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*.

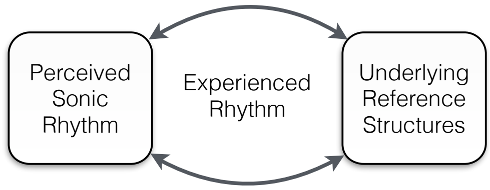

Figure 3.1: An illustration of rhythm as an interaction between perceived sonic rhythm and underlying reference structures in a musical context.

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*

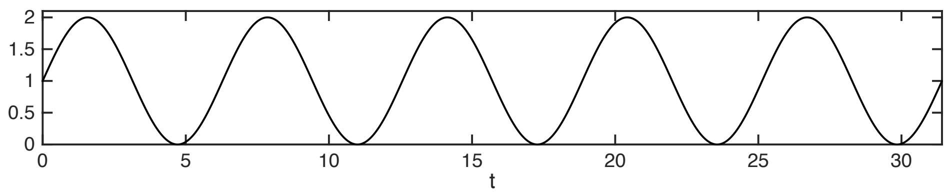

Figure 3.2: Waadeland suggests that the pulse can be represented in relation to the motion curve of a hand moving in synchrony with a metronome over time (t), where the minimal points of the pulse curve represents the pulse beats (adapted from Waadeland, 2000, figure 5.1, p. 122).

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 action-s/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).

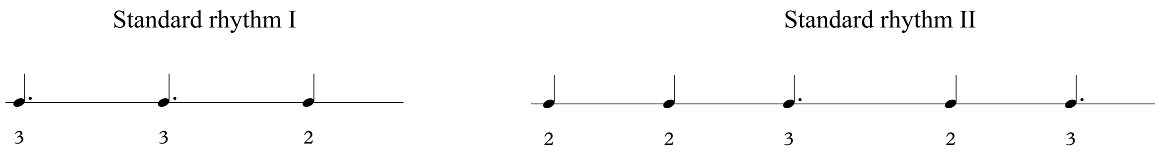

Figure 3.3: Two “basic” standard rhythms, adapted from Nielsen (1985, p. 60, ex. 8).

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).

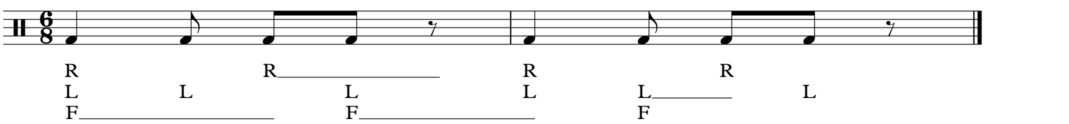

Figure 3.4: Three different metrical interpretations (R, L, and F) with a 6/8 meter as a starting point: (R) three subdivisions of every beat; (L) two subdivisions of every beat; and (F) four subdivisions of every beat.

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 context, 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); following 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.