Music–Dance

Investigating Rhythm Structures in Brazilian Samba and Norwegian Telespringar Performance

Mari Romarheim Haugen

Abstract

This thesis investigates the interaction between perceived rhythm and underlying reference structures in the experience of rhythm in music. The point of departure is that both music performers' and perceivers' body motions are relevant to the study of rhythm. Accordingly, in music genres with an intimate relationship to dance, here referred to as *music–dance*, rhythm needs to be understood in relation to the corresponding dance. The studies included in this thesis investigate rhythm patterns in sound and body motion in two music–dance styles, Brazilian samba and Norwegian telespringar, based on *motion capture* and *sound* recordings of professional musicians and dancers.

Both samba and telespringar consist of complex rhythm patterns. Samba is often characterized by so-called *systematic microtiming* at the sixteenth-note level. This was confirmed in our sound analysis of the samba groove, showing a *medium–medium–medium–long* duration pattern at the sixteenth-note level. In addition, motion analysis of the percussionist's heel tapping and the dancer's steps revealed motion patterns in synchrony with this rhythm pattern.

Telespringar, on the other hand, is often described as featuring a so-called *asymmetrical triple meter*—that is, the three beats in a measure are of uneven duration. According to previous studies, both the fiddler's foot stamping and the dancers' vertical body motions are related to this underlying meter. This relation was confirmed in the motion analysis of a fiddler's foot stamping, which revealed a very stable *long–medium–short* duration pattern at the beat level. The dancers' vertical motion patterns, however, deviated from theories suggesting that the turning points in the dancers' vertical motion curves correspond to the meter. The thesis therefore suggests an alternative interpretation with regard to the dancers' vertical motion curves—that it is the *shape* of the dancers' vertical motion that corresponds to the underlying *beat duration*, rather than the *turning points* that correspond to the underlying *beat positions*.

The main conclusion is that the underlying sixteenth-note level in samba and the underlying beat level in telespringar should not be understood as deviations from an isochronous pulse of some sort. Instead, they should be understood as inherently (and necessarily) non-isochronous, in and of themselves.

Acknowledgements

Many people have contributed to this work in various ways. First of all, a very special thank-you to my three supervisors for their mentorship: Rolf Inge Godøy, for his insightful comments and suggestions, kind support, and encouragement throughout this PhD process; Anne Danielsen, for her thorough and constructive comments on my work, for her inspiring discussions, and for so generously sharing her insights and knowledge of rhythm; and Alexander Refsum Jensenius, for invaluable suggestions and comments on my thesis in the final phase of writing.

I am grateful to all of my good colleagues in the *fourMs* research group at the University of Oslo, for interesting discussions and valuable feedback on my project at various stages. Thanks go to Kristian Nymoen, for his excellent collaboration, and for so generously sharing his knowledge of Matlab and providing several custom-made scripts for use in many of my studies; Anders Tveit, for developing a motion/sound synchronization solution for my motion capture recordings; Kathrine Frey Frøslie, for helping with the statistical analysis; and Arve Voldsund, Minho Song, and Diana Kayser for their assistance during the motion capture recordings. A special thanks to Ragnhild Torvanger Solberg, for great discussions, general support, dancing sessions inside (and outside) the lab, good company on conference trips, "juice," and friendship.

Thanks to my colleagues at the Department of Musicology, for fruitful discussions as well as pleasant lunches and coffee breaks. Thanks to my fellow PhD students Birgitte, Mikkel, Morten, Thomas, Anita, and Erlend, for valuable discussions and feedback on my work, and to postdoctoral fellows Yngvar, Astrid and Peter, for kindness and support. A special thanks to fellow PhD student and "office mate" Kai Arne Hansen (BKKF) for support, friendship, good conversations, laughs, and over-priced chocolate.

Thanks also to colleagues abroad for many inspiring discussion during conferences, workshops, and summer schools.

Thanks to my excellent copyeditor, Nils Nadeau, who improved my English text considerably.

I am also very grateful to all of the musicians and dancers who participated in the studies. This work would not have been possible without you.

Finally, thanks go to my dear friends and family. Thanks to my supportive parents, who have filled my life with music from the very beginning; to Mats, for inspiring discussions on music; and my beautiful niece Selma, for cheering for me. A special thanks to my dear sister, Ingrid, for many fruitful discussions, for her invaluable support, and most of all for always being there for me.

<div align="right">

Mari Romarheim Haugen
April, 2016

</div>

Abbreviations

Some of the abbreviations used in in this thesis:

3D   Three-dimensional.

ANOVA   Analysis of variance.

IOI   Interonset interval.

IR   Infrared.

Mocap   Motion capture.

PD   Participatory discrepancies.

SD   Standard deviation.

QTM   Qualisys Track Manager. The motion capture software accompanying the Qualisys motion capture system.

Contents

Chapter 1

Introduction

This chapter starts by presenting some initial thoughts, followed by the research questions and hypotheses. I then discuss the theoretical framework and empirical material, followed by an outline of the thesis.

1.1 Background

The experience of rhythm is intrinsically related to human body motion. This is something that I have experienced myself on several occasions, as a performer, perceiver, and music teacher. In my youth I went to a local folk music gathering in Telemark in Norway and had the privilege of dancing with one of the greatest telespringar dancers in the country. At the time I was only somewhat familiar with telespringar music, and I had not danced it before myself. Because telespringar is a couple's dance, I could follow the motion of my partner. While dancing I suddenly had another experience of the rhythm of the music, and even after I was through, I perceived the rhythm of the music differently. It was like my "feeling" of the rhythm in telespringar had changed as a result of my embodied experience of this music through dance.

Several years later I had a similar experience with samba. As part of my bachelor's-degree education at the Royal Academy of Music in Silkeborg (now in Aarhus) in Denmark, we went on a study trip to Salvador da Bahia in Brazil to study Brazilian drumming and dancing. I knew the steps of samba, but when I watched the dancing people of Salvador, I felt that my way of dancing was not exactly right, and this was also pointed out by some locals at a party: "You are jumping too much." However, after a couple of days, during a *samba-de-rhoda* session, it was like my body suddenly "got" it. It felt like it was all falling into place—the music, the dancing, the setting. As with telespringar, my experience of the rhythm of samba had changed even when I was not dancing.

Brazilian samba and Norwegian telespringar are two very different styles of music. However, the experience of "getting the rhythm" through dancing felt very much the same. These experiences, among others, made me wonder about the relationship between music and body motion in this regard. Maybe some music-related motion patterns are critical to understanding underlying structures in the music? And, if such motion patterns are present in the dances of samba and telespringar, maybe the underlying reference structures in these two styles could be investigated by studying the dance motion and musical sound together?

1.2   Rhythm and Body Motion in Samba and Telespringar

The intimate relationship between music and dance in Brazilian samba and Norwegian tele-springar (in the rest of this thesis I will refer to them simply as "samba" and "telespringar") is often highlighted in the two styles' respective literatures. Both samba and telespringar derive from oral traditions, and it is commonly understood that the music and the dance have developed together under mutual influence. Rhythm studies of these genres also emphasize the intimate relationship between music and dance (see, for example, Blom, 1981; Naveda, 2011).

Another feature that samba and telespringar share is a complex rhythm pattern. Samba is normally notated in binary meter, and recent studies have noted that samba seems to be characterized by so-called *systematic microtiming* at the level of the sixteenth note. This means that the intervals between the sixteenth notes within a beat are of uneven duration. Telespringar is normally notated in triple meter but is recognized as featuring a so-called *asymmetrical meter*, meaning that the three beats in a measure are of uneven duration. It is commonly understood that the experience of rhythm in music involves an interaction between the heard rhythm (in the actual sound) and an underlying reference structure (a mental construct in the performer/perceiver) against which the rhythm is perceived. The *pulse* represents the fundamental reference level for the production and perception of rhythm, and the *meter* is a means of organization of the pulse (London, 2012). Underlying reference structures like pulse and meter are generally thought to consist of a series of beats separated by identical intervals (Lerdahl and Jackendoff, 1983), or what is known as *isochronous meter*. However, it has been pointed out that this may not be a universal rule that applies to all styles of music (Kvifte, 2007)—in samba and telespringar, for example, it may be not only the sonic rhythm patterns but also the underlying reference structures that are non-isochronous.

In the emerging field of embodied music cognition, it has been suggested that underlying structures may also correspond to body motions such as foot stamping, upper-body swaying, and head nodding in both the production and the perception of music (Leman and Godøy, 2010). In this thesis, I have therefore approached the complex rhythm structures in samba and telespringar through both sound and motion analyses.

1.3   Research Questions and Hypotheses

The main goal of this thesis is a better understanding of the correspondence between rhythm and motion in samba and telespringar. The main research question is as follows:

> *What are the relationships between musical rhythm and human body motion in the production and perception of samba and telespringar?*

My approach to answering this question begins with my focus on underlying reference structures in the experience of musical rhythm. My assumption is that such structures are closely related to body motion and, in samba and telespringar, also intimately related to motion patterns in the corresponding dances. This suggests several sub-questions:

> *How can we identify underlying reference structures in samba and telespringar music*

and dance?

To what extent are the underlying reference structures in samba and telespringar present in the corresponding dances?

And, because I want to approach these questions empirically, it has also been relevant to pose a methodological research question:

Which methods might be used to analyze rhythm and motion correspondences in samba and telespringar?

Following from these research questions, I have developed a set of hypotheses that are linked to the following three topics: (1) rhythm; (2) telespringar and samba; and (3) methodology:

1. Rhythm:

   The underlying reference structure(s) in the experience of rhythm is intimately related to motion.

   - The underlying structure is shared among performers and perceivers within a music culture.
   - The underlying reference structure(s) can consist of non-isochronous time series.
   - The underlying reference structure(s) comprise the measure level and subdivision levels, as well as the pulse level.

2. Telespringar and samba:

   The underlying beat level in telespringar and the underlying sixteenth-note level in samba are non-isochronous.

   - The underlying reference structures in samba and telespringar are intimately related to performers' body motion.

3. Method:

   Motion capture–based analysis can provide valuable insight into the relationship between rhythm and motion in music.

   - It is possible to carry out laboratory-based motion capture studies of telespringar and samba.

1.4   Theoretical Framework

I see my work as highly interdisciplinary, combining theories and methods from the fields of music cognition, rhythm studies, and ethnomusicology. Central to it has been my access to the fourMs motion capture lab at the Department of Musicology, University of Oslo, which has allowed me to carry out empirical studies using state-of-the-art motion capture technology. My research topic concerns the intimate relationship between music and motion, as addressed by

the emerging field of *embodied music cognition* (Leman, 2008). A point of departure here is that human perception is *multimodal*—that is, all of our senses are constantly in use and work together in different ways as we experience phenomena in the world. An *ecological perspective on perception* proposes that we learn about the world by constantly interacting with it (Gibson, 1986), using all of our senses simultaneously. For example, when we drink a cup of coffee, we smell the coffee, we see its colour, we feel the weight of the cup as we lift it and the heat of the coffee as we drink it, and we hear the sound when we put the cup back on the table. Each time we drink a cup of coffee, these various sensory relationships are reinforced, and such multimodal knowledge in turn informs our perception of future cups of coffee, and of everything else in the world.

One such multimodal paradigm is that of the *action–sound coupling* (Jensenius, 2007)—that is, the fact that sounds in nature derive from the mechanical and acoustical properties of the actions and objects involved. It has been suggested that when we perceive a sound, we simultaneously mentally simulate the action that we imagine produced that sound (see Berthoz, 2000). In music cognition, such mental simulations might be related to a musical instrument, and to a performer's actual sound-producing actions. A *motor-mimetic* perspective on music cognition suggests that our sound perception also incorporates *sonic shapes* that indicate, for example, how an impulsive stroke on a drum differs from a bowed sound on a fiddle, and that these shapes can be rendered in accordance with action–sound couplings (Godøy, 2003). All of the perspectives described above suggest that music is more than a purely sonic phenomenon and in fact engages other senses in its perception.

This thesis is also related to the field of *rhythm production and perception*. As already mentioned, the experience of rhythm in music with a meter also includes an underlying reference structure. Such structures are mental constructions in performers and perceivers that assign regularity to the music and represent the framework against which we perceive sonic events (e.g., Danielsen, 2010b). I have also been influenced by rhythm studies carried out within the field of *ethnomusicology*. Here, a basic assumption has long been that there is a tacit perceived pulse—a "subjective beat" (Chernoff, 1979), "regulative beat" (Nketia, 1986), or "inner pulsation" (Kubik, 1990)—underlying the structure of the rhythm. Moreover, ethnomusicological studies highlight the fact that musicians' and dancers' body motions should be included in rhythm studies (Baily, 1985). Particularly important to this thesis is the fact that music in some music cultures does not exist apart from dancing (e.g., Grau, 1983).

This thesis can also be aligned with *empirical musicology* (Clarke and Cook, 2004), with a particular focus on observation studies and experiments. My research here has been largely based on observation studies using advanced motion capture technologies, combined with the statistical analysis of recorded sound and motion. Even though motion capture has been used in an increasing number of music studies in recent years, there are not many examples of recording sessions involving both musicians and dancers. Hence an important aspect of this project has been its interest in investigating how motion capture technologies can be used successfully to analyze rhythm structures in samba and telespringar performances.

This thesis is to a great extent focused on the two terms "music" and "motion," and on the different ways in which they are related. Though the word constellation *music and motion* implies that the "motion" is separate from the "music," I will use it for convenience's sake. Likewise, I will use the term *musical sound* when describing the purely sonic part of the music,

and *music-related motion* when describing different types of body motion that are carried out specifically in music performance and perception. For example, in a telespringar performance the musical sound is the sound from the fiddle, whereas the music-related motion includes the fiddler's sound-producing actions and foot stamping as well as the dancer's body motion.

1.5 Outline of Thesis

The thesis consists of two parts: six chapters framing the project, followed by five research papers that have been published or submitted. Figure 1.1 supplies a visual overview of the thesis and the relationships of the different parts to the three main areas of contribution. In what follows I will briefly describe the contents of the chapters.

- Chapter 2 supplies an overview of relevant topics within the field of music cognition, with a special focus on music and motion correspondences. The *sound–source* relationship and a *motor-mimetic* perspective on music perception and cognition are outlined. The importance of including dance in rhythm studies of music with an intimate relationship to dance is highlighted, as is the relevance of music culture to music perception.

- Chapter 3 is specifically concerned with rhythm. The interaction between perceived sonic rhythms and underlying reference structures is discussed in relation to certain selected levels of the underlying structure: *pulse*, *meter*, and *subdivision*. I argue that underlying reference structures can be both isochronous and non-isochronous and label the underlying subdivision level the *metrical subdivision*. I then propose the term *non-isochronous metrical subdivision* for an underlying subdivision consisting of non-isochronous beats.

- Chapter 4 gives an overview of the methods used in the thesis. First, technologies related to tracking and storing motion data, data-processing methods, and motion and sound analysis are presented. Next, the methods for analyzing and visualizing motion and sound data are discussed. The importance of combining quantitative and qualitative methods in studies of music and motion is also highlighted.

- In chapter 5, I introduce the five papers included in the thesis, including abstracts. I discuss the findings of the papers in relation to the three main areas of contribution of the thesis: rhythm, samba/telespringar, and method.

- Chapter 6 discusses the work presented in this thesis. It includes a summary of the themes dealt with in the thesis, a discussion of its findings and conclusions, and some directions for future work.

Part 2 includes the five research papers:

- Paper I has a methodological focus and investigates how infrared motion capture systems can be used in rhythm studies of telespringar.

- Paper II evaluates different accelerometer devices for dance research.

- Paper III investigates the existence of asymmetrical beat patterns in music and body motion in telespringar, based on a motion capture study.

- Paper IV investigates periodic body motions in a telespringar performance and their relationship to the underlying reference structures of the music, pointing to the existence of such structures in the form of shared tacit knowledge among the performers.

- Paper V investigates whether systematic microtiming patterns previously found in the sound of samba may also be present in samba performers' body motion, based on a motion capture study of a percussionist and a dancer.

Chapters: **Topics:**   **Papers:**

Chapter 1:
Introduction

Chapter 2:
Music Cognition

Chapter 3:
Rhythm

Chapter 4:
Method

Chapter 5:
Research Summary

Chapter 6:
General Discussion

Rhythm

Samba

Telespringar

Method

Paper V

Paper IV

Paper III

Paper II
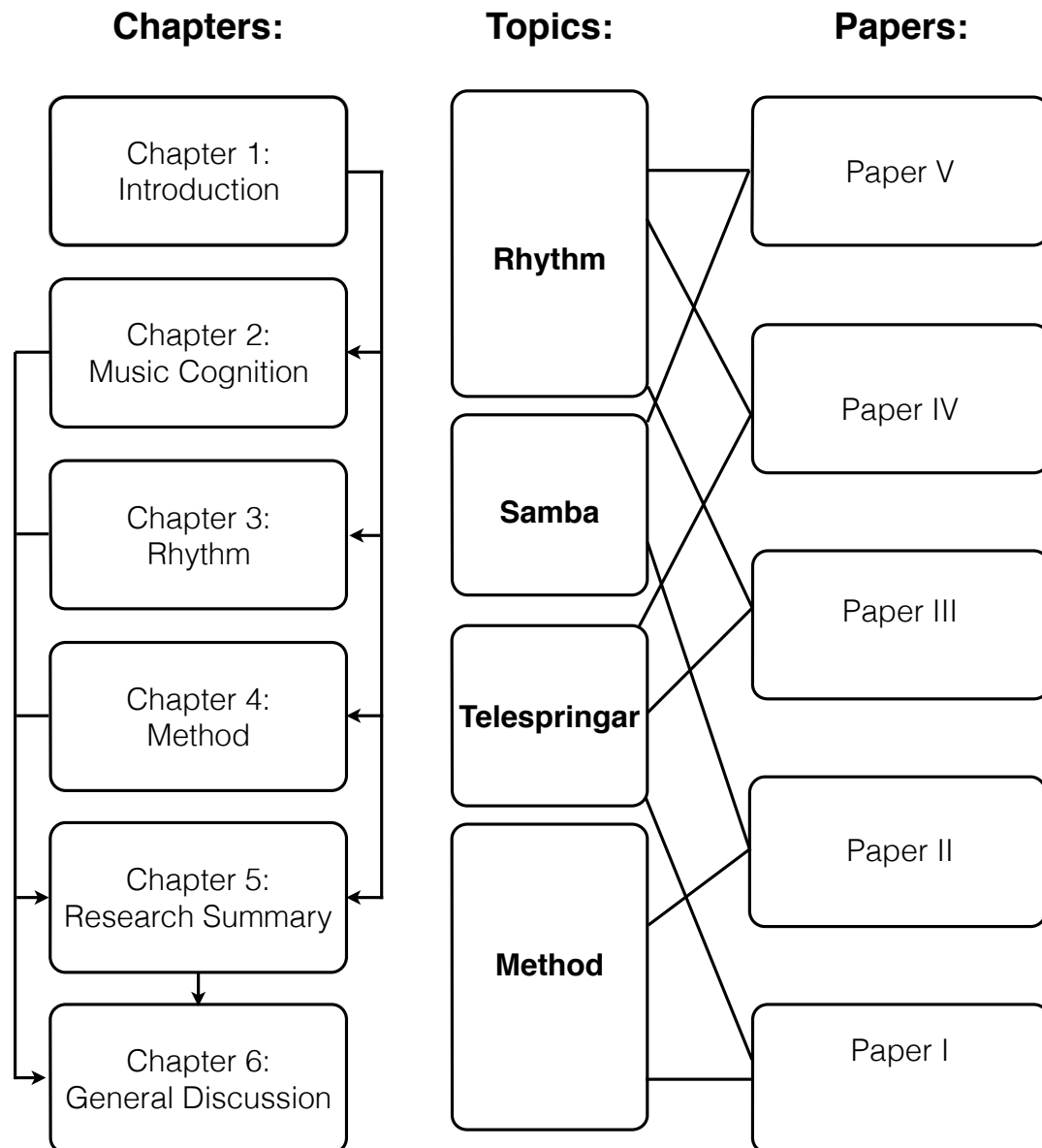
Paper I

Figure 1.1: A visual illustration of the thesis. The first part of the thesis consists of six chapters. The second part consists of five research papers. The illustration shows how the papers relate to the three main topic areas of the thesis: rhythm, telespringar/samba, and method.

Chapter 2

Music Cognition

> Music may be as much a motor event as a
> sonic event, as well as, of course, a social
> fact.
> —John Baily, 1985, p. 258.

This chapter investigates music and motion correspondences in music with an intimate relationship to dance. I present relevant music/motion factors in musical performance and the nature of their interaction, and I consider the relevance of music culture to music perception.

2.1 Introduction

Human cognition and perception are complex processes that encompass constant interactions between memory, attention, expectation, information that arrives via our senses, filtering of this information, and so on (Sethares, 2007). In other words, when we perceive the world, we do so in the context of our previous experiences (Snyder, 2000). One theoretical starting point in this regard is that we obtain knowledge about the world by constantly interacting with it with our bodies (Gibson, 1986). For example, scholars have pointed out that when we perceive a sound, we simultaneously perceive the action that we believe caused the sound (that is, the sound's source) (see, for example, Berthoz, 2000; Godøy, 2003). Hence, music is not only a sonic phenomenon but also a multimodal one that encompasses both sound and motion (see, for example, Godøy and Leman, 2010).

Music cognition also depends on the relevant *music culture* (Leman, 2008)—that is, when a group of people shares the same musical experiences, conceptions, ideals, and norms. I use *music culture* instead of *culture* in this study to emphasize that, in a musical context, our experiences—and what kind of music we are most exposed to—may be more relevant than the geographical area. For example, telespringar may be part of the larger music culture of traditional Norwegian dance music, but not all Norwegians are familiar with this kind of music, and certainly most of us are exposed to other music styles in our daily lives.

This thesis is concerned with music cultures in which music and dance are intrinsically related—that is, music styles associated with (and mutually influential upon) specific dances. Samba and telespringar are both related to specific dances and often characterized by their com-

plex rhythmic patterns, which have been linked directly to the corresponding dances (see, for example, Blom, 1981; Naveda, 2011). Investigations into these kinds of music styles may offer valuable insight into people's music perception and the relationship between rhythm, motion, and music culture in particular. I will pursue these ends from the viewpoint of *embodied music cognition* (Leman, 2008), which holds that the body's interactions with the world is inherent in cognitive processes.

2.2   Music–Dance

In some music cultures, interestingly, there is no word for "music" that does not also encompass dance, playing, singing, and the whole social event that situates them. When staying in Tanzania, Danish music teacher Steen Nielsen sought the word for "music" in Swahili, and the closest he got was *ngoma*, which primarily means "drum," "dance," and "play/party," and *kucheza*, which also includes "playing" (music, theatre, football, cards, and so on), and "dancing" (Nielsen, 1985, p. 35). Bjørkvold (1999) uses the term *ngoma* specifically for children's music culture, which often includes "playing" (both playing music and playing in general). Other examples of multimodal musical terms include *yoi*, used by the Tiwi people in northern Australia to encompass the dance, songs, and rhythms of a musical event (Grau, 1983), and *egwu* in the Igbo language of Nigeria, which comprises music, song, dance, and drama (Baily, 1985). In Danish music pedagogy, the term *kucheza* is used, as well as *Sang, Dans og Spil* [singing, dancing, and playing], or SDS, which was coined by Danish music pedagogues in the late 1980's.

While many of these terms include music with an intimate relationship to dance, they all account for the event as well. In this thesis, though, I do not attempt to include the event but instead privilege the direct music/dance relation and will refer to this particular kind of music as *music–dance* in this thesis. Music–dance not only refers to musical styles where music is *only* performed with the corresponding dance but also to musical styles where the rhythm should nevertheless be understood in relation to the corresponding dance, as is the case with samba and telespringar.

2.2.1   Music—Dance Performance

The work included in this thesis is based on investigations of *music–dance performances*—that is, recordings that include both music and dance. There are various factors in play in such a music–dance performance. First, we have the performers—the musicians producing *musical sound*, and the dancers performing *dance motion*. In some music–dance styles, the dancers may also perform sound-producing actions (for example, hand claps), but this is not relevant to the performances investigated in this thesis. Second, we have the musician's body motion—actions directly related to *sound production*, such as the bowing motion in fiddle playing, and other actions, such as upper-body swaying. Music and motion correspondences can be investigated based on the sounds and the performers' body motions in such music–dance performances. Third, we have the *interaction* between the musicians and the dancers in relation to the music culture in question. Because the way of playing and the way of dancing are rooted in the music culture, the interaction between the musicians and the dancers may also reflect shared
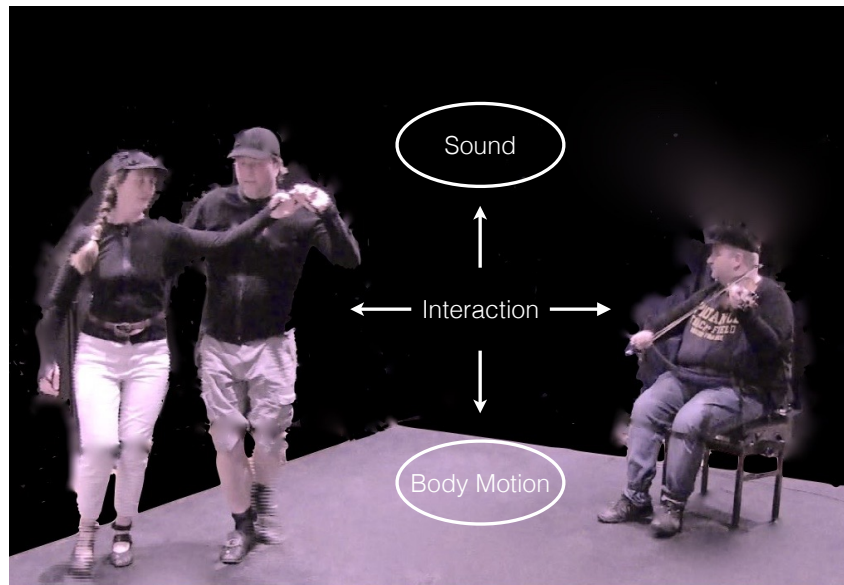
Figure 2.1: An illustration of relevant factors in play in a music–dance performance. There are interactions between musical sound and music-related body motions, and between the performers.

knowledge about the execution of the music–dance. The ways of playing or dancing, in other words, are never random, because both musicians and dancers must agree that their practice falls within the musical style. Both previous experience and tradition—that is, the history behind the musical style's performance as music–dance—inform this *common understanding* about how this musical style is usually performed.

In a music–dance performance, there are interactions between musical sound and music-related body motions, and there is an interaction between the performers—one that relates to both the produced and perceived rhythm in the actual performance (shared timing) and the common understanding (intention) derived from the shared experiential knowledge of the music culture. Figure 2.1 illustrates these elements and their interactions.

2.3 Physical Sound, Sound Perception, and Music Cognition

It is important to differentiate between the *physical sound* and the *perceived sound* (Bengtsson, 1973; Bregman, 1990). The physical sound exists in the world as a pressure wave that propagates though the air, and it can be *measured*, whereas the perceived sound refers to how we hear the physical sound. In addition, sound perception and cognition includes more than "hearing." Sethares (2007) differentiates between the *signal*, the *perceptual apparatus*, and the *cognitive* and/or *cultural framework*, noting that a computer can only recognize features that are "in the signal" (the physical sound), not those that are the result of people's perception and cognition. Sethares points out that musical concepts such as notes, beats, melodies, rhythms, and meter result from human perception and cognition and are not properties of the sonic signal (Sethares, 2007, p. 14).

It has also been pointed out that our perception of the world includes the simulation of motion. An *embodied perspective on cognition* suggests that our cognitive capacities are continually shaped by embodied interactions with the environment, meaning that *perception* and *action*

are understood as mutually dependent (see, for example, Clayton and Leante, 2013; Leman, 2008). Within an embodied framework, music-related sound and music-related motion are intimately linked (Godøy, 2010; Haga, 2008; Jensenius, 2007; Leman, 2008). Furthermore, such music-related motion can be either physical (sound-producing motion, body swaying, dance) or imagined (Leman and Godøy, 2010).

The terms *perception* and *cognition* are sometimes used interchangeably in the literature, but often perception is aligned with sensing and cognition with the processing of what is perceived. Here, *perception* refers to the "perception of something external," in the context of one's previous experiences, whereas *cognition* encompasses processes where nothing is perceived, such as when something is imagined.

2.4   Music and Motion

As we recall, *embodied music cognition* asserts that cognitive processes include the body's interaction with the world. This follows from an *ecological perspective on perception*, proposed by the American psychologist James J. Gibson (1904–1979), stating that we learn about the world by continuously interacting with it (Gibson, 1986). Gibson points out that when we perceive an object, we simultaneously perceive the action that we relate to that object. He coined the term *affordances* for the actions that people assign to the objects that they perceive (Gibson, 1977, 1986). An object may have multiple affordances—for example, spoons afford eating, but they also afford playing. In order to relate spoons to either affordance, one draws on one's experience; our knowledge about the possible relationships between objects and their affordances derives from our daily interactions with these objects in the world.

One obvious way in which sound and motion correspond in a musical context is through the musicians' *sound-producing actions* (Jensenius, 2007). For example, when a percussionist hits a *pandeiro* (a Brazilian hand frame drum with jingles) with his/her hand, a sound will be produced. In other words, there is and an *action-sound* relationship between the instrument, the hand and the performed action.

As already mentioned, it has been suggested that the *sound-source* relationship is not only relevant for sound-producing musicians but for *sound perception* in general. The so-called *motor theories of perception* suggest that when we perceive a sound, we simultaneously perceive the *source* of the sound, including a simulation of the action that *produced the sound* (see, for example, Berthoz, 2000; Godøy, 2010; Liberman and Mattingly, 1985). Accordingly, sound perception is not only a matter of feature extraction based on the sound signal but also of the process of aligning the sound to existing knowledge of sound-source relationships (Leman et al., 2008). Hence, sound perception depends on previous experiences related to how that sound is produced. Influenced by the ideas of Gibson (1986), Shove and Repp (1995) have proposed an "ecological level" of sound perception, whereby the environmental objects involved in the event are perceived directly, so that it is not only the *sound* of a bowing violinist that is heard, but a *violinist bowing* (Shove and Repp, 1995, p. 59). Even when perceivers lack direct experience with the sound-producing instrument in question, they may have other connections regarding how the sound in question is produced. Cox (2006) refers to the *mimetic hypothesis* , which observes that playing air-guitar, conducting, or singing are common to many, even when the actual

ability to play, conduct, or sing is reserved to relatively few. An air-piano study, including persons with different levels of expertise, showed clear correspondences between the participants' imitative behaviour and the sound-producing actions assumed necessary to produce the musical sound, but not surprisingly, experts were more detailed in their imitative rendering (Godøy et al., 2006).

2.4.1 Mirror Neurons

The view proposed by motor theories of perception that perception is simulated action is also supported by the discovery of so-called *mirror neurons*. Gallese et al. (1996) found that a set of neurons became active in a monkey's brain both when the monkey performed an action and when it observed the same action being carried out. Later, it was determined that a particular class of mirror neurons, called *audio-visual mirror neurons*, was activated not only when the monkey carried out or observed a sound-producing action but also when it only heard the sound related to the action (Keysers et al., 2003; Kohler et al., 2002). People do the same thing. For example, Haueisen and Knösche (2001) found that when expert pianists listened to piano music, motor-related areas of the brain associated with piano playing were activated. This suggests that pianists mentally simulate piano playing when they are listening to piano music—that is, the body motions related to the sound production are an integral part of the perception of the music. Action simulation is restricted to motions that are biologically possible—that is, imitable (Wilson and Knoblich, 2005)—which accords with the ecological perspective on perception and its claim that our experiences influence how we perceive the world.

2.4.2 Sonic Objects and Gestural Affordances

It has been suggested that the sound-source relationship in the experience of music need not include an image of an actual musical instrument. Godøy (2010) suggests that our capacity for source recognition in music perception could be termed *ecological knowledge*, and that such knowledge is acquired through the accumulation of sound-source experience. Consequently, we have a large repertoire of *images* of sound-producing actions that are evoked when perceiving musical sound, even when we cannot see the musician. Godøy notes that such sound-producing images can be categorized in relation to the *shape* of the sound that they produce, along the lines of Schaeffer's terminology for describing sonic objects according to their overall envelope of duration (Godøy, 2003, 2006). These categories can be used to describe three different *action-sound types*:

1. *Impulsive*, for example, hitting a drum or a key on a keyboard, resulting in a rapid sonic attack,

2. *Sustained*, as in bowed string instruments, resulting in a continuously changing sound, and

3. *Iterative*, a series of rapid and discontinuous motions, as in guiro playing, resulting in a series of successive attacks that tend to fuse together.

These categories enable us not only to recognize similar sound-producing actions in different contexts but also to perceive similarities between sound-producing motion and other types of motion (Godøy, 2010). Godøy (2003) further observes "a *motor-mimetic* element in music perception and cognition, meaning that we mentally imitate sound-producing actions when we listen attentively to music, or that we may imagine actively tracing or drawing the contours of the music as it unfolds" (Godøy, 2003, p. 318). A motor-mimetic perspective on music highlights the intimate relationship between music and motion and captures the fact that simulated sound-producing actions can be both directly related to playing an instrument and imitative of a sonic *shape* that can be gesturally rendered. Godøy also points out that the relation between the simulated sound-producing action and the musical sound may extend across more complex musical phrases and textures. Several free dance studies support this claim, demonstrating that dancers seem to follow salient events in the sound with their body motions (see, for example, Godøy, 2009; Haga, 2008). In some cases, Godøy claims, the reverse may also be the case—that is, "motor-mimesis can translate from visual images to sound by re-tracing the visual contours as sound-producing actions, 'sonorizing' visual images" (Godøy, 2003, p. 319). Haslinger et al. (2005) found that the observation of silent piano playing (meaningful "sound-producing" actions) activated auditory areas in the brains of pianists, implying that this relation works both ways: sound to motion and motion to sound.

2.4.3   Periodic Non-Sound-Producing Body Motion

As mentioned at the beginning of this chapter, a music–dance performance also includes body motion that is non sound producing, such as the dance motions and the musicians' foot stamping or upper-body swaying. These motions have also been called ancillary, sound accompanying, or sound supporting (see, for example, Jensenius, 2007; Van Dyck et al., 2013), but I find those terms slightly misleading, because in music–dance, many of those motions do not *support* the sound production but instead supply the rhythm in the music. Therefore I prefer the term non-sound-producing to label such motion.

Specific music-related body motions, such as foot tapping, head nodding, and upper-body swaying, are often synchronized with a periodic underlying feature of the music. In music–dance styles, many non-sound-producing actions are periodic and as important to the musical rhythm as the actual sonic rhythm (see, for example, Bengtsson, 1974; Blom, 2006; Grau, 1983; Kubik, 1990). How such periodic underlying features are perceived, how they relate to body motions, and how several people might relate to the same periodic structure in music have all been investigated using the theory of *entrainment*, to which I will turn now.

2.5   Entrainment

Entrainment theory refers to the process whereby two or more independent rhythmic processes interact and even synchronize. Clayton (2012, 2013) differentiates between three levels of entrainment in a musical context:

1. *Intra-individual* entrainment, or that which takes place within a person.

2. *Inter-individual* entrainment, or that which takes place between individuals in a group.

3. *Inter-group* entrainment, or that which takes place between different groups. (This level of entrainment is not relevant to the present study.)

The *dynamic attending theory* is concerned with perceivers' ability to entrain to an *external rhythm*. It assumes the existence of *internal oscillations* in a perceiver, known as *attending rhythms*, that can become entrained to regularities in the environment (Large and Jones, 1999). At an *intra-individual* level, dynamic attending theory has been used to explain the human ability to perceive a regular metrical level in music (see also section 3.3).

Clayton (2012) notes that the three levels of musical entrainment build upon one another. According to dynamic attending theory, we can perceive meter and coordinate our actions in relation to this structure at an inter-individual level—that is, we can entrain these actions as part of a group (Clayton, 2012, p. 51), when one person's attentional rhythms become entrained to another person's actions, and these rhythms then influence the attending individual's further behaviour. In a musical context, one individual's actions (music playing) therefore can become coordinated with the actions of another (Clayton, 2013, p. 26).

Although it has been pointed out that our ability to entrain to an external rhythm might be innate, the perception of an underlying reference structure in music—for example, meter—is also highly dependent on the *music culture* (Clayton, 2013). As pointed out above, our perception is determined by our previous experience, and by repeated regularities in our environment. For example, did Hannon and Trehub (2005) find culture-specific musical biases between adults from Bulgaria and Macedonia (who are exposed to music with a non-isochronous meter) and adults from North America (who are exposed to music with an isochronous meter). This suggests that meter perception depends on one's familiarity with the specific music culture.

2.6 Music Culture

Music culture can be defined as what arises when multiple people share a repertoire of musical concepts and practices (Baily, 1985; Blacking, 1995; Clayton et al., 2013; Snyder, 2000). In this sense, a specific music style can represent a music culture. We can talk about Afro-Brazilian samba as a music culture and encompass within it everything that makes people who are familiar with samba music recognize the given music as samba, including specific rhythmic patterns, typical phrasings, samba-specific contexts for performance, signature motion patterns, specific related dances, and so on. In addition, we can talk about sub-styles of a music culture, such as *samba de roda*. Telespringar could also be categorized as a sub-style of traditional Scandinavian dance music. We can even talk about *music–dance* as a music culture, meaning all musical styles in which music and dance are intrinsically related.

In many music–dance styles, underlying reference structures such as pulse and meter are not necessarily represented by the actual sonic events. Agawu (2003) describes how the underlying reference structure in many West and Central African dances is indicated by recurring rhythmic patterns that do not follow the underlying pulse, which is only visible in the corresponding dance (Agawu, 2003, p. 73). Similarly, Kubik (1990) investigated Brazilian drum patterns and found that the underlying pulse was not in the sound but in the musicians' and dancers' body motion. Blom points out that the underlying meter in telespringar consists of non-isochronous

sequences, and that this underlying structure should be understood in relation to the dancers' body motion and the musicians' foot stamping (Blom, 2006).

In his work on Swedish polska, as well, Kaminsky (2014) proposes that *iterative patterning*, or pattern repetition, is one of the mechanisms that prompt motion. He explains that iterative patterning can operate on two levels, *overt* and *submerged pattern implication*. In the former, the repetitive motion of dancers would align with an actual repeating sound, whereas in the latter, the regular beat is not in the sound but is "implied and understood based on socially learned cues" (Kaminsky, 2014, p. 52). He uses the example of the beginning of "Stir It Up" by Bob Marley, where a perceiver familiar with reggae would recognize the meter despite (or because of) the off-beat guitar riff. In Swedish polska, on the other hand, the metrical "cues" are not in a specific sonic rhythm:

> The same mechanisms of cultural learning that would allow a reggae fan, for instance, to nod her head on the beat simply from hearing the off-beat skank in the beginning of "Stir It Up" allows a seasoned polska dancer to hear the beginning of a tune in a given regional style and dance the appropriate dance, whether or not the music opens with overt pattern repetition. The only difference is that the first example can be explained as simple beat induction, while the second requires understanding of a more complex implied pattern. (Kaminsky, 2014, pp. 52–53)

In music–dance styles, the specific way of playing and the specific way of dancing are not random but incorporated in and integral to the type of music. Hence, the music–dance performance not only depends on the interaction between *external rhythms* in the sound or in the performers' body motion in a specific performance but also relates to underlying concepts in a specific music culture. Music–dance performances rely upon the interactions between the musicians and the dancers, and the sound and the motion, in the actual performance, but they also rely upon shared conceptions grounded in the specific *music culture*.

Clayton (2013) points out that while dynamic models can say something about the interactions in a specific performance, ethnographic models can say something about what is *intended* by the performers in the musical performance. He argues that studies of musical interaction and coordination could benefit from an interdisciplinary approach that combined these perspectives.

2.7   Gestural Renderings in Music–Dance

According to Godøy's (2003) *motor-mimetic* perspective on music, there is a relationship between simulated sound-producing actions and musical sound, and, in some situations, between simulated musical sound and visual sound-producing actions. As previously mentioned, Haslinger et al. (2005) found that the observation of silent but meaningful "sound-producing" actions activated auditory areas in the performers' brains. In music cultures with an intimate relation to dance, such meaningful silent actions may also include dance motion. Because the music and the dance have evolved together through mutual interaction, the musician's mental images may derive from both experienced sound-producing actions and images of motion patterns in the corresponding dance. Within a motor-mimetic framework, then, images of the motion patterns in the dance—the actual motion or the *shape* of the motion—may inform the

musician's playing, even when dancers are not present. In this regard, the musician does not have to have direct experience with the actual dance, but he or she must have some sense of the underlying *shape* of the dance.

In terms of the relationship between musicians and dancers in traditional Norwegian dance music, Blom (1993) points out that both groups build upon a common understanding of the musical rhythm. The musicians have to *shape* the music in such a manner that it corresponds to the motion in the dance (Blom, 1993). Because these music styles are based on oral traditions, this knowledge might be part of what musicians acquire when they learn to play these music styles.

2.8 Summary

This chapter addressed various music and motion correspondences, beginning with the relationship between sound-producing actions and sound, which seems to be relevant to sound-producing musicians and perceivers. It was pointed out that the sound-source relationship need not involve an image of an actual instrument but could instead be based on sound and motion *shapes*.

Two levels of entrainment in a musical context were discussed: *intra-individual* and *inter-individual*. The importance of acknowledging conceptions from music cultures was highlighted. Finally, it was suggested that in music–dance styles, the musicians might have a mental image of the dance while playing. The musicians and dancers may share an understanding of the underlying reference structures through commonly shaped mental images.

Chapter 3

Rhythm

> All of the rhythms that we perceive are
> rhythms which originally resulted from
> human activity
> —Paul Fraisse 1982, p. 150.

This chapter gives an overview of some central rhythm concepts. It discusses the relationship between perceived sonic rhythm and underlying reference structures on several metrical levels, including pulse, meter, and subdivision.

3.1 Introduction

As previously mentioned, the focus in this thesis is on correspondences between musical rhythm and body motion. The aim is to gain a better understanding of the experience of rhythm by investigating the relationships between sounding music, underlying reference structures, and body motion. In chapter 2 it was pointed out that musical concepts such as melody and rhythm are the products of human cognition (see section 2.3). Hence, musical rhythm, it is argued, is a meeting of sound and perceiver. Due to the nonlinear nature of our hearing system, representations of physical sound do not necessarily directly translate into how we actually perceive the sound (with regard to, for example, the determination of a rhythmic event's onset). I have therefore chosen to differentiate between *sonic rhythm* (based on the physical sound signal) and *perceived sonic rhythm* (based on our cognition of the sound signal). In addition, the experience of rhythm in a musical context also includes a set of underlying reference structures—that is, a mental construct in the performer and perceiver. Such structures do not necessarily exist in the sound signal itself but instead supply a framework against which we perceive rhythm. In other words, the experience of rhythm includes the interaction between *perceived sonic rhythms* and *underlying reference structures*. The underlying structures exist on several metrical levels, so this chapter begins with a discussion of the relationship between perceived sonic rhythm and those structures. It then explores this interaction in relation to certain specific structural levels—namely, *pulse*, *meter*, and *subdivision*.
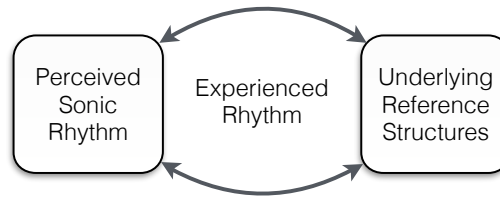
Figure 3.1: An illustration of rhythm as an interaction between perceived sonic rhythm and underlying reference structures in a musical context.

3.2   Rhythm

In chapter 2, distinctions between the physical sound, sound perception, and music cognition were discussed, in order to highlight the fact that music perception is more than the experience of sonic events (see section 2.3). Within musical rhythm, as well, one can distinguish between aspects of the sonic rhythm, the perception of the sonic rhythm, and the underlying reference structures. It is particularly interesting, from a cognitive perspective, to look at the ways in which we perceive sonic rhythms. Alternatively, measurements and analyses of the physical sound can also provide interesting insights into the experience of rhythm, which is why audio analyses inform many rhythm studies. Yet we must remember that when we measure sound, we produce a *representation* of the physical sound, whereas the *perception* of the physical sound is a rather more complex process that depends on multiple factors. This is why it is important to differentiate between sonic rhythm and perceived sonic rhythm.

In addition, as mentioned, we perceive sonic rhythm against some underlying reference structure. This chapter, then, explores the interactions between sonic rhythm (the physical sound), the perception of sonic rhythm (a cognitive process), and underlying reference structures (emerging mental constructs). Here, the term *perceived sonic rhythm* will label the sonic domain as perceived, and the term *underlying reference* structure will label the mentally constructed reference behind it (but not necessarily coinciding with or deriving from it). *Rhythm* in this text means *experienced rhythm*, including both the perception of sound and the underlying reference structures, whereas *sonic rhythm* refers to rhythms based on the physical sound signal. A presentation of the interplay between perceived sonic rhythm and underlying reference structures is illustrated in figure 3.1.

Also central to this thesis is the intimate relationship between rhythm and motion. As pointed out in chapter 2, human perception is multimodal and closely related to body motion. An *ecological perspective of perception* (Gibson, 1986) states that we learn about the world by continuously interacting with it, and the so-called *motor theory of perception* suggests that we make sense of what we hear by mentally simulate how the sounds are produced (see, for example, Berthoz, 2000; Godøy, 2010, see also section 2.4). Shove and Repp (1995) argue that musical motion is in fact *audible human motion* (Shove and Repp, 1995, p. 60). According to this view, our perception of rhythm involves not only the processing of sonic input but also the actions we relate to it. As pointed out in section 2.4, these actions can be directly sound producing (Godøy et al., 2006), or they can be mental simulations related to sonic objects (Godøy, 2003, 2006). Iyer (2002) also suggests that music perception implies an understanding

of bodily motion, and that "rhythm" is "human motion" from an ecological perspective. Body motions related to underlying reference structures in the experience of rhythm—foot tapping, body swaying, and head nodding, for example—are also frequently mentioned (e.g., Godøy and Leman, 2010) and experienced. From an embodied perspective on music cognition (Leman, 2008), both the *perceived sonic rhythm* and the *underlying reference structures* in figure 3.1 incorporate an understanding of body motion.

3.2.1 Terminology

The terms used to describe rhythm concepts are not exclusive and are often closely related to music notation, both because they emerge from studies of *music as notated* and because some rhythm studies are specifically interested in how music should be notated. Many of these terms, and the concepts behind them, have been challenged in turn by music that derives from oral traditions, where knowledge can be both implicit and embodied, as well as challenging to notate. Rhythm is also studied from different perspectives in different fields, meaning that the same terms might be used to describe different phenomena. For example, a *beat* can be conceptualized from a physical perspective (a pressure wave that propagates through the air) or a psychological perspective (something perceived). Interdisciplinary work must incorporate many of these rather multivalent terms, demand, in turn, some prefatory discussion, which follows here.

3.2.2 Perceived Sonic Rhythm and Underlying Reference Structure

The relationship between perceived sonic rhythm and underlying reference structures has been approached from various perspectives. Some theories on rhythmic structures are based on music with a written score (e.g., Clarke, 1985; Cooper and Meyer, 1960; Lerdahl and Jackendoff, 1983). Bengtsson (1973) points out that it is important to distinguish between *music as notated* and *music as performed* and *perceived*. The notation does not always give a true picture of the rhythmic structure intended by the composer or as perceived (Bengtsson, 1987, p. 75). At the end of the 1950s, the so-called Uppsala school of rhythm research started at Uppsala University under Bengtsson, Gabrielsson, and their colleagues. They sought to analyze musical rhythm based on *sound* rather than *notation* (Bengtsson et al., 1969, p. 49). One point of departure was to identify specific structural features below the musical surface that characterized a given music style; another was the hypothesis that good performances are neither mechanical nor random. Uppsala-school researchers found this to be the case for a range of musical styles and further noted that informed listeners could easily tell whether a musical dialect was being performed in the right way or with the "wrong accent" (Bengtsson, 1987, p. 74). Through the use of various apparatuses for sound recording and analysis (see, for example, Bengtsson et al., 1969), Uppsala researchers discovered systematic duration patterns that were specific to different styles of music. Their basic hypothesis was that these patterns represented an active force beneath the musical surface that organized the experience of the music's rhythm (Bengtsson, 1987, p. 74). They then introduced the concept of *systematic variations* (SYVAR) to label these consistent and recurrent patterns (Bengtsson, 1974, 1987; Bengtsson and Gabrielsson, 1980; Gabrielsson, 1982).

Charles Keil also highlights the importance of analyzing music *as performed*. Keil (1995) criticizes music theorists' emphasis on *syntax* in music analysis and argues that music is not primarily about *structure* but about *process* (Keil, 1995, p. 1). Elsewhere he claims:

> The power of music is in its participatory discrepancies, and these are basically of two kinds: processual and textural. Music, to be personally involving and socially valuable, must be "out of time" and "out of tune." (Keil, 1987, p. 275).

Keil then points out that some jazz musicians tend to play "on top" of the pulse, while others tend to be more "laid back," or behind the pulse, and that when these tendencies mix in one performing combo, the result will be a rhythm pattern with a "swing" or "groove" (Keil, 1966, p. 34). He uses *participatory discrepancy* (PD) to label those "little discrepancies within a jazz drummer's beat, between bass and drums, between rhythm section and soloist, that create 'swing' and invite us to participate" (Keil, 1987, p. 277).

In his essay "Description of grooves and syntax/process dialectics," Kvifte (2004) argues that "processual descriptions must be understood in relation to syntax" (Kvifte, 2004, p. 54). Kvifte considers *syntax* to belong to the domain of experience and points out that exactly how we categorize sonic events—as early, late, or on the beat, for example—depends on an *experienced reference*. He claims that the power of the groove does not derive from the PDs understood as process, but from the *relation between syntax and process* (Kvifte, 2004, p. 61). Kvifte sees the Uppsala school's SYVAR concept as in some ways parallel to the PD concept. However, he suggests that SYVAR might be used to describe *underlying patterns* in specific styles of music (such as a Vienna waltz), and PD to specifically describe *dynamic* groove discrepancies (such as small differences in timing between two or more performers) (Kvifte, 2004, p. 75).

Parallel to Kvifte's approach, Danielsen highlights the interaction between *virtual* non-sounding reference structures and actual sonic rhythmic events (Danielsen, 2006, 2010b), the former of which she calls *figure*, and the latter, *gesture* (Danielsen, 2010b, p. 6). Gesture is a perceived entity that implies a holistic approach to rhythm that includes every aspect of it, including its proposal for an underlying reference structure—that is, its figure.

Honing (2013) refers to sonic events that are played slightly "early" or "late" in a musical performance (recall Keil's PDs) as *expressive timing* (see also Clarke, 1985, 1987). Those sonic events are perceived as belonging to the underlying reference structure but not as perfectly matching it. Expressive timing is instead a *nuancing* of that reference structure. The Uppsala school has emphasized that expressive timing is seldom random or unintended—that is, seldom a consequence of human imperfection or lack of ability. On the contrary, it is what makes the music "come alive" (Snyder, 2000), just like Keil's PDs and Kvifte's syntax/process relations. Bengtsson (1987) notes:

> What we call perfection in good musical performances has its own form of precision that is not identical with mechanical exactitude. [...] What good performers do is not random, however, but represents another kind of precision. Good musicians

are experts in different kinds of such non-mechanical precision. (Bengtsson, 1987, p. 78)

Expressive timing is closely related to *categorization*, in that a given nuance will be perceived within the "category" of some level of the underlying structure, even if it does not match up. Also, since the categories are related to the mentally constructed underlying reference, the categorization must also be subjective. For example, one person may perceive a sonic event in a music performance within the boundaries of the underlying structure (nuance), whereas another person may perceive the same sonic event as "out of time." Consequently, categorization is closely related to memories and previous experiences (Huron, 2006), with regard to both specific songs and music-cultural capital—that is, one's familiarity with the given style of music (Trehub and Hannon, 2006). London (2012) points out, "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" (London, 2012, p. 123).

Syncopated sonic events are not perceived as nuances but as products of the subdivision level between the underlying metrical beats. The determination as to whether a sonic event that does not coincide with a metrical beat is a nuance or the product of another level of pulse results from the process of subjective categorization.

Johansson (2010) introduces the concept of *rhythmic tolerance* to describe the relationship between underlying reference structures and sonic rhythms in traditional asymmetrical pols/springar styles of Norway and Sweden. He points out that even though the inter-onset intervals between pulse-related sonic events vary from measure to measure, the groove is not necessarily perceived as unstable. The fact that these rhythm patterns are still found to be within the stylistic boundaries of the music implies an inherent flexibility to stylistic categories, he concludes. This shows that metrical interpretation is not only a matter of perceiving beat durations based on the sonic signal but is also dependent on performers' and perceivers' familiarity with the musical style. Kvifte (2007) highlights the importance of perceivers' knowledge and experience in metrical interpretation. He propose that metrical entrainment might be more of a *pattern-recognition* task—that is, learning to recognize and discriminate among a large number of (musical) patterns—than a matter of extracting metrical information from the sound signal based on certain general rules (Kvifte, 2007, p. 81).

The underlying reference structure may also be intimately related to performers' and perceivers' body motion, and in some musical cultures intimately related to dance. In his work on music from Eastern and Central Africa, Kubik (1979) differentiates between *rhythm patterns*, which refer to sonic events in the music, and *movement patterns*, which relate to both sounding and non-sounding musical phenomena. Relatedly, Chernoff (1991) notes that African musicians often avoid sounding notes on the underlying pulse. Blom (1981) demonstrates that traditional folk music and dance in Norway evolved in tandem, and that the idiosyncratic underlying reference structure of some of this music may be directly related to the vertical motions of its dancers. The experience of this particular rhythm, entails a culture-specific and implicit embodied knowledge of the underlying structures.

3.3 Pulse

The *pulse* is often described as successive mental beats that provide a fundamental reference level against which we perceive and interpret rhythm (Honing, 2012; London, 2012; Parncutt, 1994). According to Lerdahl and Jackendoff (1983), such pulse beats are temporal *points* in time and thus have no duration as such. The interval between two beats, which is consistent, they call the *time-span*. They argue that the pulse beats have to be evenly spaced out, and, accordingly, the time-spans have to be equal. Along these lines, then, a pulse consists of a regularly recurring (isochronous) series of identical imaginary time points (beats) with a distinct, stable rate of repetition, or tempo. Lerdahl and Jackendoff (1983) name this the *tactus*.

From a cognitive perspective, Honing (2013) relates the *beat* to the *tactus* via the notion of *beat induction*—that is, "the cognitive skill that allows us to hear a regular pulse in music and enables our synchronization with it" (Honing, 2012, p. 85). Honing prefers *induction* to perception because the pulse does not need to be sounded in order to be experienced: "While rhythm can be characterized as the varying pattern of durations that is perceived to be present in the music, meter involves our perception and, more importantly, (embodied) anticipation of such rhythmic patterns" (Honing, 2013, p. 380). Honing describes one aspect of the interaction between rhythm and meter as *syncopation*—that is, when a sonic musical event occurs between metrical beats.

In his essay "African influence on the music of the Americas" (1967), Waterman speaks of a *metronome sense*—an underlying pulse that is thought to be part of the perceptual equipment of both performers and perceivers of African music. When a pulse-related beat is played (or made sonic), it serves as a confirmation of this pulse, he claims (Waterman, 1967, p. 211). Povel and Essens (1985) propose that an *internal clock* might resonate with the pulse level in music, and that perceivers try to align such a clock to perceived accents in an expressed rhythmic pattern.

The theory of *dynamic attending* represents a more dynamic approach to underlying pulse. A basic assumption of the theory of dynamic attending is the possibility of *entrainment*—the process through which two oscillators, for example, self-adjust to a shared phase and/or periodicity (Clayton and Will, 2005, see also section 2.5). Dynamic attending to rhythm thus relies on two related assumptions: (1) there are internal oscillations in a perceiver, and (2) the external event's rhythm drives these internal oscillations, or the so-called attending rhythms, which entrain to the external rhythm (Large and Jones, 1999, p. 123). Dynamic attending, then, derives from the entrainment of *external rhythms* (sonic) and *internal rhythms* (mental). In response to a regular external rhythm, correspondingly regular and temporally focused peaks of attention (forming the pulse) arise in the perceiver (Large and Jones, 1999, p. 134). Large and Jones (1999) also explain how this internal oscillation generates periodic activity, or what they call *expectation*—an active temporal anticipation that, unlike a fixed clock, can itself entrain and synchronize with an external rhythm.

In a groove, Danielsen (2006) points out, the perceived sonic rhythm triggers an underlying *basic pulse* or *internal beat* that is crucial to understanding the groove (Danielsen, 2006, p. 55). Influenced by the theory of dynamic attending, Danielsen devised a *beat bin* model that challenges the conviction that the pulse consists of beats at temporal points—that is, that these beats do not have durations (Danielsen, 2010a). She even demonstrates how conflicting pulse locations might result in two beat locations that merge into one, producing, in turn, an *extended*
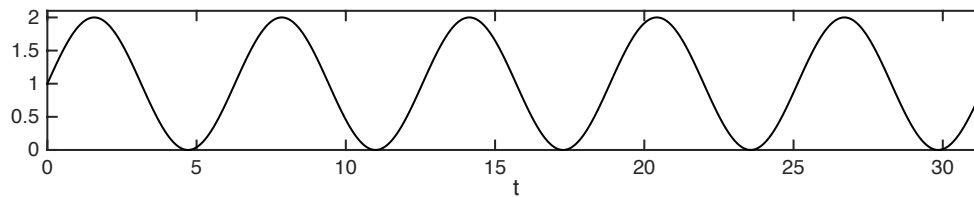
Figure 3.2: Waadeland suggests that the pulse can be represented in relation to the motion curve of a hand moving in synchrony with a metronome over time (t), where the minimal points of the pulse curve represents the pulse beats (adapted from Waadeland, 2000, figure 5.1, p. 122).

beat. [1] With regard to the theory of dynamic attending, she concludes that the attentional focus must be broadened as a consequence of multiple possible pulse locations (Danielsen, 2010a), and that the beat bin encompasses all of them (Danielsen, forthcoming).

Waadeland (2001) advocates for a continuous understanding of pulse as pulsations related to rhythmic body motions. He proposes a definition of pulse that relates to a corresponding and continuous *movement curve*, and a model of *rhythmic structure* where the underlying *pulsation* is represented by a continuous mathematical function whereby the local minimal points coincide with discrete pulse points. He describes this using a wave metaphor (see figure 3.2).

The notion that the underlying pulse must be understood in relation to performers', musicians', and dancers' body motions has been pointed out in several ethnomusicological studies. In some musical cultures, pulse motions are described as an intrinsic part of rhythm production and perception. Agawu (2003) notes that in many West and Central African dances, there are reoccurring rhythm patterns, which he calls *topoi*, that serve as an identifying signature of the dances and drumming. The key to understanding the structure of a given topos is the dance, or choreography, upon which it is based (Agawu, 2003, p. 73). In a study of Brazilian drum patterns, Kubik (1990) explains that, since the percussionists' "inner pulsation" was often not present in the sound, one had to find it in the body motion of the performers and dancers.

The fiddler's regular foot stamping in traditional Scandinavian folk music is seen as inherent to the fiddle playing, and it has been suggested that the rhythm patterns emerging from this stamping are related to the underlying pulse level of the music (see, for example, Ahlbäck, 2003; Blom, 2006; Kvifte, 1999). These patterns have likely also shaped the corresponding dance (see, for example, Bengtsson, 1974; Blom, 1981; Kvifte, 2004). In addition, studies of these music styles highlight that the underlying pulse does not have to be isochronous (see, for example, Bengtsson, 1987; Groven, 1971; Kvifte, 1999), something that seems to be supported by the performer's body motion described above. I will return to these correspondences later.

3.4   Meter

Whereas the pulse refers to successive beats, the *meter* could be described as an *organization of* those beats. Research has shown that we tend to interpret perceived sonic rhythm according to a metrical scheme. Even when we are presented with a series of isochronous and equivalent-

---

[1]Polak (2010, section 4.6), Gerischer (2006, pp. 45–46), and Iyer (2002, pp. 400–401) all mention the *flam*—a stroke in drumming that consists of a played double note whose two attacks are so close that they are perceived as one event.

sounding events, we lock them into a periodic pattern via *subjective accentuation* (Bolton, 1894; Brochard et al., 2003).

In a musical context, Lerdahl and Jackendoff (1983) find that *meter* arises when certain pulse beats are perceived as more accentuated than others within a regularly occurring *pattern*—that is, a periodic alternation between "strong" (accented) and "weak" (unaccented) beats, or a *metrical hierarchy*. In Lerdahl and Jackendoff's model, the construction of this metrical hierarchy follows mathematical rules related to music notation and depends on two or more levels (time-spans) of beats—for example, a quarter-note or dotted-half-note level. The time-spans between beats at one level must be either two or three times longer than those between beats at the next level down. In order for a given beat to be a "strong beat," they continue, it must occur in at least two levels. The first beat in a measure (as dictated by the meter) is considered the strongest and called the *downbeat*. It is the interaction of different levels of beats (or the regular alternation of strong and weak beats within a level) that generates a sense of meter (Lerdahl and Jackendoff, 1983, p. 68). Lerdahl and Jackendoff point out that perceivers tend to focus primarily on one level of metrical structure, and typically the pulse level, or tactus (Lerdahl and Jackendoff, 1983, p. 21).

Parncutt (1994) describes how a rhythmic sequence can evoke several pulse levels at the same time and introduces the term *pulse sensation* to encompass "all rhythmic levels spontaneously evoked in the mind of a perceiver" (Parncutt, 1994, p. 410). Like Lerdahl and Jackendoff, Parncutt finds that a sense of meter arises when two or more pulses are perceived at the same time—for example, the pulse at the regular quarter-note level and a pulse at the measure level in "perceived 3/4 time" (Parncutt, 1987, p. 134). Parncutt also proposes a theory of *pulse salience* to capture the strength or prominence of a given pulse sensation. In line with Lerdahl and Jackendoff, Parncutt sees the tactus level as the pulse sensation with the greatest salience (Parncutt, 1994, p. 413).

London (2012) uses the term *rhythm* to denote "patterns that are phenomenally present in the music" (London, 2012, p. 4), which seems to correspond to the way *sonic rhythm* is used in this text. He emphasizes the difference between sonic *rhythm* and *meter* and refers to the aforementioned *subjective accentuation*[2] to demonstrate that metrical accents are mental constructs that do not necessarily coincide with the pattern of accentuation in the sound: "Meter is thus more than a response to invariant features of the musical stimulus" (London, 2012, p. 13). One does not necessarily infer a meter by grafting the accentuation in a sonic rhythm onto an internal pattern of accents. Instead, one compares the perceived sonic rhythm to "a repertoire of well-known rhythmic/metric templates" (London, 2012, p. 67)—for example, a "backbeat" groove always has its dynamic accents on beats 2 and 4. A similar point was made by Kaminsky (2014), who noted that people's ability to recognize meter is based on socially learned cues related to music culture, such as, for example, when someone familiar with reggae recognizes the meter in the intro of "Stir It Up" by Bob Marley, despite (or because of) the off-beat guitar riff (see also section 2.6). To explain the perception of meter in music with an unfamiliar meter, London turns to Jones's notion of *abstraction* (Jones, 1990)—that is, the extraction of invariant aspects from the music. He points out that in music with an unfamiliar meter, a perceiver must first extract relevant periods, based on what is heard, then locate a fitting

---

[2]He suggests *subjective metricization* to empasize that what is subjective is the perceiver's sense of the underlying meter against which the clicks are heard.

metrical framework for them. For performers, on the other hand, the sense of meter has to be established internally before they start playing (London, 2012, p. 69).

In this section, we have looked at *meter* from different perspectives, some of which include both the pulse level and subdivisions of the pulse. In what follows, I use meter only to refer to a particular organization of the basic pulse—that is, the pulse level in a metrical sense, or the notated measure level—and the term *underlying structure of reference* to encompass *all* levels of underlying pulse.

3.4.1   Meter and Motion

The intimate relationship between musical meter and body motion has been documented in a number of studies, for example via the aforementioned theory of *entrainment*. Entrainment studies observe that humans seem to be able to coordinate their actions with an external auditory stimulus. Clayton (2012) describes the *intra-individual entrainment* that takes place within a person and is responsible for the perception of musical meter (see also 2.5). Our ability to synchronize our rhythmic motion to an external (sonic) rhythm has also been investigated in a number of studies using finger tapping, dancing, and other body motions, performed in time with an audible rhythm (see, for example, Repp and Su, 2013; Danielsen et al., 2015). According to existing research on pulse and body motion, body motion does not only represents a reaction to sonic rhythmic input, but it can also facilitate the processing of temporal structures (Su and Pöppel, 2012) and improve the perception of timing, or even time keeping (Manning and Schutz, 2013). In a study of spontaneous motions to metrical music, Toiviainen et al. (2010) found that several levels of the metrical hierarchy could be embodied simultaneously.

In some musical styles that are intimately related to dance, it has been suggested that meter obtains from a *tacit knowledge* shared by those who are familiar with the styles. Parallel to London's (2012) *metrical templates*, mentioned above, Blom (2006) proposes a *motor theory of rhythm*, implying that "culture specific movement styles of a social group represent shared kinaesthetic experiences embedded in its musical forms of expression, thus constituting the implicit and shared background knowledge from which socially appropriate rhythmic actions/reactions are generated" (Blom, 2006, p. 79).

Phillips-Silver and Trainor (2005) found that infants who are bounced to a sonic rhythm with an ambiguous meter prefer the metrical interpretation to which they have been bounced. In a cross-cultural study, Trehub and Hannon (2006) found that people's responses to music with meters of differing complexity were culture specific. In other words, how we experience the meter in music is related to music culture, our previous experiences, and the kind of music to which we are most exposed. This highlights the importance of incorporating performers and perceivers from different music cultures into rhythm studies.

3.4.2   "Standard Rhythms," Meter, and Music Culture

An interesting model regarding how meter is constructed in relation to a perceived sonic rhythm is that of the "standard rhythm,"[3] a genre-specific rhythmic pattern (or recurring ostinato) that identifies the meter, though it is not aligned with the pulse level of the meter. The standard

---

[3]The term "standard rhythms" is based on Steen Nielsen's (1985) *standardrytmer* (Danish).

Standard rhythm I   Standard rhythm II

3   3   2   2   2   3   2   3
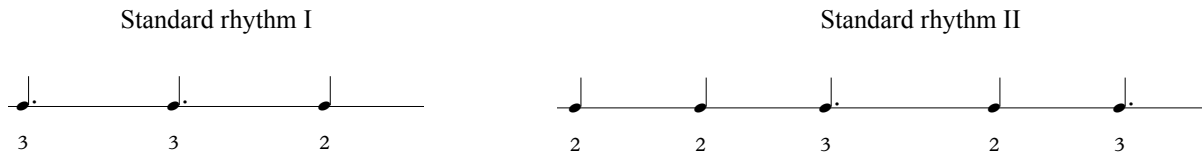
Figure 3.3: Two "basic" standard rhythms, adapted from Nielsen (1985, p. 60, ex. 8).

rhythm, which is central to some music cultures, can be understood using London's (2012) rhythmic/metric template as well (see section 3.4).

The vitality of the relationship between standard rhythm and meter has been highlighted in many studies of music and dance from West and Central Africa (see Agawu, 2003). The standard rhythms are usually played on the bell or another high-pitched instrument. Nketia (1986) describes standard rhythms using the term *time line* and relating them to groupings of (isochronous) pulse subdivisions (Nketia, 1986, pp. 131–132). In 1985, Danish music teacher Steen Nielsen (1939–2003) published *People and Music in Africa*,[4] a study partially based on his stay in Tanzania (Nielsen, 1985). Nielsen suggests two "basic" standard rhythms in the African music he experienced: Standard rhythm I, which consists of eight events (3–3–2), and Standard rhythm II, which consists of twelve events (2–2–3–2–3)—in both, the events refer to a level of subdivision (see figure 3.3). He explains that these standard rhythms can be varied in a number of ways, both by shifting them and by adding subdivisions that are otherwise silent.

London (2012) also refers to standard rhythms I and II (an *8 cycle* and a version of a *12 cycle*, respectively, in his terminology). London explains that the rotation of a 3–3–2 pattern (Standard rhythm I) occurs in a number of music styles (London, 2012, p. 152), whereas only some rotations of the 2–2–3–2–3 pattern (Standard rhythm II) are found (London, 2012, pp. 156–157). He suggests that this may be a contingency of the music style and practice or a consequence of the features of those patterns in particular.

The relationship between standard rhythms and meter points to the fact that while the experience of rhythm is an interaction between perceived or produced sonic rhythms and an underlying reference structure, the actual pattern of the underlying structure does not have to be present in the sound at all. It is, however, inherent in the production and perception of the standard rhythms.

Since rotations of the same standard rhythm are used in different music cultures, is it not enough to recognize the rhythm pattern—one must also know the music culture. The music styles referred to above have an intimate relationship to dance, and it has been pointed out that the perceived sonic rhythm should be understood in that context (Agawu, 2003; Kubik, 1990). Kaminsky (2014) explains that in some music styles, the dancer would align his or her motion according to a rhythm pattern typical for the style of music (that is, a standard rhythm), whereas in other styles of music, like Swedish polska, the cue is not in a specific sonic rhythm but in a more complex pattern (see also section 2.6). Either way, however, we find the same mechanisms of cultural learning.

---

[4]*Mennesker og musik i Afrika* (my translation).

Figure 3.4: Three different metrical interpretations (R, L, and F) with a 6/8 meter as a starting point: (R) three subdivisions of every beat; (L) two subdivisions of every beat; and (F) four subdivisions of every beat.

3.4.3 Ambiguity and Meter

London (2012) notes that many melodies can be experienced in multiple metrical contexts, because the listener can choose different sonic events with which to determine the downbeat, according to the context and performance conditions. He terms these kinds of sonic rhythms *metrical malleable* (London, 2012, pp. 99–100).

In polymetric music, as well, multiple meters can coexist. Regarding the 6/8 meter, for example, one might perceive every third eighth note as a metrical beat—that is, every metrical beat has three subdivisions. One could express this level by tapping the right hand (R) on every third eighth note. One might also perceive every *second* eighth note as a metrical beat—that is, every metrical beat has two subdivisions. One could express this level by tapping the left hand (L) on every second eighth note. Lastly, one could perceive every *fourth* eighth note as a metrical beat and express this level by tapping one's foot (F) accordingly. One could play the rhythm notated in figure 3.5 using the right hand (R), left hand (L), and foot (F) to make all of the possible metrical beats in all of the layers audible, and then one could switch between the R, L, or F layers as one's metrical reference structure. London (2012) refers to David Locke's "Gestalt flip" to describe the way in which one might switch between alternative metrical interpretations.

Which pulse level determines the meter can vary with the perceiver and the musical context. Danielsen (2015) describes how the changing sonic rhythm patterns in "Nasty Girl" by Destiny's Child alternate the main pulse between a "fast" interpretation and a "half-tempo" interpretation. Another example is the phenomenon of "turning the beat around" that appears frequently in electronic dance music (EDM), whereby an initial rhythm pattern with prominent sonic events at a given pulse level gradually shifts as the beats between the pulses become more prominent, "forcing" the perceiver to shift his or her metrical interpretation as well (Butler, 2006).

How one perceives meter can also be influenced by one's music culture or familiarity with the music that is sounding. For example, in a music culture that draws upon the aforementioned "standard rhythms," the standard 6/8 bell pattern (Nielsen's "Standard rhythm II" [1985] and Nketia's *time line* used in triple rhythm [1986]) will immediately evoke its intended meter—for example, in the R meter described above—whereas perceivers from other music cultures might interpret it differently. In fact, the reason why "turning the beat around" is so compelling in an EDM club setting is that the genre initiates expectations of prominent pulse-aligned sonic events, whereas in other musical cultures, such as the one described in Chernoff (1979), the most prominent sonic events are usually played off the beat.

Butler (2006) points out that during metrically ambiguous musical passages in a club con-

text, dancers *construe the meter* rather than absorb metrical information from the sound, and also that this act of construction occurs in and between bodies as well as in minds (Butler, 2006, p. 137). As mentioned above, Phillips-Silver and Trainor (2007) found that one's metrical interpretation is influenced by one's previous experiences, including pulse-related body motions. This accords with the view that in musical styles where music and dance have evolved together, the underlying reference structure may be both conditioned by the dance and also intrinsic to the music, even when it is detached from actual dancing (Bengtsson, 1974).

3.4.4   Non-Isochronous Meters

The preconception that underlying reference structures must consist of equally spaced beats has been criticized by several researchers (e.g., Bengtsson, 1987; Kvifte, 2004; Johansson, 2009; Polak, 2010). Furthermore, non-isochronous patterns should not be viewed as deviations from an underlying isochronous pulse but instead as coherent, alternative reference structures. London (2012), for example, uses NI (Non-Isochronous) *meters* to refer to meters with non-isochronous beat patterns, in contrast to I (Isochronous) *meters*, and he points out that I meters can have NI subdivisions (London, 2012, p. 124). Likewise, Bengtsson and Gabrielsson (1980) acknowledge that identifying a mechanical regular norm as an underlying pulse is often useful when one is measuring systematic variations, but there are clearly musical styles where the underlying pulse is *not* isochronous—for example, the Vienna waltz. (Their aforementioned concept of *systematic variations* [SYVAR], however, refers to these consistent/recurring deviations from the mechanical norm—that is, to non-isochronous structures [Bengtsson and Gabrielsson, 1980, p. 257].)

Much traditional dance music in Sweden and Norway is in so-called *asymmetrical meter*—that is, the three pulse beats in a measure are of uneven duration. The Swedish folk music collector Einar Övergaard (1871–1936), who travelled in Sweden and Norway from 1892 to 1904 to transcribe traditional music, struggled with how to notate it. He started with a $2\frac{1}{2}/4$ meter, then went to a 1/2 + 2/4 meter, then ended up using a normal 3/4 meter (Ramsten, 1982, pp. 205–206), as has been done since, despite the non-isochronous pulse beats. It was (and remains) difficult to capture the underlying reference structures of these musical styles in traditional musical notation. In an attempt to document the fluctuations, fiddler, composer, and music researcher Eivind Groven (1901–1977) carried out a rhythm study in the 1930s using an old Morse-code instrument (Groven, 1971). While listening to music recordings, he tapped the beat on his equipment, producing a paper slip whose dots represented the beats in time. He then measured the distances between the dots and calculated an average beat duration pattern for each recording. He found that tunes from different regions seem to have slightly different beat patterns, but that tunes from within a region followed the same pattern. Also along these lines, Kvifte (1999) registered beat positions by playing the perceived beats on a MIDI keyboard in time with Norwegian folk music tunes, while Johansson (2009) turned to software that showed the music's waveform and played the sound back at the same time. By moving his cursor back and forth on the waveform, he could determine where one sonic event ended and the next started.

Blom (1981) cautions that the underlying pulse in musical styles with asymmetrical meters should be understood in relation to the corresponding dances. He observes that the vertical

motions of the dancers' bodies are both periodic and apparently correlated to the underlying pulse in the music. He calls this link the *patterned libration of the body's center of gravity* and concludes that the underlying structure of each local style can be represented by these motion patterns (e.g., Blom, 1981, 1993, 2006).

Theories of a *common fast pulse* (CFP) postulate an underlying level of isochronous subdivision that provides a stable reference upon which the non-isochronous levels are constructed (Kauffman, 1980; Waterman, 1967). Kvifte (2007) is, however, critical of the use of *common fast pulse* theories to explain how non-isochronous pulses are possible. Kvifte points instead to a specific style of traditional Norwegian dance music in which both the beats and the subdivisions are uneven in their durations (Kvifte, 2007). Instead of a common reference at some level of subdivision, Kvifte proposes a *common slow pulse* (CSP) theory—that is, a common reference at some level *above* the beat level. Influenced by Blom's libration pattern hypothesis, Kvifte argues that this beat level can be discerned in relation to rhythmic body motions, meaning that we divide the common slow pulse into equal or unequal beats using our bodies. He also proposes that body motion is a better candidate for understanding the reference structures in music than a fixed clock pulse, which, in the case of these musical styles, at least, is not present in the music at either the pulse or the pulse subdivision level. All of this suggests that body motion should be incorporated into studies of non-isochronous meters, as is the case in this thesis.

3.5 Subdivisions in the Underlying Structure

The previous sections dealt with the underlying pulse level and its organization into meter as two fundamental reference points for the experience of rhythm. This final section addresses the fact that the underlying reference structure usually involves more than the pulse level and the meter. Nketia points out that subdivisions of a given pulse are also usually included in the underlying reference structure (Nketia, 1986, p. 126). Like the pulse itself, these subdivision levels are also mental constructs and not necessarily present in the sound. Yet they supply the reference against which sub-tactus events, such as, for example, *syncopations*, are perceived.

Like the pulse level, and also in line with Kvifte's (2004) *common slow pulse* theory, these underlying subdivision levels are not necessarily isochronous in nature. Polak (2010) observes that non-isochronous subdivisions (subpulses, in his terminology) are inherent in the repertoire of jembe music from Mali (Polak, 2010, section 1). Polak found that the jembe scheme generally followed a *short–medium–long* progression that seemed to be stable across performances. He also found that this non-isochronous subdivision pattern seemed to persist across a wide range of tempi and was, in fact, inherent to the metrical structure (Polak, 2010, section 7).

Jazz swing is often conceptualized as a triple shuffle with isochronous subdivisions. However, Benadon (2006) found swing ratios that ranged from 55:45 to 58:42, which is a long way from the isochronous 67:33 (Benadon, 2006, pp. 90–91). Similar results have been found in rhythm studies of Brazilian samba, which is characterized by its complex rhythm patterns atop systematic microtiming at the level of the sixteenth note. Based on field audio recordings of Brazilian percussionists in the Bahia region, Gerischer (2006) found a *medium–short–medium–long* duration pattern at sixteenth note level. Similar findings are presented by Gouyon (2007); fol-

lowing an analysis of audio recordings containing short excerpts of samba music, he found that the third and fourth sixteenth notes in samba seem to be systematically played slightly ahead of their quantized positions. These results were confirmed by Naveda et al. (2009). In addition, their findings implied a systematic delay of the first sixteenth-note position in the instruments in the low frequency region of the spectrum. According to Kvifte's (2004) suggested use of SYVAR and PD (see section 3.2.2), samba seems to be featured by both systematic variations (SYVAR) and very small differences in timing between performers (PD). But how might one differentiate between systematic microtiming (as inherent in the underlying reference structure) and (systematic) expressive timing (or PD)? Performers' body motions might represent one way to do so, and new motion capture technology allows researchers to incorporate body motion into their projects.

It is important to differentiate between expressive timing and *systematic microtiming*: the former belongs to the sonic domain—that is, to *nuances* within the framework of an underlying structure—whereas the latter arises from the underlying reference structure in musical styles with non-isochronous subdivisions. However, systematic microtiming may not be the best way to describe non-isochronous underlying subdivisions. What about those cases where a sonic rhythm is played *systematically* "ahead" or "late" in relation to an isochronous pulse? If we perceive a groove as "laid back" over the course of a whole song, it must be "late" in relation to some underlying reference. Systematic microtiming has also been used to describe what I refer to as (systematic) expressive timing in this text. The way one labels different rhythmic concepts, of course, influences the way we conceptualize rhythm as a whole. Thus we need another term to describe those occasions when the underlying subdivision level consists of non-isochronous beats that are *not* deviations from an isochronous structure. I will refer to the underlying subdivision level as the *metrical subdivision* and use the term *non-isochronous metrical subdivision* to describe its non-isochronous variation.

3.6   Summary

This chapter has presented an overview of some rhythm research with a focus on pulse and meter as underlying reference structures. It has been emphasized that one must differentiate between the rhythm structures in the sonic signal (both physical and perceived) and the underlying structures that are constructed in the minds of performers and perceivers. The interaction between perceived sonic rhythm and underlying reference structures was highlighted and discussed in relation to some levels of the underlying structure—namely, *pulse*, *meter*, and *subdivision*. Since underlying structures are mental constructs in performers and perceivers, they are not necessarily represented by sonic events in the music. Hence, various theories on how such underlying structures should be understood and analysed were presented. It was pointed out that metrical interpretation is highly dependent on performers' and perceivers' experience, and also their acquaintance with the music culture. In musical styles with an intimate relationship to dance, it was argued that dancers' body motion is of relevance to rhythm studies. Furthermore, underlying reference structures can be both *isochronous* and *non-isochronous*, and it was suggested to call the underlying subdivision level a *metrical subdivision* and to use the term *non-isochronous metrical subdivision* to describe when such an underlying subdivision consists

of non-isochronous beats.

Chapter 4

Methods

Quantitative methods allow us to observe patterns of
organization that might otherwise be difficult or
impossible to decipher. For the new empiricist, an
interest in quantitative methods has nothing to do with
science. It has everything to do with becoming a more
observant music scholar.
—David Huron 1999, p. 22.

This chapter gives an overview of the methods used in the papers included in this thesis. Different kinds of motion capture technologies, relevant motion and sound analyses (including post-processing procedures), and statistical analysis are presented. Finally, a short discussion regarding methodological limitations is given.

4.1   Introduction

In chapter 2 I pointed out that body motion is integral to the perception of musical sound, and in chapter 3 I described the interaction between sonic rhythm, perceived sonic rhythm, and the underlying reference structures against which one encounters rhythm. In some music cultures, as well, these underlying reference structures do not seem to consist of isochronous sequences, nor are they necessarily represented by actual sonic events. Based on the view that underlying reference structures may be present in performers' and perceivers' body motion during music performance, scholars have suggested that these motions should be incorporated into rhythm studies (Blom, 1981; Kvifte, 2004).

In this thesis I investigate two musical styles that are thought to feature non-isochronous underlying reference structures: telespringar and samba (see, e.g., Blom, 1981; Gerischer, 2006; Kvifte, 1999; Naveda, 2011). Since I am interested in the relationship between musical rhythm and body motion in samba and telespringar, both sound and motion data are included in my study.

This chapter looks at how sound features that are relevant to rhythm perception can be extracted from the sound signal. It also introduces various motion capture technologies and methods for analyzing motion data.

4.2   Sound Analysis

In chapters 2 and 3, I highlighted the distinction between physical and perceived sound, yet it remains relevant to investigate certain features of the sound signal itself—that is, changes in the sound signal's *waveform* (time domain) and *spectrum* (frequency domain) that might impact the perception of rhythm. Here we are primarily interested in detecting changes that are likely to be perceived as the beginnings of salient events in the sound. Snyder (2000) uses "the term *event* to describe a perceptible change in an acoustic environment" and defines *rhythm* as the occurrence of two or more *events* within the span of one's short-term memory (Snyder, 2000, p. 159). London (2012) points out that the shortest interval that we can hear or perform as an discrete part of a rhythmic figure is about one hundred milliseconds, and the upper limit of such an interval would be around five or six seconds (London, 2012). If sonic events occur at an interval that is less than one hundred milliseconds, we tend to fuse them into one event; sonic events that are more than six seconds apart are perceived as unrelated.

A sudden burst in energy often signals a new musical event (Collins, 2010). The *physical onset* of a sonic event can be defined as the point in time when the acoustic event begins (that is, when the amplitude first becomes greater than zero), whereas the *perceptual onset* can be defined as the point in time when the beginning of the sonic event is perceived (Vos and Rasch, 1981; Wright, 2008). Wright (2008) points out that the perception of a musical event is dependent on what he calls the event's *perceptual attack time* (PAT), meaning its perceived rhythmic placement. Wright illustrates how an event's perceptual onset trails behind its physical onset, and that the PAT is generally later still. In a sound that permits a sudden increase in energy, such as a highly percussive sound, the PAT can be almost the same as the onset time, whereas a sound that fades in very slowly and gradually may not have a perceptual attack time at all (Wright, 2008, pp. 19–21).

4.2.1   MIRtoolbox

The sound analysis in this thesis was primarily conducted using the MIRtoolbox (Lartillot and Toiviainen, 2007) that was developed at the University of Jyväslylä in Finland. The MIRtoolbox offers a set of functions for feature extraction from an audio signal, and from musical sound in particular. A sound signal can be visualized as an amplitude waveform, and an amplitude envelope, showing the outer shape of the waveform. As previously mentioned, an onset is the point in time when a new musical event starts, and *onset detection* describes the process of automatically spotting onsets in an audio signal. Because the audio envelope displays fluctuation in energy, physical onsets can be estimated based on this curve. Instead of detecting the physical onset, as defined above, then, the onset detection function in the MIRtoolbox performs a peak-picking procedure on the envelope curve and reports the temporal positions of those peaks.[1]

A challenge with regard to onset detection is that one does not know whether the temporal position of the peak (the local maximum) in the envelope curve is the point that is actually perceived as the sonic event's onset. However, in my research I am primarily interested in

---

[1]There are other options for onset detection in the MIRtoolbox, but this method is used in most of the studies using onset detection referred to in this thesis.
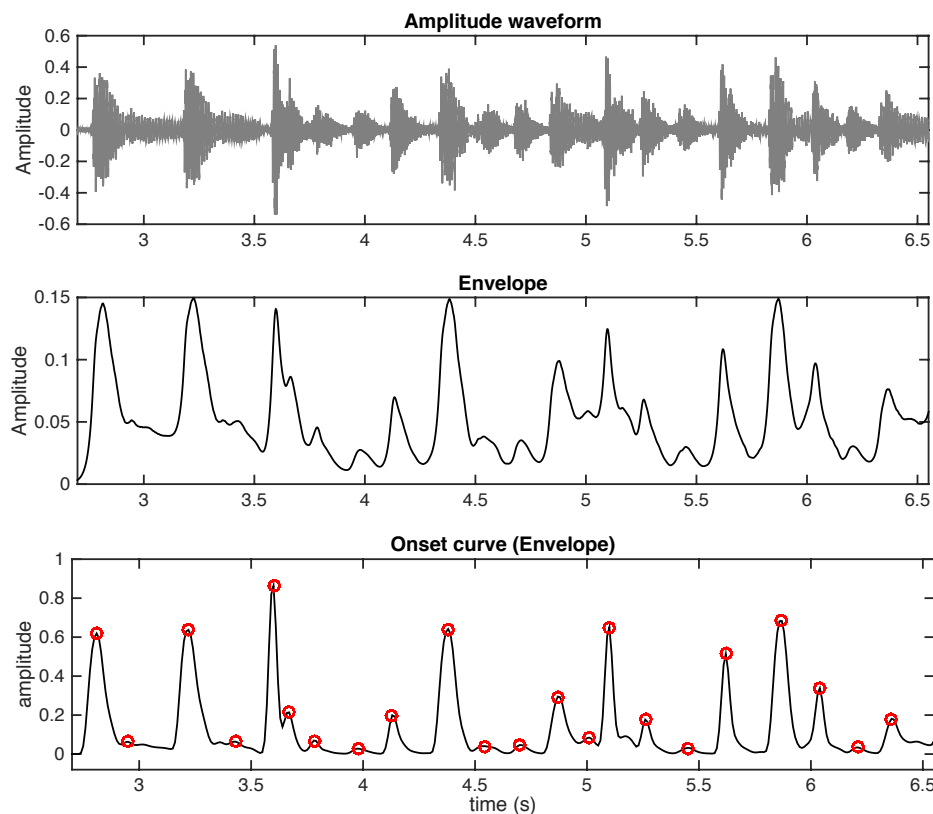
Figure 4.1: Plots showing the (top) amplitude waveform, with the corresponding (middle) envelope and (bottom) onset curve of a samba groove played on a pandeiro. Red circles on the onset curve represent the detected onsets. The second and fourth onsets on the onset curve do not represent the occurrence of new sonic events, whereas the others do.

rhythmic patterns, meaning that I want every onset to be measured in the same manner, and that I am not necessarily concerned that the estimated temporal position equal the perceptual onset. Along the lines of Wright's concept of an event's perceptual attack time, the physical onsets of percussive sounds can be quite close to the event's perceived temporal position. Consequently, the method for onset detection described above might work best when one is analyzing sounds with a pronounced sudden increase in energy. Figure 4.1 shows plots of the audio waveform, envelope, and onset curve for an excerpt of a samba groove played on a *pandeiro*. Even though the onset detection seems to function well on this sound excerpt, onset numbers 2 and 4 do not represent the beginning of new sonic events. Detected onsets that represent the occurrence of sonic events later in the groove have the same energy value, as well, meaning that setting a threshold would not have solved this problem. This demonstrates, once again, the importance of checking whether the calculated features actually represent what we perceive.

The perception of musical events can also depend on changes in the frequency domain, and these can be investigated using a *spectrogram*—that is, a representation of the distribution of energy along frequencies over time. The change of spectrum between frames can be computed via the *spectral flux*, which can be plotted over time. The peaks in the spectral flux graph indicate contrasts in the spectrogram. Prominent changes in the spectrogram, as indicated by the peaks in the spectral flux, may represent temporal positions that are relevant to the perception of rhythm. Figure 4.2 displays the spectrogram and corresponding spectral flux of the guitar riff
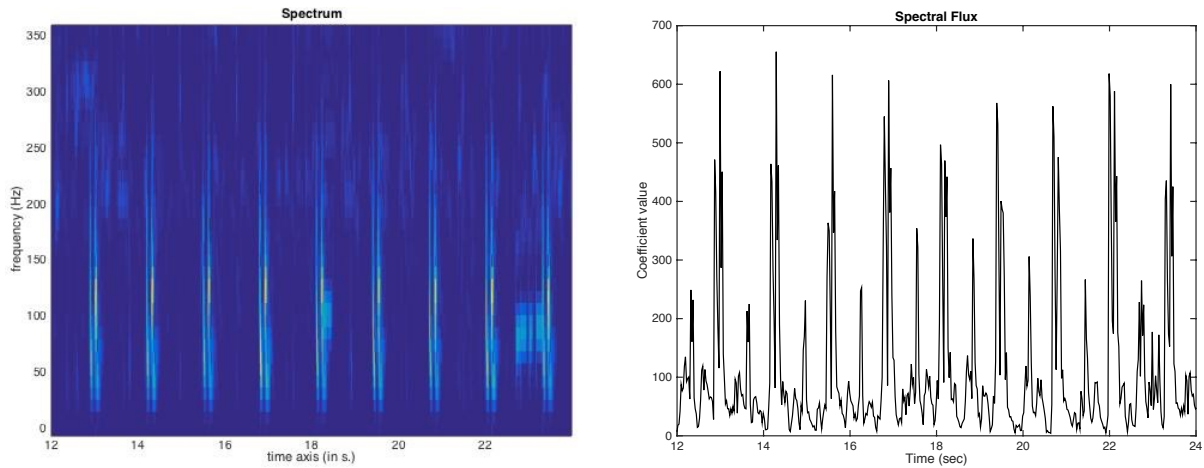
Figure 4.2: Plots showing the (left) spectrogram (0–350 Hz) with corresponding (right) spectral flux over time of the introduction to D'Angelo's "Left and Right." The highest peaks in the spectral flux indicate important changes in the spectrogram.

in the beginning of D'Angelo's "Left and Right." [2]

4.3 Motion Capture

There are a number of ways in which human motion can be described and analyzed. In a broad sense, *motion capture* (mocap) describes any process by which motion is captured, but it is commonly used to describe a *digital* means of tracking and representing motion. This is also how the term is used in this thesis. A number of mocap technologies have been used in studies of music-related motion (Burger, 2013; Nymoen, 2013) , all with different strengths and drawbacks with regard to, for example, data quality, user interface, portability, and so forth. Only a given study's particular objective can dictate which system will be best. In what follows, I discuss those most relevant to my work here: *inertial systems* and *optical systems*.

Inertial sensors typically use accelerometers and gyroscopes. They come in a range of sizes and with a range of capabilities. On one hand, inertial sensors are well suited to capturing the motions of the human body, because we move about in an environment that includes forces like gravitation. On the other hand, because this environment is also three-dimensional, spatial data is often useful. Even though position can be estimated from the accelerometer data obtained from some inertial systems, studies requiring high positional precision will benefit from systems that track position data in particular (Skogstad et al., 2011).

4.4 Optical Motion Capture Systems

Optical systems are based on various types of video cameras. In the studies referred to in this thesis, regular video cameras and infrared (IR) video cameras are used. A regular video camera is often sufficient for many scholarly efforts. A number of techniques for motion analyses and visualization of motion using video recordings have been developed in recent years (see, e.g.,

---

[2]D'Angelo (2000). Voodoo. Virgin Records.

Jensenius, 2013). In the present thesis, however, the main focus has been on analysing data from the Qualisys motion capture system. The recorded video files have only been used as a reference for checking the placement of markers, as well as playing back the audio alongside the motion capture data.

Optical infrared motion capture (mocap) systems have proven useful for many studies of music-related motion. They can capture position data in three dimensions with great precision and accuracy in both the spatial and temporal domains. Today's systems can also track motion at a very high sampling rate. *Passive* optical infrared mocap systems consist of markers with a reflective surface and cameras that both emit infrared light and record the light that is reflected from the markers. In order for the software to be able to calculate each marker's 3D position, the marker must be captured by at least two cameras simultaneously. Hence, the placement of both the cameras and the markers on the participants being recorded is of great importance. The number of cameras in use can also impact the quality of the recording. More cameras will allow the markers to be captured from more angles without being occluded; in addition, more cameras increase the capture space. *Active* mocap systems use markers that emit light themselves, which the cameras then record. One advantage to this is that, because each marker sends out unique sequences, they can be readily identified, eliminating the possibility of marker switching or confusion. Likewise, active mocap works well for capturing highly reflective objects. One disadvantage to the system is that the active markers require power, necessitating cables between the markers and receivers attached to the participants. Compared to active systems, passive systems are rather less obtrusive.

4.4.1 Qualisys

All of the motion data used in this thesis was collected in the fourMs[3] motion capture lab in the Department of Musicology at the University of Oslo, Norway, using a high quality optical infrared motion capture system from Qualisys.[4] The system consisted of nine Oqus 300 and four Oqus 400 cameras, though the earliest recordings only used the Oqus 300 cameras. The system can accommodate either active or passive markers, but only passive markers were used here. The system also enables work with a *rigid body*—that is, a fixed pattern of several markers is attached to a rigid object such as a fiddle, allowing the system to track its position and rotation rather than an assortment of otherwise unrelated markers. However, no rigid bodies were used in the studies in this thesis.

4.4.2 Recording

Before a recording can be carried out, the system must be calibrated. This process informs the system about camera placement in relation to one another and sets the origo and orientation of the Cartesian coordinate system according to which the markers are tracked. The software Qualisys Track Manager (QTM) tracks the markers in relation to X, Y, and Z axes, meaning that the position of each marker is defined by three coordinates (X, Y, and Z) in every frame. In the Qualisys system, vertical motion is usually tracked along the Z axis.

---

[3]http://fourms.uio.no
[4]http://www.qualisys.com
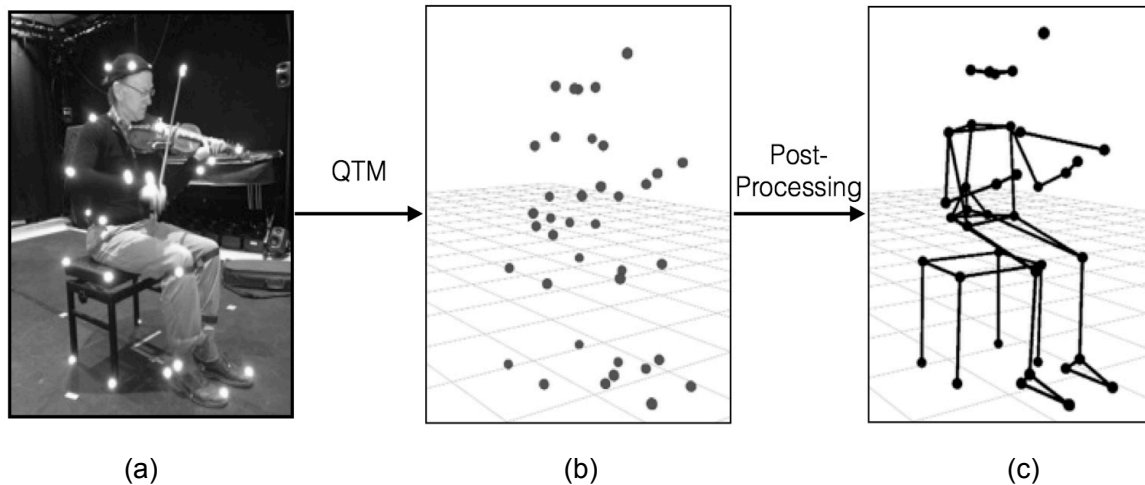
Figure 4.3: (a) A picture of a telespringar performer with passive reflective markers attached to him, (b) the markers represented by unidentified dots in a virtual 3D environment, as recognized by Qualisys Track Manager (QTM), and (c) the markers after they have been manually identified and connected by "bones".

Unlike their active counterparts, passive markers are represented by unidentified dots in a virtual 3D space (see figure 4.3b). Some mocap systems have predefined models with corresponding marker placements. In QTM the initially unidentified dots must be manually identified and labelled after the recording has been completed. Once the dots have been identified, they can be connected with "bones" to create a stick figure (see figure 4.3c). In addition to aiding the interpretation of the animation, the bones also stabilize the animation. Lastly, the stick figure also allows for the qualitative analysis of the animation.

In all of the mocap studies prepared for this thesis, we used full-body motion capture—even when we were interested in the motion of only a few markers, we conducted whole-body measurements, in case they would inspire further insight. These measurements also allowed us to create stick figures for qualitative analysis and as illustrations in presentations, and they supplied a "safety net" as we worked to identify the markers. For example, when recording dancers, it can be very hard to differentiate between a toe and a heel marker in cases where the marker disappears and then reappears as unidentified if there were no other markers for reference.

The markers were primarily attached to the participants' joints. In addition, when we were recording several participants simultaneously, we found it useful to introduce a number of control markers—that is, markers that were not intended for analysis but for differentiation of the participants. In telespringar dancing, for example, the dancers twirl around one another and hold one another's arms in different ways. Without control markers attached to the dancers' torsos (front and back) and upper and lower arms, it would have been very hard to tell them apart. The control markers were also connected with the proximate joint markers by "bones", which facilitated the identification process.

4.5 Data Processing

Although optical infrared mocap systems provide high-quality data in both the spatial and the temporal domains, it will most likely include some noise and tracking error. For example, since these recordings are based on light reflected from the markers attached to a person's body, a wobbly marker would result in noisy data, and marker occlusion would cause the marker to disappear entirely in the recording. Hence, some post-processing must occur before the data can be analyzed.

4.5.1 Gap-Filling

As previously described, the system relies on the capture of reflected light by the cameras. Thus if a marker becomes occluded or moves outside the capture space, it will disappear in the recording, resulting in so-called *dropouts* and leaving *gaps* (missing frames) in the data set. Shorter gaps can be *gap-filled* using various interpolation techniques to mathematically estimate the trajectory between known data points. Gap-filling should only be used for short gaps, since longer gaps in the data set may not be possible to estimate mathematically. The simplest method of gap-filling is *linear interpolation*, meaning that a straight line is drawn between two known data points. *Polynomial (spline) interpolation*, on the other hand, produces a curved trajectory between known data points. In the work presented here, gap-filling was principally executed in QTM using polynomial interpolation, given the "curvy" and variably paced nature of body motion. However, in the case of a "false" sudden rise or fall before a gap—for example, if a participant accidentally touched the marker before it disappeared—polynomial interpolation would most likely continue the false motion by producing an artificial bump in the trajectory, whereas linear interpolation would simply draw a straight line between the known points.

4.5.2 Smoothing

Because mocap data often reveals some noise or artefacts due to, for example, gap-filling, tracking error, unstable markers, marker occlusion, or calibration quality (Jensenius et al., 2012), *smoothing* the data is often a good idea. The *moving average* is the simplest way to smooth a data stream. It sets the output y_n to the mean of a fixed number of data points in a sequence of the input x_n. In the next step x_{n+1} the first number in the previous group is dropped and the next point in the sequence is added. This procedure is repeated for the whole sequence of numbers. In the present thesis, the *MoCap Toolbox* (Burger and Toiviainen, 2013) for Matlab was used to smooth data. The toolbox offers two filtering options, a fast *Butterworth* filter and an accurate *Savitzky-Golay* filter. Since the latter returned the best results here, it was chosen for use in most cases.

4.5.3 Sound and Motion Synchronization

When one is recording music-related motion, synchronized sound is desirable. However, Qualisys does not have a sound-recording option, so this is not an easy task. It has an integrated video-recording option, so one solution is to rely upon the sound from the video, but the video recording starts slightly after the mocap recording, resulting in an offset between them. Another

solution is to do a separate audio recording and synchronize it with the mocap recording through a sound-producing action at the start—for example, by using a clapperboard with a reflective marker attached to it. A third solution is to use the *external timebase* option in QTM, which allows for synchronization to external hardware that is not directly integrated in QTM, and this is what we did for the telespringar recordings in this thesis. We used a custom Max/MSP patch running on a separate Macintosh computer and communicating with QTM, as well as audio recording software (Reaper or Logic Pro X), also running on the Macintosh computer. When QTM was set to record, the audio recording started simultaneously, resulting in synchronized sound and motion data.
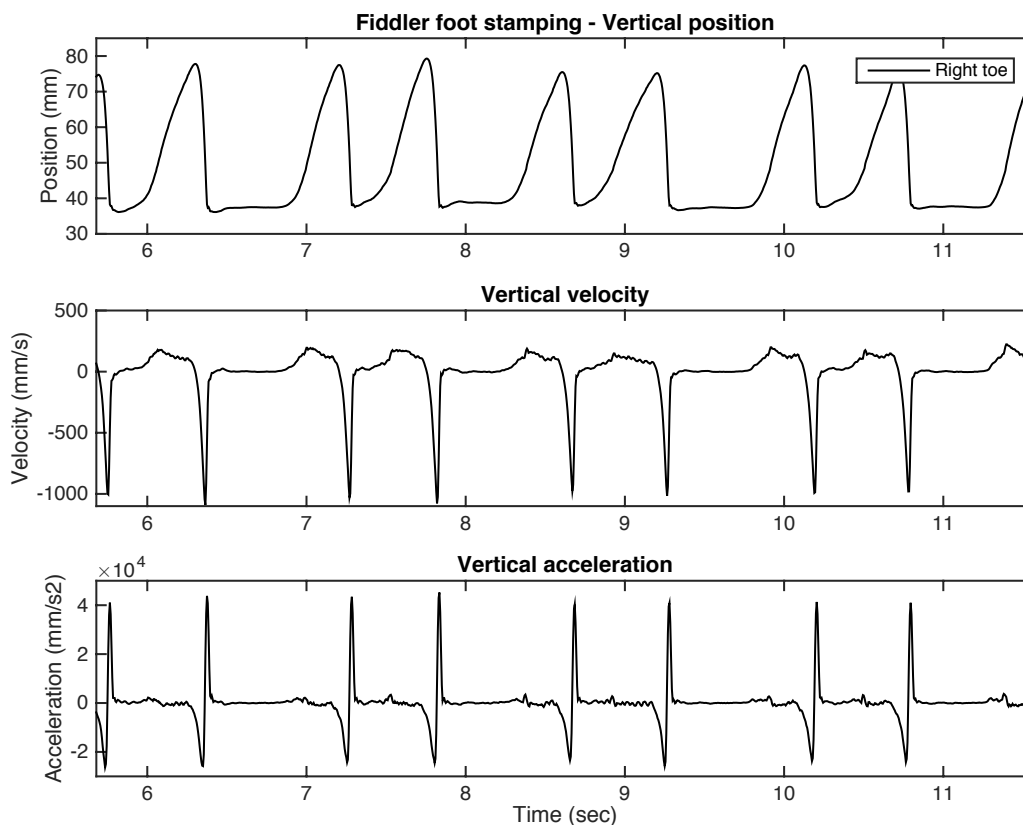
4.6   Motion Analysis

The motion analyses presented in several papers in this thesis are based on motion data collected using the Qualisys mocap system. Several of the analyses were conducted using the functions in the MoCap Toolbox for Matlab.

4.6.1   The MoCap Toolbox

The MoCap Toolbox was developed at the University of Jyväskylä in Finland (Burger and Toiviainen, 2013) and contains a set of functions written in Matlab for analyzing and visualizing motion capture data. The motion data from Qualisys (in .mat or .tsv formats) can be read directly into the toolbox, which then returns a *MoCap data structure*—that is, a data structure that contains all of the collected position data for each marker (in three dimensions)—as well as information about the number of markers, marker names (labels), frame rate, and number of frames recorded.

4.6.2   Position, velocity, and acceleration

As previously mentioned, the Qualisys system captures position data in three dimensions in accordance with a coordinate system set via calibration. In QTM, the position of the markers is measured in millimetre (mm) along three axes (X, Y, and Z). Hence, both *velocity* and *acceleration* can be calculated by deriving the position data once or twice, respectively. The MoCap Toolbox offers two options for derivation: (1) a fast version that uses differences between two successive frames and a Butterworth smoothing filter, and (2) an accurate version that uses derivation with a Savitzky-Golay FIR smoothing filter (Burger, 2013). Both methods return a data structure containing velocity or acceleration data in three dimensions. In many situations, as well, it can be useful to develop a *norm data structure* instead of a MoCap data structure, where each marker is represented by three coordinates. In order to create a norm data structure, the toolbox calculates the Euclidean norm of the vector data. This function can be applied to position, velocity, and acceleration data.

Figure 4.4: Plots showing an example of the vertical position (top) and the corresponding vertical velocity (middle) and vertical acceleration (bottom) of a fiddler's right toe over time.

4.6.3 Visualization of motion data

Many of the motion analyses are based on visualizations of the body motion of interest, so good visualization techniques are crucial. The MoCap Toolbox offers a set of plotting options. The markers can be plotted as an animation, showing the markers as dots with the bones connecting them (see figure 4.3c). The plotting option used extensively here was *time-series* plots, where the motion data (position, velocity, or acceleration) is plotted over time (see figure 4.4).

4.6.4 Custom-Made Matlab Function

This thesis focuses on rhythmic patterns in both music and motion. When investigating rhythmic patterns in body motion—for example, upper body swaying and foot stamping—the given motion's turning points are often of great interest. They can be measured manually through a visualization of the motion. But because we needed to do this in so many of our analyses, we created a Matlab script for picking local maxima and minima (peaks and troughs) in the motion data.[5] The Matlab function takes a MoCap data structure as its input and calculates the directional changes in the trajectories for a chosen marker in a chosen dimension—for example, the directional changes for the right hip marker in the vertical dimension (see figure 4.5). The

---

[5]Created by Kristian Nymoen and slightly modyfied by me.
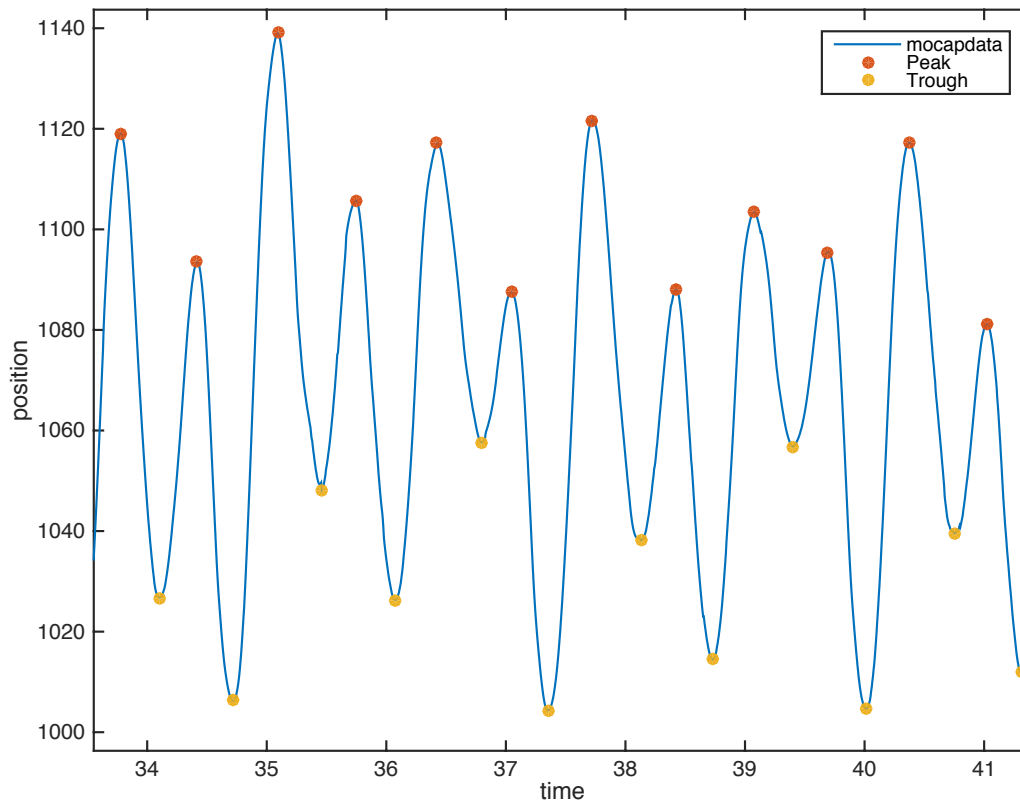
Figure 4.5: The vertical motion of a marker. The estimated directional changes obtained with the custom-made Matlab function **mcdirectionchange.m** are indicated by red (peaks) and yellow (troughs) dots on the plotted mocap data.

function returns the temporal positions of all of the detected peaks and troughs in both frames and seconds.

4.7   Statistics

Statistics is a useful tool for the presentation and analysis of empirical data. For this thesis, statistics were mainly used to represent averaged beat duration patterns and to test whether the beat durations of interest differed significantly from one another. All statistical analyses in this thesis were performed using IBM SPSS (the Statistical Package for the Social Sciences) version 22.

4.7.1   Descriptive Statistics

One way to represent beat duration patterns is via the *mean* duration of the beats—that is, the sum of the beat durations of interest divided by the number of beats in question. For example, because telespringar is in triple meter, the beat duration patterns in the telespringar studies were represented by the mean values of the first, second, and third beats, respectively. The mean value is useful because it says something about the central tendency in a series of data. In

order for the mean to be meaningful—that is, to actually express the tendency of the data—the samples have to be *normally distributed*. If the mean value is not in the approximate middle of the measured data, then it is not a viable means of describing the tendency of the data. Normal or Gaussian distributions are important in statistics and supply the premise for many statistical tests. The distribution of the data can also easily be tested through a histogram or boxplot. Mean values should always be reported with the corresponding *standard deviation* (SD). The standard deviation is a measure of the spread of the data around the mean. For *skewed distributions*, the *median* with corresponding *quartiles* should be reported instead of the mean and SD.

4.7.2 Comparing Means

In the work included in the thesis, I used statistical analyses mostly to compare means. Because normal distribution is a premise for the use of a mean, it was also a premise for these tests. The *t-test* is used to test the statistical significance of the difference between the means of two sets of data. However, since I always compared more than two sets of data (three mean beat durations in telespringar and four mean sixteenth-note durations in samba), t-tests were *not* used in these studies; instead, I applied a one-way *analysis of variance* (ANOVA). The one-way ANOVA compares the means of two or more groups.

Because I work with rhythm patterns and variable beat durations, I used one-way ANOVA to test whether the mean beat durations for different beats or subdivision types (for example, the first, second, and third beats in telespringar) differed significantly. The ANOVA only gives information as to whether there are differences between the groups, so in order to test which of the beat types' beat durations that differed from one another, I also included a *Bonferroni corrected post-hoc test* for pairwise comparison. In telespringar, for example, I therefore executed three pairwise comparisons, between the first and second, the first and third, and the second and third beats.

In the study presented in paper IV, we wanted to investigate which of three variables—dancer, tune, and beat type—was responsible for the most variation in beat duration in two telespringar recordings. For this purpose, we executed a repeated-measures three-way ANOVA, which compares the means of a dependent variable when there are three independent variables (A, B, and C) and determines which of the independent variables that is responsible for the most variance (effect). The effects of the variables are called *main effects* (A, B, and C), but the effects of the combinations of the variables—that is, the *interactions* (AB, AC, BC, and ABC)—must also be taken into account. Both the main effects and the effects of the interaction are reported.

4.8 Methodological Considerations

All of the motion data analyzed in this thesis was based on recordings executed in a motion capture lab using an optical infrared motion capture system, which required passive markers to be attached to the participants. The performances also had to be executed inside a restricted capture space. In other words, the recording situation did not resemble a normal performance situation for these music styles. Still, only professional performers participated in the studies, which hopefully compensated the influence of the lab environment (Naveda and Leman, 2008,

see also discussion in section 6.1.7). In addition, the performers were instructed to dance and play as normally as possible.

The motion and sound analyses in this thesis were based on measurements that represent both the musical sound and the performers' body motions as quantitative data. Quantitative methods allow us to study rhythmic patterns in great detail, whereas traditional musical notation, for example, is not nearly detailed enough to capture the rhythmic structures of the musical styles investigated here. If the analyses were based only on my own observation, as well, they might have been biased toward my personal experience or reflect my subjective perception. However, though quantitative data is "objective" to an extent, the features that are extracted, such as peaks in the sound and turning points in the motions, are based on qualitative considerations. Likewise, the results of the analyses were also evaluated qualitatively to ensure that the extracted features reflected perceptually relevant features. Huron (1999) points out that even though quantitative methods are useful tools in research on music and motion, they will never capture the full complexity of music. In music and motion research, then, the best result may derive from a combination of qualitative and quantitative approaches.

4.9 Summary

This chapter has presented an overview of the methods used in the studies included in this thesis. Technology related to tracking and storing motion data, data-processing methods, and motion and sound analysis was reviewed. While there are several possibilities for capturing motion, this thesis relied upon an optical infrared motion capture system from Qualisys. It included reflective markers and multiple cameras that both emitted infrared light and recorded the light reflected back from the markers.

The collected sound data was analyzed using the MIRtoolbox for Matlab (Lartillot and Toiviainen, 2007), and the collected motion data was analyzed using the MoCap Toolbox for Matlab (Burger and Toiviainen, 2013). Some statistical analysis was also used—mainly descriptive statistics and ANOVAs. It was clearly necessary to combine quantitative and qualitative methods to produce the most effective approach to the study of music and motion.

Chapter 5

Research Summary

This chapter introduces the five papers included in the thesis, including abstracts. It discusses the findings of the papers in relation to the three main areas of contribution of the thesis: rhythm, samba/telespringar, and method.

5.1 Introduction

A principal concern of this thesis is the link between music and motion in the processes of rhythm perception and production, particularly in the context of the interaction between perceived sonic rhythms and underlying reference structures. This thesis also explores the possibilities of different motion capture technologies for rhythm studies, and, among other things, the papers include insights specifically related to rhythm patterns in telespringar and samba. I am interested in the embodied knowledge that can be expressed though body motion in performers and perceivers as they experience musical rhythm via the interaction between physical rhythm (the actual sound signal), perceived sonic rhythm, and underlying reference structures.

Papers I and II have a methodological focus. Paper I discusses the use of an optical infrared motion capture system in rhythm studies of telespringar, and paper II evaluates sensors that provide accelerometer data. Papers III and IV present rhythm studies of telespringar that are based on two different motion capture recordings of professional telespringar performers. Paper V presents a rhythm study of samba that is based on a motion capture recording of two professional samba performers, a percussionist and a dancer.

5.2 Papers

5.2.1 Paper I

Studying Rhythmical Structures in Norwegian Folk Music and Dance Using Motion Capture Technology: A Case Study of Norwegian Telespringar
Haugen, M. R.
In *Musikk og tradisjon* (28), pp. 27–52, Novus forlag Oslo, 2014.

Abstract. Norwegian telespringar is often referred to as being in so-called *asymmetrical triple meter* — that is, the three beats in the measure are of uneven duration. Previous studies report that a systematic *long—medium—short* beat duration pattern seems to be a prominent feature of telespringar. This paper investigates how motion data can be incorporated into studies of rhythmical structures in Norwegian telespringar using motion capture technology. It is reported from two motion capture studies: first, a fiddler playing telespringar on a Hardanger fiddle; second, a couple dancing telespringar. Participants' movements were recorded using an advanced optical infrared motion capture system. Motion analysis of the fiddler's foot stamping confirms the *long—medium—short* beat-duration hypothesis. In addition, the fiddler's upper-body movements seem to be in synchrony with the bar level of the music. Motion analysis of the up/down movement of the body's center of gravity in telespringar dancing shows a consistent *libration pattern*. These results appear to suggest that prominent rhythmical features of telespringar are represented in both the fiddler's and the dancers' body motion. They also indicate that motion capture technology is an effective means of investigating music-related movements in telespringar.

This paper is based on a presentation I gave at a seminar titled "Technology versus Folk Music and Dance"[1] that was hosted by Norsk folkemusikklag (NFL) at Ole Bull Akademiet in 2014. It seeks to introduce marker based infrared motion capture technology and to illustrate how this technology might be used in rhythm studies of telespringar. Previous hypotheses involving the correspondences between music and motion in this style of music is taken as its point of departure, and the paper demonstrates that these music-related motions can be investigated with great precision and accuracy using today's technology. Because the paper is intended for readers with little or no knowledge of motion capture technology, it also includes details about the motion tracking system, the exportation of data, and post-processing. The main focus of this paper is on motion analysis, but it also includes a short section dealing with sound analysis.

Telespringar is one of the styles of traditional dance music in Sweden and Norway that are in so-called *asymmetrical meter*, and the paper begins with a short overview of the ways in which scholars from 1882 to the present have tried to document this quality using notation and other tools (see, e.g., Bengtsson et al., 1969; Groven, 1971; Kvifte, 1999; Mårds, 1999; Johansson, 2009; Ramsten, 1982). The intimate relationship between the music and the corresponding dance is often highlighted in related rhythm studies, thanks in large part to the pioneering work of anthropologist and ethnomusicologist Jan-Petter Blom. Based on his close observation and experience as a dancer and fiddler, Blom proposes a relationship between musical meter and the vertical motion of the dancer's center of gravity that he calls the *libration pattern* (see, e.g., Blom, 1981, 1993, 2006). Telespringar is normally played on the Hardanger fiddle, and it has also been suggested that the fiddler's foot stamping might indicate the musical meter as well (Ahlbäck, 2003; Blom, 2006; Kvifte, 1999). There might even be a relationship between the musical meter, the dancers' vertical motions, and the bowing motions involved with the Hardanger fiddle (Kvifte, 1987; Blom, 2006).

The analyses in this paper are based on two motion capture recordings, one of a fiddler

---

[1]Original title: "Teknologi versus folkemusikk og folkedans."

and another of two dancers. In addition to the various music and motion correspondences suggested by previous research, the present study also looked at the fiddler's recurring upper-body swaying. Unlike a previous motion capture study involving telespringar (Mårds, 1999), this study included a synchronized sound recording, enabling the analysis of the sound and motion relationship; it also incorporated a recording of a couple dancing in a ring formation, as they would normally do, rather than single dancers dancing in a straight line, as they did in the previous study. Lastly, the analysis of the dancers' motions was based on markers attached to their hips rather than their heads. Interestingly, the results of the analysis of the dancers' vertical motions did concur with previous findings (Mårds, 1999)—that is, the motion seems to follow a *long–short–short* duration pattern rather than a *long–medium–short* pattern. The fiddler's foot stamping, on the other hand, showed a very stable *long–medium–short* pattern, supporting the view that the meter in telespringar consists of beats of uneven duration.

The results of this study support the view that performers' body motions should be incorporated into rhythm studies of telespringar. In addition, it is clear that motion capture technology is an effective means of investigating these music-related motions in telespringar.

5.2.2 Paper II

Evaluating Input Devices for Dance Research
Haugen, M. R., and Nymoen, K.
To appear in *Post-proceedings of the 11th International Symposium on Computer Music Multidisciplinary Research*, The Interdisciplinary Centre for Computer Music Research (forthcoming).

Abstract. Recording music-related motions in ecologically valid situations can be challenging. We investigate the performance of three devices providing 3D acceleration data, namely Axixity AX3, iPhone 4s and a Wii controller tracking rhythmic motions. The devices are benchmarked against an infrared motion capture system, tested on both simple and complex music-related body motions, and evaluations are presented of the data quality and and suitability for tracking music-related motions in real-world situations. The various systems represent different trade-offs with respect to data quality, user interface and physical attributes.

As pointed out in chapter 4, marker-based infrared motion capture systems have proven very useful to the study of music-related motion, because they can provide position data with very great spatial and temporal precision and accuracy. However, because the systems include visible markers, and recordings often must be done in a motion capture lab with cameras that surround the participants, their ecological validity may be challenged. In addition, the number of participants that one can capture simultaneously is restricted due to the limited capture space and the complication of marker occlusion—in some situations, then, inertial sensors can be a viable alternative to optical motion capture systems. In paper II, we evaluated the performance of three different inertial sensor systems providing 3D motion data: Axixity AX3, iPhone 4s,

and a Wii controller. We wanted to compare the systems' abilities to capture music-related motion of different characters and complexities. Hence, we set the sensors to track various body motions performed in synchrony with both a metronome and pre-recorded samba music. The sensors tracked impulsive motion to the metronome, jumps to the metronome, simple shaker motion to samba music, and complex hip motion in samba dance to samba music. For reference, the motions were also recorded using an optical motion capture system from Qualisys. Since the data streams from the different devices were not synchronized, they had to be aligned, so impulsive motions were executed with all of the devices at once at the beginning and end of each recording, producing unambiguous peaks in all of the data steams at the same time that served as common reference points.

Investigations of the noise level from the different sensors revealed that the Wii data seemed to be noisier than the data from the AX3 and the iPhone. A plot of the raw data from the Wii revealed jagged lines, which indicated a low bit depth. This is possibly the reason why the Wii data was much noisier. In the motion analysis this was compensated for by applying a (Butterworth) smoothing filter (see section 4.5.2).

All of the devices did track the impulsive motion to the metronome and also the general tendencies of the shaker motion to the samba music. The AX3 and the Wii also tracked the jumps to the metronome. The iPhone acceleration data was limited to $\pm$ 2 G, so it did not fully reflect the magnitude of the acceleration in the impulsive motion. This is probably why the iPhone data also indicated unclear peaks in the shaker recording and could not capture the jumps in the jump recording. While none of the sensors seemed to be able to capture the complex samba dance motion, a qualitative evaluation of the fluctuation in acceleration indicated a slight increase in acceleration around the corresponding second and fourth sixteenth notes in the music.

The last part of the study looked at the extent to which the data from a single AX3 device could be temporally aligned with an audio recording without using an optical infrared system. Temporal alignment is, of course, critical to rhythm studies. Sound-producing impulsive motion were executed with the device in the beginning and end of the recording, producing unambiguous peaks in both the sound and motion data streams at the same time that served as common reference points. Using the common reference points in the two data streams, we were able to align data from AX3 to the audio recording by itself. This meant that music-related motion could be recorded in ecologically viable environments using only an audio recorder and an input device.

We concluded that, while optical infrared motion capture systems are unrivaled in terms of data quality, inertial systems may be preferable in situations where ecological validity is important. The data from the inertial systems AX3 and Wii was proven to be useful for analyzing rhythmic structures in simple music-related motions, suggesting that inertial systems can be used for capturing simple motions in real-world situations.

5.2.3   Paper III

Asymmetrical Meter in Scandinavian Folk Music and Dance: A Case Study of Norwegian Telespringar
Haugen, M. R.

Abstract. Certain traditional Norwegian and Swedish dance tunes in triple meter are referred to as being in so-called *asymmetrical meter* — that is, the three beats in the measure are of uneven duration. Norwegian telespringar is recognized for a type of asymmetrical meter featuring a systematic *long—medium—short* duration pattern at beat level. These systematic microtiming patterns are often described in terms of deviations from an underlying isochronous pulse. However, it has been argued that performers' body motion may offer a more perceptually relevant structure of reference than an abstract fixed clock pulse. This study investigates whether the asymmetrical beat patterns previously shown in telespringar music are also represented in the body motion of performers who are playing and dancing. It is reported from two motion capture studies: first, a fiddler playing telespringar on a traditional Hardanger fiddle; second, a couple dancing telespringar. Motion analysis of the fiddler's foot stamping indicates a very regular *long—medium—short* beat pattern. In addition, the fiddler's upper-body swaying and the vertical motion of the body's center of gravity in telespringar dancing are in synchrony with the bar level of the music. The fiddler's foot stamping confirm the *long—medium—short* beat duration hypothesis and support the view that the systematic microtiming features in telespringar are not a matter of deviation from an underlying isochronous pulse. Instead, they actually constitute an essential feature of telespringar.

This paper draws upon the data set from paper I to argue that the underlying pulse in telespringar should be understood as asymmetrical in and of itself, not as deviating from a pulse consisting of isochronous beats. The paper starts by noting that the pulse level in music is often externalized through body motions (for example, head nodding and foot stamping), and those motions can therefore be used in rhythm studies of music to locate the meter. Accordingly, and particularly given the close relationship between music and dance, the rhythm structures in telespringar should be understood in relation to the corresponding dance. In order to sketch the duration pattern at the beat level based on the sound, I applied a peak-picking function that detected sudden increases in signal energy (onset detection, see also section 4.2.1). This audio analysis indicated a *long–medium–short* duration pattern, though this result should be used with caution, since the sound of telespringar is very complicated, and the peaks in the audio signal are therefore ambiguous.

The body motions in question here were the fiddler's foot stamping and the dancers' vertical motion of their centers of gravity (*libration curves*). The fiddler's foot stamping was estimated based on a marker attached to his right heel, and the libration curves were estimated based on a marker attached to the dancer's lower back. The fiddler's foot stamping indicated a very stable *long–medium–short* duration pattern at the beat level, supporting the claim that the underlying pulse in telespringar consists of beats of uneven durations. The *libration curve* also seemed to

be very stable at the measure level. Previous studies (Blom, 1981, 1993, 2006) have concluded that the turning points in the dancers' libration curves correspond to the underlying meter in telespringar; however, in this case, the libration curves indicated a *long–short–short* duration pattern. That is, the duration of the first beat, based on the dancers' vertical motion, occupies an average of 47 percent of each measure. Based on this result, I explore whether the dancers' vertical motion ought to be understood as an additional metrical level—that is, 2/4 time. If so, the dancers' libration pattern could be seen as a countermotion in relation to the underlying asymmetrical meter indicated by the fiddler's foot stamping. These conclusions are at best preliminary, because they are based on but one recording of a fiddler and one recording of two dancers.

5.2.4   Paper IV

Investigating Periodic Body Motions as a Tacit Reference Structure in Norwegian Telespringar Performance
Haugen, M. R.
To appear in *Empirical Musicology Review* (forthcoming).

Abstract. The pulse level in music is often described as a series of isochronous beats that provides an underlying reference structure against which we perceive rhythmic patterns. This notion is challenged, however, by music styles that seem to feature an underlying reference structures that consists of beats of uneven duration, such as certain traditional Scandinavian dance music genres in so-called *asymmetrical meter*. This study investigates periodic body motion as a reference structure in a specific style of traditional Norwegian dance music called *telespringar*. The intimate relationship between music and motion is often highlighted in rhythm studies of telespringar, so this study encompasses both sound and motion analyses. It is based on a motion capture study of three telespringar performers; one fiddler and two dancers. Motion analysis of the fiddler's foot stamping indicates a stable *long–medium–short* duration pattern at beat level. Motion analysis of the dancers' vertical motion of the hips revealed a periodic pattern in synchrony with the measure level. Telespringar playing is very ornamented, which made it difficult to determine the precise physical onsets of musical events, and consequently it was not possible to determine a rhythmic pattern based on an analysis of the physical sound. Still, the correspondence between the fiddler's periodic foot stamping and the dancers' periodic vertical motions appeared to persist throughout the whole performance. This result implies that, rather than being primarily derived from the sound itself, the underlying rhythmic structures in telespringar depend upon a shared and embodied knowledge of the underlying asymmetrical reference structure that is implicit in the production and perception of telespringar.

This paper follows up the work presented in papers I and III but is based on a new data set: two

motion capture recordings of three professional telespringar performers, one fiddler and two dancers. Unlike the study presented in papers I and III, the fiddler and the dancers in this study were recorded simultaneously, and two different tunes were recorded as well.

In this paper, I argue that meter is not only induced by features in the musical sound, finding that underlying reference structures such as pulse and meter can also be parts of both performers' and perceivers' tacit knowledge of a music style. Moreover, I propose that these underlying structures can be intimately related to performers' body motion. For my case study, I investigate the underlying reference structure in telespringar through an analysis of performers' body motions in a telespringar performance: the fiddler's foot stamping and the dancers' vertical body motion.

According to existing hypotheses and findings, the fiddler's foot stamping and the vertical motion of the dancers' center of gravity (the *libration pattern*) are related to underlying structures in telespringar. Here, the fiddler's foot stamping revealed unambiguous acceleration peaks related to the points in time when the feet (toes) hit the floor. As previously, the duration pattern based on the fiddler's foot stamping revealed a stable *long–medium–short* pattern, indicating that the underlying structure in telespringar is asymmetrical.

I also hypothesized that beat-related sonic events correspond to the underlying reference structure in telespringar. However, the fiddle is a bowed instrument and the sound is produced through continuous bowing motion that does not generate acute amplitude peaks. While I perceived beat-related sonic events that corresponded to the fiddler's foot stamping when I listened to the recording, the onsets of those events were not possible to pinpoint based on analyses of the physical sound. This also evokes previous studies that have wondered whether one can talk about onsets as points in time in telespringar, since the related style of playing is very ornamented, and the beginnings of new musical events often include slides and grace notes (Johansson, 2010). Interestingly, people familiar with the style of telespringar have no problem identifying and dancing to its "beat."

According to Blom's *libration hypothesis* (Blom, 1981), the pattern of the dancers' vertical motion of the body's center of gravity is related to meter. More specifically, the dancers perform two down/up motions in each measure, and the first "dance beat" should be measured from the first peak in the libraiton curve to the second peak (down/up), the second "dance beat" from the second peak to the second trough (down), and the third "dance beat" from the second trough to the first peak in the next measure (up). In the study reported in paper IV, I used motion data from two dancers in two different recordings. When estimating beat duration patterns based on the turning points in the dancers' vertical motions, as described above, those patterns differed from the fiddler's foot stamping. In addition, they differed from one another. This finding is, to some extent (the first beat duration occupies close to 50 percent of the measure), in line with the findings in papers I and III, and with Mårds's findings (Mårds, 1999). However, looking at the video of the recordings, I did not get the impression that the dancers were making the countermotions that were introduced as a possibility in paper III. On the contrary, the dancers' motions, the fiddler's foot stamping, and the music itself seemed to follow exactly the same beat.

As a consequence, I sought an alternative interpretation for the dancers' libration curves and decided to demarcate them according to the fiddler's foot stamping rather than their own turning points. This showed that the *shape* of the curves was very stable when demarcated in this way,

and that the beginning of a measure seemed to fall between turning points. The duration of the first beat seemed to correspond to a small vertical S-shape; the duration of the second beat, to a deep down/up motion; and the duration of the third beat, to an up/down motion. Based on these findings, I then presented an alternative libration pattern—one that was identical to Blom's libration hypothesis but phase-shifted in its position in relation to the beat level.

I concluded from all of this that the underlying reference structure in telespringar is inherently asymmetrical, and that both the fiddler and the dancers seem to relate to this shared reference, which is indicated by the fiddler's foot stamping and corresponds to the shape of the dancers' libration curves. The fact that an unambiguous temporal beat position could not be derived from the physical sound implies that the meter is not only induced by the properties of the sound but derives from shared and embodied knowledge among the performers and perceivers who are familiar with telespringar. While this paper specifically deals with telespringar, I believe that the notion of a tacit reference structure that is closely related to body motion would be relevant to other music styles as well.

5.2.5 Paper V

Rhythmical Structures in Music and Body Movement in Samba Performance
Haugen, M. R., and Godøy, R. I.
In *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, Yonsei University, 2014.

Abstract. Samba groove is often characterized by its complex rhythmical patterns. Recent studies, based on audio recordings of samba music, report that the 3rd and the 4th 16th-notes are played slightly ahead of their corresponding quantized position, and that this seems to be a prominent feature of samba groove. Considering that samba derives from a culture where music and body motion are intrinsically related, may suggest that we should include both sound and motion data in studies of its rhythmical structures. In this paper we investigate whether the microtiming features, previously shown in samba music, may also be represented in the body motion of performers playing and dancing samba. We report from a motion capture experiment where two skilled samba performers, a percussionist and a dancer, were recorded using an advanced optical infrared motion capture system. Our audio analysis confirms the existence of systematic microtiming patterns on the 16th-note level in samba music. In addition, motion analysis of the percussionist's heel tapping and the dancer's steps revealed motion patterns in synchrony with the systematic microtiming features found in samba music. These observations support the view that the systematic micro timing of 16th-notes in samba playing is not a deviation from an underlying perceived pulse with isochronous subdivisions, but rather constitutes an essential feature of samba.

In this paper we investigated the correspondences between music and body motion in a samba performance. The aim of this paper was somewhat similar to that of paper IV—that is, to investigate the underlying reference structure as embodied knowledge—but this study specifically addresses underlying subdivisions at the level of the sixteenth note. Rhythm studies of samba, mainly based on audio recordings, report that samba seems to be characterized by *systematic microtiming* at the level of the sixteenth note (see, e.g., Gerischer, 2006; Gouyon, 2007; Naveda et al., 2009)—that is, the third and fourth sixteenth note in a given beat seem to be played slightly ahead of their quantized position. However, as I have pointed out elsewhere, it makes more sense to investigate rhythm in relation to performers' body motions than to an abstract fixed clock pulse (see, for example, section 3.4.4), and this study did just that.

It drew upon a sound and motion capture recording of two professional samba performers, a percussionist and a dancer. The percussionist played a samba groove on a traditional Brazilian hand drum called *pandeiro*, and the dancer performed a dance in *samba no pé* style. Since all of the sixteenth notes were played and consequently present in the physical sound, their temporal positions could be estimated using onset detection to locate peaks on an amplitude envelope curve. The sixteenth-note durations were then estimated based on the inter-onset-intervals (IOIs) and showed a *medium–medium–medium–long* duration pattern, confirming systematic microtiming at the sixteenth-note level. In addition, the position of the first and fourth sixteenth notes seemed to be in synchrony with the percussionist's foot motion, and the position of the first, second, and third sixteenth notes seemed to be in synchrony with the dancer's feet. We concluded from this that the systematic microtiming at the sixteenth-note level in samba did not represent a deviation from an underlying pulse with isochronous subdivisions but rather constituted an essential structural feature of samba.

5.3   Contributions

This section discusses the contributions of the papers in relation to the three main hypotheses of this thesis:

1. Rhythm:

   The underlying reference structure(s) in the experience of rhythm is intimately related to motion.

2. Telespringar and samba:

   The underlying beat level in telespringar and the underlying sixteenth-note level in samba are non-isochronous.

3. Methodology:

   Motion capture–based analysis can provide valuable insight into the relationship between rhythm and motion in music.

In the following I will present the findings of the papers in relation to the main areas: rhythm, telespringar, samba, and method.

5.3.1   Rhythm

A main point in these studies is that the underlying reference structure in the experience of rhythm is intimately related to body motion and should be regarded as an intrinsic part of the music as common knowledge among people conversant with the given musical style. The sound analysis of telespringar tunes that is presented in papers I, III, and IV did not reveal any unambiguous onsets of new events, however the underlying reference structure seemed to be externalized both by the fiddler's integrated foot stamping and also the dancers vertical motion. In the samba performance discussed in paper V, the underlying reference structure at the sixteenth-note level seemed to be present in the percussionist's and the dancer's feet motions. In addition, the percussionist performed a small hand gesture just before he started playing, and the duration of this gesture was equal to the beat level in the groove that followed. These findings suggest that the underlying structure of samba resides within the percussionist before he starts playing. Rather than following from the rhythm patterns in the sound, then, the underlying reference structures are dependent on tacit knowledge and are used, not only in process of perception, but also to *produce* the particular rhythm patterns (in both the dance and the sound). The fact that these structures are expressed through body motion also supports the view that underlying reference structures and body motion are intimately related to one another.

The underlying reference structure in the experience of musical rhythm is often presumed to consist of isochronous time series, and non-isochronous sonic rhythm patterns that do not coincide with such a structure have been referred to as deviations. The findings presented in the papers included in this thesis support the view that the underlying reference structure is not necessarily isochronous. Papers I, III, and IV investigated rhythm patterns in telespringar and showed that the fiddler's foot stamping seemed to follow a systematic non-isochronous pattern at the beat level that supports existing theories regarding a non-isochroous pulse level, or a so-called *asymmetrical meter*, in telespringar. Paper IV presents findings that indicate that the dancers relate to the same reference as that indicated by the fiddler's foot stamping. In paper V, on rhythm structures in samba, the non-isochronous pattern at the sixteenth-note level that had arisen in audio analyses was also located in the percussionist's foot stamping and the dancer's steps. These findings suggest that non-isochronous duration patterns are also present in performers' body motion, providing support for the hypothesis that the underlying reference structure at the sixteenth-note level in samba is non-isochronous.

5.3.2   Rhythm Patterns in Telespringar

Literature on telespringar states that telespringar is in a so-called *asymmetrical meter* that follows a *long–medium–short* duration pattern at the beat level. Papers I, III, and IV evaluate this claim by incorporating the performer's body motion into the analysis. As already described, the fiddler's foot stamping in the motion capture recording followed a very stable *long–medium–short* duration pattern, supporting the view that the underlying structure at the beat level in telespringar is asymmetrical. It might be worth noting that even though the fiddlers follow the same overall duration pattern at the beat level, the averages related to each beat's percentage of the entire measure differs. This indicates that the exact ratio numbers (the three beats' average duration as a percentage of the measure) should not be understood as fixed (see
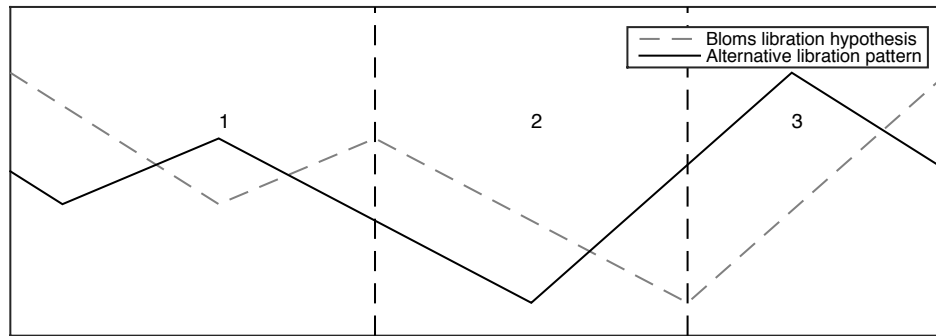
Figure 5.1: The libration hypothesis presented by Blom (1981, 1993, 2006) is represented by the dashed line in the figure, and the alternative libration pattern presented in paper V is represented by a solid line. Note that the shape of the alternative libration pattern is identical to Blom's libration hypothesis, but its position in relation to the beat level is phase shifted.

also discussion in section 6.1.4).

Paper IV represents an important contribution to the field of music and dance studies of telespringar. The analysis of the dancers' vertical motion in relation to the fiddler's regular foot stamping produced an alternative interpretation of the libration hypothesis presented by Blom (see, e.g., Blom, 1981, 1993, 2006). Blom's hypothesis suggests that certain *turning points* in the dancers' vertical motion of the body's center of gravity are related to the *beat positions* indicated by the fiddler's foot stamping. However, my findings did not support this theory but instead indicated that the turning points seemed to fall between the fiddler's foot stamping. Still, the *shape* of dancers' libration curves were very stable indicating that specific *motion shapes* seem to correspond to the *beat durations* (IOIs) between the fiddler's foot stamping (see figure 5.1). The overall shape of my alternative libration pattern is identical to Blom's but phase shifted in relation to the fiddler's foot stamping.

5.3.3 Rhythm Patterns in Samba

Previous studies have noted that systematic microtiming at the sixteenth-note level seems to be a prominent feature of the samba groove. Gerischer (2006) found a *medium–short–medium–long* duration pattern, whereas Gouyon (2007) and Naveda et al. (2009) describe how the third and fourth sixteenth notes seemed to be played slightly ahead of their quantized position. "Converting" such observations from a "deviation approach" to the pattern concept—that is, looking at patterns instead of deviations—means that at least the fourth sixteenth note is longer than 25 percent of the measure. The results in paper V indicate a *medium–medium–medium–long* duration pattern, based on both audio and motion analysis (the percussionist and dancers' foot motions). This supports the idea that the fourth sixteenth note in samba fills a prominent structural role, being both longer in duration and expressed in the percussionist's foot motion. Interestingly, the duration of the second sixteenth note in our study does differ from Gerischer's findings. This maybe due to regional differences—Gerischer did her fieldwork in the Bahia region in northeast Brazil, whereas the performers in our study were from Rio.

By including both sound and motion data in our rhythm analysis, we were able to investigate

the systematic microtiming features in samba in relation to body motion instead of a constructed timeline. The motion analysis in this study showed that the systematic microtiming pattern previously found in the sound of samba was also present in the performers' body motion. This result support the view that systematic microtiming in samba should not be understood as a deviation from an underlying pulse with isochronous subdivisions. Instead, the underlying (metrical) sixteenth-note level in samba should be understood as being non-isochronous.

5.3.4 Method

Paper II investigates the performance of inertial systems when capturing music-related motion. Unlike optical infrared motion capture systems, for example, these sensors are not based on reflected light and are not restricted by a limited capture space. They may also be more suitable for capturing multiple people simultaneously, given the challenge of marker occlusion. Although not as accurate and precise as data from an infrared motion capture system, the inertial-system findings demonstrate that many music-related motions can be captured using inertial sensors. It is true that the simplest motion seem to be captured with the highest accuracy and precision—for example, a duration pattern describing the general tendencies of a simple shaker motion could be calculated based on obtained acceleration peaks. General tendencies related to acceleration amplitude could also be observed, including the fluctuation in acceleration in samba dance.

Since optical infrared motion capture systems are unrivalled when it comes to accuracy and precision, and I only recorded two or three performers simultaneously, Qualisys was used in the studies presented in papers III, IV, and V. Although the ecological validity of a study might be undermined by the controlled environment of a motion capture lab, it has been proposed that the use of professional performers counterbalances the influence of the artificial lab environment (Naveda and Leman, 2008).

The motion capture technologies available today represent incredible possibilities for investigating music-related motions with high precision. Paper I shows how optical infrared motion capture systems can be used in investigations of music-related motions in telespringar. In musical styles characterized by non-isochronous underlying structures, it has been suggested that performers' body motions may offer a more perceptually relevant reference structure than an abstract fixed clock pulse (Kvifte, 2004). Paper I illustrates how such body motions can be measured using an optical infrared motion capture system, and how the collected motion data can contribute valuable insight into the relationship between rhythm and motion in music.

5.4 List of Publications

Papers Included in the Thesis

  I Studying Rhythmical Structures in Norwegian Folk Music and Dance Using Motion Capture Technology: A Case Study of Norwegian Telespringar
Haugen, M. R.
In *Musikk og tradisjon* (28), pp. 27–52, Novus forlag Oslo, 2014.

  II Evaluating Input Devices for Dance Research
Haugen, M. R., and Nymoen, K.
To appear in *Post-proceedings of the 11th International Symposium on Computer Music Multidisciplinary Research*, The Interdisciplinary Centre for Computer Music Research (forthcoming).

  III Asymmetrical Meter in Scandinavian Folk Music and Dance: A Case Study of Norwegian Telespringar
Haugen, M. R.
In *Proceedings of the Ninth Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*, pp. 432–436, Royal Northern College of Music, Manchester, 2015.

  IV Investigating Periodic Body Motions as a Tacit Reference Structure in Norwegian Telespringar Performance
Haugen, M. R.
To appear in *Empirical Musicology Review* (forthcoming).

  V Rhythmical Structures in Music and Body Movement in Samba Performance
Haugen, M. R., and Godøy, R. I.
In *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, Yonsei University, 2014.

Other Papers

- Danielsen, A., M. R. Haugen, and A. R. Jensenius (2015). Moving to the beat: Studying entrainment to micro-rhythmic changes in pulse by motion capture. In *Timing and Time Perception* 3(1-2), pp. 133–154.

- Nymoen, K., M. R. Haugen, and A. R. Jensenius (2015). MuMYO – Evaluating and Exploring the MYO Armband for Musical Interaction. In *Proceedings of the International Conference on New Interfaces For Musical Expression*. Louisiana State University.

- Godøy, R.I., M. Song, K. Nymoen, M. R. Haugen, and A. R. Jensenius (2016). Exploring sound-motion similarity in musical experience. In *Journal of New Music Research*.

Chapter 6

General discussion

This chapter discusses the work presented in this thesis. It includes a summary of the themes dealt with in the thesis, a discussion of its findings and conclusions, and some directions for future work.

6.1 General Discussion

This section presents a general discussion of some central topics of the thesis. It also discusses some limitations of the studies.

6.1.1 Underlying Reference Structures and Music Culture

Much of this thesis concerns the interaction between perceived sonic rhythm and underlying reference structures in the experience of rhythm in a musical context. Because such underlying structures are mental constructs of both perceivers and performers, we must understand the different factors that inform these constructs. Certain dynamic approaches have related them to regularities in the sonic signal (see, for example, Large and Jones, 1999). Others see them as very dependent on perceivers' and performers' familiarity with the relevant music culture (Clayton, 2013; Drake and El Heni, 2003; Hannon and Trehub, 2005).

In section 3.4.2, I argued for the importance of music culture in the context of so-called *standard rhythms* (also called *time line*), particularly in music styles where the underlying reference structure is not derived from sonic *pulses* in the music to which one can entrain but from specific sonic *rhythm patterns* that reveal an underlying pulse (see, for example, Agawu, 2003; Jones, 1959; Nielsen, 1985; Nketia, 1986). Likewise, the perception of meter in Scandinavian folk music is based on cultural learning, but instead of centering upon a certain standard rhythm, it involves a more complex implied pattern (Kaminsky, 2014; see discussion in section 2.6).

In section 2.2, I termed musical styles with an intimate relationship to a specific dance *music–dance*. Papers I, III, IV, and V all highlight the intimate relationship between music and dance in telespringar and samba, demonstrating, as well, that the rhythm structures in telespringar and samba should be understood in relation to the corresponding dances. This does not mean that the corresponding dance is the *only* corresponding motion that is possible in these music styles. As I noted in chapter 2, music perception and cognition encompass numerous music and motion correspondences. For example, one could overtly imitate the percussionist's

body motion while playing the *pandeiro* or trace the perceived contour of the melody in the tele-springar tune. The point is instead that, in these music–dance styles, the music and the dance have evolved together under a condition of mutual influence, so for performers and perceivers who are familiar with the style, the specific way of dancing seems to be intrinsically related to the specific way of playing, and vice versa. Naveda (2011) points out that the Europeans' early descriptions of Afro-Brazilian samba reveal a poor understanding of the relationship between the music and the dance in these music styles: "Participants without the tacit knowledge of how movements are related to sonic patterns will listen, move and understand it differently" (Naveda, 2011, p. 51). Any culturally relevant interpretation of music–dance genres must acknowledge the impact of tacit music-cultural knowledge and experience.

6.1.2 Representing Underlying Reference Structures as *Duration Patterns*

It is one thing to understand how various factors influence the mental construction of an underlying reference structure, and it is quite another to determine how to represent such structures. One of the main aspects of the work included in this thesis is that underlying structures do not necessarily consist of isochronous sequences. It is therefore misleading to refer to non-isochronous metrical sequences as "deviations" from an artificial fixed clock pulse. Kvifte (2004) suggests we think of non-isochronous time series as *patterns* instead (Kvifte, 2004, p. 62). Accordingly, I represent the underlying reference structures in both samba and telespringar as *duration patterns* in this thesis. Duration patterns have already been used to represent the underlying *asymmetrical meter* in telespringar (see, for example, Groven, 1971; Johansson, 2009; Kvifte, 1999; Mårds, 1999), as well as the fluctuation in sixteenth-note durations in samba (Gerischer, 2006), so my work aligns well with earlier efforts in this regard.

The duration patterns in papers I, III, IV, and V are all presented using both qualitative and quantitative descriptions. The quantitative version refers to ratio numbers—that is, the underlying metrical beats/subdivisions' average IOIs are presented as percentages of the measure in telespringar and beat in samba. Based on the view that these patterns are stable and cyclic, mean duration values are a sensible way to represent these structures. The fact that all metrical beats types (beat 1 durations, beat 2 durations, and so on) were normally distributed advocates for representing beat durations by their means. The corresponding low standard deviations (SDs) accompanying the averaged beat/subdivision durations also indicate that the patterns are stable over time (the use of descriptive statistics in this thesis is presented in section 4.7.1).

The qualitative descriptions of the duration patterns rely upon three categories: *short*, *medium*, and *long*. These labels are based on the quantitative representations described above and allow us to describe alternative duration patterns with slightly differing ratio numbers as part of the same *perceived* pattern (pattern recognition will be discussed further in section 6.1.4). However, this categorization descibes the mean beat durations according to one of three categories without explaining anything about how these categories are proportionally related to one another (we know that medium is shorter than long, but not how much shorter). For example, in paper IV, the metrical duration pattern of one of the telespringar tunes (derived from the fiddler's foot stamping) is represented by the ratio numbers 39–38–23. Despite the small difference between beats 1 and 2, the foot stamping is described as following a stable *long–medium–short* duration pattern. The difference between beats 1 and 2 is statistically significant and must be acknowl-

edged, but it is obvious that beat 3 is considerably shorter than the others. If we only looked at the ratio numbers, we might have the impression that the two first metrical beats are equal in duration; if we only looked at the categorizations, we might have the impression that the second metrical beat fell exactly between the first (long) and third (short) beat durations. This demonstrates that a representation of duration patterns can benefit from a combination of qualitative and quantitative descriptors.

6.1.3 Underlying Reference Structure as Metrical Shape

The calculations of the beat durations in telespringar in papers I, and III, and in samba in paper V, rely upon directional changes—that is, *turning points*—in the participants' vertical motion. In the samba recording, this is based on the percussionist's and dancer's vertical motion of their feet, and in the telespringar recordings, this is based on the fiddler's vertical foot stamping and the dancers' vertical motion of the lower back. The results presented in paper IV, on the other hand, extend this framework by demonstrating that it is not the *turning points* in the telespringar dancers' vertical motion (or libration curve) that relate to the temporal points of an underlying meter, but rather that the *shape* of the libration curve corresponds to the underlying *beat durations*. This may suggest that how we get from point to point is as important as the duration between the metrical reference points. One may ask whether certain motion trajectories *between* the metrical beats in telespringar, and the metrical subdivisions at the sixteenth-note level in samba, provide information about the *shape* of these metrical beat durations?

Thinking of underlying reference structures as *trajectories between beats* can be contextualized further using Godøy's (2003) *motor-mimetic* perspective on music perception and cognition (see also section 2.4.2), which describes an intimate relationship between musical sounds and their corresponding simulated sound-producing actions. It has been suggested that a sound might be perceived as a sonic *shape*, which might in turn imply a corresponding imagined action with a similar *shape*. If the underlying reference structure consists not only of *points in time* but also of the *trajectories between the points*, the underlying structure might also be perceived as a *shape*. Godøy's (2003) concept of *action/sound shape* correspondences may further be extended to a metrical *motion/duration shape* correspondence—encompassing not only the *duration*, but also the trajectories (as actual or simulated motion) between metrical beats. This is also related to Danielsen's (2010a) *beat bin* model. Here she argues that instead of thinking of the beats of the pulse as a series of *points* in time they should be understood as *shapes* (see section 3.3). However, instead of focusing on pulse beats as shapes, I want to direct attention to the importance of the trajectories *between* metrical beats. The common shape of the durations between beat points may work as an underlying reference trajectory, and, such trajectories may also be intimately related to body motion. This may also be related to Waadeland's (2000) "wave metaphor" which he uses to link pulse to a corresponding *motion curve* where pulse points coincide with local minimal points in the motion curve (see figure 3.2). My hypothesis, on the other hand, is that the underlying reference trajectories (as actual or simulated motion) do not necessarily have a rise/fall shape that goes from turning point to turning point in a motion curve, but that each metrical beat duration has a corresponding *metrical trajectory* with a certain *shape*. Since the music and the dance in both samba and telespringar are intrinsically related, such metrical shapes might relate to periodic motion patterns in the dance. If this is the

case, then perceivers' and performers' implicit knowledge of the underlying reference structure in samba and telespringar might incorporate knowledge about the underlying *metrical shape*.

As already mentioned, the results in paper IV support the idea of an underlying reference trajectory in telespringar, suggesting that a possible *metrical shape* may be intimately related to the vertical motion in the corresponding dance. Is it possible to find a similar relationship in samba? So far, we do not know. Previous studies, primarily based on audio analyses, have noted that the sixteenth-note level seems to be important to the samba groove (see, for example, Gerischer, 2006; Gouyon, 2007; Naveda et al., 2009). This is also supported by the findings in paper V, indicating that the non-isochronous duration pattern found in the sound of samba is also found in the percussionist's foot motion and the dancer's footsteps. Since the foot motion of the percussionist and the dancer in samba have an impulsive character, they most likely represent the underlying beat positions and not an underlying reference trajectory (equivalent to the fiddler's foot stamping in telespringar). Naveda (2011, p. 36) points out that analyses of the motion in samba dancing are rare, but it has been suggested that the pulsating rhythm originates in the torso. Accordingly, it is suggested that future work on this genre should also include the analysis of dancers' torso motion in relation to the sounding music.

6.1.4   Pattern Recognition and "Metrical Shape" in Telespringar

As discussed in section 6.1.2, the duration pattern in telespringar can be represented both as ratio numbers, according to each beat duration's percentage of the complete measure, and as belonging to the categories *long*, *medium*, or *short*. In the telespringar study in paper IV, the fiddler's foot stamping follows a stable *long–medium–short* beat duration pattern in both of the tunes that he performed (see also discussion in 6.1.2). Interestingly, the percentage ratio between the beats is not the same between the two tunes. The first is 39–36–25 and the second is 39–38–23. Why might this be?

As previously pointed out, metrical interpretation does not only depend on (or derive from) the perceived sonic rhythm, but also it also depends on culture-specific categorization. Johansson (2010) introduces the concept of *rhythmic tolerance* to describe an aspect of the relationship between underlying reference structures and sonic rhythm in traditional asymmetrical *pols/springer* styles of Norway and Sweden. He points out that even though the inter-onset intervals (IOIs) between pulse-related sonic events vary between measures, we do not necessarily perceive the groove as unstable. The fact that these rhythm patterns are still acknowledged as falling within the stylistic boundaries of the music implies a flexibility of stylistic categories, Johansson concludes.

Kvifte (2007) suggests that instead of thinking of metrical information as something that can be computed only from the sound signal using certain general rules, we should allow that the perception of meter may be more of a *pattern-recognition* task that is closely related to a perceiver's experience (Kvifte, 2007, p. 81). According to the *metrical shape* hypothesis suggested above, such pattern recognition may not only apply to duration patterns, but also to the trajectories corresponding to the beat durations. In telespringar, as well, those trajectories may be intimately related to the dancers' libration curves, as the findings of paper IV appear to indicate. Here, the dancers' libration curves were not only divided into *measures* according to the fiddler's foot stamping, but also into *beats*. The chunks corresponding to the first, second, and

third beats in each tune were plotted in the same graph, respectively. A qualitative evaluation of the dancers' libration curves at the beat level suggested that even though the average beat durations differed between the two tunes, the libration curves seemed to have the same shape. This indicates that even though the ratio between the three beats in telespringar may differ slightly, the underlying shape relating to each beat duration remains the same. In turn, this suggests that the pattern-recognition task involved in telespringar may encompass shapes related to beat durations that can be "stretched" while remaining within the stylistic categories of telespringar. More recordings must be analyzed in order to determine whether the relationship between beat durations and libration shape is in fact a general tendency of telespringar.

6.1.5 Representations of Music and Dance

One of the aims in this work has been to highlight the intimate relationship musical rhythm and body motion. I have also argued against the preconception that underlying reference structures in the experience of rhythm must be isochronous. Samba and telespringar are interesting cases in this regard, for several reasons: (1) they both derive from music cultures where music and dance are intrinsically related (I call these genres *music–dance* in chapter 2); (2) it is commonly understood that the rhythm structures in samba and telespringar music should be understood in relation to the corresponding dance (see, for example, Blom, 1981, 1993; Kvifte, 2004; Naveda, 2011; Naveda and Leman, 2008; and (3) samba and telespringar are often characterized by complex rhythm patterns—that is, *asymmetrical meter* in telespringar (see, for example, Groven, 1971; Kvifte, 1999 and *systematic microtiming* at the sixteenth-note level in samba (see, for example, Gerischer, 2006; Gouyon, 2007; Naveda et al., 2009). In order to document the non-isochronous structures in telespringar and samba, the metrical beats in telespringar and the metrical subdivisions in samba were represented as mean beat and sixteenth-note duration patterns, respectively.

As already indicated, the results from the studies presented in this thesis should not be interpreted as providing *fixed* ratios between the metrical beats or metrical subdivisions that describe the underlying reference structures in telespringar and samba, respectively (See section 3.5 for a discussion of *metrical subdivision*). The numbers presented are *representations* of the music's underlying structure, as expressed through performers' body motions. The point of these results is to demonstrate both that these music styles are featured by non-isochronous underlying reference structures, and that these structures exist as *implicit knowledge* in performers and perceivers who are familiar with the style of music. In addition, the recognition of these underlying structures relies upon an *embodied* knowledge that can only be acquired through bodily experiences.

The motion data analyzed in this thesis was mainly collected using an infrared motion capture (mocap) system. As discussed in section 4.4, these systems can provide position data with great precision and accuracy. The mocap system records light from reflective markers that were attached to the participants' bodies. For all of the recordings, whole-body measurements were acquired, though only the data related to the motion of the specific body parts of interest was included (see section 4.4.2 for a discussion of marker placement). Both samba and telespringar dances include whole-body motion, but since the aim of this thesis was to investigate periodic rhythmic structures in musical sound and music-related body motion (including dance), only

the motion assumed related to rhythmic structure was incorporated.

Our analyses of body motion in telespringar performance in papers III and IV were focused on the fiddler's integrated foot stamping and the dancers' vertical body motions (librations). The analyses of the telespringar dance were based on the vertical motion of the marker attached to the lower back of each dancer. These libration curves may not capture the "feeling" of the dance but rather represent the dancers' vertical displacement. Vertical motion, for example, is clearly related to gravity, and the extensive use of twirling encompasses rotational energy. It is probable that the dancers' experience of the dance is closely related to these forces (see also Kvifte, 1999, p. 421). The analyses presented in papers I, III and IV are not concerned with such forces, though they are clearly relevant to the feeling of the dance. However, as previously pointed out, the main aim of this thesis was to investigate music and motion correspondences in the experience of rhythm. Based on the analyses in papers I, III, and IV, the dancers' libration curves seem to be intrinsically related to the underlying reference structure in telespringar, which are shown to provide valuable insights into the experience of telespringar rhythm.

The analysis of samba dance presented in paper V is based on the dancer's vertical motion of her feet. The representation of the dancer's feet motion revealed that it seems to be synchronized with the first three sixteenth notes in each beat. This indicates that the played sixteenth notes (that is, the sonic events) in samba should not be regarded as deviations from an underlying isochronous subdivision but rather as coherent non-isochronous metrical subdivisions at the sixteenth-note level. While the analysis of the dancers' feet was clearly relevant, it was also pointed out in the chapter that the pulsation of samba groove might be integrated into the hip motion in samba dancing as well. This motion is quite complex and seems to involve a number of rapid directional changes in all three dimensions. Hence, position data might not be the best point of departure for investigating rhythm structures based on hip motions in samba. In paper II, samba hip motion was tracked using accelerometers, but the performance of the accelerometer devices was too imprecise for such a complex motion. The acceleration data based on second-order derivations of position data from Qualisys did, however, indicate a slight increase in acceleration on the second and the fourth sixteenth note in a beat. One possible solution for analyzing the complex hip motions in samba dance could be to sonify them. I will return to this in the section on future work.

6.1.6 The Musical Instruments

The telespringar tunes analyzed in papers I, III, and IV were all played on a Hardanger fiddle. As pointed out, the sound of Hardanger fiddle playing is very complex, both because the style of playing is very ornamented and because the sympathetic strings that run under the board resonate when playing on the bowed strings. Telespringar can, however, be played on a number of instruments, and the sound of a flute or singing voice might have supplied clearer onsets than the sound of a Hardanger fiddle. In the samba study in paper V, on the other hand, the music was played on the Brazilian hand drum called the *pandeiro*, which produced quite unambiguous peaks in the audio envelope, whereas a samba groove played by a large samba *bateria* may have resulted in less clear onsets. The discussion in the previous sections suggests that the perception of an underlying reference structure may depend more on performers' and perceivers' implicit knowledge (or pattern recognition) related to the corresponding music culture than on the sonic

properties of a specific performance. Accordingly, the onset clarity, or lack of such, in the sound signal should not influence the performers' metrical perception, or the execution of the dance. This cannot be tested based on the data presented here, but it would be something to look at in the future.

6.1.7 Ecological Validity

All of the analyses in this thesis are based on recordings executed in a motion capture lab, which is not an environment where music is usually performed. In addition, because I used an infrared motion capture system, reflective markers had to be attached to the participants' bodies. In order to minimize the awkwardness, I only used professional performers in these studies, because it has been pointed out that performers' expertise can compensate for the influence of an artificial environment (Naveda and Leman, 2008). For example, professional music recording can be even more unnatural than these motion capture recordings. The performers were told to do whatever they would normally do when performing, and all of them confirmed that they had done so. Consequently, the data used in the studies included in this thesis are thought to represent a typical samba and telespringar performance. Of course, future studies should also include recordings in real-world situations.

6.2 Conclusions

This thesis has investigated the relationship between sonic rhythm and underlying reference structures in samba and telespringar, based on sound and motion analyses of samba and telespringar performances. Its conclusions are as follows:

- Motion capture technologies enable the analysis of correspondences between musical sound and body motion with a degree of precision and accuracy that would have been difficult to achieve otherwise. I conclude from this that motion capture can provide valuable insights into music and motion correspondences in the experience of rhythm.

- It is possible to track simple music-related motion in real-world situations using only one accelerometer and a sound recording.

- The underlying reference structures in musical rhythm do not necessarily consist of isochronous time series.

- Our audio analysis of the samba groove confirmed so-called *systematic microtiming* of the sixteenth-note level in samba, in a *medium–medium–medium–long* duration pattern. The percussionist's foot stamping, which was synchronized with the sounding of the first and fourth sixteenth note in each beat, seemed to follow this pattern. Also, the dancers' footsteps seemed to be synchronized with the sounding first, second, and third sixteenth note in each beat. These findings support the view that the *metrical sixteenth-note level* in samba is in fact non-isochronous.

- The findings in our studies confirm the view that the underlying metrical beat level in telespringar is asymmetrical and follows a *long–medium–short* pattern. This beat duration

pattern was found in the fiddler's regular foot stamping in telespringar playing. This foot stamping also corresponded to the telespringar dancers' libration curves, though it is the libration *shape* that corresponds to the beat *durations* (implied by the fiddler's foot stamping), not the turning points in the dancers' libration curve (as suggested by previous hypotheses) that correspond to the temporal *positions* of the fiddler's foot stamping. This led to a general conclusion that metrical perception may encompass not only the *temporal positions* of the metrical beats but also the *trajectories* (as actual or simulated motion) between them.

6.3   Future Work

The analyses presented in this thesis include recordings of quite few performers, but there is always room for more. Future research should also include recordings of music performances in real-world situations. The results in paper II suggest that certain simple music-related motions can be tracked using only an accelerometer and a sound recording, meaning that these motions can be recorded in situations where music and dance are normally performed, such as concerts and parades, without the requirement of visible markers attached to the performers.

As pointed out above, investigating duration patterns was very useful to this work. However, the *metrical shape* hypothesis presented in section 6.1.3 suggests that there might exist trajectories between metrical beats that correspond to the metrical beat durations. In the future, I would like to include such trajectories alongside the beat durations. One possible solution for estimating *mean shapes* or *median shapes* based on motion trajectories between beats, in tandem with averaged duration patterns, might be a *functional data analysis* (FDA), which could also be used in future rhythm studies.

An ecological perspective on perception suggests that sound perception includes a simulation of the action that we imagine produced the sound (see section 2.4). A motor-mimetic perspective on sound perception and cognition suggests that such an action–sound relationship might be understood as a correspondence between *sonic shapes* and *motion shapes*, so that, for example, an impulsive motion corresponds to an impulsive sound. It has also been suggested that this relationship might work both ways, so that a music-related motion could be represented, and maybe also analyzed, as sound. One way to investigate the relationship between the musical sound and the corresponding dance might be to *sonify* the dance motion, and especially the periodic dance motions—for example, the hips and feet in samba, and the dancers' libration curves in telespringar.

In this thesis, the focus has been on the performers, and I argued that the rhythm structures in samba and telespringar should be understood in relation to the corresponding dances. Another perspective would be to investigate how people without any experience with samba and telespringar would choose to move to the sound of samba and telespringar. This might provide some insights into how much of the rhythm perception of samba and telespringar actually depends on tacit knowledge of the existing relationship between the sonic rhythm pattern and the motion.

Bibliography

Agawu, V. K. (2003). *Representing African music : postcolonial notes, queries, positions*. New York: Routledge.

Ahlbäck, S. (2003). About asymmetrical beat in the polska. In M. Ramsten (Ed.), *The Polish dance in Scandinavia and Poland*, Volume 17, pp. 165 – 180. Stockholm: Svenskt visarkiv.

Baily, J. (1985). Music structure and human movement. In P. Howell, I. Cross, and R. West (Eds.), *Musical Structure and Cognition*, Chapter 10, pp. 237–258. London: Academic press.

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

Bengtsson, I. (1973). *Musikvetenskap : en översikt*. Scandinavian university books. Stockholm: Esselte studium.

Bengtsson, I. (1974). On notation of time, signature and rhythm in swedish polskas. In G. Hilleström (Ed.), *Studia instrumentorum musicae popularis III*, pp. 22–31. Stockholm: Nordiska Musikförlaget.

Bengtsson, I. (1987). Notation, motion and perception: Some aspects of musical rhythm. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music*, Volume 55, pp. 69–80. Stockholm: The Royal Swedish Academy of Music.

Bengtsson, I. and A. Gabrielsson (1980). Methods for analyzing performance of musical rhythm. *Scandinavian Journal of Psychology 21*, 257–268.

Bengtsson, I., A. Gabrielsson, and S. M. Thorsén (1969). Empirisk rytmforskning: en orientering om musikvetenskaplig bakgrund, utveckling och några aktuella projekt. *Svensk tidskrift för musikforskning 51*, 49–118. Særtrykk av: Svensk tidskrift før musikforskning. 51.

Berthoz, A. (2000). *The brain's sense of movement*. Cambridge, Mass.: Harvard University Press.

Bjørkvold, J.-R. (1999). *Det musiske menneske* (5. ed.). Oslo: Freidig forlag.

Blacking, J. (1995). *Music, culture, experience*. Chicago studies in ethnomusicology. Chicago: University of Chicago Press.

Blom, J.-P. (1981). The dancing fiddle. on the expression of rhythm in hardingfele slåtter. In J.-P. Blom, S. Nyhus, and R. Sevåg (Eds.), *Norsk Folkemusikk*, Volume VII, pp. 305–312. Oslo: Universitetsforlaget.

Blom, J.-P. (1993). Rytme og frasering - forholdet til dansen. In B. Aksdal and S. Nyhus (Eds.), *Fanitullen. Innføring i norsk og samisk folkemusikk*, pp. 161–184. Oslo: Universitetsforlaget, Oslo.

Blom, J.-P. (2006). Making the music dance: Dance connotations in norwegian fiddling. In I. Russell and M. Alburger (Eds.), *Play it like it is. Fiddle and dance studies from around the Noth Atlantic*, Volume 5, pp. 75 – 86. The Elphinstone Institute, University of Aberdeen.

Bolton, T. L. (1894). Rhythm. *The American Journal of Psychology 6*(2), 145–238.

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

Brochard, R., D. Abecasis, D. Potter, R. Ragot, and C. Drake (2003). The "ticktock" of our internal clock direct brain evidence of subjective accents in isochronous sequences. *Psychological Science 14*(4), 362–366.

Burger, B. (2013). *Move the Way You Feel: Effects of Musical Features, Perceived Emotions, and Personality on Music-Induced Movement*. Ph. D. thesis, University of Jyväskylä.

Burger, B. and P. Toiviainen (2013). Mocap toolbox – A matlab toolbox for computational analysis of movement data. In R. Bresin (Ed.), *Proceedings of the 10th Sound and Music Computing Conference, (SMC)*, pp. 172–178.

Butler, M. J. (2006). *Unlocking the groove : rhythm, meter, and musical design in electronic dance music*. Profiles in popular music. Bloomington, Ind: Indiana University Press.

Chernoff, J. M. (1979). *African rhythm and African sensibility : aesthetics and social action in African musical idioms*. Chicago: University of Chicago Press.

Chernoff, J. M. (1991). The rhythmic medium in african music. *New Literary History 22*(4), 1093–1102.

Clarke, E. F. (1985). Structure and expression in rhythmic performance. In P. Howell, I. Cross, and R. West (Eds.), *Musical structure and cognition*, Chapter 9, pp. 209–236. London: Academic Press.

Clarke, E. F. (1987). Categorical rhythm perception: An ecological perspective. In A. Gabrilesson (Ed.), *Action and perception in rhythm and music*, Volume 55, pp. 19–33. Stockholm: The Royal Swedish Academy of Music.

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

Clayton, M. (2012). What is entrainment? Definition and applications in musical research. *Empirical musicology review. 7*(1–2), 49–56.

Clayton, M. (2013). Entrainment, ethnography and musical interaction. In M. Clayton, B. Dueck, and L. Leante (Eds.), *Experience and meaning in music performance*, Chapter 2, pp. 17–39. New York: Oxford University Press.

Clayton, M., B. Dueck, and L. Leante (2013). *Experience and meaning in music performance.* Oxford University Press. Includes bibliographical references and index.

Clayton, M. and L. Leante (2013). Embodiment in music performance. In M. Clayton, B. Dueck, and L. Leante (Eds.), *Experience and Meaning in Music Performance*, Chapter 9, pp. 188—207. USA: Oxford University Press.

Clayton, Martin; Sager, R. and U. Will (2005). In time with the music: the concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology 11*, 3–142.

Collins, N. (2010). *Introduction to computer music.* Chichester: Wiley.

Cooper, G. W. and L. B. Meyer (1960). *The rhythmic structure of music.* Chicago: University of Chicago Press.

Cox, A. (2006). Hearing, feeling, grasping gestures. In A. Gritten and E. King (Eds.), *Music and Gesture*, Chapter 3, pp. 45–60. Farnham: Ashgate Publishing Ltd.

Danielsen, A. (2006). *Presence and pleasure: the funk grooves of James Brown and Parliament.* Middletown, Conn.: Wesleyan University Press.

Danielsen, A. (2010a). 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*, Chapter 2, pp. 19–36. Farnham: Ashgate.

Danielsen, A. (2010b). Introduction: Rhythm in the age of digital reproduction. In A. Danielsen (Ed.), *Musical Rhythm in the Age of Digital Reproduction*, pp. 1–16. Farnham: Ashgate.

Danielsen, A. (2015). Metrical ambiguity or microrhythmic flexibility? Analysing groove in 'Nasty Girl' by Destiny's Child. In R. Appen, A. Doehring, D. Helms, and A. F. Moore (Eds.), *Song Interpretation in 21st-Century Pop Music*, pp. 53–72. Farnham, Surrey: Ashgate.

Danielsen, A. (forthcoming). Extending the beat: Beat bin meter in contemporary computer-based grooves. In H. Neuhoff and R. Polak (Eds.), *Musical Meter in Comparative Perspective. Theoretical Discourse and Case Studies from Various Cultures*. Cologne: Verlag Dohr.

Danielsen, A., M. R. Haugen, and A. R. Jensenius (2015). Moving to the beat: Studying entrainment to micro-rhythmic changes in pulse by motion capture. *Timing Time Perception 3*(1-2), 133–154.

Drake, C. and J. B. El Heni (2003). Synchronizing with music: Intercultural differences. *Annals of the New York Academy of Sciences 999*(1), 429–437.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music*, Volume 1, Chapter 6, pp. 149 – 180. New York, NY: Academic Press.

Gabrielsson, A. (1982). Perception and performance of musical rhythm. In M. Clynes (Ed.), *Music, Mind, and Brain. The Neuropsychology of Music*, pp. 159–169. New York: Plenum Press.

Gallese, V., L. Fadiga, L. Fogassi, and G. Rizzolatti (1996). Action recognition in the premotor cortex. *Brain 119*(2), 593–609.

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

Gibson, J. J. (1977). The theory of affordances. In R. Shaw and J. Bransford (Eds.), *Perceiving, acting, and knowing : toward an ecological psychology*, Chapter 3, pp. 67–82. Hillsdale, New Jersey: Erlbaum.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, N.J.: Lawrence Erlbaum. Opprinnelig utgitt: Boston : Houghton Mifflin, 1979.

Godøy, R. I. (2003). Motor-mimetic music cognition. *Leonardo 36*(4), 317–319.

Godøy, R. I. (2006). Gestural-sonorous objects: embodied extensions of schaeffer's conceptual apparatus. *Organised Sound 11*(02), 149–157.

Godøy, R. I. (2009). Geometry and effort in gestural renderings of musical sound. In M. Sales Dias, S. Gibet, M. M. Wanderley, and R. Bastos (Eds.), *Gesture-Based Human-Computer Interaction and Simulation: 7th International Gesture Workshop, GW 2007, Lisbon, Portugal, May 23-25, 2007, Revised Selected Papers*, pp. 205–215. Springer Berlin Heidelberg.

Godøy, R. I. (2010). Gestural affordances of musical sound. *Musical gestures: Sound, movement, and meaning 5*, 103–125.

Godøy, R. I., E. Haga, and A. R. Jensenius (2006). Playing "air instruments": Mimicry of sound-producing gestures by novices and experts. In S. Gibet, N. Courty, and J.-F. Kamp (Eds.), *Gesture in Human-Computer Interaction and Simulation: 6th International Gesture Workshop, GW 2005, Berder Island, France, May 18-20, 2005, Revised Selected Papers*, pp. 256 – 267. Berlin, Heidelberg: Springer Berlin Heidelberg.

Godøy, R. I. and M. Leman (2010). *Musical gestures: sound, movement, and meaning*. New York: Routledge.

Gouyon, F. (2007). Microtiming in "samba de roda" — preliminary experiments with polyphonic audio. In *SBCM 2007 Proceedings*, pp. 197–203.

Grau, A. (1983). Sing a dance. dance a song: The relationship between two types of formalised movements and music among the tiwi of melville and bathurst islands, north australia. *Dance Research: The Journal of the Society for Dance Research Vol. 1*(No.2 , Autumn), 32–44.

Groven, E. (1971). Musikkstudiar - ikkje utgjevne før. 1. rytmestudiar. In O. Fjalestad (Ed.), *Eivind Groven. Heiderskrift til 70-årsdagen 8. oktober 1971*, pp. 93–102. Oslo: Noregs boklag.

Haga, E. (2008). *Correspondences between music and body movement*. Ph. D. thesis, University of Oslo.

Hannon, E. E. and S. E. Trehub (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America 102*(35), 12639 – 12643.

Haslinger, B., P. Erhard, E. Altenmüller, U. Schroeder, H. Boecker, and A. O. Ceballos-Baumann (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Cognitive Neuroscience, Journal of 17*(2), 282 – 293.

Haueisen, J. and T. R. Knösche (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience 13*(6), 786–792.

Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences 1252*(1), 85–91.

Honing, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *Psychology of music*, Chapter 9, pp. 369–404. San Diego, Calif.: Academic Press.

Huron, D. (1999). The new empiricism: Systematic musicology in a postmodern age. In *Lecture 3 from the 1999 Ernest Bloch Lectures*, pp. 1–32. Berkeley, University of California.

Huron, D. (2006). *Sweet anticipation : music and the psychology of expectation*. Cambridge, Mass: MIT Press.

Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in african-american music. *Music Perception 19*(3), 387–414.

Jensenius, A. R. (2007). *Action - sound: developing methods and tools to study music-related body movement*. Ph. D. thesis, University of Oslo.

Jensenius, A. R. (2013). Some video abstraction techniques for displaying body movement in analysis and performance. *Leonardo 46*(1), 53–60.

Jensenius, A. R., K. Nymoen, S. A. v. D. Skogstad, and A. Voldsund (2012). A study of the noise-level in two infrared marker-based motion capture systems. In B. L. Sturm, S. Dahl, and J. Larsen (Eds.), *Proceedings of the 9th Sound and Music Computing Conference - "Illusions"*, pp. 258 – 263. Logos Verlag Berlin.

Johansson, M. (2009). *Rhythm into style: studying asymmetrical grooves in Norwegian folk music*. Ph. D. thesis, University of Oslo.

Johansson, M. (2010). The concept of rhythmic tolerance : examining flexible grooves in Scandinavian folk fiddling. In A. Danielsen (Ed.), *Musical Rhythm in the Age of Digital Reproduction*, Chapter 5, pp. 69–83. Farnham: Ashgate, cop. 2010.

Jones, A. M. (1959). *Studies in African music : 1*. London: Oxford University Press.

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.

Kaminsky, D. (2014). Total rhythm in three dimensions: Towards a motional theory of melodic dance rhythm in swedish polska music. *Dance Research 32*(1), 43–64.

Kauffman, R. (1980). African rhythm: A reassessment. *Ethnomusicology 24*(3), 393 – 415.

Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology 2*(3), 275 – 283.

Keil, C. (1995). The theory of participatory discrepancies: A progress report. *Ethnomusicology 39*(1), 1 – 19.

Keil, C. M. H. (1966). Motion and feeling through music. *The Journal of Aesthetics and Art Criticism 24*(3), 337 – 349.

Keysers, C., E. Kohler, M. A. Umiltà, L. Nanetti, L. Fogassi, and V. Gallese (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research 153*(4), 628–636.

Kohler, E., C. Keysers, M. A. Umiltà, L. Fogassi, V. Gallese, and G. Rizzolatti (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science 297*(5582), 846–848.

Kubik, G. (1979). Pattern perception and recognition in african music. In J. Blacking and J. W. Kealiinohomoku (Eds.), *The Performing Arts. Music and Dance*, pp. 221 – 249. The Hague: Mouton Publishers.

Kubik, G. (1990). Drum patterns in the "batuque" of benedito caxias. *Latin American Music Review / Revista de Música Latinoamericana 11*(2), 115 – 181.

Kvifte, T. (1987). Strøkfigurer - en side av bueteknikken i den norsk felemusikken. In *Sumlen*, Volume 1986, pp. 19 – 32. Stocholm: Årbok for vise- og folkemusikkforskning.

Kvifte, T. (1999). Fenomenet "asymmetrisk takt" i norsk og svensk folkemusikk. *Studia musicologica norvegica 25*, 387 – 430.

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

Kvifte, T. (2007). Categories and timing: On the perception of meter. *Ethnomusicology Vol. 51*(No. 1), 64 – 84.

Large, E. W. and M. R. Jones (1999). The dynamics of attending: how people track time-varying events. *Psychological review 106*(1), 119 – 159.

Lartillot, O. and P. Toiviainen (2007). A matlab toolbox for musical feature extraction from audio. In *International Conference on Digital Audio Effects*, pp. 237– 244.

Leman, M. (2008). *Embodied music cognition and mediation technology*. Cambridge, Mass.: MIT Press.

Leman, M. and R. I. Godøy (2010). Why study musical gestures? In R. I. Godøy and M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning*, Chapter 1, pp. 3 – 11. Routhledge.

Leman, M., F. Styns, and N. Bernardini (2008). Sound, sense and music mediation: a historical-philosophical perspective. In P. Polotti and D. Rocchesso (Eds.), *Sound to sense, sense to sound: a state of the art in sound and music computing*, Chapter 1, pp. 15 – 46. Berlin: Logos Verlag Berlin GmbH.

Lerdahl, F. and R. Jackendoff (1983). *A generative theory of tonal music*. Cambridge, Mass.: MIT Press.

Liberman, A. M. and I. G. Mattingly (1985). The motor theory of speech perception revised. *Cognition 21*(1), 1 – 36.

London, J. (2012). *Hearing in time: psychological aspects of musical meter*. Oxford: Oxford University Press. 2nd ed.

Manning, F. and M. Schutz (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin Review 20*(6), 1133 – 1139.

Mårds, T. (1999). *Svikt, kraft og tramp: en studie av bevegelse og kraft i folkelig dans*. M.A. thesis, Norges Idrettshøgskole.

Naveda, L. (2011). *Gesture in Samba: A cross-modal analysis of dance and music from the Afro-Brazilian culture*. Ph. D. thesis, Faculty of Arts and Philosophy, Ghent University, Belgium.

Naveda, L., F. Gouyon, C. Guedes, and M. Leman (2009). Multidimentional microtiming in samba music. In M. Pimenta, D. Keller, R. Faria, M. Queiroz, G. Ramalho, and G. Cabral (Eds.), *12th Brazilian Symposium on Computer Music*, pp. 1–12. SBCM.

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

Nielsen, S. (1985). *Mennesker og musik i Afrika*. Herning: Systime.

Nketia, J. H. K. (1986). *The music of Africa*. A Gollancz paperback. London: Gollancz.

Nymoen, K. (2013). *Methods and technologies for analysing links between musical sound and body motion*. Ph. D. thesis, University of Oslo.

Parncutt, R. (1987). The perception of pulse in musical rhythm. In A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music*, Volume 55, pp. 127–138. Stockholm: The Royal Swedish Academy of Music.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception 11*(4), 409 – 464.

Phillips-Silver, J. and L. J. Trainor (2005). Feeling the beat: Movement influences infant rhythm perception. *Science 308*(5727), 1430 – 1430.

Phillips-Silver, J. and L. J. Trainor (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition 105*(3), 533 – 546.

Polak, R. (2010). Rhythmic feel as meter: Non-isochronous beat subdivision on jembe music from mali. *Music Theory Online 16*(4).

Povel, D.-J. and P. Essens (1985). Perception of temporal patterns. *Music Perception: An Interdisciplinary Journal 2*(4), 411–440.

Ramsten, M. (1982). *Einar Øvergaards folkmusiksamling*. Stockholm: Svenskt vis-arkiv/Almqvist and Wiksell International.

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

Sethares, W. A. (2007). *Rhythm and transforms*. London: Springer.

Shove, P. and B. H. Repp (1995). Musical motion and performance: theoretical and empirical perspectives. In J. Rink (Ed.), *The Practice of performance: studies in musical interpretation*, pp. 55 – 83. Cambridge: Cambridge University Press.

Skogstad, S. A., K. Nymoen, and M. E. Høvin (2011). Comparing inertial and optical mo-cap technologies for synthesis control. In *Proceedings of SMC 2011 8th Sound and Music Computing Conference "Creativity rethinks science"*, pp. 421–426. Padova University Press.

Snyder, B. (2000). *Music and memory: an introduction*. Cambridge, Mass.: MIT Press.

Su, Y.-H. and E. Pöppel (2012). Body movements enhances the extraction of temporal structures in auditory sequences. *Psychological Research 76*(3), 373 – 382.

Toiviainen, P., G. Luck, and M. Thompson (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception: An Interdiciplinary Journal 28*(1), 59 – 70.

Trehub, S. E. and E. E. Hannon (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition 100*(1), 73–99.

Van Dyck, E., D. Moelants, M. Demey, A. Deweppe, P. Coussement, and M. Leman (2013). The impact of the bass drum on human dance movement. *Music Perception: An Interdisciplinary Journal 30*(4), 349 – 359.

Vos, J. and R. Rasch (1981). The perceptual onset of musical tones. *Perception and Psychophysics 29*(4), 323 – 335.

Waadeland, C. H. (2000). *Rhythmic movements and moveable rhythms: syntheses of expressive timing by means of rhythmic frequency modulation*. Ph. D. thesis, Norwegian University of Science and Technology.

Waadeland, C. H. (2001). "It don't mean a thing if it ain't got that swing"—Simulating expressive timing by modulated movements. *Journal of New Music Research 30*(1), 23 – 37.

Waterman, R. A. (1967). African influence on the music of the americas. In S. Tax (Ed.), *Acculturation in the Americas*, pp. 207–218. New York: Cooper Square.

Wilson, M. and G. Knoblich (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin Vol. 131*(3), 13.

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

Papers