beat" in a rhythmic sequence) tend to be higher than thresholds of "rhythmic irregularity" in terms of onset asynchrony or anisochrony (see section 3.2.2, as well as the discussion of this thesis's papers' results in section 6.1.2).

On the other hand, no one but the strictest adherents of PD theory (who are increasingly in the scholarly minority) would insist that highly quantized music cannot "groove" or that any given pattern would be "groovier" or display a "higher groove" simply by increasing its magnitude of onset asynchrony relative to an idealized isochronous metrical scheme (and thus misaligning the P-centers of simultaneous rhythmic events) in facile ways. In fact, most modern structural music-theoretical approaches to groove music (see section 2.1.2) posit that rhythms identified as "groove-based" can be played with a range of *different* microrhythmic "styles" or "feels" (pushed, laid-back, on-the-beat, swung, straight, and so on) that will be perceived as either more or less aesthetically appropriate (read: groovier) according to cultural-stylistic norms, tempo, orchestration, or personal subjective preference, among various other factors. As such, although this thesis's experiments show that musicians are able to play a given rhythmic pattern with an on-the-beat, pushed, or laid-back approach in a systematic fashion, this is not to say that they would necessarily choose to apply any of these three microrhythmic feels in real musical situations, because, as usual, context matters. Therefore, I make no value judgments as to the superior "grooviness" of either laid-back or pushed styles as opposed to on-the-beat styles, and vice versa. Rather, following the ethos of the TIME project, which accepts at face value the claims of musicians, music producers, and music scholars of groove-based musics that "timing" and other microrhythmic features are key factors in the success and appeal of good performance (Brøvig-Hanssen et al., forthcoming; Johansson et al., in preparation), the reason we explicitly instructed musicians to perform in various microrhythmic styles was chiefly for the sake of a controlled comparison of how musicians communicate intended timing via their manipulation of various acoustic features. Based on findings from certain "groove rating" studies (Frühauf et al., 2013; Janata et al., 2012; Madison & Sioros, 2014), one might be tempted to expect that, upon hearing the performances from our experiments, an average listener would likely rate on-the-beat performances as higher in "groove" than either laid-back and pushed (at least those performances that utilized timing–sound strategies with perceptually salient degrees of asynchrony/anisochrony). However, as mentioned in section 2.1.3, other studies have found that certain systematic combinations of either laid-back (Skaansar et al., 2019), pushed (Matsushita & Nomura, 2016), or dynamic microrhythmic profiles (essentially a mix of on-the-beat, laid-back, and/or pushed timing within a groove's basic pattern; Davies et al., 2013; Senn et al., 2016) are perceived by listeners as either equally high or higher in groove rating than quantized profiles. Furthermore, Senn and colleagues (2018) remind us that, more than any particular micro- or macrorhythmic structural feature of groove-based patterns, familiarity and style bias tend to correlate with groove ratings—if one does not like a genre/style (or pattern) and is unfamiliar with it, chances are one will neither derive pleasure from nor want to move to its groove.

As such, any study that manipulates onset asynchronies in musical excerpts and finds that more quantized stimuli receive higher "groove ratings" can, in fact, only claim that the absence of microtiming (read: negligible onset asynchronies) *may* be better suited to specifically induce the "urge to move" and/or "pleasure"—or any other experiential phenomenon, for that matter—under *highly specific experimental conditions* and for *specific participants* with a given musical expertise and cultural background. Generally, we must be wary of the power of reductionist formulations (and the very same could be said for inverse, equally problematic statements by proponents of strong PD theory along the lines of "microtiming is what makes music groove"). While operationalization is, of course, needed, its extreme versions risk introducing an aspect of circular reasoning to the research, to quote science philosopher Chang Hasok (2019, sec. 2.1): "if the measurement method defines the concept and there is nothing more to the meaning of the concept, the measurement method is automatically valid, as a matter of convention or even tautology." Instead, definitions and methods are more viable when they cohere with the broader concept with which they deal, as well as with empirical analyses of real-world examples of the concept. It becomes a problem, then, when operationalizing groove as a rating of "a pleasurable urge to move" leads to conclusions that exclude many types of music traditions rife with microrhythmic variety that would almost certainly be described as "highly groovy" by both enthusiasts and performers. In fact, these types of music run the ironic risk of being misclassified as "having low groove" if we were to base our conceptions of groove exclusively on a handful of psychological listening experiments conducted in the last two decades. Therefore, both music-theoretical and psychological approaches to groove should strive for definitions and methods that are able to account for the existence of a broad spectrum of microrhythmic variation in groove-based music, as well as the fact that musicians and listeners alike report their enjoyment of same. Likewise, it must be acknowledged that any singular understanding of "groove" cannot be complete and exhaustive. At the end of the day, any approach can be valid as long as one is aware that it is just that—an *approach*, not a comprehensive or absolute *account*. As long as we continue to take seriously the good-faith claims of practitioners of music-performance traditions that ascribe great importance to microrhythm (regardless of the differences in terminology) by further exploring the timing and sound nuances of groove performances, scholarly research will be better served in its ability further reveal the beauty and appeal that groove-based music has for so many of us.

6.1.5 Limitations

Ecological validity

While we strove to design the experiments with high ecological validity that resembled real musical performance without sacrificing too much control over the variables, it remains the case that the musicians performed in a novel and somewhat artificial laboratory setting. However, previous performance studies have shown that expert performers can compensate for the influence of artificial environments (Haugen, 2016; Naveda & Leman, 2008), and, similarly, our participants, all of whom were professional or semi-professional musicians with a high degree of performance experience, reported in post-interviews that they felt comfortable during the experiments, because they closely resembled commercial studio recording situations. Still, one must be wary of overgeneralizing results conducted in laboratory settings—no matter how realistic—to real-world musical performance situations, especially given the fact that expert musicians are highly adaptable in their use of temporal and/or sound features, and there are so many genre and instrumental-ensemble contexts in which the patterns they were instructed to perform are applicable. In fact, though our experiments involved typical patterns performed in a "vamp" or "loop-like" repeated fashion, as is common in many groove-based styles, they cannot be considered to belong to any particular established genre/style per se—simply looping a pattern, that is, does not constitute a "song." As discussed in section 2.3.3, since particular microrhythmic feels are often associated with certain patterns and styles, it is therefore likely that any of the tested participants would not choose to play the instructed patterns with all three systematic pre-conceived feels but rather with just one or two (for example, drum and guitar backbeat patterns tend to be associated with laid-back rather than pushed feels in medium to slow tempi, at least). At the same time, individual player preference and idiosyncrasies tend to play a part in a musician's timing–sound strategy, thus there are always exceptions to any "rule" in a given performance tradition. As such, the various strategies of systematic onset timing, intensity, duration, or frequency manipulation observed in our experiments' simple groove-based contexts may only provide *clues* to how musicians would generally approach the production (and perception) of different microrhythmic feels in real stylistic contexts that apply the same types of patterns.

Average group trends versus individual performance strategies

One methodological limitation is the extent to which insight regarding timing–sound strategies in performance can be gained via a focus on average statistical group trends (papers I and II) alone, as opposed to supplementing them with more detailed mappings of individual performances (paper III). In the former case, as mentioned in section 4.4, each musician's individual performance is first represented by an average value of each feature, then the general trends across the entire group are further represented by averages of those averages. This method invariably leads to a reduction of detail regarding the nuances of timing–sound strategies at the individual level. While such an approach is undoubtedly useful, it is important to reiterate that general trends do not necessarily translate into what actually occurs in specific individual performances. This is clearly demonstrated by the differences in the drummers' intensity results between paper II and paper III. Based on the results of the general onset and intensity trends in paper II alone, one might be tempted to conclude that an overarching timing–sound strategy, or "archetype," has been identified. But paper III showed that, in fact, a wider range of archetypical strategies can be categorized and identified within the sample. Of course, the degree of nuance or level of resolution (read: number of clusters and/or archetype categories) in such a hierarchical clustering method is itself constrained by the number of performance features included in the analysis, as well as the threshold parameter values set by the researcher (which are further limited, perhaps only by their imagination but also their capacity to interpret the total volume and scope of the results). In any case, the method used in paper III provides a more in-depth look at the diversity of timing–sound strategies demonstrated by musicians than that of papers I and II, allowing for interpretations of close hearings of individual performances backed up by the onset/intensity profile matrices and phylogenetic trees, as well as the archetype visualizations. At the same time, the greater the scope and size of the mapped-out diversity, the more challenging it becomes to identify broad trends; therefore, the two methods ideally complement each other.

Systematicity versus Intentionality

In this thesis, it was hypothesized that musicians may deliberately utilize systematic combinations of timing–sound strategies to "accentuate" the perceived lateness and earliness of laid-back and pushed rhythms, respectively. While obtained results lend evidence to this hypothesis, this is not to directly imply that musicians are consciously aware of the perceptual effects their actions have on themselves and/or their listeners; any relationship between timing and sound features, that is, could also be the result of either motor-kinesthetic or instrument-ergonomic limitations, as well as degree of difficulty. In other words, additional changes in any sound feature may simply be the unintentional byproducts of intentional onset-timing manipulation due to factors outside of the musicians' explicit control, no matter how systematically those changes are produced. At the same time, post-interview responses by many musicians in both papers I and II indicate a degree of conscious awareness regarding how the further, more deliberate, manipulation of stroke sound can help to communicate and/or achieve different microrhythmic effects. For example, laid-back approaches were commonly associated with "slower/looser/wider/larger" body movements and, conversely, pushed and on-the-beat approaches with "faster/tighter/narrower/smaller" movements. Specific techniques for lengthening or shortening stroke attack/decay duration were also identified, such as "sweeping"/"sliding out" versus "muting" of strings for guitarists and regulating the degree of stick/pedal "bounce" for drummers.

As such, though musicians themselves are likely not directly and consciously aware of the perceptual effects that duration, intensity, or frequency have on the timing of rhythmic events, whether in terms of the P-center of sounds or in terms of the detectability of onset asynchronies, this does not preclude the possibility that they can hear or feel the effects that longer/shorter, louder/softer, or brighter/darker sounds might have when further applied to late/early strokes. As suggested in paper II (conclusions), it may be that they are able to intuitively utilize certain combinations of onset and intensity/duration/frequency in order to better achieve a given microrhythmic feel. Still, these combinations may differ from musician to musician, and their perceptual effects upon listeners may also vary from individual to individual.

6.2 Summary and Conclusions

This thesis investigated the interaction between timing and sound features in the expression of microrhythmic feel in simple groove-based patterns, based on audio analyses of instructed timing performance experiments. Overall, findings from this thesis's papers provide further evidence that the production of "timing" in groove-based music involves more than simple onset relations between instruments and/or timing references. While onset (and/or peak) manipulation may still be a primary, salient vehicle expressing different microrhythmic feels in groove-based music, durational, dynamic, and frequency-related nuances of strokes played with variable onset timing may also play a significant role in how "timing" is communicated and perceived by performing musicians and listeners. As such, in groove-based performance, "timing" isn't everything—instead, "timing–sound" is more likely the case. This insight calls for a more holistic approach to the aesthetics of microrhythm and also suggests that more systematic microrhythmic analyses of performance will be better served by the investigation of a wider range of acoustic features involved in the production of groove-based microrhythm.

Some general conclusions are as follows:

- In order to distinguish preconceived laid-back (behind-the-beat), on-the-beat, and pushed (ahead-of-the-beat) microrhythmic feels from one another in simple groove-based contexts, skilled instrumentalists are on average able to systematically control the timing (onset) location of their rhythmic events to a high degree and also tend to systematically manipulate the sound features of duration, intensity (SPL), and/or frequency (SC).

- The magnitudes of onset asynchrony/anisochrony in performance can be subtle, often verging on thresholds of perceptual discriminability. Given that sound features have been found to interact with timing in previous perceptual studies, the perceived lateness, earliness, or on-the-beatness of rhythmic events played against foundational reference rhythms in a groove may be modulated (either augmented or attenuated) by musicians when they concomitantly manipulate the sound features of rhythmic events. Although a wider range of nuanced timing–sound combinations can be found at the individual participant level, average group timing–sound strategies found across all participants in each instrumentalist group indicate more generally how sound features tend to be utilized in communicating desired microrhythmic feels.

- Musical context influences microrhythm—that is, the sound and shape of a regular external timing reference can also affect the production of microrhythmic feel. To date, the effect of the timing reference was analyzed only for the drummers, and on average the short click-like metronome reference led to greater onset NMA than the instrumental reference for all drum instruments in the on-the-beat condition. The type of pattern instructed can also affect timing–sound strategies, as the drummers showed a tendency to play the more complex backbeat pattern (2) louder (kick and snare) and/or earlier (hi-hat) than the simpler pattern (1). More detailed study is needed, however.

A final note: while the analyses of all the musicians' performances in our experiments suggest that sound parameters are systematically used to differentiate between preconceived microrhythmic feels in terms of statistical differences between measured physical signals, and heuristic JND thresholds or P-center findings provided by adjacent psychoacoustic literature may provide clues as to how noticeable those differences may be to the average listener, the actual perceptual effects that these various timing–sound strategies will have on listeners (or musicians themselves) still needs to be investigated more directly via follow-up listening experiments. Ultimately, however, how these kinds of performances will be perceived by listeners will likely vary widely, depending not only on the physical magnitudes of timing–sound parameter differences but also on various subjective factors such as familiarity with the aesthetic rules of a given genre, personal preference/taste, and rhythmic acuity, among others. This reception will also surely depend on which given instrumental layer one chooses as the “primary” reference for establishing a subjective sense of pulse and meter at any given time, which, in turn, might be influenced by the listening context and experience of the listener (which are, naturally, not the same as those of the musicians in actual performance or as the idealized calculated versions to which we must invariably resort in analytical investigations).

6.3 Future Work

There are several avenues for future study opened up by this thesis’s work, particularly regarding the further analysis of the ample audio and mocap data collected from all three performance experiments (guitar/bass/drummers). In what follows, I shall present a few.

Mapping investigations of the individual timing–sound strategies of guitarists and bassists could be easily undertaken, as was already done for the drummers, using the hierarchical clustering and phylogenetic visualization method developed in this thesis. Furthermore, the method could be applied to the analysis of popular commercial recordings to compare and contrast microrhythmic feels typical of influential musicians, providing further insight into idiomatic timing–sound strategies of both individual performers as well as broader genres/styles.

Follow-up listening/perceptual experiments, such as those featured in “groove-rating” investigations of the type reviewed in section 2.1.3, are another promising path. Tests wherein different qualitative features (“urge to move,” “pleasure,” and so on) could be correlated with the microrhythmic feels in contexts that have not yet been systematically compared and studied might address solo versus ensemble performance or simple versus more complex patterns. Also, expanding upon this thesis’s main hypothesis that sound features are important to the perception of timing, it would be particularly interesting to present listeners with stimuli based on performances where both onset/peak location and intensity, duration, and/or frequency were systematically manipulated to test whether they affected subjective groove ratings. At the same time, artificially manipulating baseline performances in any timing or sound feature may not be as ecological, since the results of this thesis’s experiments indicate that, when aiming to play with a given timing feel, musicians often alter the sound of the rhythmic events in particular (though not always). Therefore, an alternative would be to present listeners with stimuli from real performances that naturally demonstrate differences in timing and sound features in a balanced fashion.

Further P-center experiments can be conducted as well, including follow-up tests on the various instrumental sounds produced by the musicians, using archetypical timing–sound combinations from each microrhythmic style condition. One could test whether the produced timing–sound combinations lead to average P-center locations and shapes (spread) as anticipated by both previous P-center literature and the theoretical speculations we have outlined in the papers—for example, whether a representative laid-back guitar stroke shaped with a slower attack (“swept stroke”) and darker tone would result in a later and wider P-center than one with a faster attack (“swift stroke”) and brighter tone.

Asynchrony/anisochrony and pulse-attribution JND threshold investigations of the types reviewed in chapter 3 represent another important avenue of future research. Just as with music perception studies in general (Schutz & Gillard, 2020; Schutz & Vaisberg, 2014), groove studies could also benefit from more up-to-date and ecologically valid heuristic psychoacoustic perceptual thresholds regarding a broader range of instrumental sounds and musical rhythmic contexts beyond simpler monotonic isochronous sequences involving singular sounds, including ones involving more complex musical and polyphonic settings. Novel tests along the same lines could look at whether listeners also produced timing–sound combinations corresponding to those of musicians in a live performance context, as in our experiments.

Mechanistic investigations could also be conducted using the ample motion-capture data I collected from this thesis’s experiments. They would shed further light on how sounds with combined timing and sound strategies are, in fact, mechanistically achieved (what types of sound-producing hand/wrist, arm, or leg movements are used to produce laid-back versus pushed strokes, and so on). Related to questions of systematicity versus intentionality (see section 6.1.5), it may that be certain types of actions enable musicians to more successfully achieve a desired pushed or laid-back microrhythmic style against a timing reference rhythm. That is, rather than playing in an intentional manner to produce a certain perceptual effect, they might employ sounds that allow them to better achieve those styles, which in turn produce certain perceptual effects. For example, it may be more difficult for a guitarist to achieve a laid-back feel by playing stronger staccato notes (fast attacks, short durations, and high intensity) that require quick, precise, and sharp movements than it would be to play this feel with softer legato notes (slow attack, long duration, lower intensity) that conversely require slower, looser, and broader movements. In addition, further experiments where musicians are instructed to play patterns with various specific combinations of timing and sound strategies may elucidate how successful they are at producing using opposite (or non-intuitive) strategies.

Finally, upon completion of the TIME project toward the end of 2021, project stakeholders plan to make the full data sets from both audio and motion-capture material captured in this thesis's experiments publicly available, and we encourage other groove scholars to further test our hypotheses, as well as attempt to reproduce our results.

Bibliography

Abel, M. (2014). *Groove: An aesthetic of measured time*. Brill. https://doi.org/10.1163/9789004242944

Algom, D., & Fitousi, D. (2016). Half a century of research on garner interference and the separability-integrality distinction. *Psychological Bulletin*, *142*(12), 1352–1383. https://doi.org/10.1037/bul0000072

Bechtold, T. A., & Senn, O. (2018). Articulation and dynamics influence the perceptual attack time of saxophone sounds. *Frontiers in Psychology*, *9*, 1–11. https://doi.org/10.3389/fpsyg.2018.01692

Benadon, F. (2006). Slicing the beat: Jazz eighth-notes as expressive microrhythm. *Ethnomusicology*, *50*(1), 73–98.

Bengtsson, I., Gabrielsson, A., & Thorsén, S.-M. (1969). Empirisk rytmforskning: En orientering om musikvetenskaplig bakgrund, utveckling och några aktuella projekt. *Svensk Tidskrift För Musikforskning*, *51*, 49–118.

Berliner, P. (1994). *Thinking in jazz: The infinite art of improvisation*. University of Chicago Press.

Bilmes, J. A. (1993). *Timing is of the essence: Perceptual and computational techniques for representing, learning, and reproducing expressive timing in percussive rhythm*. Ph.D. dissertation, Massachusetts Institute of Technology.

Bjerke, K. Y. (2010). Timbral relationships and microrhythmic tension: Shaping the groove experience through sound. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 85–101). Ashgate/Routledge.

Bøhler, K. K. (2013). *Grooves, pleasures, and politics in salsa Cubana: The musicality of Cuban politics—and the politics of salsa Cubana*. Ph.D. dissertation, University of Oslo.

Bowman, R. (1995). The Stax sound: A musicological analysis. *Popular Music*, *14*(3), 285–320.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* MIT Press.

Brøvig-Hanssen, R., Sandvik, B., Aareskjold-Drecker, J. M., & Danielsen, A. (forthcoming). A grid in flux: Sound and timing in electronic dance music. *Music Theory Spectrum*.

Burnim, M. V., & Maultsby, P. K. (2015). *African American music: An introduction* (2nd ed.). Routledge.

Butterfield, M. W. (2006). The power of anacrusis: Engendered feeling in groove-based musics. *Music Theory Online*, *12*(4), 1–17. https://doi.org/10.30535/mto.12.4.2

Butterfield, M. W. (2010). Participatory discrepancies and the perception of beats in jazz. *Music Perception*, *27*(3), 157–176. https://doi.org/10.1525/mp.2010.27.3.157

Butterfield, M. W. (2011). Why do jazz musicians swing their eighth notes? *Music Theory Spectrum*, *33*(1), 3–26. https://doi.org/10.1525/mts.2011.33.1.3

Câmara, G. S. (2016). *Swing in early funk and jazz-funk (1967–1971): Micro-rhythmic and macro-structural investigations*. Master's thesis, University of Oslo. https://doi.org/10.13140/RG.2.2.20129.20325

Câmara, G. S., & Danielsen, A. (2019). Groove. In A. Rehding & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory* (pp. 271–294). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190454746.013.17

Carlsen, K., & Witek, M. A. G. (2010). Simultaneous rhythmic events with different schematic affiliations: Microtiming and dynamic attending in two contemporary r&b grooves. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 51–68). Ashgate/

Routledge.

Chang, H. (2019). Operationalism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/operationalism

Clarke, E. F. (1987). Categorical rhythm perception: An ecological perspective. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 19–34). Royal Swedish Academy of Music.

Clarke, E. F. (1989). The perception of expressive timing in music. *Psychological Research, 51*(1), 2–9. https://doi.org/10.1007/BF00309269

Clarke, E. F., & Cook, N. (2004). *Empirical musicology: Aims, methods, prospects.* Oxford University Press.

Clarke, K. R., Somerfield, P. J., & Gorley, R. N. (2008). Testing of null hypotheses in exploratory community analyses: similarity profiles and biota-environment linkage. *Journal of Experimental Marine Biology and Ecology*, *366*(12), 55–69. https://doi.org/10.1016/j.jembe.2008.07.009

Collier, G. L., & Collier, J. L. (2002). A study of timing in two Louis Armstrong solos. *Music Perception*, *19*(3), 463–483. http://www.jstor.org/stable/10.1525/mp.2002.19.3.463

Collins, N. (2006). Investigating computational models of perceptual attack time. *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC9)*, 923--929.

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. University of Chicago Press. http://catalog.hathitrust.org/api/volumes/oclc/1139523.html

D'Amario, S., Daffern, H., & Bailes, F. (2019). Perception of synchronization in singing ensembles. *PloS One*, *14*(6), e0218162. https://doi.org/10.1371/journal.pone.0218162

Danielsen, A. (2006). *Presence and pleasure: The funk grooves of James Brown and Parliament*. Wesleyan University Press. https://doi.org/10.5860/choice.44-5554

Danielsen, A. (2010). Here, there and everywhere: Three accounts of pulse in D'Angelo's "Left and right." In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 19–34). Ashgate/Routledge.

Danielsen, A. (2012). The sound of crossover: Micro-rhythm and sonic pleasure in Michael Jackson's "Don't stop 'til you get enough". *Popular Music and Society*, *35*(2), 151–168. https://doi.org/10.1080/03007766.2011.616298

Danielsen, A. (2018). Pulse as dynamic attending: Analysing beat bin metre in neo soul grooves. In C. Scotto, K. M. Smith, & J. Brackett (Eds.), *The Routledge companion to popular music analysis: Expanding approaches* (pp. 179–189). Routledge. https://doi.org/10.4324/9781315544700

Danielsen, A., Haugen, M. R., & Jensenius, A. R. (2015). Moving to the beat: Studying entrainment to micro-rhythmic changes in pulse by motion capture. *Timing and Time Perception*, *3*, 133–154. https://doi.org/10.1163/22134468-00002043

Danielsen, A., Nymoen, K., Anderson, E., Câmara, G. S., Langerød, M. T., Thompson, M. R., & London, J. (2019). Where is the beat in that note? Effects of attack, duration, and frequency on the perceived timing of musical and quasi-musical sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(3), 402–418. https://doi.org/10.1037/xhp0000611

Danielsen, A., Waadeland, C. H., Sundt, H. G., & Witek, M. A. G. (2015). Effects of instructed timing and tempo on snare drum sound in drum kit performance. *Journal of the Acoustical Society of America*, *138*(4), 2301–2316. https://doi.org/10.1121/1.4930950

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

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32*(3), 341–365. https://doi.org/10.1068/p3370

Donnadieu, S. (2007). Mental representation of the timbre of complex sounds. In J. W. Beauchamp (Ed.), *Analysis, synthesis, and perception of musical sounds* (pp. 272–319). Springer New York. https://doi.org/10.1007/978-0-387-32576-7_8

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception and Psychophysics*, *53*(1), 25–33. https://doi.org/10.3758/BF03211712

Drake, C., & Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception and Psychophysics*, *54*(3), 277–286. https://doi.org/10.3758/BF03205262

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception*, *10*(3), 343–378.

Ehrlé, N., & Samson, S. (2005). Auditory discrimination of anisochrony: Influence of the tempo and musical backgrounds of listeners. *Brain and Cognition*, *58*(1), 133–147. https://doi.org/10.1016/j.bandc.2004.09.014

Ellis, M. C. (1991). An analysis of swing subdivision and asynchronization in three jazz saxophonists. *Perceptual and Motor Skills*, *73*, 707–713.

Fischer, R., & Milfont, T. L. (2010). Standardization in psychological research. *International Journal of Psychological Research*, *3*(1), 88–96. https://doi.org/10.21500/20112084.852

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (1st ed., pp. 149–180). Academic Press.

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

Frane, A. V., & Shams, L. (2017). Effects of tempo, swing density, and listener's drumming experience, on swing detection thresholds for drum rhythms. *Journal of the Acoustical Society of America*, *141*(6), 4200–4208. https://doi.org/10.1121/1.4984285

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America*, *98*(5), 2524–2531. https://doi.org/10.1121/1.413218

Friberg, A., & Sundberg, J. (1994). Just Noticable Difference in duration, pitch and sound level in a musical context. *Proceedings of 3rd International Conference for Music Perception and Cognition, Liège 1994 :*, 339–340.

Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, *19*(3), 333–349. https://doi.org/10.1525/mp.2002.19.3.333

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

Fujii, S., Hirashima, M., Kudo, K., Ohtsuki, T., Nakamura, Y., & Oda, S. (2011). Synchronization error of drum kit playing with a metronome at different tempi by professional drummers. *Music Perception*, *28*(5), 491–503. https://doi.org/10.1525/mp.2011.28.5.491

Gerischer, C. (2006). O suingue baiano: Rhythmic feeling and microrhythmic phenomena in Brazilian percussion. *Ethnomusicology*, *50*(1), 99–119.

Gjerdingen, R. O. (1989). Meter as a mode of attending: A network simulation of attentional rhythmicity in music. *Integral*, *3*, 67–92.

Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America*, *110*(1), 563–572. https://doi.org/10.1121/1.1376133

Goebl, W., & Parncutt, R. (2002). The influence of relative intensity on the perception of onset asynchronies. *ICMPC7, July 2002*, 1–4. papers://eba95c07-21c6-4247-a2f7-57339f1ca15a/Paper/p11980

Goebl, W., & Parncutt, R. (2003). Asynchrony versus intensity as cues for melody. *Proceedings of the 5th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM-2003), September*, 376–380.

Gordon, J. W. (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America*, *82*(1), 88–105. https://doi.org/10.1121/1.395441

Haas, W. B. De. (2007). *The role of tempo in groove and swing timing*. Master's thesis, University of Amsterdam. http://www.mcg.uva.nl/papers/Haas-MSc-Thesis(2007).pdf

Haugen, M. R. (2016). *Music-dance. Investigating rhythm structures in Brazilian samba and Norwegian telespringar performance*. Ph.D. dissertation, University of Oslo.

Haugen, M. R., & Danielsen, A. (2020). Effect of tempo on relative note durations in a performed samba groove. *Journal of New Music Research*, *49*(4), 349–361. https://doi.org/10.1080/09298215.2020.1767655

Hibi, S. (1983). Rhythm perception in repetitive sound sequence. *Journal of the Acoustical Society of Japan (E)*, *4*(2), 83–95. https://doi.org/10.1250/ast.4.83

Hirsh, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America*, *31*(6), 759–767. https://doi.org/10.1121/1.1907782

Hofmann, A., Wesolowski, B. C., & Goebl, W. (2017). The tight-interlocked rhythm section: Production and perception of synchronisation in jazz trio performance. *Journal of New Music Research*, *46*(4), 329–341. https://doi.org/10.1080/09298215.2017.1355394

Honing, H., & de Haas, W. B. (2008). Swing once more: Relating timing and tempo in expert jazz drumming. *Music Perception*, *25*(5), 471–476. https://doi.org/10.1525/mp.2008.25.5.471

Hove, M. J., Keller, P. E., & Krumhansl, C. L. (2007). Sensorimotor synchronization with chords containing tone-onset asynchronies. *Perception and Psychophysics*, *69*(5), 699–708. https://doi.org/10.3758/BF03193772

Hove, M. J., Marie, C., Bruce, I. C., & Trainor, L. J. (2014). Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(28), 10383–10388. https://doi.org/10.1073/pnas.1402039111

Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception*, *19*(3), 387–414. https://doi.org/10.1525/mp.2002.19.3.387

Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, *27*(3), 359–370. https://doi.org/10.1016/j.cub.2016.12.031

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, *141*(1), 54–75. https://doi.org/10.1037/a0024208

Jesteadt, W., Wier, C. C., & Green, D. M. (1977). Intensity discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America*, *61*(1), 169–177. https://doi.org/10.1121/1.381278

Johansson, M. (2010). *Rhythm into style: Studying asymmetrical grooves in Norwegian folk music*. Ph.D. dissertation, University of Oslo.

Johansson, M., Brøvig-Hanssen, R., Sandvik, B., Klette-Bøhler, K., & Danielsen, A. (in preparation). Shaping rhythm: Timing and sound in five rhythmic genres. *Ethnomusicology Review*.

Johnson, D. M., Watson, C. S., & Jensen, J. K. (1987). Individual differences in auditory capabilities. I. *Journal of the Acoustical Society of America*, *81*(2), 427–438.

https://doi.org/10.1121/1.394907

Jones, D. L. (2017). *Fathom Toolbox for MATLAB: Software for multivariate ecological and oceanographic data analysis*. College of Marine Science, University of South Florida. https://www.marine.usf.edu/research/matlab-resources/.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*(5), 323–355. https://doi.org/10.1037/0033-295X.83.5.323

Jones, M. R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology: A Journal of Research in Music Cognition*, *9*(2), 193–228. https://doi.org/10.1037/h0094147

Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, *2*(3), 275–283. https://doi.org/10.1525/can.1987.2.3.02a00010

Keil, C., & Feld, S. (2005). *Music grooves: Essays and dialogues*. Fenestra Books. https://doi.org/10.2307/3052737

Kernfeld, B. (2003). Groove (i). In *Grove Music Online*. https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.

Kilchenmann, L., & Senn, O. (2015). Microtiming in swing and funk affects the body movement behavior of music expert listeners. *Frontiers in Psychology*, *6*, 1232. https://doi.org/10.3389/fpsyg.2015.01232

Kilchenmann, L., & Senn, O. (2011). “Play in time, but don’t play time”: Analyzing timing profiles in drum performances. In A. Williamon, D. Edwards, & L. Bartel (Eds.), *Proceedings of the International Symposium on Performance Science 2011* (pp. 593–598). European Association of Conservatoires.

Kvifte, T. (2004). Description of grooves and syntax/process dialectics. *Studia Musicologica Norvegica*, *30*, 54–77.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159. https://doi.org/10.1037/0033-295X.106.1.119

Lartillot, O., Nymoen, K., Câmara, G. S., & Danielsen, A. (2021). Computational localization of attack regions through a direct observation of the audio waveform. Journal of the Acoustical Society of America, 149(1), 723–736. https://doi.org/10.1121/10.0003374

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A MATLAB Toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & D. Reinhold (Eds.), *Data analysis, machine learning and applications: Studies in classification, data analysis, and knowledge organization* (pp. 261–268). Springer.

Legendre, P., & Legendre, L. (2012). *Numerical ecology: Developments in environmental modeling*. Elsevier.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press.

Locke, D. (1982). Principles of offbeat timing and cross-rhythm in southern Eve dance drumming. *Ethnomusicology*. https://doi.org/10.2307/851524

London, J. (2001). Rhythm. In *Grove Music Online*. http://www.oxfordmusiconline.com/grovemusic/documentID/omo-9781561592630-e-0000045963.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter* (2nd ed.). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199744374.001.0001

London, J., Nymoen, K., Langerød, M. T., Thompson, M. R., Code, D. L., & Danielsen, A. (2019). A comparison of methods for investigating the perceptual center of musical sounds. *Attention, Perception, and Psychophysics*, *81*(6), 2088–2101. https://doi.org/10.3758/s13414-019-01747-y

London, J., Polak, R., & Jacoby, N. (2017). Rhythm histograms and musical meter: A corpus study of Malian percussion music. *Psychonomic Bulletin and Review*, *24*(2), 474–480. https://doi.org/10.3758/s13423-016-1093-7

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, *1*(4), 424–441. https://doi.org/10.2307/40285271

Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, *24*(2), 201–208. https://doi.org/10.1525/mp.2006.24.2.201

Madison, G. (2014). Sensori-motor synchronisation variability decreases as the number of metrical levels in the stimulus signal increases. *Acta Psychologica*, *147*, 10–16. https://doi.org/10.1016/j.actpsy.2013.10.002

Madison, G., Gouyon, F., Ullén, F., & Hörnström, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1578–1594. https://doi.org/10.1037/a0024323

Madison, G., & Merker, B. (2002). On the limits of anisochrony in pulse attribution. *Psychological Research*, *66*(3), 201–207. https://doi.org/10.1007/s00426-001-0085-y

Madison, G., & Sioros, G. (2014). What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00894

Matsushita, S., & Nomura, S. (2016). The asymmetrical influence of timing asynchrony of bass guitar and drum sounds on groove. *Music Perception*, *34*(2), 123–131. https://doi.org/10.1525/mp.2016.34.2.123

Melara, R. D., & Marks, L. E. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception and Psychophysics*, *48*(2), 169–178. https://doi.org/10.3758/BF03207084

Merker, B. (2014). Groove or swing as distributed rhythmic consonance: Introducing the groove matrix. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00454

Merker, B., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, *45*(1), 4–17. https://doi.org/10.1016/j.cortex.2008.06.011

Monson, I. T. (1996). *Saying something: Jazz improvisation and interaction*. University of Chicago Press.

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, *83*(5), 405–408. https://doi.org/10.1037/0033-295X.83.5.405

Naveda, L., Gouyon, F., Guedes, C., & Leman, M. (2011). Microtiming patterns and interactions with musical properties in samba music. *Journal of New Music Research*, *40*(3), 225–238. https://doi.org/10.1080/09298215.2011.603833

Naveda, L., & Leman, M. (2008). Sonification of samba dance using periodic pattern analysis. *Proceedings of the 4th International Conference on Digital Arts, Artech08*, 16–26.

Neuhoff, H., Polak, R., & Fischinger, T. (2017). Perception and evaluation of timing ptterns in drum ensemble music from Mali. *Music Perception*, *34*(4), 438–451. https://doi.org/10.1525/mp.2017.34.4.438

Nketia, J. H. K. (1974). *The music of Africa*. W. W. Norton.

Nymoen, K., Danielsen, A., & London, J. (2017). Validating attack phase descriptors obtained by the timbre toolbox and mirtoolbox. *Proceedings of the 14th Sound and Music Computing Conference 2017*, 214–219.

Palmer, C. (1996). On the assignment of structure in music performance. *Music Perception*, *14*(1), 23–56. https://doi.org/10.2307/40285708

Pastore, R. E., Harris, L. B., & Kaplan, J. K. (1982). Temporal order identification: Some parameter dependencies. *Journal of the Acoustical Society of America*, *71*(2), 430–436. https://doi.org/10.1121/1.387446

Peterson, S. M. (2008). *Something real: Rap, resistance, and the music of the Soulquarians*. Ph.D. dissertation, University of Oregon.

Polak, R. (2010). Rhythmic feel as meter: Non-isochronous beat subdivision in jembe music from Mali. *Music Theory Online*, *16*(4). https://doi.org/10.30535/mto.16.4.4

Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., & London, J. (2018). Rhythmic prototypes across cultures. *Music Perception*, *36*(1), 1–23. https://doi.org/10.1525/mp.2018.36.1.1

Polak, R., & London, J. (2014). Timing and meter in mande drumming from Mali. *Music Theory Online*, *20*(1). https://doi.org/10.30535/mto.20.1.1

Polak, R., London, J., & Jacoby, N. (2016). Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: A corpus study of malian jembe drumming. *Frontiers in Neuroscience, 10*, 1–11. https://doi.org/10.3389/fnins.2016.00285

Poudrier, È., & Repp, B. H. (2013). Can musicians track two different beats simultaneously? *Music Perception*, *30*(4), 369–390. https://doi.org/10.1525/mp.2013.30.4.369

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*(4), 411–440. https://doi.org/10.2307/40285311

Pressing, J. (2002). Black Atlantic rhythm: Its computational and transcultural foundations. *Music Perception*, *19*(3), 285–310. https://doi.org/10.1525/mp.2002.19.3.285

Prögler, J. A. (1995). Searching for swing: Participatory discrepancies in the jazz rhythm section. *Ethnomusicology*, *39*(1), 21–54. http://www.jstor.org/stable/852199

Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acta Acoustica United with Acoustica*, *43*(2), 121–131.

Reinholdsson, P. (1987). Approaching jazz performances empirically: Some reflections on methods and problems. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 105–125). Royal Swedish Academy of Music.

Repp, B. H. (1998a). Obligatory “expectations” of expressive timing induced by perception of musical structure. *Psychological Research*, *61*(1), 33–43. https://doi.org/10.1007/s004260050011

Repp, B. H. (1998b). The detectability of local deviations from a typical expressive timing pattern. *Music Perception*, *15*(3), 265–289. https://doi.org/10.2307/40285768

Repp, B. H. (2003). Phase attraction in sensorimotor synchronization with auditory sequences: Effects of single and periodic distractors on synchronization accuracy. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 290–309. https://doi.org/10.1037/0096-1523.29.2.290

Repp, B. H. (2004). On the nature of phase attraction in sensorimotor synchronization with interleaved auditory sequences. *Human Movement Science*, *23*(3–4 special issue), 389–413. https://doi.org/10.1016/j.humov.2004.08.014

Repp, B. H. (2005). A review of the tapping literature. *Psychonomic Bulletin and Review*, *12*(6), 969–992.

Repp, B. H. (2006). Does an auditory distractor sequence affect self-paced tapping? *Acta Psychologica*, *121*(1), 81–107. https://doi.org/10.1016/j.actpsy.2005.06.006

Repp, B. H., London, J., & Keller, P. E. (2012). Distortions in reproduction of two-interval rhythms: When the “attractor ratio” is not exactly 1:2. *Music Perception*, *30*(2), 205–223. https://doi.org/10.1525/mp.2012.30.2.205

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin and Review*, *20*(3), 403–452. https://doi.org/10.3758/s13423-012-0371-2

Roholt, T. C. (2014). *Groove: A phenomenology of rhythmic nuance*. Bloomsbury Academic.

Rose, R. F. (1989). *An analysis of timing in jazz rhythm section performance*. DMA thesis, University of Texas at Austin. http://search.proquest.com.ezproxy.ecu.edu.au/docview/303815471/abstract/5464F1E8014F4525PQ/1?accountid=10675

Rossing, T. D., Moore, R. F., & Wheeler, P. A. (2002). *The science of sound*. Addison Wesley.

Schubert, E., & Wolfe, J. (2006). Does timbral brightness scale with frequency and spectral centroid? *Acta Acoustica United with Acoustica*, *92*(5), 820–825.

Schulze, H.-H. (1989). Categorical perception of rhythmic patterns. *Psychological Research*, *51*(1), 10–15. https://doi.org/10.1007/BF00309270

Schutz, M., & Gillard, J. (2020). On the generalization of tones: A detailed exploration of non-speech auditory perception stimuli. *Scientific Reports*, *10*(1), 9520. https://doi.org/10.1038/s41598-020-63132-2

Schutz, M., & Vaisberg, J. M. (2014). Surveying the temporal structure of sounds used in Music Perception. *Music Perception*. https://doi.org/10.1525/MP.2014.31.3.288

Scott, S. K. (1998). The point of P-centres. *Psychological Research*, *61*(1), 4–11. https://doi.org/10.1007/PL00008162

Senn, O., Bechtold, T., Rose, D., Câmara, G. S., Düvel, N., Jerjen, R., Kilchenmann, L., Hoesl, F., Baldassarre, A., & Alessandri, E. (2020). Experience of groove questionnaire: Instrument development and initial validation. *Music Perception*, *38*(1), 46–65. https://doi.org/10.1525/mp.2020.38.1.46

Senn, O., Bullerjahn, C., Kilchenmann, L., & von Georgi, R. (2017). Rhythmic density affects listeners' emotional response to microtiming. *Frontiers in Psychology*, *8*, 1709. https://doi.org/10.3389/fpsyg.2017.01709

Senn, O., Kilchenmann, L., Bechtold, T., & Hoesl, F. (2018). Groove in drum patterns as a function of both rhythmic properties and listeners' attitudes. *PloS One*, *13*(6), e0199604–e0199604. https://doi.org/10.1371/journal.pone.0199604

Senn, O., Kilchenmann, L., von Georgi, R., & Bullerjahn, C. (2016). The effect of expert performance microtiming on listeners' experience of groove in swing or funk music. *Frontiers in Psychology*, *7*, 1487. https://doi.org/10.3389/fpsyg.2016.01487

Senn, O., Rose, D., Bechtold, T., Kilchenmann, L., Hoesl, F., Jerjen, R., Baldassarre, A., & Alessandri, E. (2019). Preliminaries to a psychological model of musical groove. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01228

Seton, J. C. (1989). *A psychophysical investigation of auditory rhythmic beat perception*. Ph.D. dissertation, University of York.

Sioros, G., Câmara, G. S., & Danielsen, A. (2019). Mapping timing strategies in drum performance. *20th International Society for Music Information Retrieval Conference*, 776–783. https://doi.org/10.5281/zenodo.3527926

Sioros, G., Miron, M., Davies, M., Gouyon, F., & Madison, G. (2014). Syncopation creates the sensation of groove in synthesized music examples. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01036

Skaansar, J. F., Laeng, B., & Danielsen, A. (2019). Microtiming and mental effort. *Music Perception: An Interdisciplinary Journal*, *37*(2), 111–133. https://doi.org/10.1525/mp.2019.37.2.111

Snyder, B. (2000). *Music and memory: An introduction*. MIT Press.

Stadnicki, D. A. (2017). Play like Jay: Pedagogies of drum kit performance after J Dilla. *Journal of Popular Music Education*, *1*(3), 253–280. https://doi.org/10.1386/jpme.1.3.253_1

Stewart, A. (2000). "Funky drummer": New Orleans, James Brown and the rhythmic transformation of American popular music. *Popular Music*, *19*(3), 293–318.

Stover, C. (2009). *A theory of flexible rhythm spaces for diasporic African music*. Ph.D. dissertation, University of Washington.

Tekman, H. G. (1997). Interactions of perceived intensity, duration, and pitch in pure tone sequences. *Music Perception: An Interdisciplinary Journal*, *14*(3), 281–294. https://doi.org/10.2307/40285722

Tekman, H. G. (2001). Accenting and detection of timing variations in tone sequences: Different kinds of accents have different effects. *Perception and Psychophysics*, *63*(3), 514–523. https://doi.org/10.3758/BF03194417

Tekman, H. G. (2002). Perceptual integration of timing and intensity variations in the perception of musical accents. *Journal of General Psychology*, *129*(2), 181–191. https://doi.org/10.1080/00221300209603137

ten Hoopen, G., Boelaarts, L., Gruisen, A., Apon, I., Donders, K., Mul, N., & Akerboom, S. (1994). The detection of anisochrony in monaural and interaural sound sequences. *Perception and Psychophysics*, *56*(1), 110–120. https://doi.org/10.3758/BF03211694

Villing, R. C. (2010). *Hearing the moment: Measures and models of the perceptual centre*. Ph.D. dissertation, National University of Ireland.

Vos, J., & Rasch, R. (1981). The perceptual onset of musical tones. *Perception and Psychophysics*, *29*(4), 323–335. https://doi.org/10.3758/BF03207341

Vos, P. G., Mates, J., & van Kruysbergen, N. W. (1995). The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies. *Quarterly Journal of Experimental Psychology Section A*, *48*(4), 1024–1040. https://doi.org/10.1080/14640749508401427

Wilson, O. (1974). The significance of the relationship between Afro-American music and West African music. *Black Perspective in Music*, *2*(1), 3–22. https://doi.org/10.2307/1214144

Witek, M. A. G. (2017). Filling in: Syncopation, pleasure and distributed embodiment in groove. *Music Analysis*, *36*(1), 138–160. https://doi.org/10.1111/musa.12082

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, *9*(4). https://doi.org/10.1371/journal.pone.0094446

Wohlschläger, A., & Koch, R. (2002). Synchronization error: An error in time perception. In P. Desain & L. Windsor (Eds.), *Rhythm perception and production* (pp. 115–127). Swets and Zeitlinger.

Wojtczak, M., Beim, J. A., Micheyl, C., & Oxenham, A. J. (2012). Perception of across-frequency asynchrony and the role of cochlear delays. *Journal of the Acoustical Society of America*, *131*(1), 363–377. https://doi.org/10.1121/1.3665995

Wojtczak, M., Mehta, A. H., & Oxenham, A. J. (2017). Rhythm judgments reveal a frequency asymmetry in the perception and neural coding of sound synchrony. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(5), 1201–1206. https://doi.org/10.1073/pnas.1615669114

Wright, M. J. (2008). *The shape of an instant: Measuring and modeling perceptual attack time with probability density functions*. Ph.D. dissertation, Stanford University.

Yee, W., Holleran, S., & Jones, M. R. (1994). Sensitivity to event timing in regular and irregular sequences: Influences of musical skill. *Perception and Psychophysics*, *56*(4), 461–471. https://doi.org/10.3758/BF03206737

Zera, J., & Green, D. M. (1993). Detecting temporal asynchrony with asynchronous standards. *Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.406816

Zuckerkandl, V. (1956). *Sound and symbol: Music and the external world*. Pantheon Books.

Papers

I Effects of instructed timing on electric guitar and bass sound in groove performance.
Guilherme Schmidt Câmara, Kristian Nymoen, Olivier Lartillot, and Anne Danielsen.
In *The Journal of the Acoustical Society of America* 147(2), 1028-1041, 2020.

II Timing Is Everything . . . or Is It? Effects of Instructed Timing Style, Reference, and Pattern on Drum Kit Sound in Groove-Based Performance.
Guilherme Schmidt Câmara, Kristian Nymoen, Olivier Lartillot, and Anne Danielsen.
In *Music Perception* 38 (1), 1–26, 2020.

III Mapping Timing and Intensity Strategies in Drum-Kit Performance of a Simple Back-Beat Pattern.
Guilherme Schmidt Câmara, Georgios Sioros, and Anne Danielsen.
Submitted to *The Journal of New Music Research*.

Paper I

Effects of instructed timing on electric guitar and bass sound in groove performance

Guilherme Schmidt Câmara, Kristian Nymoen, Olivier Lartillot, and Anne Danielsen

In *The Journal of the Acoustical Society of America* 147(2), 1028–1041 (2020).

https://doi.org/10.1121/10.0000724

JASA ARTICLE

Effects of instructed timing on electric guitar and bass sound in groove performance

Guilherme Schmidt Câmara,[a] Kristian Nymoen, Olivier Lartillot, and Anne Danielsen
RITMO Center for Interdisciplinary Studies in Time, Rhythm and Motion, Department of Musicology, University of Oslo, Oslo, Norway

ABSTRACT:
This paper reports on two experiments that investigated the expressive means through which musicians well versed in groove-based music signal the intended timing of a rhythmic event. Data were collected from 21 expert electric guitarists and 21 bassists, who were instructed to perform a simple rhythmic pattern in three different timing styles—"laid-back," "on-the-beat," and "pushed"—in tandem with a metronome. As expected, onset and peak timing locations corresponded to the instructed timing styles for both instruments. Regarding sound, results for guitarists revealed systematic differences across participants in the duration and brightness [spectral centroid (SC)] of the guitar strokes played using these different timing styles. In general, laid-back strokes were played with a longer duration and a lower SC relative to on-the-beat and pushed strokes. Results for the bassists indicated systematic differences in intensity (sound-pressure level): pushed strokes were played with higher intensity than on-the-beat and laid-back strokes. These results lend further credence to the hypothesis that both temporal and sound-related features are important indications of the intended timing of a rhythmic event, and together these features offer deeper insight into the ways in which musicians communicate at the microrhythmic level in groove-based music. *© 2020 Acoustical Society of America.*
https://doi.org/10.1121/10.0000724

(Received 26 August 2019; revised 18 December 2019; accepted 24 January 2020; published online 11 February 2020)
[Editor: Tamara Smyth] Pages: 1028–1041

I. INTRODUCTION

In groove-based music, where rhythm is of paramount importance, timing is essential to musicians interested in conveying different types of "feel" to a performance. Musicians can achieve these different timing feels or styles by, among other things, subtly altering the temporal location of events at the "microrhythmic" metrical level (on the order of 10–40 ms) by playing either slightly early ("pushed") or slightly late ("laid-back") in relation to other players' rhythm, a metronomic beat reference, or simply their own internal pulse (Butterfield, 2006, 2011; Câmara, 2016; Câmara and Danielsen, 2019; Danielsen, 2006, 2010, 2018; Frane and Shams, 2017; Friberg and Sundström, 2002; Iyer, 2002; Kilchenmann and Senn, 2011, 2015; Senn *et al.*, 2017).

"Groove" has been conceptualized in various ways in recent decades. In musicology and ethnomusicology, it has denoted the individual rhythmic patterns that comprise a given style (Kernfeld, 2003), the overall "rhythm matrix" comprised by all instruments within a performance (Iyer, 2002; Monson, 1996) or an experienced aesthetic quality, or "feel," stemming from the various rhythmic interactions between/within instruments of an ensemble (Butterfield, 2006; Keil, 1987). The adjective "groove-based" has been typically reserved for musical genres derived from African-American performance traditions that share a range of rhythmic properties (Câmara and Danielsen, 2019; Pressing, 2002). Most recently, however, groove has been operationalized by music psychologists as the aspect of any music, regardless of cultural origin, that elicits various experiential phenomena, such as the "urge to move" and "pleasure," in particular (Madison, 2006; Janata *et al.*, 2012). While early studies found either negative or no effects of microrhythmic asynchronies on subjective ratings of various assumed "groove features" (Davies *et al.*, 2013; Frühauf *et al.*, 2013; Madison *et al.*, 2011; Madison and Sioros, 2014), recent studies have found that stimuli with microtiming profiles derived directly from or resembling those of original performed music do not necessarily obtain lower ratings than stimuli with artificially reduced onset asynchronies (Kilchenmann and Senn, 2015; Senn *et al.*, 2016; Senn *et al.*, 2017; Skansaar *et al.*, 2019). Regardless, in this article we subscribe to the musicological meanings of groove, broadly denotative of the rhythmic structural aspects of music historically designated as groove-based, and thus make no assumption regarding the degree to which onset asynchronies elicit higher or lower ratings of any operationalized groove feature.

Traditionally, research into microrhythm in groove-based music has often focused on measuring and mapping relationships between discrete timing points in rhythmic events such as onsets and offsets (see Câmara and Danielsen, 2019). Recently, however, the role of other sound parameters, such as timbre, loudness, and duration, and their interaction with timing have been acknowledged as fundamental to both the production and perception of microrhythmic expressivity in performance (Butterfield, 2011; Danielsen *et al.*, 2015).

In particular, studies of a sound's P-center or perceptual "moment of occurrence" (Morton *et al.*, 1976) or perceptual attack time (PAT; Gordon, 1987)—involving both natural speech and musical sounds—have shown that, in a rhythmic

---

[a] Electronic mail: g.s.camara@imv.uio.no

JASA
https://doi.org/10.1121/10.0000724

isochronous context, the perceived temporal location of a sound event is a complex matter that involves more than onset timing. In fact, it is clear that a sound's P-center is contingent upon various acoustic properties, including the perceived total duration of the sound and its attack duration (the length of time from onset to maximum amplitude peak, often referred to as "rise time" or "attack time"). Virtually all research on the P-centers of musical sounds (whether acoustic or artificial/synthesized) has found that the shorter the attack duration of a sound, the earlier its average P-center will be heard relative to its onset, and, conversely, the longer the attack, the later the P-center (Bechtold and Senn, 2018; Danielsen *et al.*, 2019; Gordon, 1987; Villing, 2010; Wright, 2008; Vos *et al.*, 1995). Most studies also suggest a similar effect related to total duration whereby a longer duration leads to a later P-center (Danielsen *et al.*, 2019; Scott, 1998; Seton, 1989; Vos *et al.*, 1995). A recent study by Danielsen and colleagues (2019) found that positively correlated combinations of attack and total duration (short attack/short duration, long attack/long duration) caused a "redundancy gain" whereby both factors shifted P-centers either earlier or later in time, respectively, whereas negatively related combinations (short attack/long duration, long attack/short duration) caused a "redundancy loss" whereby, because one factor tends to shift the P-center earlier while the other one shifts it later, the cumulative shifting effect was either attenuated or canceled out. This relation implies a potential sensory-perceptual interference between the two factors.

A few studies have shown that frequency/pitch also affects timing perception; yet, conflicting results have inhibited a consensus regarding its effect on the P-center. Seton (1989) found that high-frequency tones generally led to later P-centers than did middle-frequency tones. However, Danielsen and colleagues (2019) found that in a musical stimulus set, sounds lower in frequency/darker in timbre (as subjectively determined by the researchers) generally led to later P-centers but only when those sounds were of longer duration. Gordon (1987) also found that for sounds with shorter attack durations, the P-center was primarily determined by amplitude cues, whereas for sounds with longer attack durations, the P-center was also influenced by spectral features. Danielsen and colleagues (2019) argued that their finding that the impact of frequency was stronger when the duration was longer corroborated Gordon's finding that only sounds with longer attack durations were influenced by spectral features since sounds with longer rise times also tend to be long, in general.

Even fewer studies have focused on the effect of loudness/intensity on the P-center. Seton (1989) examined the effect on the P-center of loudness presentation on a sawtooth tone using an adjustment paradigm and found no effect. However, results for the mixed-level presentation compared to the blocked-level presentation revealed a dependency on loudness that resulted in smaller P-center shifts (that is, shifts earlier in time) for higher intensity levels. Gordon (1987) also found that a saxophone sound played with greater intensity led to an earlier P-center than one played with lesser intensity. Bechtold and Senn (2018) confirmed this result and further found that for saxophone sounds, both greater articulation (a stronger "tongue attack") and higher dynamics (a louder sound) led to an earlier P-center, on average, in comparison to sounds with lesser articulation and lower dynamics.

In groove performance, timing onset asynchronies between instruments may often be quite subtle, verging on the threshold of discriminability (in musical contexts, about 30 ms of discrepancy; see Butterfield, 2010; Clarke, 1987). Therefore, the modulation of the sound of a given instrument with a certain intended timing via features known to affect the P-center, such as duration, frequency, and intensity, can help to communicate intended timing. In other words, the exploitation of potential redundancy gains across the dimensions of timing and sound might convey an intentional early, late, or on-the-beat timing feel. One example of this would be to increase the duration of a note played intentionally late in order to enhance its laid-back character or, conversely, to shorten an intentionally early stroke to enhance its pushed-ness.

Research into such systematic covariances between timing and acoustic factors in musical performance is very scarce. In a microtiming performance study of snare drum sound with ten drummers, Danielsen and colleagues (2015) found that when they were instructed to play a standard backbeat-based pattern intentionally behind the beat alongside a metronome set at 96 beats per minute (bpm), the general trend across all participants was to play strokes with a higher intensity ("louder") relative to an "on-beat" condition. At the individual level, most of the drummers (seven out of ten) also played their behind-the-beat strokes with a significantly lower spectral centroid (SC; "darker") relative to those played on-the-beat. The study's findings suggest that in order to signal the intended timing of a sound event as microrhythmically early or late, musicians systematically manipulated not only the onset timing of their strokes but also other sound parameters. Several performance experiments with pianists have also shown that when they were instructed to emphasize a melody tone in a polyphonic piano performance, they played it both louder and earlier than the other voices (Goebl, 2001; Palmer, 1996; Repp, 1996). Goebl and Parncutt (2002) also found that the relative perceptual salience of two tones in a piano chord depended primarily on their relative intensity rather than the asynchrony between their onset timing locations.

Numerous performance studies have found a systematic relationship between intensity and duration whereby beats accented with greater intensity also tend to be lengthened in performance (Clarke, 1989; Dahl, 2000, 2004; Drake and Palmer, 1993; Gabrielsson, 1974, 1999; Waadeland, 2006). Regarding perception, Woodrow (1909) drew attention to the similar function of relative duration and intensity in the formation of musical accents, and this has been confirmed in more recent perception studies (Povel and Okkerman, 1981; Tekman, 2002; Windsor, 1993). Garner interference studies (see Algom and Fitousi, 2016) have also found interaction

https://doi.org/10.1121/10.0000724

effects between auditory dimensions of pitch/timbre and intensity (Melara and Marks, 1990) and timing and duration (Tekman, 2002). The latter study, for example, found that when listeners attended to timing, an agogic accent (a perceptual accent created via delayed timing rather than dynamics) was easier to classify if it were combined with increased intensity (where the positive correlation of timing and intensity led to a redundancy gain).

To sum up, several experiments with both music performance and music/audio perception point to the intimate relationship between the temporal and auditory aspects of microrhythm as well as the integration of timing and various perceptual dimensions of sound.

While it is crucial to communicate in performance whether a stroke is meant to be early, late, or on the expected beat position, timing tolerance varies with genre and context so that an actual early or late onset positioning of an intended sound event, measured in relation to a beat reference, might, in fact, be perceived by the listener as subsumed within the "beat-bin" for on-the-beat playing—that is, the perceived temporal window in which a given stroke will tend to be heard as synchronous with the perceived beat Danielsen, 2010; Danielsen *et al.*, 2019). It is therefore clear that the communication of early or late timing would benefit from something more than the simple positioning of the onset timing of a sounded event early or late in relation to an expected metrical beat.

The present study intends to further test the hypothesis that sonic features are important in the production of intended timing in a musical context by extending the inquiry to additional salient rhythm-section instrumentalists typically found in groove-based ensembles: namely, electric bassists and guitarists. Perhaps just as much as drummers, rhythm guitarists and bassists are keenly attuned to microrhythmic expression and spend years practicing the incorporation of different degrees of push and pull, as aesthetically appropriate, into a given groove. More so than the drums, however, the guitar and bass afford a greater potential for manipulating sound features other than intensity, such as frequency/pitch and timbre, by virtue of being harmonic/polyphonic pitched string instruments and duration by virtue of their ability to create sustained sounds by allowing their strings to vibrate post-impulse.

We hypothesize that in the process of achieving early, late, or on-the-beat microrhythmic timing, instrumentalists leave sonic "stamps" on the sounds of their own instruments that are systematically related to these intended timing profiles. We explore this possibility by measuring changes in the duration, loudness [sound pressure level (SPL)], and brightness (SC) of sound events after we have instructed instrumentalists to play a basic pattern under different timing dictates, so as to address the following question: To what extent are there systematic differences in the acoustic signal between guitar and bass strokes played with different intended timings across subjects? In accord with findings from previous performance and P-center studies, we expect that, on average, laid-back strokes will have longer duration/lower SC/lower SPL and pushed strokes shorter duration/higher SC/higher SPL relative to on-beat strokes.

II. METHOD

A. Participants, tasks

1. Experiment 1 (guitar)

Twenty-one electric guitarists (three female), 22–50 years of age [mean = 33 years, standard deviation (SD) = 7 years], took part in the experiment. All were semi-professional or professional musicians well versed in either rock or jazz and/or soul/funk and recruited from either local performance scenes or conservatory/academic institutions with between 9 and 40 years of instrument performance experience [mean = 20 years, SD = 8 years]. All participants were paid an honorarium to take part in the experiment.

The guitar participants were instructed to play a simple "backbeat" pattern—ubiquitous in groove-based styles and very familiar to guitarists—comprised of two alternating chords with specific four-note voicings that coincided with the second and fourth quarter-note beats of a 4/4 meter [see Fig. 1(a)] with a plectrum of their choosing. They performed along with a metronome track at a tempo deemed comfortable from a pilot of the experiment (96 bpm) in four different timing-style conditions:

(1) in as *natural* a manner as possible (condition: natural);
(2) in a *laid-back* manner, or behind the beat of the metronome (condition: laid-back);
(3) in a *pushed* manner, or ahead of the beat of the metronome (condition: pushed);
(4) synchronized with, or *on*-the-beat of the metronome (condition: on)

2. Experiment 2 (bass)

Twenty-one professional or semi-professional electric bassists (all male), 23–49 years of age (mean = 32 years, SD = 7 years), recruited from the same scenes as the guitarists, with between 6 and 38 years of instrument performance experience (mean = 28 years, SD = 10 years) participated. All participants were paid an honorarium to take part in the experiment.

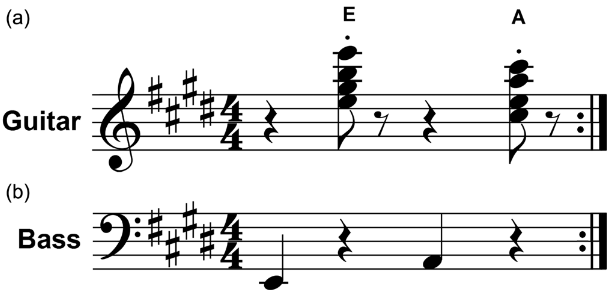

FIG. 1. (a) Two-chord backbeat pattern for guitarists. (b) Two-note "downbeat" pattern for bassists.

JASA
https://doi.org/10.1121/10.0000724

The bass participants were instructed to perform a musical figure more appropriate for bass accompaniment—a "downbeat" pattern comprised of two alternating single notes coinciding with the first and third quarter-note beats of a 4/4 meter [see Fig. 1(b)]—with their fingers. As in experiment 1, the participants performed the pattern along with a metronome track (96 bpm) in the four different timing style conditions.

3. Apparatus and procedure

The recordings were done at the Motion Capture Laboratory at the Department of Musicology, University of Oslo, Oslo, Norway, in the spring of 2018. The guitarists were provided with a 2002 Fender Telecaster (Scottsdale, AZ), and the bassists were provided with a 2016 Fender Precision Bass (Ensenada, Mexico). Both instruments' pickup tone settings were preconfigured and not alterable for the duration of the experiment.

To ensure that participant monitoring of the feedback sound would be as ecologically realistic as possible while keeping atmospheric leakage and signal noise to a minimum, we routed the signal output from the instruments through analog amplifier emulators (guitar, Tech 21 Blonde V2; bass, Tech 21 VT-Bass DI; New York, NY). Amplified signals were sent through a splitter, where one signal went directly into an RME BabyFace Pro sound card (Audio AG, Haimhausen, Germany) to be recorded with the audio software Reaper 5.77 (Cockos Inc., San Francisco, CA) at a sampling frequency of 44.1 kHz and 24-bit resolution. The other signal was routed to an analog mixer (Yamaha MG10XU, Hamamatsu, Japan) and then fed into headphones (Beyerdynamic DT-990 Pro, Heilbronn, Germany) that were supplied to the participants for monitoring. The playback metronome reference track was routed into the mixer from the recording software and then into the headphones, which, combined with the direct monitoring of the instrument signals, ensured a "near zero" latency monitoring situation in which the delay between what performers heard in playback and what they played on their instruments was negligible (1–2 samples $\approx$ 0.02 ms).

For the metronome-beat reference track, we used two woodblock sounds. One was pitched higher and located on the first quarter-note beat locations (SC = 2370 Hz, duration from onset to offset = 40 ms), and the other pitched lower and located on the remaining second, third, and fourth quarter-note beat locations (SC = 1565 Hz, duration = 50 ms). Both sounds had very short and impulsive attack transients (duration from onset to amplitude peak $\approx$ 2 ms) with a gradual decay toward the signal offset. The woodblock sounds were perceptually aligned to the 4/4 beats of the isochronous software grid based on mean P-center results from a separate click-alignment (CA) experiment (see the Appendix for details). Based on these results, the high-pitched sound was placed 0 ms from the initial grid beats (that is, no displacement), thereby clearly demarcating the start of each 4/4 measure, and the onsets of the low-pitched woodblock sounds were located 4 ms early in relation to the second, third, and fourth grid beats.

Participants were positioned on a stool in the middle of a large studio with a headset, and the experimenter was located in the same room at a control desk. Participants were given time to acquaint themselves with the instrumental setup and reported when they were ready to begin. The experiment had a repeated measures design. At the beginning of each session, the "natural" timing condition was assigned first, primarily as a warm-up in order to allow participants to acquaint themselves to the following timing conditions (laid-back, "on," and pushed), which were subsequently randomized. Each timing-condition task lasted for 67.5 s (27 bars). Participants began to play as soon as they had entrained themselves to the metronome. If a participant was dissatisfied with their performance during or after a condition trial, they were invited to repeat it as many times as they wanted.

After the performances, we conducted short interviews to gain feedback related to the experimental setup, as well as insight into the various performance strategies that had been applied to satisfy the different timing condition tasks. In total, the performance and interview session lasted around 45 min.

B. Audio analysis

1. Duration and location of guitar and bass strokes

In order to calculate values for the various sound features (described in Sec. II B 2), we first had to define the total duration of each individual stroke, which meant determining its beginning (onset) and ending (offset). These time points were calculated based on algorithms from the MIR Toolbox (version 1.8) audio analysis package (Lartillot *et al.*, 2008), and defined as follows:

(1) Peak: the global maximum amplitude of the envelope of a given stroke, where the envelope is calculated as the sum of bin amplitudes across the columns of a spectrogram (*mirevents* with the "attacks" option and default parameter settings; frame length = 30 ms, frame hop factor = 2%);
(2) Onset: the starting point of the attack phase, where the attack phase is a line fitted on the audio waveform through an optimization of both the slope and the maximum amplitude (*mirevents* with both "Attacks" and "Waveform" options, WaveformThreshold = 3%);
(3) Offset: the ending point of the decay phase, computed using the same approach as the onset (*mirevents* with both "Decays" and "Waveform" options, WaveformThreshold = 3%).

2. Selection of sound descriptors

In addition to the two temporal location descriptors of onset and peak (defined above), we selected three sound descriptors for the main analyses, defined as follows:

(1) Duration: elapsed time interval of the signal, measured in milliseconds;
(2) SC: the weighted mean of the frequencies present in the signal, determined using a Fourier transform, with their magnitudes as the weights, measured in Hz;

https://doi.org/10.1121/10.0000724

(3) SPL: the unweighted root-mean-square (rms) amplitude of the signal, measured in dB, with a 0 dB reference given as the average rms amplitude of all strokes in all timing conditions.

Following Danielsen and colleagues (2015), SPL and SC were chosen as audio descriptors as both intensity and frequency have been found to affect perceived timing of sounds. SPL is highly correlated to perceived loudness and remains a widely used metric in perceptual studies (Rossing *et al.*, 2002). SC is a readily quantifiable frequency descriptor and purportedly accounts for the perceived brightness or sharpness of sounds (Donnadieu, 1987; Schubert and Wolfe, 2006).

An additional descriptor included in this study is "duration," because, as mentioned in the Introduction, various studies report that the total elapsed time of a sound, as well as the duration of its attack segment, in particular, has been shown to have a significant effect on the P-center, whereby longer total and attack durations tend to shift the experienced P-center later in time.

We decided to use the maximum-energy peak time point to separate strokes into their attack and decay segments (see Fig. 2). Each of the three sound descriptors was therefore calculated for the following segments:

(a) Attack: interval between onset and maximum peak;
(b) Decay: interval between maximum peak and offset;
(c) Total: interval between onset and offset.

This calculation yielded nine descriptors for the analysis: attack duration, decay duration, total duration, attack SPL, decay SPL, total SPL, attack SC, decay SC, and total SC.

If the attack duration is defined as the elapsed time from signal onset to maximum amplitude peak, the waveform/spectrogram analysis of the signal from a typical guitar stroke in this experiment reveals that it often encompasses the approximate time interval between the onsets of the impulse transients from the first and last note/string of the chord struck. For the electric bass, the attack duration tends to be simply the time it takes to reach maximum amplitude following a single string impulse. Spectrally, both instruments' attack segments demonstrate impulsive transients with higher noise and wider frequency ranges caused by the striking of the strings. The decay duration, on the other hand, encompasses the signal energy post-impulse for both instruments. Spectrally, the decay segment signal is markedly more harmonic (that is, less noisy), sustained, and free of impulse transients. Decay length tends to be affected both by the amount of striking force transferred from either the plectrum or the fingers to the strings and how long the strings are allowed to vibrate, which performers may further regulate by dampening the vibrated strings with their hand ("muting") or releasing the pressed frets on the neck.

3. Preprocessing and statistical analysis

Because the natural timing condition was used mainly as a practice/adjustment task, all data for these series were omitted. For the remaining timing condition data, all strokes were cropped into individual segments via custom scripts in MATLAB (version R2018a) (MathWorks, Natick, MA). Audio recordings were first cropped into individual segments according to the grid points corresponding to the pattern from 400 ms before the grid point to 900 ms after the grid point. Then, the parts of the audio signal that belonged to the previous and following strokes were silenced. Segments that contained no signal around the grid point were marked as missing strokes and removed. Typically, participants would wait one or two measures before beginning to play in order to entrain to the metronome, resulting in 2–4 missing strokes per recording. Timing mistakes, defined as strokes with onsets more than a sixteenth note (156 ms) early or late in relation to the metric grid, were also marked and removed automatically. These events were not considered as instances of early/late microtiming relative to the beat but rather as qualitatively different beat-level syncopations. The automated

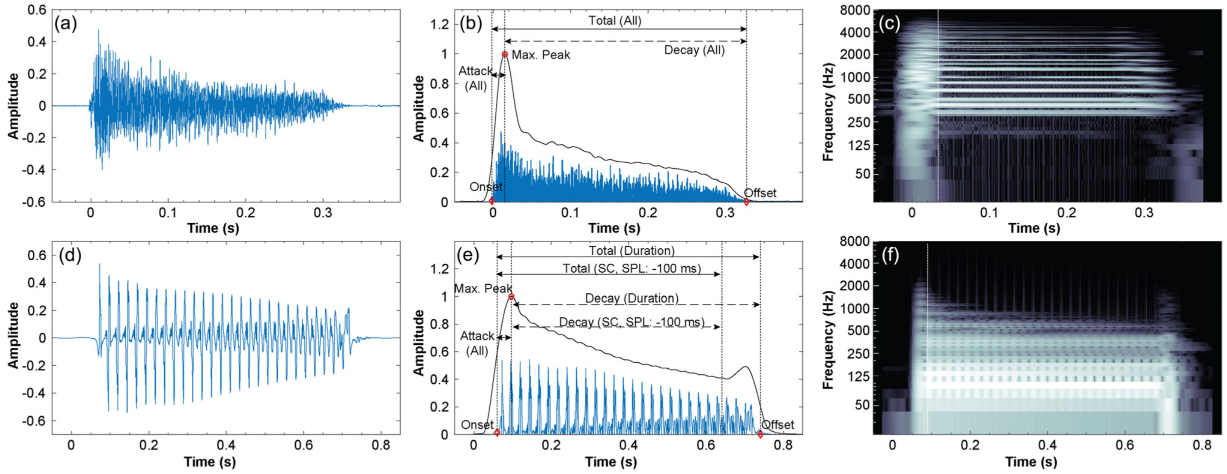

FIG. 2. (Color online) Waveforms of a typical guitar (a) and bass (d) stroke from the experiment. Schematic illustrations of the rectified waveforms and envelopes from onset/offset and attack/decay/total segments were extracted for guitar (b) and bass (e) strokes. Spectrograms display the approximate divide between attack/decay segmentation (white vertical lines) for guitar (c) and bass (f).

process was validated by plotting the waveform for each segment and inspecting the silenced audio parts and the marked missing strokes. Out of all 37 422 total possible strokes captured per experiment [strokes (2) × measures (27) × timing style condition (3) × participant (21) × audio descriptor (11)], missing strokes accounted for ca. 5.3% of the guitar and 5.8% of the bass data, and timing mistakes for ca. 2.5% of the guitar and 1.2% of the bass data. The total amount of invalid missing strokes/timing mistakes was highly comparable between timing conditions (guitar, pushed = 2.6%, on = 2.6%, laid-back = 2.7%; bass, pushed = 2.3%, on = 2.3%, laid-back = 2.4%).

Some bass strokes were stopped actively by the player's slapping or hitting of the string. The resulting pickup sound at the stroke's offset would greatly impact the calculated values for SC and SPL (total and decay) if it were included. Since our main interest was in how these features change over the entire decay rather than only the offset, we removed the last 100 ms from all participants' bass strokes, regardless of whether they featured offset pickups or not, before calculating SC and SPL. Subsequently, the various (attack, decay, and total) brightness (SC), intensity (SPL), and temporal (duration) audio descriptors were calculated for each individual stroke via appropriate MIR Toolbox functions (*mircentroid*, *mirrms*, *mironset/mirevents*).

Before our statistical analysis, extreme outliers were defined as values more than three times the interquartile range away from the median of each timing condition, participant, and audio descriptor separately and excluded. In total, out of all valid strokes captured (guitar, 34 496; bass, 34 859), extreme outliers removed accounted for ca. 0.30% of the guitar and 0.55% of the bass data (guitar, pushed = 0.06%, on = 0.09%, laid-back = 0.15%; bass, pushed = 0.26%, on = 0.13%, laid-back = 0.16%).

Then, to test whether participants systematically manipulated temporal location and sound descriptors between instructed timing style conditions, tests of differences across all participants ($N = 21$) were conducted with instructed timing style as the independent variable and onset, peak, duration, SC, and SPL as the dependent variables, using mean values per participant and timing style condition. To check for normality, within-subjects standardized distributions (average of all strokes across timing style conditions for a given participant subtracted from each individual timing style condition value for the same participant; see Fischer and Milfont, 2010) were manually screened via histograms and *Q-Q* plots for each dependent variable separately. For those series that showed normal standardized distributions, one-way repeated-measures analyses of variances (RMANOVAs) were conducted. For standardized distributions showing departure from normality, corresponding non-parametric tests of differences of medians (Friedman) were conducted instead. All pairwise *post hoc* comparison tests (paired *t*-test or Wilcoxon) were Bonferroni corrected for multiple comparisons (three pairs). Violations of sphericity (Mauchly's test) were corrected using Greenhouse-Geisser.

Finally, in order to gauge the degree to which the sound descriptors generally covary with each other, correlation tests for all instrument strokes across participants and style conditions were conducted between the nine sound descriptors. For descriptor pairs with normal standardized distributions, Pearson's correlations tests were run; for pairs with at least one variable showing departure from normality, Spearman's correlations were run. All statistical analyses were performed using SPSS (version 25, IBM, Inc., New York).

III. RESULTS

An overview of results from the tests of difference (RMANOVA/Friedman) across all participants for the temporal location (onset, peak) and sound (duration, SC, SPL) descriptors can be found in Table I. Average descriptive statistics for all audio descriptors across participants for both instruments can be found in Table II. In Sec. III A, we report detailed results for the different measures.

TABLE I. RMANOVA/Friedman tests on the effect of the timing condition on all audio descriptors across participants for guitarists ($N = 21$) and bassists ($N = 21$). Significant *p*-values appear in bold. DV, dependent variable; err., error.

		Guitar							Bass						
DV	Segment	Statistic		df	df (err.)	*p*	Effect size		Statistic		df	df (err.)	*p*	Effect size	
Onset	—	*F*	40.387	1.262	25.230	<**0.001**[a]	η_p^2	0.669	*F*	105.052	1.143	22.859	<**0.001**[a]	η_p^2	0.840
Peak	—	*F*	62.648	1.328	26.560	<**0.001**[a]	η_p^2	0.758	*F*	95.554	1.164	23.272	<**0.001**[a]	η_p^2	0.827
Duration	Attack	χ^2	12.361	2.000	0.000	**0.002**	*W*	0.294	*F*	0.151	2.000	40.000	0.860	η_p^2	0.008
	Decay	*F*	10.586	2.000	40.000	<**0.001**	η_p^2	0.346	*F*	1.832	1.366	27.327	0.186	η_p^2	0.084
	Total	*F*	15.642	2.000	40.000	<**0.001**	η_p^2	0.439	*F*	1.838	1.347	26.950	0.186	η_p^2	0.084
SC	Attack	χ^2	10.296	2.000	40.000	**0.006**	*W*	0.245	*F*	0.110	2.000	40.000	0.896	η_p^2	0.005
	Decay	*F*	8.619	2.000	40.000	<**0.001**	η_p^2	0.301	χ^2	6.381	2.000	0.000	**0.041**[b]	*W*	0.152
	Total	*F*	9.806	2.000	40.000	<**0.001**	η_p^2	0.329	χ^2	0.286	2.000	0.000	0.867	*W*	0.007
SPL	Attack	*F*	0.441	2.000	40.000	0.646	η_p^2	0.022	*F*	15.292	1.446	28.918	<**0.001**[a]	η_p^2	0.433
	Decay	*F*	0.329	2.000	40.000	0.721	η_p^2	0.016	*F*	14.421	1.520	30.401	<**0.001**[a]	η_p^2	0.419
	Total	*F*	0.223	2.000	40.000	0.801	η_p^2	0.011	*F*	14.664	1.520	30.395	<**0.001**[a]	η_p^2	0.423

[a] Greenhouse-Geisser corrected.

[b] *Post hoc* tests revealed no significant differences between pairwise comparisons (laid-back vs on, pushed vs on, laid-back vs pushed).

TABLE II. Descriptive statistics for all audio descriptors for guitar and bass strokes across all participants for guitarists ($N = 21$) and bassists ($N = 21$). SDs/median absolute deviations appear in parentheses. Note: *Mdn* = Median.

		Guitar				Bass		
		Pushed	On	Laid-back		Pushed	On	Laid-back
Onset (ms)	*M*	−53 (22)	−16 (9)	19 (37)	*M*	−63 (28)	−17 (12)	27 (23)
Peak (ms)	*M*	−28 (23)	7 (10)	51 (33)	*M*	−31 (27)	15 (11)	59 (23)
Attack duration (ms)	*Mdn*	22 (5)	24 (5)	29 (7)	*M*	32 (5)	32 (6)	32 (6)
Decay duration (ms)	*M*	269 (92)	263 (82)	288 (86)	*M*	707 (109)	672 (24)	692 (52)
Total duration (ms)	*M*	295 (93)	286 (84)	321 (88)	*M*	739 (112)	704 (25)	724 (54)
Attack SC (Hz)	*Mdn*	1580 (60)	1540 (60)	1510 (70)	*M*	896 (142)	897 (207)	909 (220)
Decay SC (Hz)	*M*	1806 (168)	1797 (137)	1730 (160)	*Mdn*	563 (133)	570 (151)	517 (139)
Total SC (Hz)	*M*	1756 (151)	1746 (124)	1703 (150)	*Mdn*	270 (18)	268 (23)	271 (20)
Attack SPL (dB)	*M*	−7.86 (2.93)	−8.15 (2.96)	−8.10 (3.41)	*M*	−10.67 (4.37)	−12.10 (4.00)	−13.18 (3.54)
Decay SPL (dB)	*M*	−1.24 (2.70)	−1.13 (2.37)	−1.30 (3.08)	*M*	1.53 (4.24)	0.03 (3.67)	−0.70 (3.17)
Total SPL (dB)	*M*	−0.31 (2.68)	−0.29 (2.42)	−0.42 (3.09)	*M*	1.80 (4.22)	0.31 (3.68)	−0.44 (3.18)

A. Effects of instructed timing on onset, peak, duration, SC, and SPL across participants

1. Experiment 1—Guitar

a. Onset/peak. Results revealed a significant main effect of instructed timing on onset and peak location [see Fig. 3(a)]. *Post hoc* pairwise comparisons further confirmed significant differences at $p < 0.001$ between all timing conditions for onset and peak). These group statistics results suggest that participants produced both the onset and the peak of their strokes, on average, earlier or later in time, respectively, compared to the average on-the-beat position. The descriptive statistics of the microtiming onset and peak profiles for each individual guitarist further show that each participants' average timing also corresponded to the given timing instructions in terms of either onset or peak, or both (see Table V in the supplementary material[1]).

b. Duration. Results show a significant main effect of instructed timing on duration for all stroke segments. *Post hoc* analysis revealed significantly longer laid-back strokes relative to on-the-beat for the attack ($p = 0.006$), decay ($p = 0.001$), and total ($p < 0.001$) segments, as well relative to pushed for the decay ($p = 0.028$) and total ($p = 0.01$) segments. A trend toward significance was found for attack duration ($p = 0.082$). No difference was found between pushed and on-the-beat strokes for all segments [attack, $p = 0.501$; decay, $p = 0.486$; total, $p = 0.337$; see Fig. 4(a)].

c. SC. Results show a significant main effect of instructed timing on SC for all stroke segments. *Post hoc* analysis revealed significantly lower SC for laid-back strokes relative to on-the-beat strokes for all segments (attack, $p = 0.045$; decay, $p = 0.027$; total, $p = 0.\ 01$) as well as relative to pushed strokes (attack, $p = 0.012$; decay, $p < 0.001$; total, $p < 0.001$). No difference was found between pushed and on-the-beat strokes for all segments [attack/decay/total, $p = 1.000$; see Fig. 4(b)].

d. SPL. No significant main effect of instructed timing on SPL was found for any stroke segment segments [see Fig. 4(c)].

2. Experiment 2—Bass

a. Onset/peak. Results revealed a significant main effect of instructed timing on onset and peak location. *Post*

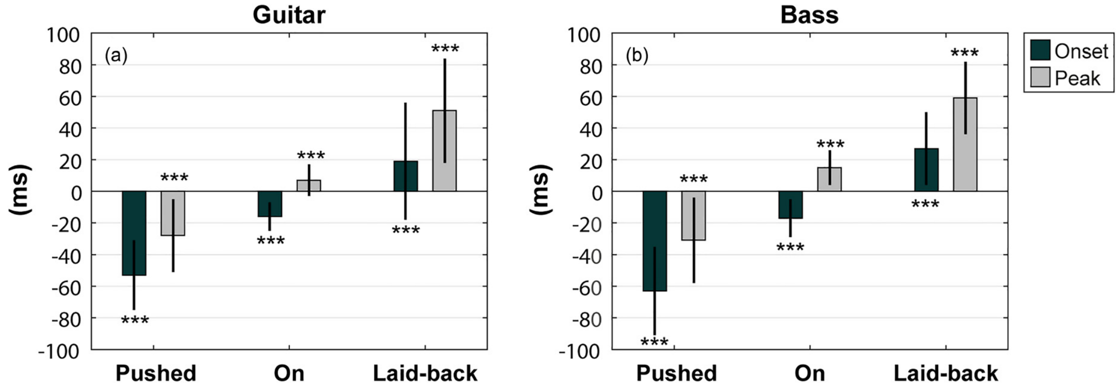

FIG. 3. (Color online) Mean stroke timing of onset and peak across participants relative to the metronome beat reference in all timing conditions for guitarists (a) ($N = 21$) and bassists (b) ($N = 21$). Significance (marked with asterisks) relates to differences within onset or peak conditions separately. Error bars indicate one SD; ***, $p < 0.001$.

JASA
https://doi.org/10.1121/10.0000724

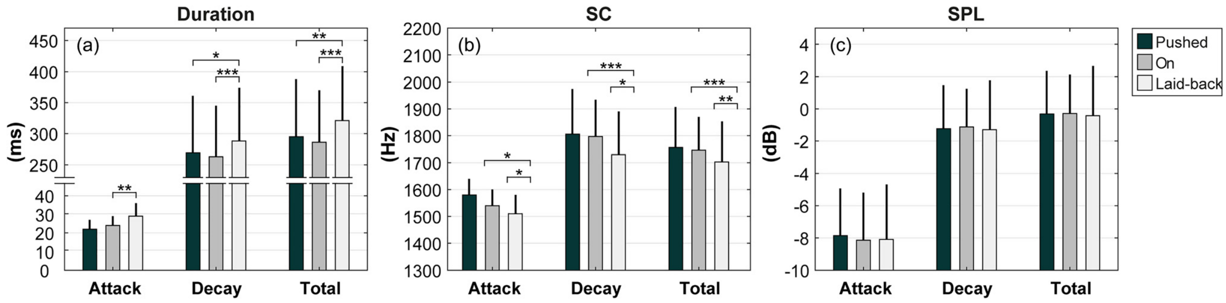

FIG. 4. (Color online) Average guitar stroke segment duration (a), SC (b), and SPL (c) across all participants ($N = 21$). All values and error bars represent mean values and one SD, with the exception of attack duration and attack SC, representing median values and median absolute deviation. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

hoc pairwise comparisons further confirmed significant differences at $p < 0.001$ between all timing conditions for onset and peak [see Fig. 3(b)]. Both the group statistics and descriptive statistics of the microtiming onset and peak profiles for each individual bassist also show that the average timing corresponded to the given timing instructions in terms of either onset or peak, or both (see Table VI in the supplementary material[1]).

b. Duration. No significant main effect of instructed timing on duration was found for any stroke segment segments [see Fig. 5(a)].

c. SC. Although a significant main effect of instructed timing on SC was found on the decay segment, *post hoc* analysis revealed no significant differences between all pairwise comparisons [laid-back vs on, pushed vs on, laid-back vs pushed; $p > 0.05$; see Fig. 5(b)].

d. SPL. Results showed a significant main effect of instructed timing on SPL for all stroke segments. *Post hoc* analysis revealed significantly louder pushed strokes relative to on-the-beat for all segments (attack/decay/total, $p < 0.001$) as well as relative to laid-back strokes (attack/decay/total, $p < 0.001$). No difference was found between laid-back and on-the-beat strokes [$p = 0.111$; decay, $p = 0.296$; total, $p = 0.296$; see Fig. 5(c)].

B. Correlation between stroke duration, SC, and SPL across conditions and participants

Table III gives an overview of correlation results between all sound descriptor pairs. For pairs with normal standardized distributions, the correlation coefficient reported is Pearson's R, whereas for pairs with at least one variable showing departure from normality, Spearman's ρ is reported.

1. Within sound descriptors

For duration and SC, the results for correlation tests between stroke segments within the same sound descriptor category showed that attack segments generally correlated only weakly (sig. correlation coefficient/effect size ±0.1–0.3) or moderately (±0.3–0.5) with the decay and total segments in both guitar and bass. For SPL, on the other hand, attack, decay, and total segments correlated strongly (>0.5) and positively with each other.

2. Between sound descriptors

For the guitar strokes, correlation results between the different sound descriptors revealed weak to moderate significant correlations among all three sound descriptor categories. For the moderate correlations, the general trends were (1) the greater the duration, the lower the SC; (2) the greater the SPL, the longer the duration; and (3) the greater the SPL, the higher the SC.

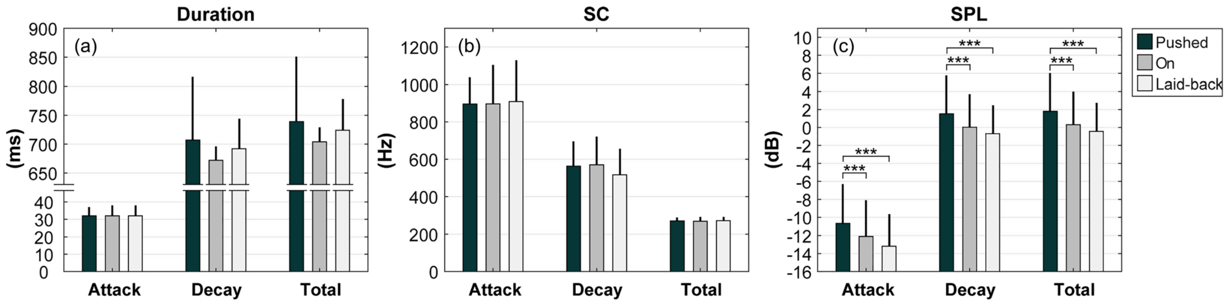

FIG. 5. (Color online) Average bass stroke segment duration (a), SC (b), and SPL (c) across all participants ($N = 21$). All values and error bars represent mean values and one SD, with the exception of decay/total SC, representing median values and median absolute deviation. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

https://doi.org/10.1121/10.0000724

JASA

TABLE III. Pearson/Spearman correlations between all sound descriptors for all guitar (lower-left triangle) and bass (upper-right triangle) strokes across style conditions and participants. Cor., correlation coefficient; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

			Bass								
			Attack duration	Decay duration	Total duration	Attack SC	Decay SC	Total SC	Attack SPL	Decay SPL	Total SPL
Guitar	Attack duration	Cor.		0.190^{***}	0.282^{***}	-0.249^{***}	0.020	0.024	-0.201^{***}	-0.161^{***}	-0.168^{***}
	Decay duration	Cor.	0.114^{***}		0.995^{***}	-0.068^{***}	0.037^{*}	0.102^{***}	-0.070^{***}	0.110^{***}	0.092^{***}
	Total duration	Cor.	0.320^{***}	0.985^{***}		-0.092^{***}	0.036^{*}	0.109^{***}	-0.086^{***}	0.093^{***}	0.075^{***}
	Attack SC	Cor.	-0.452^{***}	-0.013	-0.115^{***}		0.896^{***}	-0.057^{***}	0.128^{***}	0.021	0.030
	Decay SC	Cor.	-0.354^{***}	-0.272^{***}	-0.325^{***}	0.473^{***}		0.005	0.223^{***}	0.091^{***}	0.106^{***}
	Total SC	Cor.	-0.207^{***}	-0.315^{***}	-0.340^{***}	0.324^{***}	0.894^{***}		-0.595^{***}	-0.731^{***}	-0.715^{***}
	Attack SPL	Cor.	0.095^{***}	0.134^{***}	0.115^{***}	-0.058^{***}	0.384^{***}	0.433^{***}		0.923^{***}	0.937^{***}
	Decay SPL	Cor.	-0.140^{***}	0.325^{***}	0.274^{***}	0.032	0.237^{***}	0.265^{***}	0.809^{***}		0.999^{***}
	Total SPL	Cor.	-0.101^{***}	0.296^{***}	0.249^{***}	0.010	0.277^{***}	0.312^{***}	0.874^{***}	0.992^{***}	

For bass strokes, only pairings between the categories of SPL and SC revealed significant strong correlations, with the general trend being that the greater the attack/decay/total SPL, the lower the total SC. All other correlation pairs revealed only weak correlations (correlation coefficient < 0.30; see Table III).

IV. DISCUSSION

The analyses show that there were systematic differences in the acoustic signal among strokes played with different instructed timing styles in at least one sound descriptor category for both bassists and guitarists. The results support our main hypothesis that expert musicians systematically manipulate not only temporal locations of strokes but also sound parameters in order to communicate different intended microtiming feels ("timing styles"). In the following, we discuss these findings in more detail.

In terms of the temporal location of strokes, the descriptive statistics of the individual participants' timing show that when asked to play in a laid-back or pushed manner, each guitarist and bassist produced either the onset and/or the peak of their strokes on average earlier or later in time, respectively, compared to their own corresponding average on-the-beat position. The RMANOVA test results found that, across all participants, the difference in mean onset/peak timing of laid-back and pushed strokes relative to the on-beat condition was significant for both instruments, and approximately 40 ms in value. This corresponds roughly to the duration of a sixty-fourth note displacement (39.1 ms) from the metronome reference at 96 bpm, which indicates a refined degree of intentional control at the microrhythmic level. Such a value is above the just noticeable difference (JND) threshold for inter-onset interval (IOI) displacement in an isochronous sequence reported by Friberg and Sundberg (1995), which was constantly relative at approximately 2.5% of IOI for tone sequences above 250 ms. In this experiment, the IOI between the target guitar stroke locations and metronome beat reference was roughly 625 ms, meaning that the theoretical JND for detecting guitar strokes as off the beat would be approximately 16 ms. While such a finding may sound mundane, it is, in fact, impressive. Despite many ethnographic claims regarding the fine timing precision of musicians (Berliner, 1994; Monson, 1996; Doffman, 2019), as well as some skepticism about qualitative accounts of such purported feats of accuracy (Butterfield, 2010), actual quantitative data that prove the success and degree of this precision are relatively scarce. This provides further evidence that, at least in the performance of simple groove-based patterns at medium tempi, professional musicians are able to voluntarily control the timing of their rhythm on the order of tens of milliseconds and, furthermore, are able to produce perceptibly different stroke timing locations for laid-back, on-the-beat, and pushed feels.

On the one hand, the fact that the mean onset timing of the on-the-beat strokes tended to anticipate the metronome reference across participants for both instruments (guitar, $M = -16$ ms, SD = 9 ms; bass, $M = -17$ ms, SD = 12 ms) may be related to the effect of the phenomenon of "negative mean asynchrony" (NMA)—that is, the tendency to tap slightly ahead of a series of metronome clicks that emerges in tapping synchronization tasks. Results from various studies show that NMA ranges vary from 20 to 80 ms for untrained subjects to 0–30 ms for musicians (Repp, 2005; Repp and Dogget, 2007; Repp and Su, 2013). Because tapping with a finger in synchrony to a beat is a mode of motor synchronization not unlike manually producing strokes on a guitar or bass, NMA may also explain the consistently early on-the-beat onsets of instrument performance. On the other hand, the fact that the mean peak of the on-the-beat strokes was closer to the metronome onsets and not negative/anticipatory (guitar, $M = 7$ ms, SD = 2 ms; bass, $M = 15$ ms, SD = 2 ms) lends weight to a P-center interpretation: Just as the onset of a sound is not necessarily the moment in time at which it is perceived as synchronous in relation to another sound in an isochronous context, it would seem that guitarists and bassists use some other, later moment closer to the maximum peak as a target when synchronizing their strokes to the metronome reference sound.

Regarding the accuracy of strokes in the different timing conditions, mean onset/peak variability was lower in the on-beat conditions relative to laid-back and pushed, perhaps because most of the participants reported practicing synchronous, on-the-beat playing more than asynchronous, laid-back or pushed playing over the course of their careers. Many also described the latter two conditions as more difficult to apply in an ecologically valid fashion in the context of the experiment. On-the-beat playing was also described as more straightforward in terms of defining an acceptable degree of deviation from the metronome beats (very slight) relative to laid-back and pushed playing, which were considered more ambiguous timing categories that allowed for a bigger timing window or degree of asynchrony from the beat reference, while still sounding "correct." In fact, before and during the experiment, participants would often request extra practice runs prior to the laid-back or pushed conditions in order to decide upon the aesthetically appropriate amount of asynchrony—too far behind or ahead of the beat risked sounding wrong or syncopated ("offbeat"), whereas too near risked not feeling different enough from strict on-the-beat playing. This negotiation indicates that participants heard the beat-bin range (Danielsen, 2010; Danielsen *et al.*, 2019) of pushed and laid-back strokes as much wider than that of on-the-beat strokes. Furthermore, during successful laid-back or pushed trials, participants reported that the degree of asynchrony did not have to be too consistent over the course of a condition recording but instead could fluctuate somewhat while still feeling "right"—in short, a factor that would contribute to greater overall variability. Such flexibility of timing seems to be a hallmark of groove-based performance, and different degrees of deliberate asynchrony from the beat are like shades of color on a musician's palette of expressive devices, affording a myriad of temporal "landscapes" (Câmara and Danielsen, 2019).

As for the relationship between timing and the sound descriptors, across participant differences between instructed timing style conditions were found in duration and frequency for guitar and in intensity for bass. For the guitarists (experiment I), we found a significant effect of timing instruction on duration for all stroke segments, with the general trend that laid-back strokes were played with a slower attack and longer decay/total duration than on-the-beat and pushed strokes. No significant difference was found for duration between pushed and on-the-beat, which shared faster attacks and shorter decay profiles. In guitar performance with a plectrum, polytonal chords are necessarily played as arpeggios, where one note/string is struck at a time in succession. Therefore, guitarists may control attack duration to some extent by decreasing or increasing the amount of time that the plectrum is in contact with the strings during the stroke impulse, and decay duration by decreasing or increasing the amount of time the strings are allowed to vibrate thereafter, before muting. A majority of the participants reported that the strategies they used to achieve the laid-back timing style involved applying "slower/looser/wider/larger" movements to strokes; some of the guitarists even mentioned deliberately lengthening attack durations by "sweeping" (arpeggiating) the strings in a slower manner, or lengthening decay durations by "sliding out" the endings of the chords, or a combination of both techniques. Conversely, they reported "faster/tighter/narrower/smaller" movements to achieving the on and pushed feels, as well as the use of shorter stroke durations. These reported strategies corroborate the signal analysis results.

These guitar duration results are in line with results from P-center studies, which show that longer durations lead to later P-centers. That is, while the quintessential character of backbeat guitar playing in most groove-based styles is relatively short and percussive, comparatively longer attack and decay durations might, in fact, augment the "behind-the-beat" quality of laid-back strokes, while comparatively shorter strokes might frame on-the-beat and pushed timing conditions as either more in sync or earlier, respectively, relative to the metronome. If the primary means of achieving microrhythmic laid-back performance is to communicate late timing in relation to a beat reference, longer durations would tend to further delay the experienced P-center location of a sound. Therefore, a *late timing + slow attack and long duration* combination would appear to represent the ideal performance strategy for guitarists to further convey a laid-back condition. Conversely, an *early timing + fast attack and short duration* combination would be the ideal strategy to convey a pushed condition. Although previous P-center findings suggest that shorter attacks lead to earlier P-centers, pushed strokes were not statistically shorter than on-beat strokes—there may have been physical limitations to how fast players could strike the notes/strings of the chord to produce ever shorter attack durations. As to why the decay (and thus the total duration) of the pushed strokes was not shorter, it could be that such a short stroke was not aesthetically appealing.

Regarding whether the differences in stroke duration between these timing style conditions would be perceivable to the average listener, Johnson and colleagues (1987) found that the Weber fraction (WF, or JND threshold divided by stimuli duration) of a 1000 Hz tone ranged between 0.09 and 0.16. Based on the average WF of these heuristics (0.125), the theoretical JND threshold for detecting a difference in duration from an average total on-the-beat stroke of 286 ms would be approximately 36 ms. Under ideal listening conditions, then, it is likely that, at the very least, the laid-back guitar strokes would be heard as different (that is, longer) from the on-the-beat strokes in terms of total duration because they differed by 35 ms (±6), on average.

We also found significant differences in SC in guitar strokes for all segments (attack, decay, total), where laid-back strokes were played with lower SC compared to both on-the-beat and pushed strokes. Once again, no difference between pushed and on-the-beat conditions was found. Studies have generally determined that the higher the SC of a sound, the "brighter" or "sharper" it tends to be perceived, and, conversely, the lower the SC, the "darker" or "duller" it tends to be perceived (Schubert and Wolfe, 2006). For guitarists, then, a general strategy of combining *late timing + dark timbre* would appear to best complement laid-back strokes, and a strategy of combining *early timing + bright*

https://doi.org/10.1121/10.0000724

timbre would appear to best complement pushed and on-the-beat strokes. Regarding P-center, there is currently no consensus on the effect of pitch/frequency, but two studies with musical sound stimuli (Danielsen *et al.*, 2019; Gordon, 1987) found that lower frequencies lead to later P-centers and higher frequencies lead to earlier P-centers, but primarily in sounds with slower attack durations and/or longer total durations—conditions met by the guitar strokes in this experiment.

Differences in SC may have been achieved by guitarists via either: greater or lesser emphasis upon either the lower-pitched notes (that is, the darker strings) or the higher-pitched notes (the brighter strings) of the chord during arpeggiation, achieved by the selective aiming of the plectrum toward, or the partial muting of the fretting hand of, the particular strings in question; or striking the strings at different locations of the fretboard. The closer to the center a string is struck (toward the "neck"), the greater the amplitude of the fundamental harmonics relative to the upper partials and the "rounder" or duller the resultant sound. Conversely, the further away from the center a string is struck (toward the "bridge"), the lower the amplitude of the fundamental harmonics and the "thinner" or brighter the resultant sound (Rossing *et al.*, 2002, p. 218). For the laid-back timing condition, then, guitarists might have either played closer to the neck or placed more emphasis on the lower-pitched strings; combined with the downstroke plectrum technique also described by most of the participants (that is, the highest pitches/strings struck last in a chord), this emphasis would mean that the moment in time in which the SC was highest would be later in the attack segment, eliciting a later P-center, in turn. For on-the-beat and pushed strokes, opposite tactics (more emphasis on the higher-pitched strings/playing closer to the bridge) would contribute to the effect of relatively brighter strokes, potentially eliciting an earlier P-center that would effectively communicate strokes meant to either synchronize with or anticipate the metronome beat reference.

Unfortunately, no perceptual studies have been conducted specifically on the JNDs of the SC between two sounds, so it is difficult to assess whether the quantified differences in SC between style conditions would translate to perceptible differences in brightness.

We found moderate negative correlations between guitar stroke duration and SC: the longer a segment duration, the lower the SC tended to be. Throughout the course of a typical guitar stroke, the SC tends to rise from onset until peak, and then gradually fall toward offset. Longer durations would, in most cases, entail longer decays so that the lower SC decay portions of the sound would comprise a larger portion of the overall sound spectrum.

No significant differences in SPL were found between the timing style conditions for the guitarists. This result is somewhat surprising, given that existing research on drummers (Danielsen *et al.*, 2015) and this study's results regarding bassists indicate that both groups utilize intensity to convey different timing feels. For guitarists, apparently, manipulating intensity was either not ecological in the context of the performance task or less efficient/redundant for further enhancing asynchronous timing feel, given that duration and timbre were already being utilized to that advantage.

Turning to the sound descriptor analysis results for the bassists (experiment II), we found significant differences in SPL between timing style conditions for all stroke segments (attack, decay, total), where pushed strokes were played with significantly greater intensity relative to both on-the-beat and laid-back strokes. No significant differences were found between laid-back and on-the-beat strokes. Post-interview feedback from participants corresponded to these results in that several people reported playing "harder" strokes to achieve a pushed feel and "softer" strokes to achieve a laid-back feel. Pushed playing was also reported to require greater mental effort and bodily tension compared to either laid-back or on, meaning that greater intensity might have been an artifact of increased performance difficulty. However, melody-lead performance studies with pianists have also shown that, when they are asked to emphasize melody notes, pianists tend to play them both earlier in time and louder relative to the other tones (Goebl and Parncutt, 2002). In addition, the P-center tends to be experienced as earlier in sounds presented at higher rather than lower levels of intensity (Bechtold and Senn, 2018; Gordon, 1987; Seton, 1989). Therefore, our bassists may have chosen to emphasize the microrhythmic earliness of a pushed stroke in relation to the metronome reference by increasing intensity, which elicits both an earlier P-center and ensures that the bass's stroke is not masked by the metronome and heard as leading rather than following it.

Perception studies have also shown a relationship between dynamic accents and perceived timing. Tekman (2002), for example, found that timing differences were easier to detect in tone sequences when coupled with greater intensity. Although he only investigated the interaction between tones with delayed timing and intensity, it is possible that the conclusion applies to tones with anticipated timing as well, meaning that pushed bass strokes that are played with a greater intensity (that is, louder) would likewise facilitate their perception as actually timed early. As for the fact that on-the-beat and laid-back strokes were played with lower intensity, it may be the case that softer strokes helped communicate the condition that the bass is to be heard as either synchronous/blended with or subordinate to/following the metronome—or, in an actual musical context, any other instrument that supplies the main beat reference such as the drums.

Regarding JNDs of intensity, several studies have found that intensity discrimination of sinusoid tones is not, in fact, dependent on signal frequency but rather on sensation level or perceived loudness of stimuli presentation, whereby the louder the presentation level, the lower the JND of intensity of a given tone (Elliot *et al.*, 1966; Jesteadt *et al.*, 1977; Johnson *et al.*, 1987). At louder presentation levels of 70–80 dB (not uncommon in live music performance contexts), typical reported JNDs tend to fall around 0.4–0.8 dB.

Using such a heuristic for the bass and considering only the total SPL of strokes [because attack, decay, and total SPL strongly correlated with one another ($R > 0.9$)], the differences in intensity between pushed and on-the-beat ($+1.49\,\text{dB} \pm 0.29$) and pushed and laid-back ($+2.24\,\text{dB} \pm 0.51$) strokes are likely readily perceptible to the average listener at louder presentation levels.

The absence of differences in duration and SC for the bassists may have several explanations, mostly involving the constraints of the experiment's design and the instrument's construction. Regarding attack duration, while the guitarists had several tones to space out incrementally, the bassists may have been hard-pressed to regulate the location of the maximum amplitude peak in relation to the tone onset through playing technique alone, which results in a more-or-less fixed interval between onset and peak. Regarding decay and total duration, we may have lost critical information due to an artifact of our offset trimming: The clearly demarcated onsets of the metronome on beats two and four of the metrical grid served as a cue to maintain precise quarter-note intervals. In comparison, the backbeat pattern given to the guitarists involved relatively shorter eighth-note durations, the ends of which were unaccented by any metronome reference. As such, these durations allowed for greater ambiguity regarding how long one should play the guitar strokes and thus allowed for multiple offset or decay/total duration interpretations. Regarding the lack of SC differences in the bass results, either the bassists were relatively consistent in where they struck the strings on the fretboard across timing style conditions, or the range of tone brightness' afforded by bridge vs neck locations are not pronounced enough to allow for production of sufficiently different SCs.

Particularly notable from the results, in general, is the fact that when asked to communicate an intentional timing feel, our instrumentalists all seemed to use the sound materials available to them at the time—that is, they "do what they can with what they have." Bassists were limited to manipulating the intensity of tones, in addition to their onset timing, and therefore seized this possibility as an additional means of communicating the timing feel, whereas guitarists had greater access to at least three sound parameters and chose overwhelmingly (across all participants) to utilize at least two of them (duration and SC).

V. CONCLUSION

The results show that there were systematic differences in the duration and brightness (SC) of a series of guitar chords, and intensity (SPL) of a series of single bass tones, played with different timing instructions (laid-back, pushed, and on-the-beat) across all participants. For guitarists, the overarching sound manipulation strategy to achieve laid-back timing was one of *late* = *slower attack* + *longer duration* + *darker timbre*. Conversely, the overall strategy in order to achieve on-the-beat and pushed timing was an *on-beat and early* = *faster attack* + *shorter duration* + *brighter timbre*. These findings accord with research into the P-centers of musical sounds in the sense that the guitarists would seem to be exploiting the delayed P-center effects of longer duration and lower frequency to convey laid-back timing and, conversely, the anticipatory P-center effects of shorter duration and higher frequency to convey pushed and on-beat timing. For the bassists, the overall timing strategies were *early* = *louder* and, conversely, *on-beat and late* = *softer*, evoking the strategy of intensity manipulation to emphasize onset asynchronies found in a previous instructed timing study of snare drum performance (Danielsen *et al.*, 2015) and a perceptual intensity–timing interaction study (Goebl and Parncutt, 2002).

Our findings lend support to the hypothesis that both temporal and sound-related aspects are important to musicians in order to communicate an intended timing feel in performance. As such, future empirical studies concerned with measuring and interpreting the role of microtiming expression in groove-based music would do well to consider the durational, dynamic, and timbral aspects of events in addition to onset timing and potential interactions between them.

In future research, we plan to conduct perception experiments to determine whether listeners are able to identify bass and guitar strokes as laid-back, on-the-beat, or pushed on the basis of their sound alone. In addition, we recorded motion-capture data during the aforementioned experiment in order to investigate whether musicians also systematically utilize different movement trajectories to produce the differences in sound observed in this study.

ACKNOWLEDGMENTS

The authors wish to thank all the participating guitarists and bassists: Georgios Sioros, Victor Gonzáles Sánchez, and Martin Torvik Langerød (University of Oslo, Norway) for their assistance with the experimental setup; Carl Haakon Waadeland (Norwegian University of Science and Technology, Norway) for his help with the experimental design and recruitment of participants; Dag-Erik Eilertsen (University of Oslo, Norway) for his assistance with SPSS scripts for the statistical analysis; and Justin London (Carleton College, Northfield, MN, U.S.A.) for his comments on an early version of this manuscript. We are also grateful to the anonymous reviewers for their interesting and valuable comments and suggestions. This work was partially supported by the Research Council of Norway through its Centers of Excellence scheme, Project No. 262762 and the TIME (Timing and Sound in Musical Microrhythm) project, Grant No. 249817.

APPENDIX: BEHAVIORAL MEASUREMENT OF METRONOME SOUNDS' PERCEPTUAL CENTERS

For a rhythmic tone sequence to be perceptually isochronous, the intervals between the sounds' perceived moment of occurrence (Morton *et al.*, 1976) must be even. We thus conducted a short experiment to measure the perceptual center (P-center) of the metronome woodblock tones. These P-centers were then aligned to the 0 ms locations of the isochronous software beat grid in experiments 1 and 2. The P-centers were determined using a CA task

TABLE IV. Acoustic measurements and descriptive statistics of P-center results for metronome woodblock sounds: P-center average locations relative to the physical onset of each stimulus ($N = 10$). PC, P-center.

Pitch	Attack duration (ms)	Total duration (ms)	SC (Hz)	PC *M* (ms)	PC SD (ms)
High	2	40	2370	0	2
Low	2	50	1565	−4	7

developed by Danielsen, London and Nymoen (see Danielsen *et al.*, 2019; London *et al.*, 2019). Ten participants aligned a probe click sound (1 ms attack duration, 3000 Hz center frequency) with each of the two target stimuli three times for a total of six trials. For details regarding methods and procedure, see Danielsen *et al.* (2019). Responses were averaged across the three trials of each sound for each participant. The mean of means for each sound was calculated to produce a P-center location for each stimulus; P-centers are reported in milliseconds relative to the physical onset of the stimulus in Table IV.

---

[1] See supplementary material at https://doi.org/10.1121/10.0000724 for onset and peak descriptive statistics of each individual guitar and bass participant (Table V).

Algom, D., and Fitousi, D. (**2016**). "Half a century of research on Garner interference and the separability–integrality distinction," Psychol. Bull. **142**(12), 1352–1383.

Bechtold, T. A., and Senn, O. (**2018**). "Articulation and dynamics influence the perceptual attack time of saxophone sounds," Front. Psychol. **9**, 1692.

Berliner, P. (**1994**). *Thinking in Jazz: The Infinite Art of Improvisation* (University of Chicago Press, Chicago).

Butterfield, M. W. (**2006**). "The power of anacrusis: Engendered feeling in groove-based musics," Music Theory Online **12**(4).

Butterfield, M. W. (**2010**). "Participatory discrepancies and the perception of beats in jazz," Music Percept. **27**(3), 157–175.

Butterfield, M. W. (**2011**). "Why do jazz musicians swing their eighth notes?," Music Theory Spect. **33**(1), 3–26.

Câmara, G. S. (**2016**). "Swing in early funk and jazz-funk (1967–1971): Micro-rhythmic and macro-structural investigations," Master's dissertation, University of Oslo, Oslo, Norway.

Câmara, G. S., and Danielsen, A. (**2019**). "Groove," in *The Oxford Handbook of Critical Concepts in Music Theory*, edited by A. Rehding and S. Rings (Oxford University Press, New York), pp. 271–294.

Clarke, E. F. (**1987**). "Categorical rhythm perception: An ecological perspective," in *Action and Perception in Rhythm and Music*, edited by A. Gabrielsson (Royal Swedish Academy of Music, Stockholm), pp. 19–34.

Clarke, E. F. (**1989**). "Generative principles in music performance," in *Generative Processes in Music*, edited by J. A. Sloboda (Clardendon, Oxford, UK), pp. 1–26.

Dahl, S. (**2000**). "The playing of an accent—Preliminary observations from temporal and kinematic analysis of percussionists," J. New Music Res. **29**, 225–233.

Dahl, S. (**2004**). "Playing the accent—Comparing striking velocity and timing in an ostinato rhythm performed by four drummers," Acta Acoust. Acoust. **90**(4), 762–776.

Danielsen, A. (**2006**). *Presence and Pleasure: The Funk Grooves of James Brown and Parliament* (Wesleyan University Press, Middletown, CT), Chap. 5.

Danielsen, A. (**2010**). "Here, there and everywhere: Three accounts of pulse in D'Angelo's 'Left and right'," in *Musical Rhythm in the Age of Digital Reproduction*, edited by A. Danielsen (Ashgate, Farnham, UK), pp. 19–36.

Danielsen, A. (**2018**). "Pulse as dynamic attending: Analysing beat bin metre in neo soul grooves," in *The Routledge Companion to Popular Music Analysis: Expanding Approaches*, edited by C. Scotto, K. M. Smith, and J. Brackett (Routledge, London), pp. 179–189.

Danielsen, A., Nymoen, K., Anderson, E., Câmara, G. S., Langerød, M. T., Thompson, M., and London, J. (**2019**). "Where is the beat in that note? Effects of attack, duration and frequency on the perceived timing of musical and quasi-musical sounds," J. Exp. Psychol. Hum. Percept. Perform. **45**(3), 402–418.

Danielsen, A., Waadeland, C. H., Sundt, H. G., and Witek, M. A. G. (**2015**). "Effects of instructed timing and tempo on snare drum sound in drum kit performance," J. Acoust. Soc. Am. **138**(4), 2301–2316.

Davies, M., Madison, G., Silva, P., and Gouyon, F. (**2013**). "The effect of microtiming deviations on the perception of groove in short rhythms," Music Percept. **30**, 497–510.

Doffman, M. (**2019**). "Practical time consciousness in performance," in *Music and Consciousness II*, edited by D. Clarke, E. Clarke, and R. Herbert (Oxford University Press, Oxford, UK), pp. 170–186.

Donnadieu, S. (**1987**). "Mental representation of the timbre of complex sounds," in *Analysis, Synthesis, and Perception of Musical Sounds: The Sound of Music*, edited by J. W. Beauchamp (Springer, New York), pp. 272–319.

Drake, C., and Palmer, C. (**1993**). "Accent structures in music performance," Music Percept. **10**, 343–378.

Elliott, D. N., Riach, W. D., Sheposh, J. P., and Trahiotis, C. (**1966**). "Discrimination performance of high school sophomores on a battery of auditory tests," Acta. Otolaryngol. Suppl. **216**, 1–59.

Fischer, R., and Milfont, T. L. (**2010**). "Standardization in psychological research," Int. J. Psychol. Res. **3**(1), 88–96.

Frane, A. V., and Shams, L. (**2017**). "Effects of tempo, swing density, and listener's drumming experience on swing eetection thresholds for drum rhythms," J. Acoust. Soc. Am. **141**, 4200–4208.

Friberg, A., and Sundberg, J. (**1995**). "Time discrimination in a monotonic, isochronous sequence," J. Acoust. Soc. Am. **98**(5), 2524–2531.

Friberg, A., and Sundström, A. (**2002**). "Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern," Music Percept. **19**, 333–349.

Frühauf, J., Kopiez, R., and Platz, F. (**2013**). "Music on the timing grid: The influence of microtiming on the perceived groove quality of a simple drum pattern performance," Music. Sci. **17**, 246–260.

Gabrielsson, A. (**1974**). "Performance of rhythm patterns," Scand. J. Psychol. **15**, 63–72.

Gabrielsson, A. (**1999**). "The performance of music," in *The Psychology of Music*, edited by D. Deutsch (Academic, New York), pp. 501–602.

Goebl, W. (**2001**). "Melody lead in piano performance: Expressive device or artifact?," J. Acoust. Soc. Am. **110**(1), 563–572.

Goebl, W., and Parncutt, R. (**2002**). "The influence of relative intensity on the perception of onset asynchronies," in *Proceedings from ICMPC7, Sydney, Australia*, edited by C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (Causal Productions for AMPS, Adelaide), pp. 613–616.

Gordon, J. W. (**1987**). "The perceptual attack time of musical tones," J. Acoust. Soc. Am. **82**(1), 88–105.

Iyer, V. (**2002**). "Embodied mind, situated cognition, and expressive microtiming in African-American music," Music Percept. **19**, 387–414.

Janata, P., Tomic, S. T., and Haberman, J. M. (**2012**). "Sensorimotor coupling in music and the psychology of the groove," J. Exp. Psychol. **141**, 54–75.

Jesteadt, W., Wier, C. C., and Green, D. M. (**1977**). "Intensity discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **61**, 169–177.

Johnson, D. M., Watson, C. S., and Jensen, J. K. (**1987**). "Individual differences in auditory capabilities," J. Acoust. Soc. Am. **81**, 427–438.

Keil, C. (**1987**). "Participatory discrepancies and the power of music," Cult. Anthropol. **2**, 275–283.

Kernfeld, B. (**2003**). "Groove (i)," Grove Music Online, available at https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-2000582400 (Last viewed 11 Dec. 2019).

Kilchenmann, L., and Senn, O. (**2011**). " 'Play in time, but don't play time': Analyzing timing profiles in drum performance," in *Proceedings of the International Symposium on Performance Science 2011*, edited by A. Williamon, D. Edwards, and L. Bartel (European Association of Conservatoires, Utrecht, The Netherlands), pp. 593–598.

Kilchenmann, L., and Senn, O. (**2015**). "Microtiming in swing and funk affects the body movement behavior of music expert listeners," Front. Psych. **6**, 1232.

Lartillot, O., Toiviainen, P., and Eerola, T. (**2008**). "A MATLAB Toolbox for music information retrieval," in *Data Analysis, Machine Learning and Applications, Studies in Classification, Data Analysis, and Knowledge Organization*, edited by C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker (Springer, Berlin), pp. 261–268.

London, J., Nymoen, K., Langerød, M. T., Thompson, M. R., Code, D. L., and Danielsen, A. (**2019**). "A comparison of methods for investigating the perceptual center of musical sounds," Atten. Percept. Psychophys. **81**(6), 2088–2101.

Madison, G. (**2006**). "Experiencing groove induced by Music: Consistency and phenomenology," Music Percept. **24**(2), 201–208.

Madison, G., Gouyon, F., Ullén, F., and Hörnström, K. (**2011**). "Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres," J. Exp. Psychol. **37**, 1578–1594.

Madison, G., and Sioros, G. (**2014**). "What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it," Front. Psychol. **5**, 894.

Melara, R. D., and Marks, L. E. (**1990**). "Interaction among auditory dimensions: Timbre, pitch, and loudness," Attn. Percept. Psychophys. **48**, 169–178.

Monson, I. (**1996**). *Saying Something: Jazz Improvisation and Interaction* (University of Chicago Press, Chicago), Chap. 3.

Morton, J., Marcus, S., and Frankish, C. (**1976**). "Perceptual centers (P-centers)," Psychol. Rev. **83**(5), 405–408.

Palmer, C. (**1996**). "On the assignment of structure in music performance," Music Percept. **14**(1), 23–56.

Povel, D. J., and Okkerman, H. (**1981**). "Accents in equitone sequences," Attn. Percept. Psychophys. **30**, 565–572.

Pressing, J. (**2002**). "Black Atlantic rhythm: Its computational and transcultural foundations," Music Percept. **19**, 285–310.

Repp, B. H. (**1996**). "Patterns of note onset asynchronies in expressive piano performance," J. Acoust. Soc. Am. **100**(6), 3917–3932.

Repp, B. H. (**2005**). "Sensorimotor synchronization: A review of the tapping literature," Psychon. Bull. Rev. **12**(6), 969–992.

Repp, B. H., and Doggett, R. (**2007**). "Tapping to a very slow beat: A comparison of musicians and non-musicians," Music Percept. **24**(4), 367–376.

Repp, B. H., and Su, Y.-H. (**2013**). "Sensorimotor synchronization: A review of recent research (2006–2012)," Psychon. Bull. Rev. **20**(3), 403–452.

Rossing, T. D., Moore, F. R., and Wheeler, P. A. (**2002**). *The Science of Sound*, 3rd ed. (Addison Wesley, San Francisco).

Schubert, E., and Wolfe, J. (**2006**). "Does timbral brightness scale with frequency and spectral centroid?," Acta Acust. Acust. **92**, 820–825.

Scott, S. K. (**1998**). "The point of P-centres," Psychol. Res. **61**, 4–11.

Senn, O., Bullerjahn, C., Kilchenmann, L., and von Georgi, R. (**2017**). "Rhythmic density affects listeners' emotional response to microtiming," Front. Psychol. **8**(10), 1709.

Senn, O., Kilchenmann, L., von Georgi, R., and Bullerjahn, C. (**2016**). "The effect of expert performance microtiming on listeners' experience of groove in swing or funk music," Front. Psychol. **7**, 1487.

Seton, J. C. (**1989**). "A psychophysical investigation of auditory rhythmic beat perception," Ph.D. thesis, University of York, York, UK.

Skansaar, J., Laeng, B., and Danielsen, A. (**2019**). "Microtiming and mental effort: Onset asynchronies in musical rhythm modulate pupil size," Music Percept. **37**(2), 111–133.

Tekman, H. G. (**2002**). "Perceptual integration of timing and intensity variations in the perception of musical accents," J. Gen. Psychol. **129**, 181–191.

Villing, R. (**2010**). "Hearing the moment: Measures and models of the perceptual centre," Ph.D. dissertation, National University of Ireland, Maynooth, Ireland.

Vos, P. G., Mates, J., and van Kruysbergen, N. W. (**1995**). "The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies," Quart. J. Exp. Psychol. **48**(A4), 1024–1040.

Waadeland, C. H. (**2006**). "Strategies in empirical studies of swing groove," Stud. Musicolog. Norv. **32**, 169–191.

Windsor, W. L. (**1993**). "Dynamic accents and the categorical perception of metre," Psychol. Music. **21**(2), 127–140.

Woodrow, H. (**1909**). "A quantitative study of rhythm," Arch. Psychol. **14**, 1–66.

Wright, M. J. (**2008**). "The Shape of an Instant: Measuring and Modeling Perceptual Attack Time with Probability Density Functions," Ph.D. dissertation, Stanford University, Stanford, CA.

Paper II

Timing is everything . . . or is it? Effects of instructed timing style, reference, and pattern on drum kit sound in groove-based performance

Guilherme Schmidt Câmara, Kristian Nymoen, Olivier Lartillot, and Anne Danielsen

In *Music Perception* 38 (1), 1–26 (2020)

https://doi.org/10.1525/mp.2020.38.1.1

TIMING IS EVERYTHING...OR IS IT? EFFECTS OF INSTRUCTED TIMING STYLE, REFERENCE, AND PATTERN ON DRUM KIT SOUND IN GROOVE-BASED PERFORMANCE

GUILHERME SCHMIDT CÂMARA, KRISTIAN NYMOEN,
OLIVIER LARTILLOT, & ANNE DANIELSEN
University of Oslo, Oslo, Norway

THIS STUDY REPORTS ON AN EXPERIMENT THAT tested whether drummers systematically manipulated not only onset but also duration and/or intensity of strokes in order to achieve different timing styles. Twenty-two professional drummers performed two patterns (a simple "back-beat" and a complex variation) on a drum kit (hi-hat, snare, kick) in three different timing styles (laid-back, pushed, on-beat), in tandem with two timing references (metronome and instrumental backing track). As expected, onset location corresponded to the instructed timing styles for all instruments. The instrumental reference led to more pronounced timing profiles than the metronome (pushed strokes earlier, laid-back strokes later). Also, overall the metronome reference led to earlier mean onsets than the instrumental reference, possibly related to the "negative mean asynchrony" phenomenon. Regarding sound, results revealed systematic differences across participants in the duration (snare) and intensity (snare and hi-hat) of strokes played using the different timing styles. Pattern also had an impact: drummers generally played the rhythmically more complex pattern 2 louder than the simpler pattern 1 (snare and kick). Overall, our results lend further evidence to the hypothesis that both temporal and sound-related features contribute to the indication of the timing of a rhythmic event in groove-based performance.

Received: January 17, 2020, accepted June 26, 2020.

Key words: drumming performance, microtiming, intensity, duration, groove-based music

IN GROOVE-BASED MUSIC, IT IS ASSUMED THAT musicians can achieve different timing "feels" or "styles" in performance by subtly altering the temporal location of the onset of events at the "micro-rhythmic" metrical level (on the order of about 10–40 milliseconds) by either playing slightly early ("pushed") or late ("laid-back") in relation to other players' rhythm, a metronomic beat reference, or simply their own internal pulse (Butterfield, 2006, 2011; Câmara, 2016; Câmara & Danielsen, 2018; Danielsen, 2006, 2010, 2018; Iyer, 2002; Keil 1987, 1995; Kilchenmann & Senn, 2011, 2015; Senn, Bullerjahn, Kilchenmann, & von Georgi, 2017). Recently, the role of other sound parameters, such as loudness, duration, or timbre, and their interaction with timing have been found to be equally fundamental to microrhythmic expressivity in groove-based music (Câmara, Nymoen, Lartillot, & Danielsen, 2020; Danielsen, Waadeland, Sundt, & Witek, 2015). The present study intends to further test the hypothesis that musicians systematically manipulate acoustic sound features other than onset in order to produce different timing feels (laid-back, on-beat, and pushed) in a musical context.

In musicology and ethnomusicology, scholars have typically used the term "groove" to denote either the individual patterns that comprise a given style (Kernfield, 2003), the overall "rhythm matrix" comprised by all the instruments within a performance ("the groove" in a tune) (Iyer, 2002; Monson, 1996), or an aesthetic quality or "feel" stemming from the various rhythmic relations between or within the instruments of an ensemble in performance, either as a result of microtiming expression (Keil, 1987) or the interaction between microtiming and macrostructural features (Butterfield, 2006, 2011; Câmara, 2016; Danielsen, 2006, 2010). Related, the adjective "groove-based" has tended to denote music from genres derived from African American performance traditions that share a range of common rhythmic features (Câmara & Danielsen, 2018; Pressing, 2002). Most recently, however, "groove" has been operationalized by music psychologists as the aspect of any music, regardless of cultural origin, that elicits various experiential phenomena such as the "urge to move" (Madison, 2006) and "pleasure" (Janata, Tomic, & Haberman, 2012) in particular. In this article, however, we subscribe to the traditional musicological meanings of groove as broadly denotative of the rhythmic structural aspects of musics historically designated as "groove-based" and thus make

Music Perception, VOLUME 38, ISSUE 1, PP. 1–26, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312. © 2020 BY THE REGENTS OF THE UNIVERSITY OF CALIFORNIA ALL RIGHTS RESERVED. PLEASE DIRECT ALL REQUESTS FOR PERMISSION TO PHOTOCOPY OR REPRODUCE ARTICLE CONTENT THROUGH THE UNIVERSITY OF CALIFORNIA PRESS'S REPRINTS AND PERMISSIONS WEB PAGE, HTTPS://WWW.UCPRESS.EDU/JOURNALS/REPRINTS-PERMISSIONS. DOI: https://doi.org/10.1525/MP.2020.38.1.1

no assumption regarding the degree to which onset asynchronies elicit higher or lower ratings of any operationalized groove feature.[1]

Regarding timing in performance, researchers generally assume that expert musicians are able to control the onset of events to a precise degree at the microrhythmic level. One commonly investigated instance of microtiming expression that is thought to contribute to the qualitative feel of grooves is "swing," or the use of asymmetric long-short duration patterns in consecutive pairs of on- and off-beat notes, usually at the eighth-note or sixteenth-note metrical subdivision level. Different degrees of duration ratios in these long-short patterns (i.e., "swing ratios") have been theorized to convey various qualities of "motional energy," or "the force of momentum with which some musical events are directed toward others" (Butterfield, 2011, p. 4), which, in combination with manipulations of intensity and articulation can range from "relaxed" and "continuative" to "forward driving" and "choppy" (see also Butterfield, 2006). Investigations of commercial and field recordings have found that rhythm sections and solo instrumentalists apply varying degrees of swing ratios to either eighth notes in jazz (Benadon, 2006; Butterfield, 2011; Friberg & Sundstrom, 2002; Rose, 1989) or sixteenth notes in jazz-funk, funk, hip-hop (Butterfield, 2006; Câmara, 2016; Danielsen, 2006; Frane, 2017), samba (Gerischer, 2006; Haugen & Godøy, 2014), and djembe music (Polak, 2010). Performance experiments have also shown that drummers and percussionists alter the amount of swing depending on genre, tempo, and/or individual player preference (Haas, 2007; Honing & Haas, 2008; Haugen & Danielsen, 2020).

Another form of microtiming expression thought to contribute to overall timing feel in groove-based music involves the rhythmic interaction between and within the various instruments in an ensemble. It is considered a hallmark of technical proficiency in African American–derived groove performance practice to be able to play flexibly around a timing reference (colloquially referred to as the "beat") in a controlled fashion while maintaining a steady tempo—either behind the beat ("laid-back"), on the beat, or ahead of the beat ("pushed") (Berliner, 1994; Câmara et al., 2020; Danielsen et al., 2015; Keil 1987, 1995; Kilchenmann & Senn 2011, 2015; Monson, 1996). Such timing strategies may be referred to as groove-timing "feels" or "styles." In drumming practice, in particular, while a certain timing feel may be more commonly associated with a given rhythmic pattern or genre—such as the presence of a slight delay (laid-back timing) in the snare strokes of "back-beat" patterns (see Frane, 2017; Iyer, 2002)—particular drummers also tend to develop highly individualized strategies (Dahl, 2011; Waadeland, 2006), and any given pattern may be played with different timing feels depending on personal preference or aesthetic contextual considerations (Butterfield, 2006).

Only a handful of studies have directly investigated the relation between purported groove feels in performance and microtiming profiles via controlled laboratory experiments. Kilchenmann and Senn (2011) instructed two drummers to play the same jazz-rock rhythmic pattern with a laid-back, on-beat, and pushed timing feel along with a metronome and found that both drummers displayed distinctive onset microtiming patterns for each feel. Individual characteristics were also found—one drummer anticipated all the timing conditions in relation to the metronome (described in performance parlance as an "ahead" or "pushy" player), while the other drummer played the pushed and on-the-beat feels ahead of but the laid-back feel behind the metronome (conversely, then, a more "laid-back" player). In a performance study with ten drummers of a standard "back-beat" rock pattern at three different tempi (64, 96, and 148 beats per minute), Danielsen and colleagues (2015) also investigated the degree to which drummers systematically manipulated onset microtiming profiles of the snare drum in order to distinguish different timing feels. They further hypothesized that drummers would systematically manipulate additional sound features, such as intensity and timbre (spectral centroid), in order to produce these feels. In terms of onset timing, they found that regardless of whether drummers displayed either "pushy" or "laid-back" tendencies in relation to the metronome, all were able to produce onsets in the laid-back and pushed conditions significantly behind and ahead of their own average on-beat timing. In other words, all were able to clearly distinguish between the different timing feels in terms of microtiming onset profiles.

Regarding the hypothesized relationship between the sound characteristics of the drum strokes and their onset timing, Danielsen and colleagues (2015) found that at the medium tempo (96 bpm), drummers showed

---

[1] While early studies found either negative or negligible effects of onset asynchronies on subjective ratings of various assumed "groove" features (Davies, Madison, Silva, & Gouyon, 2013; Frühauf, Kopiez, & Platz, 2013; Madison, Gouyon, Ullén, & Hörnström, 2011; Madison & Sioros, 2014), recent studies have found that stimuli with microtiming profiles derived directly from, or resembling those of, original performed music do not necessarily obtain lower ratings than stimuli with artificially reduced onset asynchronies (Kilchenmann & Senn, 2015; Senn, Kilchenmann, von Georgi, & Bullerjahn, 2016; Senn, 2017; Skansaar, Laeng, & Danielsen, 2019).

a tendency to play their strokes with greater intensity (louder) in the laid-back condition relative to the on-beat condition. At the individual participant level, a majority of the drummers (seven out of ten) was additionally found to play these behind-the-beat strokes with a lower spectral centroid (darker). Regarding the effect of tempo, differences between timing style conditions were greatly diminished at the faster tempo, likely due to motoric limitations. However, the drummers played louder across all timing styles in the fast and medium tempo categories relative to the slow tempo category. In a similar instructed-timing style experiment with electric guitarists and bassists, Câmara and colleagues (2020) found that, in addition to manipulating onset location, the guitarists lengthened the durations of their strokes and used a slightly darker timbre when playing in a laid-back fashion, and the bassists played strokes with greater intensity when playing in a pushed fashion. Overall, findings from the latter two studies show that, in order to produce sound events as microrhythmically early or late, musicians systematically manipulated not only the onset timing of strokes but also other parameters of sound as well.

Findings from psychoacoustic studies with pure tones or clicks have suggested that thresholds for asynchrony detection (whether two sounds are heard as synchronous) can be as low as 2 ms (Hirsh, 1959; Zera & Green, 1993), and that the threshold for temporal order (correctly identifying which tone comes first/second) is generally higher, at around 20 ms (Hirsh, 1959). However, in real music, instrumental sounds tend to be complex and have overlapping spectra and/or unequal duration or loudness, all of which can lead to masking effects of various degrees that likely increase these thresholds. Butterfield (2010) tested the extent to which listeners could correctly identify the temporal order of bass and drum sounds in swing jazz excerpts with asynchrony manipulations of up to 30 ms, and found that most were not able to do so above a chance level of 50 percent. The tasks did not, however, assess whether participants were able to detect simply the presence of asynchrony between instruments. Goebl and Parncutt (2002), on the other hand, investigated the effect of intensity on onset asynchrony detection in piano tone dyads of different pitches (high/low) with equal offset locations and found that, regardless of pitch, when the tones were presented with 27 ms onset asynchrony, if the earlier tones were presented as louder ("early + loud" condition or, in classical piano parlance, "melody lead"), only 20 to 30 percent of participants detected the asynchrony, whereas if the earlier tones were presented as softer ("late + loud"), 90 to 100 percent of participants heard the tones as asynchronous. In other words, at 27 ms asynchrony, it was more difficult to detect the asynchrony when the earlier tone in the dyad was louder. This difficulty was attributed to a forward-masking effect or decreased sensitivity to synchrony due to familiarity with early and loud tone combinations. When the magnitude of asynchrony was increased to 54 ms, however, the chances of detecting asynchrony in the early + loud combinations increased to 40 to 50 percent, and in the late + loud combinations it increased to 100 percent, indicating that larger onset asynchronies generally enhanced detectability in both earlier and later directions.

Since the magnitudes of onset asynchronies in performance can be rather subtle, often verging on thresholds of perceptual discriminability, the intended timing feel of a given performance may be further augmented by musicians when they concomitantly manipulate sound features of rhythmic events that have been found to interact with timing at a perceptual level. Studies of a sound's P-center, or perceptual "moment of occurrence" (Morton, Marcus, & Frankish, 1976) or "perceptual attack time" (Gordon, 1987), for example, have shown that, at least in isochronous contexts, the perceived temporal location of a sound event is a complex matter contingent upon more than just onset timing. Practically all studies on the P-centers of *musical* sounds have found that the faster a sound's attack (onset to maximum amplitude peak) and total (onset to offset) duration, the earlier its average P-center relative to its onset, and, conversely, the longer the duration (both slow attack and long total duration), the later the P-center (Bechtold & Senn, 2018; Danielsen et al., 2019; Gordon, 1987; London et al., 2019; Scott, 1998; Seton, 1989; Villing, 2010; Vos, Mates, & Kruysbergen, 1995; Vos & Rasch, 1981; Wright, 2008). A recent study by Danielsen and colleagues (2019) also found that positively correlated combinations of attack and total duration (fast attack/short duration, slow attack/long duration) caused a "redundancy gain" whereby both factors shifted P-centers either earlier or later in time, respectively. Conversely, negatively related combinations (short attack/long duration, long attack/short duration) caused a "redundancy loss" whereby, because one factor tends to shift the P-center earlier while the other one shifts it later, the cumulative shifting effect was either attenuated or canceled out.

Loudness/intensity has also been shown to affect P-center location, though this has been underinvestigated and findings have been less conclusive. While Seton (1989) found no clear effect of intensity on a sawtooth tone, Gordon (1987) found that a saxophone sound played with greater intensity led to an earlier P-center than one played with lesser intensity. Bechtold and Senn (2018) confirmed this result and further found that, for saxophone sounds, both greater articulation (a stronger "tongue attack") and higher dynamics (a louder sound) led to an earlier P-center on average in comparison to sounds with lesser articulation and lower dynamics. Altogether, findings of the effects of duration and intensity on P-center suggest that timing and sound interact at a perceptual level. Therefore, in the production of rhythms, potential redundancy gains across dimensions of timing and sound may help to further convey an intentional early, late, or on-beat timing feel. That is, the laid-back character of a late stroke may be enhanced if it is also played with a longer duration, just as the pushed-ness of an early stroke would be enhanced if it were played with a shorter duration.

Other perception studies have also revealed interactions among perceived timing, duration, and intensity of events in performance. Various experiments with classical pianists have shown that, when they are instructed to emphasize a melody tone in a polyphonic piano performance, they tend to play it both louder and earlier than the other voices (Goebl, 2001; Palmer, 1996; Repp, 1996). As mentioned, Goebl and Parncutt (2002) found that the relative perceptual salience of two tones in a piano dyad depended on both their relative intensity and the asynchrony between their onset timing locations. Yet other studies have found a systematic relationship between intensity and duration, whereby beats accented with greater intensity also tend to be lengthened in performance (Clarke, 1989; Dahl, 2004; Drake & Palmer, 1993; Gabrielsson, 1999; Waadeland, 2006). Regarding perceptual interactions between timing and intensity, Tekman (2002) found that it was easier to correctly identify a tone as being late in relation to another tone when it was both positioned late in terms of onset timing and played with greater intensity, as opposed to a tone that was simply positioned late but played at an equal intensity to another tone. This result was theorized as underlying a higher-order, semantic-level interaction (as opposed to a lower-order, sensory-level interaction) between perceptual dimensions of timing and intensity.

To sum up, several experiments with both music performance and music/audio perception point to the intimate relationship between the temporal and auditory aspects of microrhythm, as well as the integration of timing and various perceptual dimensions of sound.

The present study intends to further test the hypothesis that musicians systematically manipulate acoustic sound features other than onset in order to produce different timing feels in a musical context. Specifically, we aim to replicate some of the findings of Danielsen and colleagues (2015) regarding stroke onset and intensity production, but extend the scope of our investigation to an entire drum kit (snare drum, kick drum, and hi-hat cymbals), increase the number of participants, measure a new sound feature (duration), and explore the effects of two novel factors that may further influence the production of timing in performance: reference stimuli and rhythmic pattern. The rationale for presenting a timing reference to drummers with different sound stimuli is based on findings from sensorimotor tapping studies that suggest that musicians tend to display less negative mean asynchrony ("NMA")—that is, synchronize more accurately to a reference—when presented with musical stimuli (see Repp, 2005; Repp & Su, 2013). In addition, while expert drummers are generally accustomed to playing along to metronomes comprised of click-like percussive stimuli, especially in a studio recording session, an instrumental track—a timing reference that more ecologically simulates a real musical production context (a trio ensemble)—will further test the effect of sound stimuli presentation for synchronization in drumming production, or, more specifically, whether the NMA disappears in a more ecological experimental context. We also chose to provide drummers with two different groove-based rhythms of differing rhythmic complexity and event density, in order to investigate contextual effects of stylistic pattern on production of intentional timing feel—that is, to gauge whether onset and sound-feature-manipulation profiles will differ for simpler groove-based patterns with no syncopation/pick-ups in relation to patterns with additional salient off-beat events. Finally, we also measure a new stroke feature, that of duration, since perceptual studies have shown that the attack and total temporal extent of a sound influences its P-center.

Overall, we hypothesize that, in the process of achieving distinctive early, late, or on-beat microrhythmic timing profiles in performance, instrumentalists leave sonic "stamps" on the sounds of their own instruments that are systematically related to these timing styles. We explore this possibility by measuring changes in duration and loudness (sound pressure level [SPL]), in addition to onsets, of sound events after instructing instrumentalists to play under different timing constraints, addressing the following question: "To what extent are there systematic differences in the acoustic signal between drum strokes played with different (a) intended timing styles along to (b) two different timing references in (c) two different groove patterns across subjects?" We hypothesize that NMA will be lower when

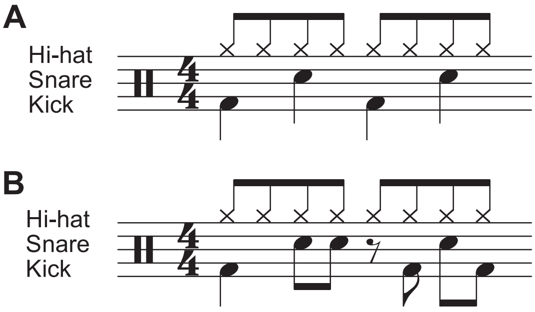

FIGURE 1. A) Pattern 1; B) Pattern 2.

the drummers are playing to the more ecological instrumental reference, and we also expect that the musical context will yield an effect, but we have no specific hypothesis regarding what this might be.

Method

PARTICIPANTS

Twenty-two male drummers, 22–64 years of age ($M = 36$, $SD = 11$) participated in the study. All were active part-time or full-time musicians recruited from either local universities/conservatories or commercial performance scenes, and they had between 4 and 40 years of professional performance experience ($M = 16$, $SD = 11$). All were familiar with either jazz, funk/soul, or rock. All participants were paid an honorarium to take part in the experiment.

TASKS

The drummers performed two rhythmic patterns at a 96 bpm medium tempo deemed to be comfortable based upon previous experiments (Câmara et al., 2020; Danielsen et al., 2015). Pattern 1 (Figure 1A) was a simplified version of the so-called "back-beat" pattern that is ubiquitous in popular groove-based music. Pattern 2 (Figure 1B) was a slightly more complex variation of the back-beat groove that included additional off-beat events: an extra eighth-note stroke on the "two-and" metrical position for the snare and a syncopated eighth-note on the "three-and," as well as a pick-up on the "four-and," for the kick drum.

Participants were presented with two categories of performance conditions:

A. Timing reference conditions
   Play each pattern (1 and 2) along to:
   a) a *metronome*, comprised of woodblock sounds (condition: Metronome)
   b) an *instrumental* backing track, comprised of guitar and bass sounds (condition: Instrumental)

FIGURE 2. Drum-kit setup. Drum-kit covered in black textile in order to avoid problems with reflections in the motion-capture recordings.

B. Timing style conditions
   In each of the timing reference contexts listed above, play each pattern (1 and 2):
   0) in as *natural* a manner as possible (condition: Natural)
   1) in a *laid-back* manner, or behind the beat relative to the timing reference (condition: Laid-back)
   2) in a *pushed* manner, or ahead of the beat relative to the timing reference (condition: Pushed)
   3) in an *on-beat* manner, or synchronized to the timing reference (condition: On)

APPARATUS AND MATERIALS

We did the recordings at the Motion Capture Laboratory, Department of Musicology, University of Oslo, Oslo, Norway, in the spring of 2018. For our drum instrumental setup, we used the following equipment: an acoustic metal snare drum, 7 inches deep and 14 inches wide (Gretsch, USA), with an Emperor X drumhead (Remo, USA) and a thin plastic muffle ring; a 21-inch kick drum (Gretsch), with a FA batter drumhead (Remo); and a hi-hat stand (Pearl, USA) with 14-inch cymbals (Yamaha, Japan) (see Figure 2). Pilot tests of the sound recordings revealed that close-microphone techniques with regular microphones led to too much leakage/bleed between hi-hat and snare signals, so these instruments were recorded with C411 contact microphones instead (AKG, Austria). For the kick drum, we used a Beta 52 (Shure, USA) microphone.

The microphone signals were sent into a BabyFace Pro sound card (RME, Germany) and recorded in the
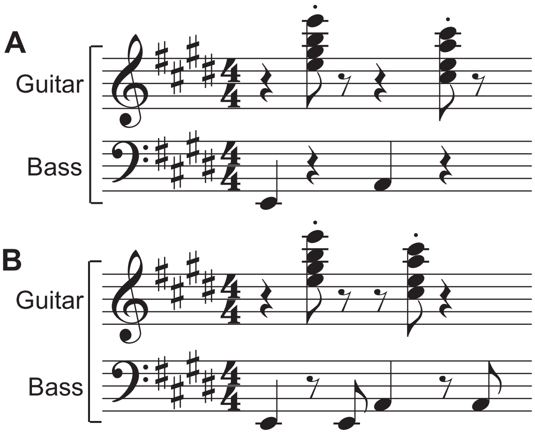
FIGURE 3. Instrumental backing track presented to participants for A) Pattern 1, and B) Pattern 2.

audio software Reaper 5.77 (Cockos Inc., NY) at a sampling frequency of 44.1 kHz and 24-bit resolution. Playback of the two timing-reference tracks was routed to a MG10XU analog mixer (Yamaha) and then fed into 250-Ohm-resistance DT 990 Pro headphones (Beyer Dynamic, Germany), which were then given to the participants for monitoring.

For the metronome beat reference track, we used two woodblock sounds (Cubase 8 Halion LE library, Steinberg/Yamaha, Germany), one pitched higher and located on the first quarter-note beat locations (spectral centroid [SC] = ca. 2370 Hz) and the other pitched lower and located on the remaining second, third, and fourth beats locations (SC = ca. 1565 Hz). Both sounds had very short, impulsive attack transients (attack duration from signal onset to max. amplitude peak ≈ 2 ms), with a gradual decay toward the signal offset (total duration ≈ 40 - 50 ms). The metronome stimuli were aligned to the software grid based on mean P-center results from our previous instructed timing study on guitarists and bassists (see Câmara et al., 2020).

The instrumental backing-track reference was comprised of electric guitar and electric bass sounds originally recorded in our audio analysis study of sound–microtiming interaction in guitar and bass (for details, see Câmara et al., 2020). For pattern 1, this was comprised of alternating quarter-note tones on bass (fundamental frequencies = E and A, total duration ≈ 610 ms) located on beats 1 and 3, and a two-chord guitar back-beat pattern (E and A, total duration ≈ 250 ms) located on beats 2 and 4 (Figure 3A). For pattern 2, an analogously complex and aesthetically appropriate version of the backing track was given, where the guitar pattern was syncopated on the second chord on the “4-and” beat, and the alternating quarter-note bass tones were always preceded by eighth-note pick-ups (total duration ≈ 300 ms) (Figure 3B). While we had no P-center data on these guitar and bass sounds, since they all had fast attacks (≈ 15–25 ms) and previous experiments have shown that P-center of plucked string instruments tend to be close to stimuli onset (Danielsen et al., 2019), all the instrumental reference stimuli were simply aligned to the software grid based on their calculated onset values (see “Audio Analysis” sub-section below for onset calculation details).

PROCEDURE

Before the experiment began, we encouraged all drummers to acquaint themselves with the drum-kit set up and warm up by playing freely for at least 10 minutes or so. Once the experiment began, in the beginning of each pattern and timing reference block, the “natural” timing condition was always given first, as a further warm up to allow participants to accustom themselves to the pattern and timing reference combination, as well as the tempo. Then, the remaining timing style conditions (laid-back, on, and pushed) were given, in randomized order. A rest period was allowed between each condition for as long as participants deemed necessary. Once started, each condition lasted for approximately 67.5 s (27 measures), and participants began to play as soon as they had entrained with the timing reference track. If a participant was dissatisfied with their performance during or after a condition trial, they were invited to repeat it as many times as they wanted. This resulted in 224 possible hi-hat strokes for both patterns, 54 snare and kick drum strokes for Pattern 1, and 81 for Pattern 2, played per condition per participant. After the performances, a short performer interview provided feedback related to the experimental setup and insight into the various performance strategies applied to achieve the different timing condition tasks. In total, the performance and interview session lasted between 45 to 60 min.

AUDIO ANALYSIS

Determination of drum stroke segments. First, we estimated the location of the onset and offset of each drum stroke directly from the audio waveform using an algorithm developed by Lartillot and colleagues (2020). The onset is the starting point of the attack phase, where the attack phase is a line fitted on the audio waveform through an optimization of both the slope and the maximum amplitude (*mirevents* with “Attacks” option set to “Waveform” and parameter settings: WaveformThreshold = 3 percent [snare/kick] and 10 percent [hi-hat]). Similarly, the offset is the ending point of the decay

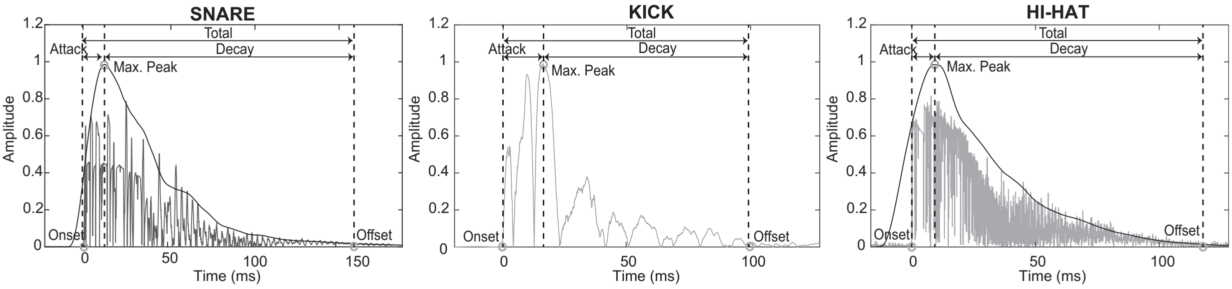

FIGURE 4. Schematic illustration of a typical rectified snare, and hi-hat stroke signal recorded with a contact microphone; and kick stroke signal recorded with a dynamic microphone. Onset/offset locations are marked with circles (calculated directly from waveforms), and attack/decay/total segments are delineated by dashed lines (calculated from envelopes [solid line] for snare and hi-hat, and directly from waveforms for kick).

phase. Then, we separated each stroke signal into attack and decay segments based on the time point of the maximum global amplitude ("peak"). Thus, for the analysis (described below), each stroke was partitioned into three segments: 1) attack (interval between onset and maximum peak); 2) decay (interval between maximum peak and offset); 3) total (interval between onset and offset).

The dynamic microphone recordings of the kick drum contain singular short and transient peaks that are easily extracted from the waveform itself. The snare and hi-hat contact microphone signals, however, display atypical, irregular transient peak patterns during the attack phase (see Figure 4). Therefore, we decided to extract their peak location from a smoothed signal envelope, calculated as the sum of bin amplitudes across the columns of a spectrogram using the *mirevents* algorithm (loc. cit) with "Attacks" option set to "Slope" (parameter settings: frame length = 30 ms [snare/kick] and 100 ms [hi-hat]; frame hop factor = 2 percent).[2]

Selection of sound descriptors. We selected three sound descriptors for the main analyses, defined as follows: 1) onset (the starting point of the stroke attack segment (see above), measured in milliseconds); 2) duration (the elapsed time interval of the stroke segment, measured in milliseconds); 3) sound pressure level (SPL; the unweighted root-mean-square (rms) amplitude of the signal, measured in dB, with a 0 dB reference given as the average rms amplitude of all strokes in all timing conditions).

Aside from the onset descriptor, we also calculated duration and SPL for each stroke segment (attack/decay/total—described above). In all, then, we analyzed seven descriptors for each drum instrument: onset, attack duration, decay duration, total duration, attack SPL, decay SPL, and total SPL.

Following Danielsen and colleagues (2015)[3] and Câmara and colleagues (2020), we used sound pressure level (SPL) as the audio descriptor to describe stroke intensity. SPL is highly correlated to perceived loudness and remains a widely used metric in perceptual studies (Rossing, Moore, & Wheeler, 2002). Duration was also chosen as various aforementioned studies report that both the attack and the total elapsed time of a sound in particular have been shown to have a significant effect on its P-center, whereby longer durations tend to shift the experienced P-center later in time.

DATA PROCESSING AND STATISTICAL ANALYSIS

We excluded the data from two participants from the analysis—audio signals from one participant suffered from distortion due to technical issues during recording, and we deemed another participant unable to complete the instructed tasks based on reports from the post-experiment interview. Because the natural timing condition was used mainly as a practice/adjustment task, we also omitted all data for these series. Subsequently, we gathered data from 5,040 recorded series of drum strokes (drum instruments [3] × timing style [3] × reference [2] × pattern [2] × participant [20] × audio descriptor [7]). We cropped all of the strokes into individual segments via custom scripts in Matlab

---

[2] The *mirevents* algorithm with both the "Waveform" and "Slope" options are available in the MIR Toolbox audio analysis package [ver. 1.8] (Lartillot, Toiviainen, & Eerola, 2008).

[3] In contrast to Danielsen and colleagues (2015), we chose to not investigate the brightness of drum strokes via spectral centroid (SC) since we opted to use contact microphones to capture the snare and hi-hat signals. While contact microphones provide recordings from which superior stroke onset and SPL information can be extracted due to better signal source isolation, they do not reproduce the *timbre* of drum sounds as faithfully as those of dynamic microphones. As such, any comparisons of their perceptual brightness via SC would not be entirely ecological without further investigation.

version R2018a (Mathworks, USA). First, we cropped audio recordings into individual segments according to the grid points corresponding to the patterns from 400 ms before to 900 ms after the grid point. Then, we silenced the parts of the audio signal that belonged to the previous and following strokes. Segments that contained no signal around the grid point were marked as missing strokes and removed. Typically, participants would wait one or two measures before beginning to play in order to entrain to the metronome, resulting in 2 to 4 missing strokes per recording. Timing mistakes, defined as strokes with onsets of more than a sixteenth note (156 ms) early or late in relation to the metric grid, were also marked and removed automatically. We considered these events not as instances of early/late microtiming relative to the beat but rather as qualitatively different beat-level syncopations. We validated this automated process by plotting the waveform for each segment and inspecting the silenced audio parts and the marked missing strokes. A few missing strokes and timing mistakes were not detected automatically and were removed manually. Out of all total possible strokes captured per instrument, the total amount of manual and automatically marked invalid strokes accounted for about 9.4 percent of the snare, 8.7 percent of the kick, and 12.8 percent of the hi-hat data.

Before our statistical analysis, we defined extreme outliers as values that were more than three times the interquartile range away from the median of each instrument, timing condition, participant, and audio descriptor separately, and excluded them. Out of all valid strokes captured in total (i.e., excluding invalid missing strokes and timing mistakes), extreme outliers accounted for about 1.0 percent of the snare, 2.0 percent of the kick, and 0.8 percent of the hi-hat drum data. To check for normality, we within-subjects standardized the data (subtracted the average of all strokes across timing style conditions for a given participant from each individual timing style condition value for the same participant; see Fischer & Milfont, 2010) and manually screened the residuals via histograms and Q-Q plots for each dependent variable and instrument separately. None showed departure from normality.

Even though all participants reported familiarity with playing along to both a metronome and an instrumental backing track, we needed to confirm whether participants were in fact able to physically place their strokes either early, ahead, or synchronous with the reference track when instructed to do so—that is, to accomplish the instructed timing style tasks in terms of simple onset timing location for at least one instrument at a time. To do so, we first compared the average (arithmetic mean) profile of onset between the laid-back and pushed series with the on-beat series for each drum and individual. Then, in order to gauge the overall trends of all drummers' timing-sound manipulation in the production of timing feel, we ran three-way RMANOVAs for each instrument and sound descriptor (onset, duration, and SPL) separately across all participants ($N = 20$), with timing style, reference, and pattern as the independent variables. Violations of sphericity (Mauchly's test) were corrected using Greenhouse-Geisser. Post hoc paired-samples t-tests were performed where significant main effects or interaction were found and were Bonferroni corrected for multiple comparisons.

We also ran supplementary paired samples t-tests to investigate potential effects of individual notes between patterns where deemed appropriate (Bonferroni corrected for multiple comparisons). All statistical analyses were performed using SPSS ver. 25 (IBM, Inc., New York).

Results

TASK SUCCESS

Our examination of the 240 onset data series (participant [20] × timing style [3] × reference [2] × pattern [2]) for each percussion instrument revealed that, for both timing references and patterns, the mean onset location of strokes corresponded to the given timing-style instructions in at least one out of three instruments (hi-hat, snare, kick). That is to say, when asked to play in a laid-back or pushed manner, the drummers successfully produced, on average, onsets earlier and later in time, respectively, in relation to the corresponding on-beat series in either the snare, kick, or hi-hat. Descriptive statistics of the microtiming onset profiles of all series by all drummers can be found in the Appendix.

EFFECTS AND INTERACTIONS OF TIMING STYLE, REFERENCE, AND PATTERN

We conducted three-way RMANOVAs for each stroke segment (attack, decay, and total) and each drum instrument separately with timing Style, Reference, and Pattern (independent variables) and onset, duration, SC, and SPL (dependent variables). An overview of these RMANOVA results across all participants ($N = 20$) for all sound descriptors can be found in Table 1, and descriptive statistics for mean and standard deviations across participants for all timing-style, reference, and pattern conditions can be found in Table 2. In the following section, we elaborate upon significant results

TABLE 1. *Results of Three-way RMANOVAs For All Audio Descriptors ($N = 20$)*

		Snare					Kick					Hi-hat				
DV	IV	*F*	*df*	*df*(error)	*p*	η_p^2	*F*	*df*	*df*(error)	*p*	η_p^2	*F*	*df*	*df*(error)	*p*	η_p^2
Onset	S	91.347	1.157	21.992	< **.001**	0.828	55.879	1.114	21.164	< **.001**	0.746	55.198	1.200	22.798	< **.001**	0.744
	R	3.368	1.000	19.000	.082	0.151	2.868	1.000	19.000	.107	0.131	3.729	1.000	19.000	.069	0.164
	P	3.350	1.000	19.000	.083	0.150	1.360	1.000	19.000	.258	0.067	9.978	1.000	19.000	**.005**	0.344
	S×R	13.129	1.838	34.922	< **.001**	0.409	19.763	1.727	32.809	< **.001**	0.510	15.121	1.816	34.500	< **.001**	0.443
	S×P	3.842	1.179	22.407	.057	0.168	2.786	1.408	26.746	.095	0.128	2.061	1.253	23.798	.162	0.098
	R×P	1.646	1.000	19.000	.215	0.080	0.008	1.000	19.000	.931	0.000	0.631	1.000	19.000	.437	0.032
	S×R×P	3.112	1.640	31.168	.068	0.141	4.281	1.459	27.728	.034	0.184	5.370	1.796	34.133	**.011**	0.220
Attack Duration	S	1.058	2.000	38.000	.357	0.053	1.106	2.000	38.000	.341	0.055	2.239	2.000	38.000	.120	0.105
	R	0.286	1.000	19.000	.599	0.015	0.426	1.000	19.000	.522	0.022	2.560	1.000	19.000	.126	0.119
	P	0.552	1.000	19.000	.467	0.028	3.305	1.000	19.000	.085	0.148	0.213	1.000	19.000	.649	0.011
	S×R	0.035	2.000	38.000	.966	0.002	0.097	2.000	38.000	.908	0.005	0.073	2.000	38.000	.930	0.004
	S×P	0.599	2.000	38.000	.555	0.031	0.615	2.000	38.000	.546	0.031	1.177	2.000	38.000	.319	0.058
	R×P	1.479	1.000	19.000	.286	0.072	0.558	1.000	19.000	.247	0.029	0.922	1.000	19.000	.874	0.046
	S×R×P	1.004	2.000	38.000	.376	0.050	0.984	2.000	38.000	.383	0.049	0.966	2.000	38.000	.390	0.048
Decay Duration	S	5.369	1.550	29.457	**.016**	0.220	5.158	2.000	38.000	.010	0.214	0.911	2.000	38.000	.411	0.046
	R	3.031	1.000	19.000	.098	0.138	0.061	1.000	19.000	.808	0.003	0.017	1.000	19.000	.898	0.001
	P	19.972	1.000	19.000	< **.001**	0.512	1.042	1.000	19.000	.320	0.052	1.910	1.000	19.000	.183	0.091
	S×R	0.359	1.627	30.921	.657	0.019	1.047	2.000	38.000	.361	0.052	0.380	2.000	38.000	.687	0.020
	S×P	1.211	1.700	32.299	.305	0.060	2.368	2.000	38.000	.107	0.111	0.819	2.000	38.000	.448	0.041
	R×P	0.158	1.000	19.000	.874	0.008	2.863	1.000	19.000	.928	0.131	1.704	1.000	19.000	.135	0.082
	S×R×P	0.525	1.450	27.548	.540	0.027	0.641	2.000	38.000	.532	0.033	0.713	2.000	38.000	.497	0.036
Total Duration	S	5.752	2.000	38.000	**.007**	0.232	4.808	1.543	29.323	**.023**	0.202	1.459	2.000	38.000	.245	0.071
	R	3.186	1.000	19.000	.090	0.144	0.064	1.000	19.000	.802	0.003	0.529	1.000	19.000	.476	0.027
	P	18.964	1.000	19.000	< **.001**	0.500	1.471	1.000	19.000	.240	0.072	1.704	1.000	19.000	.207	0.082
	S×R	0.359	2.000	38.000	.701	0.019	0.856	1.926	36.595	.430	0.043	0.238	2.000	38.000	.789	0.012
	S×P	1.086	2.000	38.000	.348	0.054	2.794	1.564	29.719	.089	0.128	0.967	2.000	38.000	.389	0.048
	R×P	0.086	1.000	19.000	.239	0.005	2.628	1.000	19.000	.697	0.122	1.828	1.000	19.000	.349	0.088
	S×R×P	0.635	2.000	38.000	.536	0.032	0.656	1.988	37.776	.524	0.033	1.163	2.000	38.000	.323	0.058

(continued)

TABLE 1. ***(continued)***

		Snare					Kick					Hi-hat				
DV	IV	*F*	*df*	*df*(error)	*p*	η_p^2	*F*	*df*	*df*(error)	*p*	η_p^2	*F*	*df*	*df*(error)	*p*	η_p^2
Attack SPL	S	4.480	2.000	38.000	**.018**	0.191	1.410	2.000	38.000	.257	0.069	5.648	2.000	38.000	**.007**	0.229
	R	0.642	1.000	19.000	.433	0.033	3.438	1.000	19.000	.079	0.153	4.396	1.000	19.000	**.050**	0.188
	P	11.241	1.000	19.000	.003	0.372	132.893	1.000	19.000	< **.001**	0.875	3.559	1.000	19.000	.075	0.158
	S×R	0.382	2.000	38.000	.685	0.020	0.826	2.000	38.000	.445	0.042	1.313	2.000	38.000	.281	0.065
	S×P	5.733	2.000	38.000	.007	0.232	1.172	2.000	38.000	.321	0.058	2.422	2.000	38.000	.102	0.113
	R×P	0.145	1.000	19.000	.221	0.008	4.242	1.000	19.000	.969	0.182	0.002	1.000	19.000	.852	0.000
	S×R×P	1.023	2.000	38.000	.369	0.051	1.402	2.000	38.000	.259	0.069	2.035	2.000	38.000	.145	0.097
Decay SPL	S	3.696	2.000	38.000	.034	0.163	1.220	2.000	38.000	.306	0.060	6.868	1.548	29.410	**.006**	0.266
	R	0.084	1.000	19.000	.775	0.004	1.206	1.000	19.000	.286	0.060	4.703	1.000	19.000	.043	0.198
	P	3.789	1.000	19.000	.067	0.166	42.669	1.000	19.000	< **.001**	0.692	3.210	1.000	19.000	.089	0.145
	S×R	1.531	2.000	38.000	.229	0.075	0.199	2.000	38.000	.820	0.010	1.506	1.912	36.327	.236	0.073
	S×P	5.861	2.000	38.000	**.006**	0.236	1.945	2.000	38.000	.157	0.093	1.904	1.905	36.201	.165	0.091
	R×P	0.620	1.000	19.000	.685	0.032	1.080	1.000	19.000	.594	0.054	0.132	1.000	19.000	.748	0.007
	S×R×P	0.405	2.000	38.000	.670	0.021	0.284	2.000	38.000	.754	0.015	0.925	1.795	34.111	.397	0.046
Total SPL	S	4.000	1.530	29.075	**.039**	0.174	1.028	2.000	38.000	.367	0.051	7.494	1.488	28.274	**.005**	0.283
	R	0.141	1.000	19.000	.712	0.007	2.490	1.000	19.000	.131	0.116	12.023	1.000	19.000	**.003**	0.388
	P	4.679	1.000	19.000	.043	0.198	98.083	1.000	19.000	< **.001**	0.838	4.024	1.000	19.000	.059	0.175
	S×R	1.368	1.101	20.921	.259	0.067	0.462	2.000	38.000	.633	0.024	1.145	1.846	35.075	.326	0.057
	S×P	6.140	1.171	22.243	**.017**	0.244	2.435	2.000	38.000	.101	0.114	2.028	1.772	33.665	.152	0.096
	R×P	0.523	1.000	19.000	.708	0.027	2.562	1.000	19.000	.145	0.119	0.008	1.000	19.000	.967	0.000
	S×R×P	0.485	1.119	21.253	.514	0.025	0.711	2.000	38.000	.498	0.036	1.919	1.685	32.007	.168	0.092

Note: DV = dependent variable, IV = independent variable, S = style, R = reference, P = pattern. Significant results in bold ($p < .05$).

TABLE 2. *Descriptive Statistics Across Timing Style, References, and Patterns For All Audio Descriptors (N = 20)*

		Onset (ms)		Attack Duration (ms)		Decay Duration (ms)		Total Duration (ms)		Attack SPL (dB)		Decay SPL (dB)		Total SPL (dB)	
	IV	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
Snare	L	17	14	7	1	145	17	152	18	-7.86	3.65	-1.12	3.30	-0.25	3.33
	O	-10	9	7	1	139	21	145	21	-7.91	3.60	-1.36	3.14	-0.46	3.18
	P	-52	20	6	1	140	22	146	22	-8.23	3.98	-1.55	3.53	-0.66	3.56
	Mtr.	-16	7	7	1	142	19	149	19	-8.05	3.87	-1.36	3.47	-0.48	3.50
	Ins.	-14	9	7	1	140	20	147	20	-7.95	3.61	-1.32	3.16	-0.43	3.19
	Pat. 1	-13	7	7	1	146	19	152	19	-8.55	3.47	-1.73	2.80	-0.87	2.87
	Pat. 2	-16	8	6	1	137	21	143	21	-7.45	4.11	-0.96	3.94	-0.04	3.94
Kick	L	7	14	13	3	113	20	126	21	-5.69	5.08	-5.12	5.42	-2.21	5.05
	O	-14	10	13	3	117	24	130	25	-5.79	4.87	-5.07	5.26	-2.19	4.88
	P	-53	22	13	3	119	23	132	23	-5.93	4.96	-5.38	5.45	-2.41	5.04
	Mtr.	-21	8	13	3	116	22	129	22	-5.71	5.06	-5.11	5.42	-2.19	5.06
	Ins.	-19	8	13	3	116	23	129	24	-5.89	4.87	-5.27	5.29	-2.34	4.89
	Pat. 1	-19	9	13	3	114	24	127	25	-7.44	4.94	-6.59	5.24	-3.75	4.86
	Pat. 2	-21	8	13	3	118	23	132	24	-4.17	5.06	-3.79	5.62	-0.79	5.17
Hi-hat	L	15	17	25	11	120	32	145	38	-3.62	3.51	-3.50	3.57	-0.29	3.50
	O	-9	10	21	11	118	33	139	40	-4.52	4.08	-4.05	4.32	-1.05	4.19
	P	-46	21	24	13	121	33	145	40	-3.77	3.67	-3.30	3.66	-0.27	3.53
	Mtr.	-14	8	24	11	120	32	144	38	-3.86	3.67	-3.54	3.72	-0.43	3.64
	Ins.	-12	8	23	11	120	33	142	39	-4.09	3.73	-3.69	3.92	-0.65	3.78
	Pat. 1	-11	7	17	13	129	38	123	35	-11.90	6.73	-2.29	4.29	-3.40	3.77
	Pat. 2	-16	8	23	10	117	31	140	37	-4.15	3.76	-3.83	3.95	-0.75	3.84

Note: IV = independent variable, L = Laid-back, O = On, P = Pushed, Mtr. = Metronome, Ins. = Instrumental, Pat. = Pattern.

TABLE 3. *Pairwise Comparisons of Main Effect of Style on Average Onset Location For Each Drum Instrument (N = 20)*

	Onset (ms)											
	Snare				Kick				Hi-hat			
Style	Δ*M*	*SD*	*p*	*d*	Δ*M*)	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*
Laid-back – On	27	13	< .001	2.00	20	13	< .001	1.61	24	16	< .001	1.52
Pushed – On	−42	22	< .001	−1.97	−39	27	< .001	−1.55	−37	24	< .001	−1.52
Laid-back – Pushed	69	31	< .001	2.23	60	36	< .001	1.74	61	34	< .001	−1.76

only. All reported onset and duration values are rounded up to 1 ms, and SPL values to 0.01 dBs.

Onset. Results showed a significant main effect of Style on onset location for hi-hat, snare, and kick drum. Post hoc pairwise comparisons revealed significant differences in mean onset between all three timing-style pairs for all instruments (laid-back vs. on, pushed vs. on, laid-back vs. pushed; see Table 3). These comparisons showed that the mean difference in onset timing between the pushed and on timing-style conditions (deltaPvO) was greater than the mean onset difference between the laid-back strokes and on-beat conditions (deltaLvO). To test whether this difference was significant, follow-up paired-samples *t*-tests were conducted. They confirmed that deltaPvO was indeed significantly larger than deltaLvO for all instruments: snare, $M = +15$ ms, $SD = 18$ ms, $p = .001$, $d = 0.83$; kick, $M = +19$ ms, $SD = 17$ ms, $p < .001$, $d = 0.95$; hi-hat, $M = +19$ ms, $SD = 17$ ms, $p = .017$, $d = 0.62$.

While we found no significant main effect of Reference on onset location for any of the instruments, we found a significant interaction between Style and Reference for all instruments such that timing style had a greater effect in the instrumental than in the metronome condition. Figure 5 illustrates the mean onset timing across participants in all style, reference, and pattern conditions. Paired-samples *t*-tests revealed that, for all

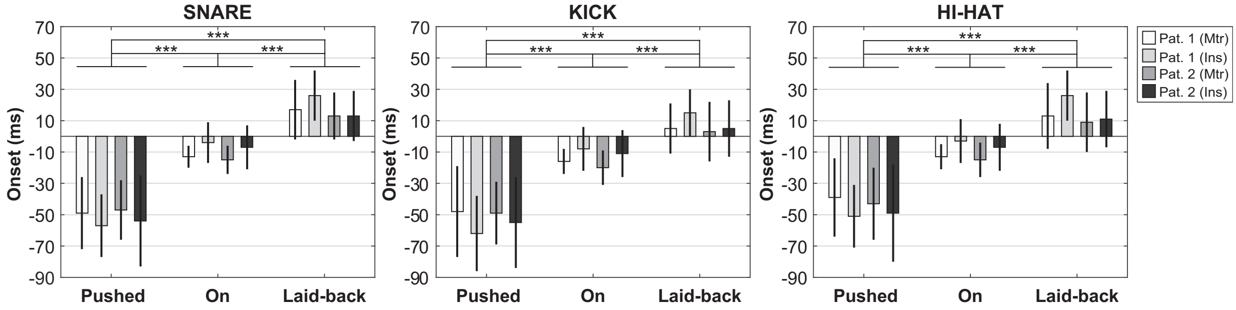

FIGURE 5. Mean stroke onset location (ms) for all drum instruments across participants ($N = 20$) in all style, reference, and pattern conditions. Error bars indicate 1 *SD*. Note: Pat. = pattern, Mtr = metronome reference, Ins = instrumental reference.

TABLE 4. *Paired Samples t-tests of Average Onset Location Between Instrumental and Metronome References In All Timing Style Conditions Across Patterns, For Each Drum (N = 20)*

Onset (ms)

	Reference	Snare				Kick				Hi-hat			
	Instrumental – Metronome	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*
Style	Laid-back	5	10	.046	0.48	6	9	.009	0.65	7	11	.016	0.59
	On	8	10	.002	0.82	8	9	.001	0.91	9	10	.001	0.93
	Pushed	−7	15	.036	−0.51	−10	15	.007	−0.68	−10	16	.014	−0.61

TABLE 5. *Pairwise Comparisons of Main Effect of Style on Snare Duration (N = 20)*

	Snare											
	Attack Duration (ms)				Decay Duration (ms)				Total Duration (ms)			
Style	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*
Laid-back – On	0	1	.970	0.23	7	10	.031	0.64	7	11	.029	0.64
Pushed – On	0	1	1.000	-0.11	1	7	1.000	0.21	1	7	.141	0.47
Laid-back – Pushed	0	1	.613	0.29	5	7	.141	0.47	6	11	.101	0.51

instruments, mean onset timing was significantly later in the instrumental reference compared to metronome across patterns in the laid-back and on conditions. For the pushed conditions, however, we found the opposite result, where mean onset timing was earlier in the instrumental reference compared to the metronome (see Table 4).

A significant main effect of Pattern was found on onset for only the hi-hat, and post hoc pairwise comparisons revealed an earlier mean onset for Pattern 1 compared to Pattern 2 ($M = -5$ ms, SD = 7 ms, $p =$ .005, $d = 0.71$).

A significant interaction on onset was also found between Style × Reference × Pattern for the kick and the hi-hat. Overall, based on examination of the plots, there was a lesser effect of timing style when drummers played the more complex pattern (Pattern 2), and no difference between metronome and instrumental reference track when playing this pattern in a laid-back timing style.

Duration. A significant main effect of Style on duration was found for the snare drum in the decay and total duration segments. Post hoc pairwise comparisons for the snare revealed significantly longer mean decay and total duration for laid-back strokes than for on-beat strokes, but no significant differences in any stroke segment between either pushed and on-beat or laid-back and pushed strokes (see Table 5). Figure 6 further illustrates the results for mean duration for

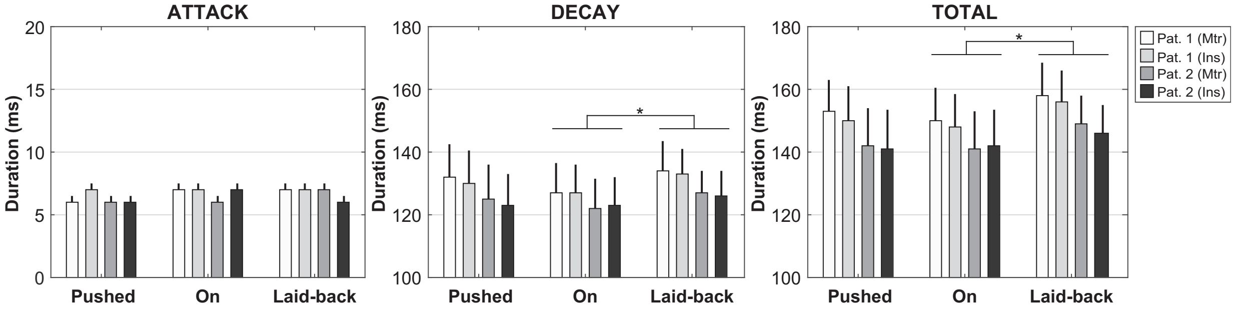

FIGURE 6. Mean snare stroke duration (all segments) across participants ($N = 20$) in all style, reference, and pattern conditions. Error bars indicate 0.5 *SD*. Note: Pat. = pattern, Mtr = metronome reference, Ins = instrumental reference.

TABLE 6. *Paired-samples t-tests Between Snare Average Note Duration in Patterns 1 and 2 (N = 20)*

	Pattern		Decay Duration (ms)				Total Duration (ms)			
	1	2	ΔM	SD	p	d	ΔM	SD	p	d
Note	"two"	"two"	9	10	.004	0.91	9	10	.001	0.92
	"two"	"two-and"	13	9	< .001	1.40	12	9	< .001	1.33
	"two"	"four"	13	10	.207	0.51	5	10	.224	0.50
	"four"	"two"	5	10	.006	0.88	9	10	.005	0.89
	"four"	"two-and"	13	10	< .001	1.32	13	10	< .001	1.26
	"four"	"four"	5	11	.228	0.50	5	11	.244	0.49

TABLE 7. *Pairwise Comparisons of Main Effect of Style on Snare Duration (N = 20)*

	Kick											
	Attack Duration (ms)				Decay Duration (ms)				Total Duration (ms)			
Style	ΔM	SD	p	d	ΔM	SD	p	d	ΔM	SD	p	d
Laid-back – On	0	1	1.000	0.11	0	1	1.000	0.11	-3	8	.333	-0.37
Pushed – On	0	1	1.000	0.35	2	5	.306	0.38	2	5	.211	0.43
Laid-back – Pushed	0	1	1.000	-0.19	5	9	.043	-0.60	-5	9	.053	-0.58

all snare-stroke segments in all style, reference, and pattern conditions.

We also found a main effect of Pattern on duration for the snare. Post hoc pairwise comparisons revealed significantly longer durations in Pattern 1 than in Pattern 2 for both decay ($M = +9$ ms, $SD = 9$ ms, $p < .001$, $d = 1.0$) and total ($M = +9$ ms, $SD = 9$ ms, $p < .001$, $d = 0.97$) segments. Since the main difference between the patterns was structural, we wanted to further investigate whether this effect was driven by the difference in individual note composition between two pattern conditions. We therefore ran a supplementary paired-samples *t*-test between the average duration of the individual notes across participants in each pattern. The test revealed that not only was the decay and total (but not attack) duration of the extra eighth note in the double stroke of Pattern 2 shorter than the single stroke of Pattern 1, but also almost of all the strokes in Pattern 2 were significantly shorter than the corresponding strokes of Pattern 1 (see Table 6).

A significant main effect of Style on duration was found for the kick drum. Post hoc pairwise comparisons revealed significantly shorter mean durations in the laid-back than the pushed strokes for the decay segment, but only a trend toward significance for total duration. (see Table 7). We found no significant differences between laid-back and on-beat strokes or pushed and on-beat strokes for any stroke segment. Figure 7 illustrates the results for mean duration for all kick drum stroke segments in all style, reference, and pattern conditions.

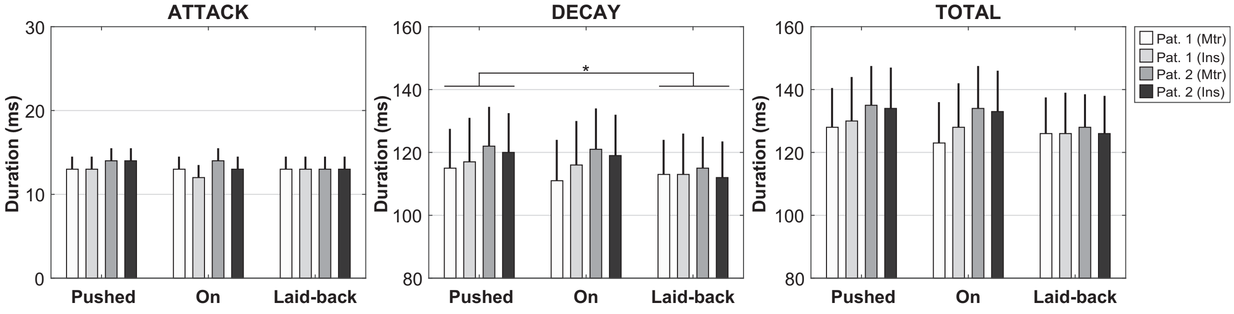

FIGURE 7. Mean kick stroke duration (all segments) across participants ($N = 20$) in all style, reference, and pattern conditions. Error bars indicate 0.5 *SD*. Note: Pat. = pattern, Mtr = metronome reference, Ins = instrumental reference.

TABLE 8. *Pairwise Comparisons of Main Effect of Style on Snare SPL (N = 20)*

	Snare											
	Attack SPL (dB)				Decay SPL (dB)				Total SPL (dB)			
Style	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*
Laid-back – On	0.04	0.46	1.000	0.10	0.25	0.54	.166	0.46	0.21	0.49	.202	0.43
Pushed – On	-0.32	0.73	.185	-0.44	-0.19	0.85	1.000	-0.22	-0.20	0.80	.823	-0.25
Laid-back – Pushed	0.37	0.58	.033	0.63	0.43	0.73	.046	0.60	0.41	0.63	.026	0.65

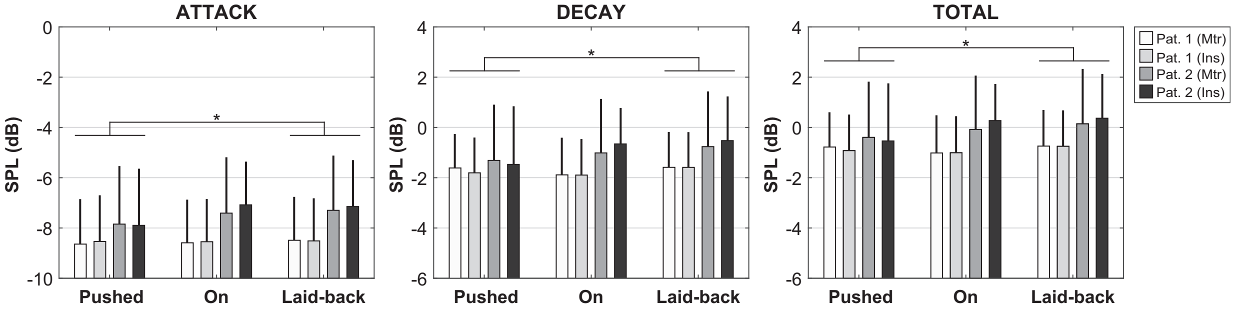

FIGURE 8. Mean snare stroke SPL (all segments) across participants ($N = 20$) in all style, reference, and pattern conditions. Error bars indicate 0.5 *SD*. Note: Pat. = pattern, Mtr = metronome reference, Ins = instrumental reference.

We found no main effects of duration for the hi-hat, nor any interactions for any stroke segment in any of the percussion instruments.

Sound Pressure Level (SPL). For the snare drum, a significant main effect of style on SPL appeared in all stroke segments (attack/decay/total). Post hoc pairwise comparisons revealed significantly higher mean attack, decay and total SPL in the laid-back compared to pushed strokes, but no significant differences in any segment between laid-back and on-beat strokes or pushed and on-beat strokes (see Table 8). Figure 8 illustrates the results for mean SPL for all snare segments in all style, reference, and pattern conditions.

We also found a significant main effect of Pattern on SPL for the snare. The post hoc pairwise comparison shows that mean SPL was higher in Pattern 2 than in Pattern 1 in the attack ($M = +1.11$ dB, $SD = 1.48$ dB, $p = .003$, $d = 0.75$) and total ($M = +0.83$ dB, $SD = 1.72$ dB, $p = .043$, $d = 0.48$) stroke segments.

In order to gauge whether the structural differences in note composition between patterns may have had an effect on snare SPL, we ran a supplementary paired-samples *t*-test between the average SPL of the notes in

TABLE 9. *Paired-samples t-tests Between Snare Average Note SPL in Patterns 1 and 2 (N = 20)*

	Pattern		Attack SPL (dB)				Total SPL (dB)			
	1	2	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*
Note	"two"	"two"	-1.01	1.44	.033	-0.70	-0.80	1.73	.319	-0.46
	"two"	"two-and"	-1.08	1.51	.028	-0.72	-0.67	1.72	.579	-0.39
	"two"	"four"	-1.22	1.51	.011	-0.81	-1.01	1.73	.102	-0.59
	"four"	"two"	-1.02	1.43	.031	-0.71	-0.80	1.72	.309	-0.46
	"four"	"two-and"	-1.09	1.50	.026	-0.73	-0.68	1.72	.562	-0.39
	"four"	"four"	-1.23	1.50	.01	-0.82	-1.02	1.73	.098	-0.59

TABLE 10. *Paired Samples t-tests of Average Snare SPL Between Pattern 1 and Pattern 2 in All Style Conditions (N = 20)*

	Snare												
	Pattern	Attack SPL (dB)				Decay SPL (dB)				Total SPL (dB)			
	1 – 2	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*
Style	Laid-back	-1.28	1.34	< .001	-0.95	-0.94	1.71	.023	-0.55	-1.00	1.63	.013	-0.62
	On	-1.32	1.07	< .001	-1.23	-1.06	1.26	.001	-0.84	-1.11	1.21	.001	-0.91
	Pushed	-0.72	2.10	.142	-0.34	-0.32	2.46	-.571	-0.13	-0.38	2.38	.480	-0.16

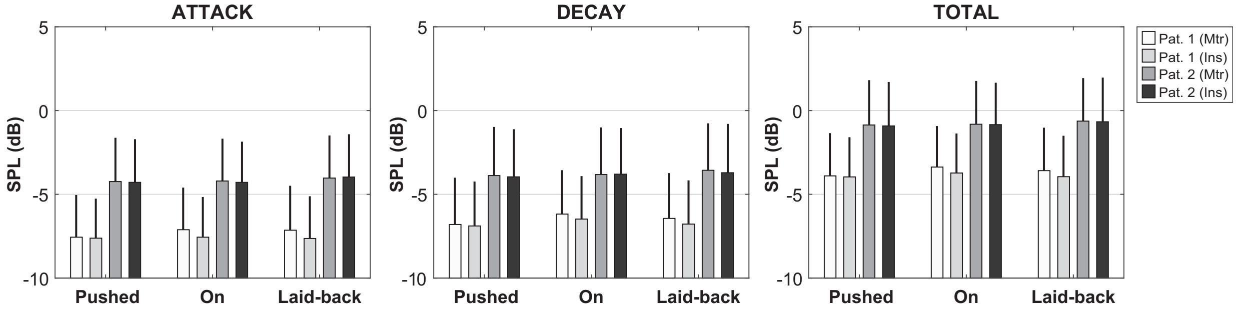

FIGURE 9. Mean kick stroke SPL (all segments) across participants (*N* = 20) in all style, reference, and pattern conditions. Error bars indicate 0.5 *SD*. Note: Pat. = pattern, Mtr = metronome reference, Ins = instrumental reference.

each pattern, as we did with duration. The tests revealed that while attack SPL of all notes in Pattern 2 were significantly higher than the corresponding notes of Pattern 1, the results of the total SPL comparisons showed no differences between the patterns (see Table 9).

We also found a significant interaction between Style and Pattern for all snare SPL stroke segments, such that there was a stronger effect of Style on SPL in Pattern 2 than in Pattern 1. Paired-samples *t*-tests revealed that mean SPL (all stroke segments) was significantly higher in Pattern 2 than in Pattern for laid-back strokes and on-beat strokes, but showed no difference for pushed strokes (see Table 10).

For the kick drum, only a main effect of Pattern on SPL was found. Post hoc pairwise comparisons showed that all stroke segments were played with higher SPL in Pattern 2 than in Pattern 1 at $p < .001$ (see Figure 9): attack, $M = +3.27$ dB, $SD = 1.27$ dB, $d = 2.58$; decay, $M = +2.81$ dB, $SD = 1.92$ dB, $d = 1.46$; total, $M = +2.96$ dB, $SD = 1.34$ dB, $d = 2.21$.

We ran a supplementary paired-samples t-test between the average SPL of the notes in each pattern. The test revealed that, for all stroke segments, not only were the differing kick notes from Pattern 2 ("three-and," "four-and") louder than all the notes in Pattern 1 ("one," "three"), but also all of the notes in Pattern 2 were played significantly louder than all of the notes in Pattern 1 (see Table 11).

For the hi-hat, we found a significant main effect of Style on SPL in all stroke segments. Post hoc pairwise

TABLE 11. *Paired-samples t-tests Between Kick Average Note SPL in Patterns 1 and 2 (N = 20)*

	Pattern		Attack SPL (dB)				Decay SPL (dB)				Total SPL (dB)			
	1	2	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*
Note	"one"	"one"	-3.83	1.22	< .001	-3.14	-3.88	1.22	< .001	-3.18	-3.85	1.20	< .001	-3.22
	"one"	"three-and"	-3.27	1.36	< .001	-2.40	-3.31	1.29	< .001	-2.57	-3.29	1.30	< .001	-2.54
	"one"	"four-and"	-1.58	2.08	.018	-0.76	-1.60	1.89	.008	-0.84	-1.60	1.93	.009	-0.83
	"three"	"one"	-3.85	1.25	< .001	-3.08	-3.91	1.23	< .001	-3.19	-3.88	1.21	< .001	-3.20
	"three"	"three-and"	-3.30	1.40	< .001	-2.36	-3.35	1.30	< .001	-2.57	-3.32	1.31	< .001	-2.53
	"three"	"four-and"	-1.60	2.10	.018	-0.76	-1.63	1.90	.007	-0.86	-1.62	1.94	.008	-0.84

TABLE 12. *Pairwise Comparisons of Main Effect of Style on Hi-hat SPL (N = 20)*

Hi-hat												
	Attack SPL (dB)				Decay SPL (dB)				Total SPL (dB)			
Style	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*	Δ*M*	*SD*	*p*	*d*
Laid-back – On	0.90	1.54	.051	0.58	0.55	1.16	.138	0.48	0.76	1.29	.049	0.59
Pushed – On	0.75	1.21	.037	0.62	0.75	0.90	.004	0.84	0.79	0.97	.005	0.81
Laid-back – Pushed	0.15	1.05	1.000	0.14	-0.20	0.72	.661	-0.28	-0.03	0.78	1.000	-0.03

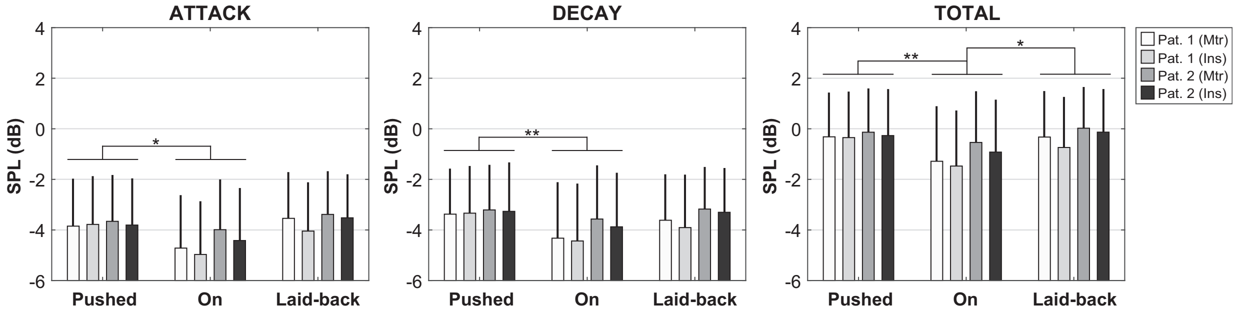

FIGURE 10. Mean hi-hat stroke SPL (all segments) across participants ($N = 20$) in all style, reference, and pattern conditions. Error bars indicate 0.5 *SD*. Note: Pat. = pattern, Mtr = metronome reference, Ins = instrumental reference.

comparisons revealed significantly higher mean SPL in the pushed condition compared to the on-beat condition, as well as significantly higher SPL for laid-back compared to on-beat for the total segment, a trend toward significance for the attack segment, but no significant difference for the decay segment. The difference in SPL between the pushed and laid-back conditions was not significant for any of the segments (see Table 12).

We also found a significant main effect of Reference on SPL for the hi-hat in all stroke segments. Post hoc pairwise comparisons revealed higher mean SPL in the metronome compared to the instrumental condition (attack, $M = +0.23$ dB, $SD = 0.50$ dB, $p = .050$, $d = 0.47$; decay, $M = +0.14$ dB, $SD = 0.29$ dB, $p = .043$, $d = 0.48$.; total, $M = +0.22$ dB, $SD = 0.28$ dB, $p = .003$, $d = 0.78$). No interactions were found on SPL for the hi-hat. Figure 10 illustrates mean SPL for all hi-hat segments in all style, reference, and pattern conditions.

Discussion

EFFECTS OF TIMING STYLE ON ONSET LOCATION

Drummers show high degree of onset timing control. Regarding onset location timing, results show that, on average, all twenty drummers were able to perform the tasks—they played laid-back and pushed strokes slightly later and slightly earlier relative to the on-beat timing style condition. These results accord with those of previous instructed timing-style performance studies with drummers by Kilchenmann and Senn (2011) and Danielsen and colleagues (2015), and with guitarists and bassists by Câmara and colleagues (2020), and they provide further evidence of musicians' high degree of intentional control of onset manipulation in order to produce different timing feels in groove-based performance. Here, the average onset difference of all instruments between on-beat and laid-back strokes across pattern and timing reference was found to range between 21 to 27 ms, and for pushed compared to on-beat strokes, between -37 to -42 ms. In music theory terms, at the experiment tempo (96 bpm), these averages correspond to durational differences between roughly a 128th note (19.6 ms) and a 64th note (39.1 ms) to distinguish the average microtiming onset profile of laid-back and pushed from on-beat, indicating a high degree of control at the microrhythmic level. This is not to say, however, that drummers consciously operate with these minute canonical note categories in mind, only that they are able to produce onsets flexibly around the beat.

As to whether the average listener would be able to distinguish these different feels based on their onset timing profiles alone, most music performance studies, as mentioned previously, have either speculated or found indirect evidence that the threshold for detection of asynchrony between two sound events in a musical context is around 30 milliseconds (Butterfield, 2011; Goebl & Parncutt, 2002). If this is correct, then it is indeed likely, based on our experiment results, that one would be able to detect both laid-back and pushed drum strokes as asynchronous relative to either the timing reference stimuli or the onset timing of the on-beat conditions, with the likelihood of detecting pushed strokes being even greater due to their larger average onset asynchrony magnitudes.

As was the case in a previous study with guitarists and bassists (Câmara et al., 2020), the variability (*SD*) of the mean onset location for all instruments was numerically higher in the pushed (24 ms) and laid-back (17 ms) styles than it was in the on-beat (12 ms) condition. This may be explained by the fact that the majority of the participants reported having practiced on-beat synchrony with a timing reference more than they had practiced laid-back and pushed styles. It may also be related to the fact that participants described laid-back and pushed as more ambiguous timing style categories that allowed for a greater range of onset locations while still sounding aesthetically "correct." This suggests that drummers, like guitarists and bassists, also regarded the "beat bin," or temporal range in which sounds are perceived as synchronous with the beat (Danielsen, 2010; Danielsen et al., 2019) in the pushed and laid-back styles as broader than that of on-beat strokes.

Asymmetry of onset asynchrony magnitude between pushed and laid-back styles. A curious result was that, when they were playing with a pushed feel, the drummers produced greater asynchrony relative to both the timing references and the average on-beat condition onset location than when they were playing with a laid-back feel. On the one hand, the greater values for the pushed strokes may be because participants described the pushed feel as more difficult and unfamiliar, which is demonstrated as well by the greater variability in average pushed stroke onsets. As such, the increased difficulty could have led to the exaggeration of the earliness of pushed strokes over the lateness of laid-back strokes. On the other hand, drummers may consider laid-back timing approaches to be aesthetically amenable to subtler degrees of lateness, whereas they may consider more pronounced earliness to be acceptable in pushed performances. That is, a performance might sound sloppy or loose, rather than simply laid-back, if the strokes are exaggeratedly late, whereas in a pushed timing feel, the threshold for early strokes sounding rushed rather than pushed may be larger, allowing for greater earliness magnitudes. To our knowledge, however, no one has yet specifically investigated whether onset asynchrony thresholds are different in systematically off-beat late vs. early events relative to an external timing reference in a rhythmic context.

Timing reference and onset NMA. As expected, in the on-beat condition the average onset timing of all drummers anticipated the timing reference. We found this anticipatory tendency, or negative mean asynchrony (NMA), in both reference conditions for all instruments. At the same time, however, the results show relatively low overall NMA values in the range of -10 to -20 ms. These results accord with findings from in-phase sensorimotor synchronization tapping studies (see Repp, 2005; Repp & Su, 2013) showing that highly trained musicians, and especially drummers, tend to display lower NMA than nonmusicians, typically in the range of 0 to 20 ms. Moreover, Fujii and colleagues (2011) found that, in a simple on-beat synchronization performance task by drummers to a metronome, NMA was lowest for the hi-hat, followed by the snare and the kick drum in the medium tempo condition (120 bpm). This also resonates with our findings, where the hi-hat displayed the lowest NMA values in the on-beat conditions. The reason drummers tend to produce lower NMA with the hi-hat cymbal may be related to the fact that it is typically considered the main timekeeping element of the drum-kit and is typically played with the strongest/dominant hand (right for right-handed players).

The metronome reference yielded greater NMA than the instrumental reference, with the snare and the hi-hat demonstrating near-negligible anticipation in the instrumental (-4 ± 13 ms and -3 ± 14 ms, respectively). As of yet, there is no consensus as to what leads to NMA in sensorimotor-synchronization tasks. One potential explanation for the instrumental reference leading to lower overall NMA, however, is that the guitar and bass sounds used in the instrumental reference track had longer attack and total durations than the faster and shorter metronome woodblock sounds. Accordingly, as predicted by the findings of research into the perceptual centers of musical sounds (Danielsen et al., 2019; Gordon, 1987; Villing, 2010), the synchronized target location for drummers would be slightly later in the instrumental reference. Regardless, our findings accord with previous studies showing that synchronization to more ecological musical stimuli tends to lead to less NMA than synchronization to metronomic stimuli (Dixon, Goebl, & Cambouropoulos, 2006; Repp, 2008; Wohlschläger & Koch, 2000).

EFFECTS OF TIMING STYLE ON SOUND SHAPE: DURATION AND INTENSITY

Timing style and duration: Longer snare and shorter kick in the laid-back condition. Regarding duration, we found significant main effects of timing style conditions for the snare and kick drums. While drummers showed a tendency to play snare strokes in the laid-back timing style condition slightly longer than they did in the on-beat condition, we found the opposite effect for the kick drum, curiously, whereby laid-back strokes were played slightly shorter than on-beat strokes. In both cases, only the duration of either the decay or total stroke was lengthened; attack length, as defined in this study, was much less alterable due to the constraints of the instrument type. (It is difficult, if not impossible, to manipulate the attack rise time of an impulsive strike of either a drumstick or foot pedal beater on a drum skin membrane as it occurs in a very short time interval where the amplitude rises very quickly.) Decay length, on the other hand, could have been lengthened by either striking the drum membrane with a greater intensity or allowing the membrane to vibrate for a longer time or both. However, a Pearson's correlation test revealed only significant weak correlations at $p < .001$ between stroke duration and intensity (snare: decay, $R = .07$, total, $R = .04$; kick: decay, $R = .09$). Instead, then, for the snare, longer decay/total snare stroke length may have been achieved by allowing the drum stick to continue to bounce lightly on the skin after the first stroke impulse for a longer time (a so-called "normal" stroke technique, with stick rebound, as opposed to a "controlled" stroke, where the stick is stopped and held firmly above the skin right after impulse (Dahl & Altenmüller, 2008). Similarly, shorter kick strokes may have been achieved by keeping the foot pedal beater pressed firmly against the drum-head after striking (often referred to as "burying the beater"), effectively curtailing the resultant sound slightly, as opposed to allowing it to bounce back freely and allowing the membrane to vibrate longer.

The snare result accords with a previous instructed timing style study that revealed that guitarists also lengthened the duration of their laid-back strokes (Câmara et al., 2020). Though the magnitude of stroke lengthening by the drummers was subtler than that of the guitarists (ca. 7 ms vs. 30 ms significant difference from the on-beat style condition), P-center studies show ample evidence that longer durations lead to later P-centers relative to signal onset. Therefore, as was the case with the guitarists' tests, a *late and long* stroke may further augment the "behind-the-beat" character of laid-back strokes, whereas shorter durations encourage the listener to experience on-beat and pushed strokes as either more in sync or earlier relative to timing reference stimuli, thereby enhancing their synchronous or early timing character, respectively. Reported strategies also corroborated these signal analysis results, in that several of the drummers described applying "slower" movements, aiming for "longer tones with more sustain," and holding the drum stick with a "looser" grip to achieve a laid-back feel, and conversely using "faster/smaller" movements and tones with "less sustain" via a "tighter grip" to achieve on-beat and pushed feels.

As to why drummers would utilize shorter kick strokes, which potentially elicit the experience of an earlier P-center in the laid-back condition, we might look at the onset location differences between kick and snare in the various timing style conditions. The overall tendency for drummers in the laid-back style conditions was to position the kick slightly earlier (+7 ms) than the snare (+17 ms) relative to the timing reference. The ensuing average inter-instrument onset difference between kick and snare (10 ms, $SD = 7$ ms) in the laid-back condition is significant in itself, $t(6.415)$, $p < .001$, and also significantly greater than the equivalent differences in the on-beat (4 ± 3 ms) $t(4.841)$, $p < .001$, and pushed (1 ± 8 m), $t(3.630)$, $p = .002$, style conditions. Some participants, in fact, reported that when they were playing in the laid-back style condition, they consciously implemented a strategy of aiming the kick drum closer to the timing reference (more on the beat) while delaying both snare and hi-hat slightly. By utilizing a strategy of relatively *earlier and shorter kick + later and longer snare*, then, the perception of an even longer interonset interval between kick and snare might be enhanced due to the additional effects of duration on P-center, thereby enhancing the overall laid-back timing feel of the performance as such.

Timing style and sound pressure level: Late and loud vs. early and loud. A significant main effect of timing style on sound pressure level (SPL) was found for the snare—drummers played laid-back strokes with higher decay/total SPL than they played pushed strokes. This was essentially a replication of the intensity findings of Danielsen and colleagues (2015), where drummers on average played laid-back strokes with the greatest intensity compared to the on-beat condition, and provides further evidence that intensity is a vital feature of timing style production. Interview responses from the present study's twenty drummers also confirm the previous study's reported association between laid-back feel and playing "heavier" by "giving more weight" to snare strokes. In addition, the drummers reported positioning themselves for the laid-back condition by leaning more backward or away from the snare and/or lifting the stick higher in preparation for a stroke, both of which may result in greater intensity (more distance allows for a higher striking velocity) as well as later onset (particularly when combined with a "flam" technique during simultaneous snare and hi-hat strokes—if both strokes fall toward the instruments at the same time but the snare begins higher up, it will land after the hi-hat). On the other hand, the drummers associated the pushed condition with "lighter/softer" or "thinner" strokes that were played with the body and hands "positioned closer to the drums," and if a flam technique was used, the snare stroke was aimed to fall on the drum before the hi-hat (thanks to a lower snare stick height).

As with duration, however, we found no single directionality of the effect of timing style on one sound feature across the different percussion instruments, as the hi-hat indicated the drummers' tendency to play both pushed and laid-back strokes with higher total SPL compared to on-beat strokes. In groove-based music performance, as mentioned, the hi-hat cymbal is widely considered to be the main "timekeeper" of the drum kit (and the entire ensemble in live performance contexts), because it clearly and consistently externalizes the density referent of the groove pattern—that is, its smallest practical metrical subdivision level (Nketia, 1974)—which, in our experiment, was manifested in both patterns as a stream of eighth notes. To clearly convey the idea that this timekeeper is meant to be heard as pushing against the timing reference or ensemble, then, drummers would want to accent it with greater intensity in order to increase its perceptual salience. While Goebl and Parncutt (2002) suggest that it is more difficult to detect the presence of an asynchrony in early and loud combinations due to a potential forward masking effect, the greater the onset magnitude between the two sounds, the higher the chance of detecting an asynchrony between them. We may recall that the onset asynchrony between pushed strokes and timing reference was found to be greater than between laid-back strokes and timing reference for all of the instruments. For the hi-hat, the average onset asynchrony of pushed strokes relative to timing reference was -46 (± 21) ms. This is substantially above the -27 ms asynchrony condition in Goebl and Parncutt's (2002) experiment, where detectability of asynchrony in early and loud tone combinations was 20–30%, and closer to the -54 ms asynchrony condition, where it was higher at 40–50%. It therefore follows that increasing the magnitude of onset asynchrony of hi-hat strokes in the pushed condition would counteract the risk of reduced detectability of asynchrony in early and loud combinations due to forward masking. That is to say, when hi-hat strokes are played louder, the earlier they are played, the higher the chance they may be heard as asynchronously pushing against the timing reference layer, rather than masking and potentially supplanting it.

On the other hand, if an early and loud tone combination is harder to detect, then, as Danielsen and colleagues (2015) suggest, a *late and loud* stroke would appear to facilitate the perception of asynchrony between tones. In fact, Tekman (2002) found that it was easier to correctly classify a tone as being late in relation to another tone when it was both positioned late and coupled with a greater intensity of delivery, compared to a tone that was simply positioned late at the same intensity as another tone. In the laid-back condition, then, amplifying the hi-hat or the snare with a slightly greater intensity would help them stand out more clearly as late strokes in relation to the timing references. In addition, since the risk of forward masking may not be present in late and loud combinations, it may be that lesser degrees of asynchrony are needed to convey an intentionally late stroke against a timing reference when they are also played with greater intensity, which may be related to the overall lower average onset differences produced between laid-back and timing reference for the hi-hat and snare at least.

EFFECTS OF MUSICAL CONTEXT: REFERENCE AND PATTERN

Instrumental reference produces more extreme early and late onset timing. Timing reference stimuli had an amplifying effect on the magnitude of asynchrony—laid-back strokes were played even later (and, conversely, pushed strokes were played even earlier) in the instrumental reference compared to the metronome. This effect was most salient in Pattern 1 (see Figure 5) and may be related to the spectral features and/or duration of the timing reference stimuli, since the instrumental guitar and bass sounds were both longer than those of the woodblock metronome stimuli and spectrally more distinct from the drum kit's sounds as well. When producing an asynchrony between two short, percussive, and highly impulsive sounds such as the woodblock and the snare/kick/hi-hat, the threshold for onset asynchrony may be reduced, because two click-like sounds in close proximity may stand out more to the listener/performer. On the other hand, when producing an asynchrony between an impulsive drum sound with faster attack and shorter duration and a guitar or bass sound with a relatively slower attack and longer duration, a greater magnitude may be required to prevent potential spectral, durational, or dynamic masking effects.

Effects of reference on sound pressure level of time-keeper (hi-hat). Timing reference stimuli for the drummers also had an effect on the dynamics of their strokes: overall, they played the hi-hat more loudly to the metronome than they did to the instrumental stimuli. This may be related to differences in their interpretations of the tasks. Typically, drummers will play alone to a metronome either as a way to practice (to hone timing skills) or when they are recording a backing track in a studio session for a song, which is later overdubbed by other instruments. When they play solo with a metronome, then, they want most of all to ensure that they are synchronized with it, and by playing the hi-hat (the time-keeping element that coincides with all of the metronome's sounds) louder, they may be trying to enhance their ability to do so. Unsurprisingly, participants reported that they tended to play more "mechanically" to the metronome. When they played to the instrumental reference, on the other hand, which evokes a live or studio trio rhythm section, they entered a more "dialogical mode" (Chernoff, 1979) of performance, whereby they played together with the other instruments rather than in strict synchronization to a time-keeper. In this context, it is possible that the drummers did not need to emphasize the timekeeper hi-hat so consistently and could focus instead on balancing their drum kit sounds with those of the recorded ensemble in a fashion they deemed appropriate to the musical context.

Effects of pattern: Musical contextual/aesthetic considerations. Pattern was found to have several effects on the onset, duration, and SPL of the different percussion instruments. Regarding onset timing, a main effect of pattern was found for the hi-hat, where the more complex rhythms of Pattern 2 led to slightly earlier strokes than did Pattern 1. As there were no differences between the hi-hat patterns themselves, this effect suggests that the greater density of notes stemming from the extra kick and snare off-beat strokes may have led drummers to play Pattern 2 in a "pushier" fashion overall. Just as drummers can be either "pushier" players or display a tendency to anticipate the timing reference even when playing laid back (Kilchenmann & Senn, 2011), certain patterns can elicit "pushier" performances by drummers in general, particularly when they are combined with greater pattern density and complexity.

As to duration, the drummers also played the snare drum with shorter durations (decay/total)—in the more complex and busier Pattern 2 than in the simpler and sparser Pattern 1. The main structural difference between the snare patterns was that, in Pattern 1, only a single eighth note stroke appeared on metrical position "two"; in pattern 2, on the other hand, a double eighth note stroke spanned the positions "two" and "two-and." This double stroke may have compressed the time between the two strokes, leading to shorter average note durations. A supplementary paired-sampled *t*-test revealed that not only were the duration of the double strokes in Pattern 2 shorter than the equivalent single stroke in Pattern 1 but also almost all of the notes in Pattern 2 were significantly shorter than those in Pattern 1. This suggests that the reason why drummers played shorter strokes in Pattern 2 may be the overall greater density and proximity of the notes of all the instruments in that pattern, rather than simply the extra snare note alone.

As to effects of pattern on SPL, both snare (attack) and kick drum (all segments) were played more loudly in Pattern 2. For the snare, we found a further interaction between timing style and pattern as well—the drummers played the snare strokes more loudly in Pattern 2 in the laid-back and on-beat conditions but not the pushed condition, which recalls the snare result of the main effect of timing style on SPL (the laid-back condition was louder than the pushed). A follow-up note-by-note analysis of stroke SPL also revealed that, for both snare and kick drums, almost all of the notes of

Pattern 2 were louder than all of the notes in Pattern 1. If greater effort leads to greater intensity, the relative greater difficulty reported by participants when playing laid-back and pushed feels in Pattern 2 as opposed to Pattern 1 may have contributed here. The reason may also be aesthetic in nature—perhaps the drummers considered the extra intensity to project a more "driving" or "energetic" rhythm in relation to the simpler Pattern 1. Overall, then, it would appear that musical context influences microtiming profiles—that is, the nature of the pattern and timing reference in a given musical context may further affect how timing feels are expressed in drumming performances.

Systematicity and intentionality. While drummers were found to systematically manipulate intensity or duration of strokes to express different timing styles, this does not necessarily imply intent on the drummers' part. That is, it is fully possible that changes in these sound features were simply byproducts of onset timing manipulation due to motor limitation effects outside of their explicit control. At the same time, the possibility of intent should not be precluded, especially since drummers at times described applying strategies that both implicitly and explicitly mentioned a focus on intensity or duration. Also, although drummers are likely not consciously aware of the perceptual effects that duration and intensity have on either the P-center of sounds or on the detectability of onset asynchronies, this does not mean they cannot hear or feel the effects that longer/shorter or louder/softer sounds might have when further applied to late/early strokes. As such, it may be that they are able to intuitively utilize certain combinations of onset and intensity/duration in order to better achieve a given timing style.

Summary and Conclusions

Our findings show that drummers systematically manipulated not only the onset location but also the intensity and/or duration of the various drum instruments when instructed to perform groove-based patterns in a laid-back, on-beat, or pushed fashion. Drummers produced distinctive average stroke onset profiles for each timing style, with the pushed condition showing a tendency to be slightly more anticipated than the laid-back condition was delayed, potentially as a result of the increased difficulty of the pushed condition reported by drummers, or a decreased sensitivity to early as opposed to late onset asynchrony. Systematic differences in the shape of the acoustic signal for strokes played with different timing styles were also found in at least one measured sound descriptor (duration and SPL) for all the instruments in the drum kit. Drummers tended to play snare strokes in the laid-back condition louder and longer on average, a timing/sound strategy that might enhance the perceived lateness of strokes due to the increased detectability of late and loud asynchronies and the P-center delaying effects of longer durations. Kick drum strokes, on the other hand, were played shorter on average in the laid-back condition, which, when viewed in light of the longer concomitant snare strokes, would amplify the perceived time *interval* between the drum strokes themselves, rather than simply enhance the delayed or anticipated character of single strokes in relation to a timing reference. Lastly, the hi-hat showed a tendency to be played louder in both asynchronous conditions that was potentially related to its role as a timekeeper—that is, greater intensity may increase its perceptual salience and thus help to highlight other intentionally produced asynchronies in relation to an external timing reference.

Musical context also influenced timing style as effects of both timing reference (metronome vs. instrumental backing track) and pattern (Pattern 1 ["simple"] vs. Pattern 2 ["complex"]) were found on either onset location, duration, or intensity across the different percussion instruments. Timing reference primarily impacted the average onset profiles of the instructed timing styles with instrumental sounds leading to more pronounced early and late timing, perhaps to ensure that asynchronies were not spectrally or dynamically masked by the wider instrumental bass and guitar sounds, with their relatively slower attacks and longer durations. The metronome led to greater NMA in the on-beat style condition for all instruments, possibly prompted by the earlier P-centers of the metronome stimuli, with their shorter durations and faster attacks. Pattern was also found to affect performance—for example, snare and kick strokes were played louder, and hi-hat strokes earlier in Pattern 2 than in Pattern 1. Since the main difference between the two instructed patterns was structural, Pattern 2's effects may be attributed to the greater note density and off-beat character of the snare and kick.

Overall, this study's findings show further evidence that the production of "timing" in groove-based music involves more than simple onset asynchrony relations between instruments and/or timing references. While, in terms of effect sizes, onset manipulation may still seem to be the salient vehicle for drummers when expressing different timing feels in groove-based music, we should not ignore the durational and dynamic nuances of strokes played with onset asynchronies. We might thus consider that more specific terms such as "timing-sound styles" or "microrhythmic feels" may better convey the potential multiple range of acoustic features involved in the production of performed rhythmic events, especially since "timing" is so heavily connoted with "onset" to the point where they are virtually synonymous. However, "timing" still holds strong currency within music performance communities and scholarly institutions, and therefore it may also be possible to simply expand the concept of "timing" to encompass the concomitant manipulation of temporal and sound features.

Regardless of what we call this phenomenon, future empirical studies concerned with measuring and interpreting the role of timing expression in music would do well to consider also the sonic shape of rhythmic events, not to mention the aesthetic and stylistic contexts in which they are produced. Music psychological studies seeking to measure "groove" ratings (operationalized as the aspect of the music that elicits the "urge to move" and/or pleasure) while running into conflicting results regarding the role of microtiming, for example, might consider manipulating not only onset but also intensity and duration of instrumental stimuli, preferably with baseline conditions that resemble timing/sound profiles from actual musical performance examples as closely as possible. Music producers concerned with "humanizing" computer-programmed grooves with artificial onset asynchrony manipulations that may otherwise sound mechanical or deadpan may also benefit from scholarly studies that further categorize or model intensity and durational profiles obtained from real performances. As Albhy Galuten, producer of the Bee Gee's seminal disco-funk track "Staying Alive" from 1977, put it, "Everyone knows that it's more about feel than accuracy in drum tracks"—though he used a drum loop in the hit's construction, it was extracted from a real performance, so, he insisted, "it felt really great—very insistent but not machinelike . . . [i]t had a human feel" (quoted in Grogan, 2018, p. 128). In the opinion of the present authors, whatever creates the particular aesthetic appeal that actual human performance elicits in listeners—that ephemeral rhythmic "feel" of groove-based music—must surely involve both the onset microtiming profiles of instruments and the manner in which they are "texturally" shaped.

In future research, we plan to conduct perception experiments to determine whether listeners are able to identify drum strokes as laid-back, on-beat, or pushed on the basis of their sound alone. In addition, we recorded motion-capture data during this experiment that we will analyze to determine whether drummers also systematically utilize different movement trajectories to produce the differences in duration and intensity we observed in this study.

Author Note

The authors wish to thank all the drummers for their participation; Georgios Sioros, Victor Gonzáles Sánchez, and Martin Torvik Langerød (University of Oslo, Norway) for their assistance with the experimental setup; Carl Haakon Waadeland (Norwegian University of Science and Technology, Norway) for his help with the experimental design and recruitment of participants; Dag-Erik Eilertsen (University of Oslo, Norway) for his assistance with SPSS scripts for the statistical analysis; Justin London (Carleton College, USA) and Olivier Senn (Hochschule Luzern, Switzerland) for their comments on earlier versions of this manuscript. We are also grateful to the anonymous reviewers for their interesting and valuable comments and suggestions. This work was partially supported by the Research Council of Norway through its Centers of Excellence scheme, project number 262762, and the TIME (Timing and Sound in Musical Microrhythm) project, grant number 249817.

Correspondence concerning this article should be addressed to Guilherme Schmidt Câmara, RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Box 1133 Blindern, NO-0318 Oslo, Norway. E-mail: g.s.camara@imv.uio.no

References

Bechtold, T. A., & Senn, O. (2018). Articulation and dynamics influence the perceptual attack time of saxophone sounds. *Frontiers in Psychology*, *9*, 1692.

Benadon, F. (2006). Slicing the beat: Jazz eighth notes as expressive microrhythm. *Ethnomusicology*, *50*, 73–98.

Berliner, P. (1994). *Thinking in jazz: The infinite art of improvisation*. Chicago, IL: University of Chicago Press.

Butterfield, M. W. (2006). The power of anacrusis: Engendered feeling in groove-based musics. *Music Theory Online*, *12*.

Butterfield, M. W. (2010). Participatory discrepancies and the perception of beats in jazz. *Music Perception*, *27*, 157–175.

Butterfield, M. W. (2011). Why do jazz musicians swing their eighth notes? *Music Theory Spectrum*, *33*, 3–26.

Câmara, G. S. (2016). *Swing in early funk and jazz-funk (1967–1971): Micro-rhythmic and macro-structural investigations* (Master's thesis). University of Oslo, Oslo, Norway.

Câmara, G. S., & Danielsen, A. (2018). Groove. In A. Rehding & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory* (pp. 271–294). Oxford, UK: Oxford University Press.

Câmara, G. S., Nymoen, K., Lartillot, O., & Danielsen, A. (2020). Effects of instructed timing on electric guitar and bass sound in groove performance. *Journal of the Acoustical Society of America*, *147*, 1028–1041.

Chernoff, J. M. (1979). *African rhythm and African sensibility: Aesthetics and social action in African musician idioms*. Chicago, IL: University of Chicago Press.

Clarke, E. F. (1989). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music* (pp. 1–26). Oxford, UK: Clarendon Press.

Dahl, S. (2004). Playing the accent: Comparing striking velocity and timing in an ostinato rhythm performed by four drummers. *Acta Acoustica United with Acoustica*, *90*, 762–776.

Dahl, S. (2011). Striking movements: A survey of motion analysis of percussionists. *Acoustical Science and Technology*, *32*, 168–173.

Dahl, S., & Altenmüller, E. (2008). Motor control in drumming: Influence of movement pattern on contact force and sound characteristics. *Journal of the Acoustical Society of America*, *123*, 3122.

Danielsen, A. (2006). *Presence and pleasure: The funk grooves of James Brown and Parliament*. Middletown, CT: Wesleyan University Press.

Danielsen, A. (2010). Here, there and everywhere: Three accounts of pulse in D'Angelo's 'Left and Right.' In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 19–36). Farnham, UK: Ashgate.

Danielsen, A. (2018). Pulse as dynamic attending: Analysing beat bin metre in neo soul grooves. In C. Scotto, K. M. Smith, & J. Brackett (Eds.), *The Routledge companion to popular music analysis: Expanding approaches* (pp. 179–189). London, UK: Routledge.

Danielsen, A., Nymoen, K., Anderson, E., Câmara, G. S., Langerød, M. T., Thompson, M., & London, J. (2019). Where is the beat in that note? Effects of attack, duration and frequency on the perceived timing of musical and quasi-musical sounds. *Journal of Experimental Psychology, Human Perception and Performance*, *45*, 402–418.

Danielsen, A., Waadeland, C. H., Sundt, H. G., & Witek, M. A. G. (2015). Effects of instructed timing and tempo on snare drum sound in drum kit performance. *Journal of the Acoustical Society of America*, *138*, 2301–2316.

Davies, M., Madison, G., Silva, P., & Gouyon, F. (2013). The effect of microtiming deviations on the perception of groove in short rhythms. *Music Perception*, *30*, 497–510.

Dixon, S., Goebl, W., & Cambouropoulous, E. (2006). Perceptual smoothness of tempo in expressively performed music. *Music Perception*, *23*, 195–214.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception*, *10*, 343–378.

Fischer, R., & Milfont, T. L. (2010). Standardization in psychological research. *International Journal of Psychological Research*, *3*, 88–96.

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

Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, *19*, 333–349.

Frühauf, J., Kopiez, R., & Platz, F. (2013). Music on the timing grid: The influence of microtiming on the perceived groove quality of a simple drum pattern performance. *Musicae Scientiae*, *17*, 246–260.

Fujii, S., Hirashima, M., Kudo, K., Ohtsuki, T., Nakamura Y., & Oda, S. (2011). Synchronization error of drum kit playing with a metronome at different tempi by professional drummers. *Music Perception*, *28*, 491–503.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (pp. 501–602). New York: Academic Press.

Gerischer, C. (2006). O suingue Baiano: Rhythmic feeling and microrhythmic phenomena in Brazilian percussion. *Ethnomusicology*, *50*, 99–119.

Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America*, *110*, 563–572.

Goebl, W., & Parncutt, R. (2002). The influence of relative intensity on the perception of onset asynchronies. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings from the 7th International Conference on Music Cognition and Perception* (pp. 613–616). Adelaide, Australia: Causal Productions for AMPS.

Gordon, J. W. (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America*, *82*, 88–105.

Grogan, J. (2018). *Origins of a song: 202 true inspirations behind the world's greatest lyrics*. Kennebunkport, ME: Cedar Mill Press.

Haas, W. B. D. (2007). *The role of tempo in groove and swing timing* (Master's thesis). Utrecht University, Utrecht, Netherlands.

Haugen, M. R., & Godøy, R. I. (2014). Rhythmical structures in music and body movement in samba performance. In M. K. Song (Ed.), *Proceedings of the ICMPC-APSCOM 2014 Joint Conference: 13th Biennial International Conference for Music Perception and Cognition and 5th Triennial Conference of the Asia Pacific Society for the Cognitive Sciences of Music* (pp. 46–52). Seoul, Korea: Yonsei University.

Haugen, M. R., & Danielsen, A. (2020). Effect of tempo on relative durations in a performed samba groove. *Journal of New Music Research,* DOI: 10.1080/09298215.2020.1767655

Hirsh, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America*, *31*, 759–767.

Honing, H., & Haas, W. B. D. (2008). Swing once more: Relating timing and tempo in expert jazz drumming. *Music Perception*, *25*, 471–476.

Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception*, *19*, 387–414.

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology*, *141*, 54–75.

Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, *2*, 275–283.

Keil, C. (1995). The theory of participatory discrepancies: A progress report. *Ethnomusicology*, *39*, 1–19.

Kernfeld, B. (2003). Groove (i). *Grove Music Online*. Retrieved December 11, 2019 from https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-2000582400

Kilchenmann, L., & Senn, O. (2011). “Play in time, but don’t play time”: Analyzing timing profiles in drum performance. In A. Williamon, D. Edwards, & L. Bartel (Eds.), *Proceedings of the International Symposium on Performance Science* (pp. 593–598). European Association of Conservatoires, Utrecht, Netherlands.

Kilchenmann, L., & Senn, O. (2015). Microtiming in swing and funk affects the body movement behavior of music expert listeners. *Frontiers in Psychology*, *6*, 1232.

Lartillot, O., Nymoen, K., Câmara, G. S., & Danielsen, A. (2020). *Precise localization of attack phases through a direct observation of audio waveform*. Manuscript submitted for publication.

Lartillot, O., Toiviainen, P., Eerola, T. (2008). A Matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data analysis, machine learning and applications, studies in classification, data analysis, and knowledge organization* (pp. 261–268). Berlin: Springer-Verlag.

London, J., Nymoen, K., Langerød, M. T., Thompson, M. R., Code, D. L., & Danielsen, A. (2019). A comparison of methods for investigating the perceptual center of musical Sounds. *Attention, Perception and Psychophysics*, *81*, 2088–2101.

Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, *24*, 201–208.

Madison, G., Gouyon, F., Ullén, F., & Hörnström, K. (2011). Modelling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology*, *37*, 1578–1594.

Madison, G., & Sioros, G. (2014). What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it. *Frontiers in Psychology*, *5*, 894.

Monson, I. (1996). *Saying something: Jazz improvisation and interaction*. Chicago, IL: University of Chicago Press.

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, *83*, 405–408.

Nketia, J. H. K. (1974). *The music of Africa*. New York: W. W. Norton.

Palmer, C. (1996). On the assignment of structure in music performance. *Music Perception*, *14*, 23–56.

Polak, R. (2010). Rhythmic feel as meter: Non-isochronous beat subdivision in jembe music from Mali. *Music Theory Online*, *16*. http://www.mtosmt.org/issues/mto.10.16.4/mto.10.16.4.polak.html

Pressing, J. 2002. Black Atlantic rhythm: Its computational and transcultural foundations. *Music Perception*, *19*, 285–310.

Repp, B. H. (1996). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America*, *100*, 3917–3932.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin Review*, *12*, 969–992.

Repp, B. H. (2008). Metrical subdivision results in subjective slowing of the beat. *Music Perception*, *26*, 19–39.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin Review*, *20*, 403–452.

Rose, R. F. (1989). *An analysis of timing in jazz rhythm section performance* (Doctoral dissertation). University of Texas at Austin.

Rossing, T. D., Moore, F. R., & Wheeler, P. A. (2002). *The science of sound* (3rd ed.). San Francisco, CA: Addison Wesley.

Scott, S. K. (1998). The point of P-centres. *Psychological Research*, *61*, 4–11.

Senn, O., Bullerjahn, C., Kilchenmann, L., & Von Georgi, R. (2017). Rhythmic density affects listeners’ emotional response to microtiming. *Frontiers in Psychology*, *8*, 1709.

Senn, O., Kilchenmann, L., Von Georgi, R., & Bullerjahn, C. (2016). The effect of expert performance microtiming on listeners’ experience of groove in swing or funk music. *Frontiers in Psychology*, *7*, 1487.

Seton, J. C. (1989). *A psychophysical investigation of auditory rhythmic beat perception* (Doctoral dissertation). University of York, York, United Kingdom.

Skansaar, J., Laeng, B., & Danielsen, A. (2019). Microtiming and mental effort: Onset asynchronies in musical rhythm modulate pupil size. *Music Perception*, *37*, 111–133.

Tekman, H. G. (2002). Perceptual integration of timing and intensity variations in the perception of musical accents. *Journal of General Psychology*, *129*, 181–191.

Villing, R. (2010). *Hearing the moment: Measures and models of the perceptual centre* (Doctoral dissertation). National University of Ireland, Dublin, Ireland.

Vos, P. G., Mates, J. & Van Kruysbergen, N. W. (1995). The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies. *Quarterly Journal of Experimental Psychology*, *48*, 1024–1040.

Vos, J., & Rasch, R. (1981). The perceptual onset of musical tones. *Perception and Psychophysics*, *29*, 323–335.

Waadeland, C. H. (2006). Strategies in empirical studies of swing groove. *Studio Musicologica Norvegica*, *32*, 169–191.

Wohlschläger, A., & Koch, R. (2000). Synchronization error: An error in time perception. In P. Desain & L. Windsor (Eds.), *Rhythm perception and production* (pp. 115–127). Lisse, The Netherlands: Swets and Zeitlinger.

Wright, M. J. (2008). *The shape of an instant: Measuring and modeling perceptual attack time with probability density functions* (Doctoral dissertation). Stanford University.

Zera, J., & Green, D. M. (1993). Detecting temporal onset and offset asynchrony in multicomponent complexes. *Journal of the Acoustical Society of America*, *93*, 1038–1052

APPENDIX. *Mean Onset Location and Standard Deviation For All Individual Participants Per Condition*

		Pattern 1						Pattern 2					
		Metronome			Instrumental			Metronome			Instrumental		
Pt.		Snare *M* (SD)	Kick *M* (SD)	Hi-hat *M* (SD)	Snare *M* (SD)	Kick *M* (SD)	Hi-hat *M* (SD)	Snare *M (SD)*	Kick *M (SD)*	Hi-hat *M (SD)*	Snare *M (SD)*	Kick *M (SD)*	Hi-hat *M (SD)*
1	*L*	46 (20)	16 (19)	12 (21)	43 (33)	18 (22)	24 (21)	19 (18)	-2 (21)	-11 (27)	3 (26)	-3 (26)	5 (26)
	O	-10 (12)	-15 (13)	-8 (11)	15 (13)	13 (12)	20 (11)	-8 (10)	-20 (10)	-10 (11)	-11 (15)	-22 (15)	-13 (15)
	P	-51 (43)	-41 (40)	-28 (48)	-82 (29)	-81 (29)	-84 (33)	-22 (21)	-21 (20)	-15 (21)	-54 (33)	-48 (29)	-42 (28)
2	*L*	23 (24)	14 (21)	33 (24)	29 (19)	6 (13)	27 (19)	-2 (21)	-16 (19)	3 (23)	-26 (15)	-39 (16)	-22 (17)
	O	-25 (18)	-31 (19)	-27 (17)	-13 (26)	-21 (23)	-17 (23)	-38 (17)	-46 (14)	-40 (15)	-40 (16)	-49 (16)	-42 (17)
	P	-65 (29)	-45 (17)	-32 (20)	-64 (25)	-69 (28)	-52 (21)	-51 (21)	-51 (22)	-39 (25)	-45 (26)	-45 (20)	-28 (23)
3	*L*	19 (16)	-3 (14)	2 (14)	33 (13)	7 (10)	21 (10)	6 (15)	-20 (16)	-16 (14)	14 (22)	-5 (22)	0 (21)
	O	-14 (9)	-19 (10)	-10 (9)	-11 (14)	-17 (14)	-6 (13)	-16 (8)	-20 (9)	-11 (9)	-5 (12)	-9 (13)	0 (13)
	P	-63 (13)	-35 (9)	-24 (11)	-55 (20)	-34 (23)	-14 (24)	-38 (16)	-21 (17)	-1 (21)	-12 (14)	-3 (13)	19 (18)
4	*L*	10 (23)	1 (22)	9 (27)	28 (19)	12 (15)	26 (20)	6 (11)	6 (11)	11 (12)	2 (12)	5 (10)	5 (18)
	O	0 (10)	-3 (10)	3 (11)	3 (15)	0 (15)	9 (14)	-10 (11)	-12 (11)	-9 (12)	1 (11)	-1 (13)	4 (13)
	P	-65 (13)	-73 (12)	-65 (14)	-72 (17)	-78 (13)	-70 (17)	-56 (14)	-58 (13)	-53 (14)	-60 (14)	-61 (15)	-59 (21)
5	*L*	9 (7)	-3 (8)	3 (8)	18 (9)	6 (7)	15 (8)	5 (8)	-2 (7)	3 (9)	17 (8)	12 (8)	17 (8)
	O	-9 (8)	-14 (7)	-10 (8)	-15 (23)	-19 (23)	-16 (22)	-9 (8)	-18 (9)	-15 (10)	3 (9)	-3 (8)	-4 (11)
	P	-42 (16)	-20 (15)	-18 (15)	-48 (29)	-36 (29)	-30 (29)	-28 (10)	-28 (11)	-27 (13)	-36 (26)	-42 (24)	-39 (29)
6	*L*	32 (10)	16 (9)	17 (12)	42 (9)	30 (12)	41 (11)	18 (17)	5 (14)	16 (17)	22 (17)	9 (16)	21 (18)
	O	-6 (5)	-6 (8)	-4 (9)	11 (9)	5 (10)	13 (10)	-11 (11)	-15 (11)	-6 (12)	8 (8)	3 (7)	14 (9)
	P	-10 (13)	0 (12)	6 (18)	-45 (27)	-47 (25)	-38 (26)	-58 (22)	-57 (20)	-54 (19)	-68 (26)	-75 (22)	-65 (24)
7	*L*	-1 (19)	-34 (17)	-35 (25)	24 (10)	0 (11)	9 (13)	11 (14)	-29 (13)	-25 (13)	13 (14)	-12 (13)	-12 (15)
	O	-6 (7)	-21 (7)	-22 (12)	-10 (18)	-25 (16)	-21 (18)	-19 (7)	-32 (9)	-34 (15)	-16 (15)	-26 (15)	-30 (18)
	P	-23 (10)	-41 (9)	-39 (12)	-25 (14)	-47 (14)	-43 (15)	-29 (12)	-45 (11)	-45 (14)	-14 (15)	-30 (17)	-37 (20)
8	*L*	21 (11)	1 (9)	8 (22)	17 (8)	5 (9)	10 (20)	19 (12)	9 (12)	19 (15)	23 (8)	15 (10)	-6 (21)
	O	-15 (11)	-18 (11)	-18 (21)	-12 (15)	-17 (15)	-13 (20)	-11 (8)	-13 (11)	-34 (22)	2 (7)	1 (8)	-11 (14)
	P	-68 (25)	-86 (25)	-83 (32)	-53 (19)	-67 (20)	-65 (26)	-52 (22)	-54 (18)	-73 (22)	-65 (22)	-62 (18)	-87 (28)
9	*L*	0 (6)	0 (7)	9 (8)	-4 (9)	-6 (8)	1 (9)	22 (10)	-4 (10)	0 (11)	-3 (16)	-20 (15)	-18 (15)
	O	-10 (7)	-13 (7)	-9 (7)	-28 (14)	-28 (15)	-24 (14)	-11 (7)	-11 (7)	-7 (7)	-21 (11)	-17 (12)	-15 (11)
	P	-32 (18)	-18 (16)	-5 (19)	-66 (16)	-64 (15)	-48 (18)	-42 (13)	-29 (13)	-16 (17)	-71 (15)	-54 (13)	-38 (19)
10	*L*	19 (12)	12 (11)	27 (11)	20 (23)	17 (19)	33 (18)	9 (11)	6 (11)	14 (10)	25 (12)	13 (14)	28 (13)
	O	-15 (11)	-19 (11)	-11 (11)	-1 (11)	-5 (12)	2 (11)	-14 (5)	-20 (7)	-10 (7)	-14 (20)	-23 (23)	-11 (21)
	P	-80 (16)	-79 (14)	-66 (15)	-76 (15)	-80 (16)	-61 (16)	-62 (18)	-68 (19)	-59 (19)	-84 (18)	-88 (17)	-79 (19)
11	*L*	-6 (13)	-11 (16)	1 (16)	18 (11)	8 (12)	20 (12)	5 (12)	1 (13)	9 (11)	-3 (15)	-9 (14)	3 (15)
	O	-28 (15)	-31 (16)	-22 (14)	-24 (22)	-31 (22)	-19 (22)	-35 (12)	-43 (12)	-31 (10)	-39 (14)	-45 (15)	-34 (15)
	P	-41 (19)	-44 (17)	-34 (17)	-33 (28)	-40 (28)	-32 (27)	-50 (15)	-55 (17)	-49 (14)	-61 (19)	-68 (16)	-59 (17)
12	*L*	37 (14)	21 (16)	22 (17)	35 (11)	34 (10)	32 (11)	45 (13)	45 (15)	49 (13)	16 (15)	20 (18)	26 (17)
	O	-7 (11)	-10 (8)	-9 (11)	6 (10)	4 (11)	3 (13)	-11 (8)	-13 (10)	-7 (9)	3 (8)	0 (9)	7 (10)
	P	-86 (23)	-82 (23)	-71 (24)	-56 (36)	-57 (31)	-57 (33)	-56 (18)	-62 (17)	-56 (17)	-91 (29)	-100 (22)	-89 (29)

(continued)

Appendix. *(continued)*

		Pattern 1						Pattern 2					
		Metronome			Instrumental			Metronome			Instrumental		
Pt.		Snare *M* (SD)	Kick *M* (SD)	Hi-hat *M* (SD)	Snare *M* (SD)	Kick *M* (SD)	Hi-hat *M* (SD)	Snare *M (SD)*	Kick *M (SD)*	Hi-hat *M (SD)*	Snare *M (SD)*	Kick *M (SD)*	Hi-hat *M (SD)*
13	*L*	49 (13)	39 (15)	40 (17)	61 (14)	49 (14)	52 (16)	40 (13)	26 (13)	31 (14)	39 (12)	29 (11)	34 (12)
	O	-18 (12)	-23 (12)	-25 (12)	13 (12)	5 (12)	4 (12)	-25 (16)	-30 (17)	-27 (17)	12 (9)	8 (9)	11 (10)
	P	-60 (17)	-54 (20)	-56 (20)	-78 (19)	-83 (26)	-86 (18)	-74 (18)	-78 (17)	-76 (18)	-82 (12)	-87 (10)	-84 (13)
14	*L*	7 (10)	-1 (10)	6 (10)	3 (17)	-9 (17)	0 (17)	0 (13)	-10 (11)	0 (12)	12 (8)	6 (7)	14 (9)
	O	-7 (10)	-7 (7)	-4 (9)	-25 (27)	-29 (27)	-25 (26)	-8 (8)	-11 (8)	-4 (8)	-4 (10)	-6 (9)	0 (9)
	P	-17 (10)	-15 (14)	-7 (12)	-48 (17)	-49 (18)	-44 (18)	-27 (11)	-26 (12)	-21 (11)	-34 (22)	-35 (19)	-30 (20)
15	*L*	23 (8)	19 (9)	27 (9)	40 (9)	29 (9)	43 (8)	15 (9)	5 (9)	17 (10)	15 (8)	8 (9)	18 (9)
	O	-5 (7)	-12 (6)	-4 (7)	4 (5)	6 (7)	8 (6)	-8 (6)	-10 (7)	-5 (6)	4 (6)	1 (6)	7 (5)
	P	-33 (8)	-37 (9)	-25 (11)	-67 (15)	-68 (17)	-58 (17)	-33 (6)	-32 (7)	-26 (8)	-53 (18)	-51 (16)	-47 (18)
16	*L*	52 (27)	17 (32)	71 (30)	46 (25)	32 (17)	60 (23)	45 (15)	41 (20)	51 (18)	48 (20)	42 (15)	56 (19)
	O	-24 (7)	-23 (10)	-18 (11)	-5 (11)	-3 (13)	3 (13)	-19 (10)	-20 (9)	-11 (11)	-12 (10)	-13 (10)	-5 (10)
	P	-91 (18)	-125 (24)	-63 (24)	-95 (17)	-131 (36)	-73 (23)	-104 (18)	-100 (15)	-92 (16)	-115 (16)	-114 (16)	-104 (16)
17	*L*	-4 (9)	-4 (7)	6 (10)	15 (11)	3 (9)	18 (12)	-3 (8)	-11 (8)	-1 (9)	5 (7)	0 (9)	7 (9)
	O	-18 (10)	-24 (10)	-21 (11)	6 (8)	-6 (6)	-2 (12)	-21 (10)	-25 (12)	-19 (10)	-5 (8)	-12 (9)	-4 (10)
	P	-23 (11)	-29 (13)	-22 (11)	-7 (18)	-15 (15)	-9 (17)	-36 (10)	-45 (11)	-36 (11)	-4 (12)	-15 (13)	-4 (13)
18	*L*	-5 (9)	-5 (9)	5 (12)	18 (16)	13 (17)	26 (21)	7 (9)	10 (9)	9 (10)	16 (10)	16 (9)	19 (9)
	O	-13 (9)	-15 (7)	-8 (8)	-3 (9)	-2 (10)	4 (9)	-10 (9)	-10 (8)	-6 (9)	-9 (7)	-5 (8)	-4 (8)
	P	-55 (12)	-56 (13)	-50 (13)	-65 (18)	-66 (20)	-60 (19)	-57 (14)	-57 (13)	-56 (13)	-70 (13)	-67 (11)	-65 (14)
19	*L*	12 (10)	12 (11)	9 (16)	33 (11)	29 (10)	38 (11)	12 (13)	4 (14)	14 (13)	17 (8)	10 (9)	17 (8)
	O	-16 (9)	-16 (10)	-18 (17)	7 (8)	7 (7)	10 (9)	-9 (7)	-13 (6)	-6 (8)	-1 (6)	-5 (6)	-5 (12)
	P	-48 (12)	-50 (13)	-63 (19)	-56 (24)	-63 (20)	-55 (20)	-44 (8)	-51 (9)	-41 (8)	-34 (12)	-35 (13)	-31 (13)
20	*L*	-8 (11)	-12 (9)	-5 (11)	11 (9)	8 (8)	14 (9)	-10 (12)	-14 (12)	-8 (12)	1 (9)	-3 (10)	-1 (13)
	O	-9 (7)	-8 (8)	-6 (8)	6 (5)	6 (7)	12 (7)	-10 (7)	-13 (10)	-10 (9)	-2 (10)	-6 (11)	0 (11)
	P	-29 (7)	-30 (12)	-26 (9)	-58 (35)	-57 (36)	-53 (34)	-26 (6)	-34 (8)	-26 (7)	-20 (14)	-21 (15)	-21 (16)

Note: Pt. = Participant, *L* = Laid-back, *O* = On, *P* = Pushed.

Paper III

Mapping timing and intensity strategies in drum-kit performance of a simple back-beat pattern

Guilherme Schmidt Câmara, Georgios Sioros, and Anne Danielsen

Submitted to *The Journal of New Music Research*

Appendix

Sound examples included in this thesis’ papers can be accessed here:

https://drive.google.com/drive/folders/1v3gykFwnKFz1smzy41olKyQOGtXODYFr?usp=sharing