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

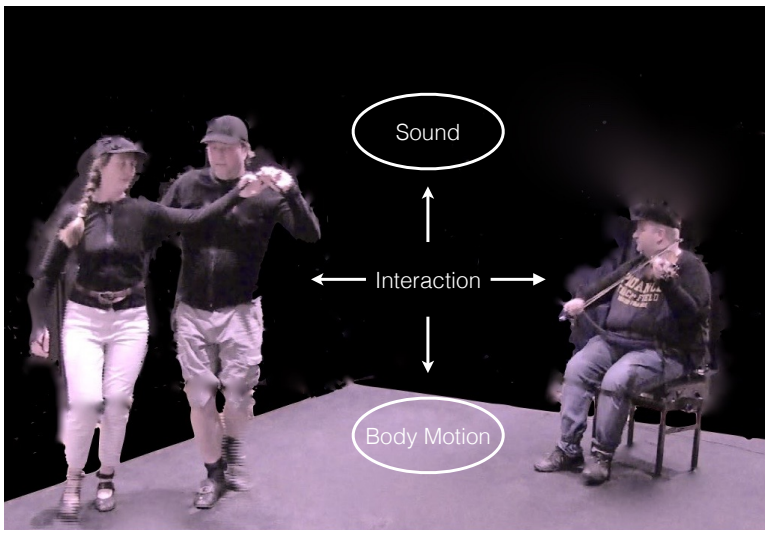

Figure 2.1: An illustration of relevant factors in play in a music–dance performance. There are interactions between musical sound and music-related body motions, and between the performers.

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 with 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.