This is the pre-print of the chapter: Jensenius, A. R., Wanderley, M. M., Godøy, R. I., & Leman, M. (2010). Musical gestures: Concepts and methods in research. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 12–35). New York: Routledge.

Chapter 2.

Musical gestures: concepts and methods in research

Alexander Refsum Jensenius
Marcelo M. Wanderley
Rolf Inge Godøy
Marc Leman

1. Introduction

In the last decade, cognitive science underwent a change of paradigm by bringing human movement into the focus of research. Concepts such as 'embodiment' and 'enactive' have been proposed as core concepts reflecting the role of the human body in complex processes such as action and perception, and the interaction of mind and physical environment (Varela et al., 1991; Noë, 2004). In music research, human movement has often been related with the notion of gesture. The reason is that many musical activities (performance, conducting, dancing) involve body movements that evoke meanings, and therefore these movements are called gestures. In Camurri et al. (2005), musical gestures are addressed from the viewpoint of their expressive character. However, there are many ways in which music-related body movements can be approached, measured, described and applied. Accordingly, there are many ways in which musical gestures are meaningful. Given the different contexts in which gestures appear, and their close relationship to movement and meaning, one may be tempted to say that the notion of gesture is too broad, ill-defined and perhaps too vague. Yet the use of this notion is very convenient in modern music research, because it allows making a bridge between movement and meaning. A closer look at the term gesture reveals its potential as a core notion that provides access to central issues in action/perception processes and in mind/environment interactions.

The chapter starts with a review of some current definitions of gesture. The second part presents a conceptual framework for differentiating various *functional* aspects of gestures in music performance. The third part presents a brief overview of some methodological approaches that can be used in gesture research.

2. Musicians' and dancers' gestures

Musical gestures, that is, human body movement that goes along with sounding music, can be divided into two main categories, namely the gestures of those that produce the sounds (the musicians), and the gestures of those that perceive the sounds (the listeners or dancers). Obviously, the musicians also listen to musical sounds, but their role is nevertheless somewhat more specific in that they are involved in the creation of sounds, whereas listeners or dancers basically respond to these sounds. When dancers are connected with a computer system that produces music based on features of the dance movements, then also dancers can be considered musicians because they generate the sounds.

Obviously, musicians and dancers each have their own sub-categories of musical gestures. For example, the gestures of musicians may be categorized as sound-producing, communicative, ancillary or sound-facilitating, and sound-accompanying. Musician's gestures have received an increasing research interest in recent years, e.g. in (Wanderley 2001) (see also Chapter 3). Conductors' gestures and gestures that are more or less purely communicative, are also included in this main category of musicians' gestures (see Chapter 11). The gestures of dancers may be thought of as having sound-accompanying characteristics, since they often follow or contrast with the musical sound. Dance movements in relation to sound have been studied in e.g. (Hodgins 1992, Haga 2008) (see also Chapters 7 and 10).

When speaking about the musical activity of musicians and dancers, it is tempting to call the involved embodiment 'gestures', rather than 'movements'. The main reason for doing this is that the notion of gesture somehow blurs the distinction between movement and meaning. *Movement* denotes physical displacement of an object in space, whereas *meaning* denotes the mental activation of an experience. The notion of gesture somehow covers both aspects and therefore bypasses the Cartesian divide between matter and mind. In that sense, the notion of gesture provides a tool that allows a more straightforward crossing of the traditional boundary between the physical and the mental world. The crossing of this boundary is at the core of the entire embodiment paradigm and it forms the strength of the current extension from disembodied music cognition to embodied music cognition (Leman, 2007). In this context, *action* can be understood as coherent chunks of gestures, or delimited segments of human movement having an intentional aspect.

However, the term gesture has been used with so many significations, that an overview of some different types of gesture definitions is needed. This overview may help to clarify the terminology used in the rest of this chapter and the book as a whole.

3. Gesture Definitions

Based on the work of Zhao (2001) and McNeill (2000), it is possible to define a general framework that considers gestures from the viewpoint of communication, control and metaphor. *(i) Communication* is involved when gestures work as vehicles of meaning in social interaction. This use of the term is common in linguistics, behavioral psychology, and social anthropology. (ii) *Control* is involved when gestures work as elements of a system, such as in the control of computational and interactive systems. This is common in the fields of human-computer interaction (HCI), computer music, and similar areas. *(iii) Metaphor* is involved when gestures work as concepts that project physical movement, sound, or other types of perception to cultural topics. This use of the term is common in cognitive science, psychology, musicology, etc. The following sections will present examples of gesture definitions within each of these three main categories.

3.1 Gesture as Communication

In recent work on gesture, *gesture* is often used to denote bodily actions that are associated with speech, particularly hand movements and facial expressions. This definition of gesture as "visible action as utterance" (Kendon 2004) is most commonly used in linguistics, psychology and behavioral studies. In particular, Kendon (1972) used the term *body motion* and later used the term *gesticulation* (Kendon 1980), before finally settling on the word *gesture* (Kendon 1982). Through a series of observation studies of people's storytelling, McNeill (1992, 2005) showed how hand movements and facial expressions are not just random movements accompanying speech, but are actually an integral part of communication. McNeill (1992) provides a taxonomy for different gestural functions based on the five types of nonverbal behavior outlined earlier by Ekman (1969): *(i) Iconics* represent a particular feature of an object, and can be described in terms of the shape and spatial extent of the gesture. Iconic gestures are often used to illustrate an action, for example imitating a knocking movement with a hand while saying "knocking on the door". *(ii) Metaphorics* are similar to iconics, but represent an abstract feature of an object. An example of a metaphoric gesture may be to say "something happened" while holding up the hands to refer to "something". *(iii) Beats* occur together with spoken words to highlight discontinuities and stress specific words. Beats are typically carried out as in/out or up/down movements, and may be seen as emphasizing the most important words in a narrative. *(iv) Deictics* indicates a point in space, for example pointing in a specific direction while saying "over there". *(v) Emblems* are stereotypical patterns with agreed meaning, such as the goodbye or OK sign.

McNeill's theory of gesture is built on the idea that gestures coexist with speech. This is not to say that they have to co-occur, but rather that gestures and speech are co-

expressive, or co-articulatory. In this respect, McNeill adopts Damasio's saying that "language is inseparable from imagery" (McNeill 2000: 57), and argues that mental imagery is embodied in the gestures that co-occur with speech. To explain the relationships between gesture and speech, McNeill (1992) presented what he calls the *Kendon continuum*, based on the typology of gestures suggested in (Kendon 1982): *gesticulation*, *emblems*, *pantomime* and *sign language*. As shown in Figure 1, this continuum covers two extremes, namely, *gesticulation* is used to denote the types of gestures that only co-occur with speech, and *sign language* to denote the types of gestures that are linguistically self-contained.

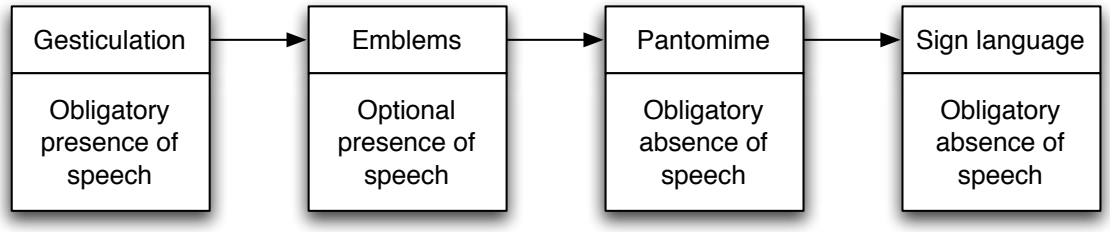

Figure 1: McNeill's (2005) Kendon continuum of gestures, and how they relate to speech.

Goldin-Meadow (2003) follows a similar line of thought but the difference is that she uses the term gesture to denote only hand movement, and leaves out other types of body movement, including facial expression. She argues that gestures may not only *support* but also *contradict* speech. For example, it may be possible to spot when people are lying because their facial expression and body movements contradict what they are saying. McNeill (2005) therefore suggests that studying *overt* gestures may reveal interesting aspects of our *covert* mental activity.

Clearly, these definitions focus on linguistic communicative aspects of gestures. As such, the term *gesture* does not refer to body movement or expression *per se*, but rather to the intended or perceived *meaning* of the movement or expression, often in accompaniment with verbal utterances. In that respect, Feyereisen and de Lannoy (1991) use a slightly wider definition when they say (p.3): "To some extent, any movement or change in position of a body segment may be considered a gesture. Accordingly, the very notion of gesture refers to a great variety of phenomena. In an extended sense, the term gesture encompasses gestures that are used in various professions and that often involve tool use, e.g. a carpenter's hammering or sawing gestures. In such a perspective, gestures are mainly actions before becoming means of communication." The latter definition implies that the term gesture may not be restricted to human-human communication, but may also be used for describing the communication between humans and machines.

3.2 Gesture for Control

In the field of human-computer interaction (HCI), there is now a considerable amount of research on how to use various kinds of body movement as input to computers. In contrast to human-human communication, it is obvious that computers traditionally have had comparably limited sensing capabilities. For example, Kurtenbach and Hulteen (1990) state that (p.310) "A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on it's way to hitting a key is neither observed nor significant. All that matters is which key was pressed". Pressing the key is highlighted as the meaning-bearing component, while the rest of the movement of the person is considered irrelevant. It is an example of a computer-centric approach to interaction design, in the sense that the constraints of the computer define the interaction.

However, a more recent trend in the HCI community is to focus on creating computer systems that can sense a broader range of human expressions. Here the challenge is to develop sensor and computer vision solutions, and corresponding computational algorithms, which understand the gestures (here used in the communication sense) in a continuous stream of movement. While humans have few problems separating a hand gesture (e.g. waving goodbye) from other types of movement (e.g. waving away a fly), this is much more problematic for computers. This is not only due to the remarkable capacity of *visual scene analysis* in humans, but is also due to the fact that we understand the intended meaning of the gesture based on the context and on our life-long experience of multimodal communication.

One approach to make computers understand the meaning of human gestures is to create means for extracting the *expressiveness* of body movement. Camurri et al. (2001) introduce the term *expressive gesture* to denote aspects of body movement that convey information about affect and emotion (p.1): "It seems likely that expressiveness in gestures is conveyed by a set of temporal/spatial characteristics that operate more or less independent from the denotative meanings (if any) of those gestures. In that sense, gestures can be conceived as the vehicles that carry these expressive characteristics and it is likely that expressiveness as such subsumes certain universal patterns and general rules". Gesture is here used to denote what can be observed, namely body movement, and it implies that a set of expressive characteristics can be extracted from these movements (see more on this in chapter 10). As such, this definition is quite different from the abovementioned definitions found in linguistics where the term gesture often refers to the actual meaning of the movement, and where the movement as such is subordinate to this meaning.

A similar definition of gesture may be found in the computer music literature, albeit with reference to sound. For example, Cadoz (1988, p.64) states that "If we call first of all, 'gesture' all physical behavior, besides vocal transmission, by which a human being

informs or transforms his immediate environment, we may then say that there can be no music without gesture, for music is not exclusively vocal." Miranda and Wanderley (2006, p.5) state that "gesture is used in a broad sense to mean any human action used to generate sounds. The term refers to actions such as grasping, manipulation, and noncontact movements, as well as to general voluntary body movements."

An important difference between the control definitions and the communication definitions presented in the previous section, is that the former is mainly focusing on manipulative gestures (Quek et al. 2002, p.172) and the latter on empty-handed gestures. Manipulative gestures thus denote gestures that are based on a physical contact, or what may also be called ergotic, haptic, or instrumental contact. In contrast, empty-handed gestures have been called semaphoric, free, semiotic, or naked gestures (Miranda and Wanderley 2006).

3.3 Gesture as Metaphor

While in the two previous sections, the notion of gesture refers to some kind of physical body movement; this section considers the use of the notion of gesture in a metaphorical sense. For example, Metois (1997, p.16) states that " [B]oth [physical and auditory gestures] present the ability to communicate musical intentions at a higher level than an audio wave form. The similarity of their level of abstraction motivated the author to label them both as Musical Gestures." Interestingly, musical gesture is here used to denote the combined sensations of physical movement and sound. This is along the lines of how several musicologists have thought about musical gesture in recent decades. From a popular music research perspective, Middleton (1993), referring to Coker's (1972) discussion of affections and emotions that could be associated with gestures, writes (p.177): "[H]ow we feel and how we understand musical sounds is organised through processual shapes which seem to be analogous to physical gestures." Middleton further argues that the idea of gestures in music should be founded on the concept of rhythm. This seems similar to Todd's (1995, p.1941) idea of relationships between musical sound and body movement. Todd claims, without actually using the word gesture, that musical movement is similar to, and imitates, movement in physical space.

A similar way of thinking about gesture as a mental entity that can be evoked from musical sound, is suggested by Hatten (2004, p.95), who argues that a musical gesture is "significant energetic shaping through time". His theory of musical gesture is based on what he calls gestural competency, which arises from physical (i.e. biological and cognitive) and social (i.e. cultural and multi-stylistic) experience. Hatten (2003) states that "Musical gesture is biologically and culturally grounded in communicative human movement. Gesture draws upon the close interaction (and intermodality) of a range of human perceptual and motor systems to synthesize the energetic shaping of motion

through time into significant events with unique expressive force. " Hatten is here making reference to the experience of musical gesture "within" music, either through the score or the musical sound, but he seems not to refer to the body movement or the generating actions that create the sound.

François Delalande (1988) defines *musical gesture* as the intersection of observable actions and mental images. He further argues that musical gestures may be studied at various levels, ranging from the purely functional to the purely symbolic, using the terms *effective*, *accompanying*[1] and *figurative* gestures (Cadoz and Wanderley 2000, pp.77-78). The term *effective gesture* denotes what we call sound-producing gestures, while the term *accompanying gestures* is used for the movements that support the effective gestures in various ways. Delalande suggests the term *figurative gesture* to refer to a mental image that is not directly related to any physical movement, but which may be conveyed through sound.

A somewhat analogous definition to those of Delalande and Hatten is suggested in (Gritten and King 2006: xx) who state that "[A] gesture is a movement or change in state that becomes marked as significant by an agent. This is to say that for movement or sound to be(come) gesture, it must be taken intentionally by an interpreter, who may or may not be involved in the actual sound production of a performance, in such a manner as to donate it with the trappings of human significance." The definition implies that there is a flow of communication between the performer and the perceiver, and movement becomes a gesture only if it is understood as such by the perceiver. An interesting question then arises: Does an action have to be carried out consciously to be seen as a gesture? In human communication, Kendon has argued that gestures have to be carried out consciously since they are intentional (Kendon 2004, p.15). Hatten, on the other hand, argues that musical gestures may be performed unconsciously but still be valid as gestures if they are observed as significant by the perceiver (Gritten and King 2006, p.162). We would assume that there are also ambiguous cases where one person may perceive an action as intentional and another person may see it as unintentional.

3.4 Terminological considerations

The gesture definitions presented so far range from using gesture as more or less equivalent to body movement, to using gesture in a metaphorical sense to describe some emergent qualities in musical sound. There are also several other types of gesture definitions that have not been discussed, such as the concepts of *articulatory* or *phonological* gesture sometimes encountered in linguistics (Liberman and Mattingly

[1] Cadoz and Wanderley (2000) translated the French term *geste accompagnateur* used by (Delalande 1988) to the English term *accompanist gesture*. We are using the term *accompanying*, so as not to confuse such movements with those of an accompanist.

1985). There are other research fields focused on body movements that seldom use the term gesture, such as *kinesiology* and *biomechanics*. A similar situation can be found in some of the music literature, where terms like *motion/movement* (Shaffer 1980; Gabrielsson 1985; Clarke 1993; Davidson 1993), *expressive movement* (Pierce and Pierce 1989; Davidson 1994) or *corporeal articulations* (Leman, 2007) have been used to denote various types of gestures.

Based on the above viewpoints, it seems straightforward to define musical gesture as an action pattern, that produces music, that is encoded in music, or that is made in response to music. Qualifications to the term musical gesture can be added whenever needed to avoid misunderstandings. For example, one can speak about sound-producing gestures, sound-modifying gestures, sound-accompanying gestures, sonic gestures, playing gestures and so on. The essential point is to have a terminological apparatus that is sufficiently specific for differentiating subtle aspects of gestures when dealing with music. The main advantage in using the term gesture is that it surpasses the Cartesian divide between physics and mind. As mentioned above, we may think of movement as the changing of a physical position of a body part or an object, which can be objectively measured. The notion of gesture goes beyond this purely physical aspect in that it involves an action as a movement unit, or a chunk, which may be planned, and goal-directed, and perceived as a holistic entity (Buxton 1986).

4. Concepts for Studying Musical Gestures

Musical gestures may be studied from different viewpoints, such as the subjective, objective and communicative viewpoint (see Chapter LemanXXX), or the phenomenological, biomechanical and functionalist viewpoint (Ramstein (1991). The subjective phenomenological level has a focus on the descriptive aspects of gestures, such as describing gestures in terms of its cinematic (e.g. the speed), spatial (the size of space), and temporal dimensions (e.g. frequency range). The objective or intrinsic level has a focus on the conditions for gesture generation, such as various biomechanical and motor control constraints (see Chapter 9). The communicative or functional level has a focus on the purpose of a movement or action in a certain context, e.g. whether it is sound-producing, sound-modifying, and so on. All these three levels operate within a performance environment in which spatial aspects of musical gestures are constrained. Below, we first define the role of the performance environment. Then we go deeper into the communicative/functional level.

4.1 Spatial Aspects of Musical Gestures

The performance environment in which musical gestures are carried out can be conceived in terms of three concepts, namely *scene*, *position* and *space*, as illustrated in Figure 2. First, the *performance scene* may be thought of as a physical space that both

the performer and the perceiver recognize as one in which a performance is carried out. This is evident as both performers and perceivers tend to change their attention when the performer enters the performance scene. In a typical concert situation the performance scene is clearly defined as a part of the *stage*, a location everyone's attention is naturally focused on due to the seating and general expectation of what is to come. But a performance scene may also refer to a social construct, and may thus be created anywhere. A typical example is how street musicians claim a part of the pavement as their performance scene, which people walking past will usually respect.

Figure 2: The *performance scene* is the imagined area in which performance can happen. The *home position* is the position where the musician is sitting (or standing) at ease before starting to perform. The *start position* is where the performance starts from, and the *performance position* is the position(s) of the musician during performance.

Concerning the *performance positions*, it is possible to define the *home-position* of a performer to be the resting position in which the performer sits or stands before starting to act (Sacks and Schegloff 2002). In a musical context, and particularly in Western classical music, this can be understood as when a musician is standing or sitting at ease with the instrument before starting to perform. When in home position, the perceiver will usually know that the performance has not yet begun and will wait until the performer moves into *start position* before expecting any sound to be produced. Finally, the *performance position* is the one from which the performance action originates.

Reference can then be made to a set of *performance spaces*, or a *personal space*. Laban's (1963) term *kinesphere* denotes an imaginary box surrounding a person, which defines the maximum movement possibilities from a certain point in space. Laban argues that the kinesphere is a mental construct that one is always aware of in interaction with the environment and with others. Within the kinesphere, it is possible to further distinct between different *performance spaces* or *gesture spaces*, i.e. imaginary bounding boxes for various types of musical gestures. For example, when playing the piano the performer has a well-defined sound-producing gesture space in the visual part of the keyboard, as indicated in Figure 3. This gesture space can usually be observed by

both the performer and the perceiver, and makes it possible to identify where the sound-producing gestures are being carried out.

Figure 3 also indicates the performance spaces of other types of musical gestures. The identification of these spaces is needed in order to have a clear understanding of where different types of gestures should be carried out in relation to an object (e.g. an instrument). This knowledge of performance spaces for various types of musical gestures is helpful for setting up expectations when perceiving a performance. This is why audiences may be surprised if a musician happens to perform outside of such conventional performance spaces, for example by playing with the fingers on the strings of the piano instead of on the piano keyboard.

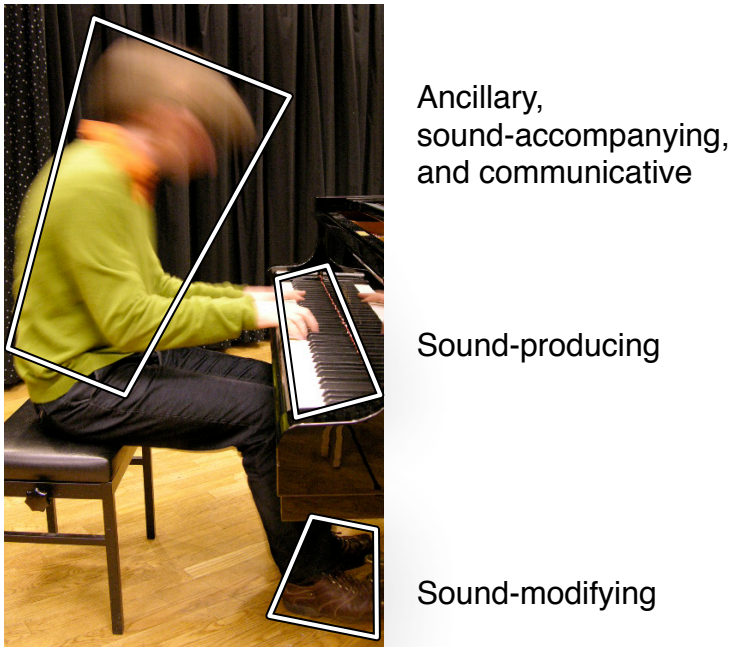

Figure 3: The *gesture space* can be seen as an imaginary box surrounding the space in which performance movements can be carried out. Here the gesture spaces for various musical movements are indicated.

Furthermore, a sound-producing gesture can be defined as an excitatory action, i.e. an action of setting some object (e.g. parts of an instrument) into motion by hitting, stroking, or bowing. Godøy (2008) has suggested that a sound-producing gesture can be seen as consisting of an *excitation phase*, a phase where there is contact with, and energy transfer to, the instrument, combined with a preceding *prefix*, a movement trajectory to the point of contact, and a succeeding *suffix*, a movement trajectory away from the point of contact, such as depicted in Figure 4. The prefix is the part of a sound-producing gesture quite simply because the effector (finger, hand, arm) has to move from an initial position to the contact position, but is also important for defining the quality of the excitation. The suffix is the return to equilibrium, or the initial position, after the excitation. Adapting Kendon's terminology, one could talk of a *gesture unit*,

defined as a *goal-directed movement excursion*, which starts and ends in a home position (Kendon 2004, p.111).

It should also be noted that gestures might be nested, in the sense that several actions that follow each other may be perceived as one coherent gesture. For example, playing a scale run on a piano may be seen as a series of separate actions if the focus is on the finger movements, but can also be perceived as one coherent gesture if the focus is on the movement of the hand or the upper body. Apparently, humans are able to perceive many concurrent actions simultaneously, and these concurrent action layers are often an integral part of the musical texture (see Chapter 5).

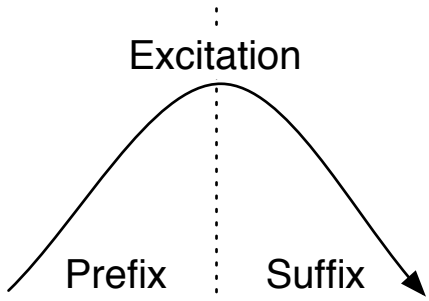

Figure 4: A sound-producing gesture may be seen as having an excitation phase surrounded by a prefix and suffix. These three are closely connected and are important for both its performance and its perception.

4.2 Functional Aspects of Musical Gestures

To understand more about the functions of various musical gestures, it is straightforward to discern four functional categories of musical gestures, based on work by Gibet (1987), Cadoz (1988), Delalande (1988) and Wanderley and Depalle (2004), namely, sound-producing gestures, communicative gestures, sound-facilitating gestures and sound-accompanying gestures.

- *Sound-producing gestures* are the ones that are effective in producing sound. They can be further subdivided into gestures of *excitation* and *modification*. Sound-producing gestures are called *instrumental gestures* in (Cadoz 1988), and *effective gestures* in (Delalande 1988).
- *Communicative gestures* are intended mainly for communication. As will be discussed later in this chapter such movements can be subdivided into *performer–performer* or *performer–perceiver* types of communication. Communicative gestures are called *semiotic gestures* in (Cadoz and Wanderley 2000). Several of these can also be considered *gestures* in the way Kendon (2004) and McNeill (1992) use the term.
- *Sound-facilitating gestures* support the sound-producing gestures in various ways. As will be discussed and exemplified in a later section, such gestures can be

subdivided into *support*, *phrasing* and *entrained* gestures. Sound-facilitating gestures are called *accompanying gestures* in (Delalande 1988), *non-obvious performer gestures* in (Wanderley 1999), and *ancillary gestures* in (Wanderley and Depalle 2004).

- *Sound-accompanying gestures* are not involved in the sound production itself, but follow the music. They can be *sound-tracing*, i.e. following the contour of sonic elements (Godøy et al 2006a), or they can *mimic* the sound-producing gestures (Godøy et al 2006b).

Figure 5 shows an illustration of different types of musical gestures in piano performance. Note that the different categories are not meant to be mutually exclusive, as several gestures have multiple functions. For example, hitting a final chord followed by a theatrical lift can be seen as having sound-producing and sound-facilitating as well as communicative functions. This functional multiplicity is illustrated in the *dimension spaces* in Figure 6. Dimension spaces are commonly used to analyze interactive systems (Graham et al. 2000), and have also been used to analyze the functionality of digital musical instruments (Birnbaum et al. 2005). Here they are used to visualize how the gestures of a musician and a dancer cover different functions.

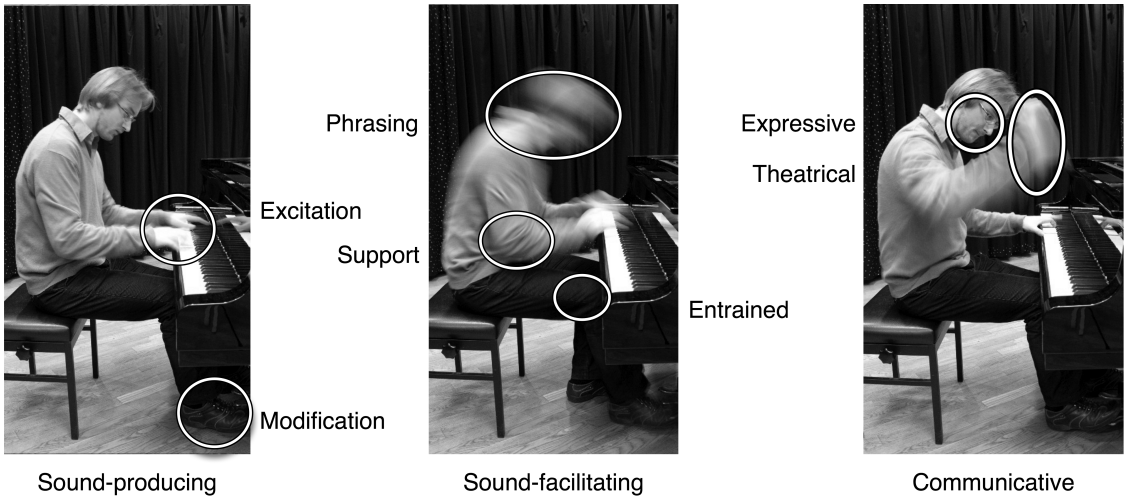

Figure 5: Examples of where different types of musical gestures (sound-producing, sound-facilitating and communicative) may be found in piano performance.

Based on the typology presented by Cadoz (1988), the sound-producing gestures can be divided in *excitation* and *modification gestures*. Excitation gestures may be further subdivided into *impulsive*, *sustained* and *iterative* actions, each having distinct energy profiles. Excitation gesture are either *direct* or *indirect*, depending on whether or not there is an object between the sound-producing element of the instrument and the object carrying out the excitation. For example, the actions of playing the harp or congas are *direct* since fingers and hands are directly in control of the resonating objects (strings and drum membrane). But there are also many *indirect* instrumental actions that involve

one or more objects in the interaction, for example the bow in string instruments, the key mechanism on the piano, or sticks for drums.

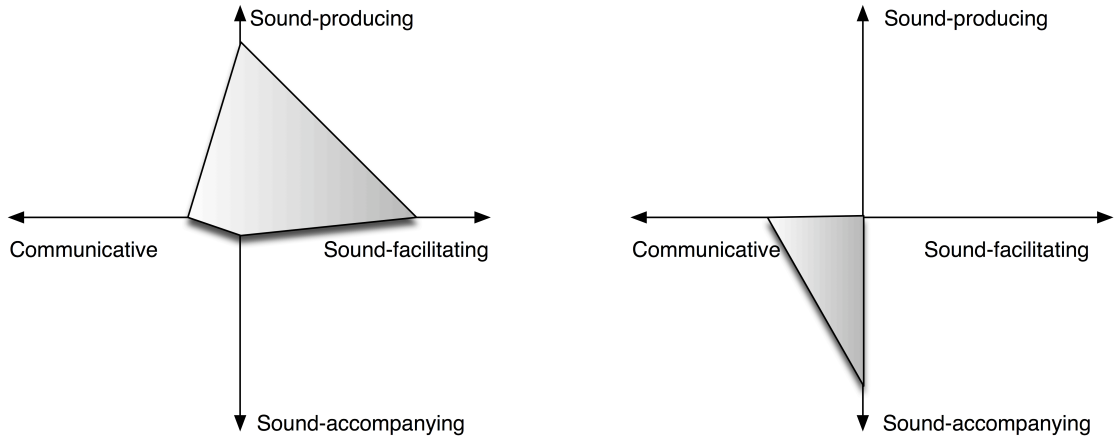

Figure 6: Dimension spaces illustrating how the gestures of a musician (left) and a dancer (right) may be seen as having different functions. Here the musician's movements have a high level of sound-producing and sound-facilitating function, while the dancer's movements have a high level of sound-accompanying and some communicative function.

Modification gestures do not actually produce sounds themselves, but they modify the quality of the sound. Cadoz (1988) suggests to further subdivide these gestures in (i) parametric gestures that continuously change a parameter, such as bow pressure in violin playing, and (ii) structural gestures that modify or change the structure of the object, such as placing a mute on a trumpet.

Most musical instruments are played with both excitation and modification gestures (Kvifte 1989). These two gesture functions may to a certain extent be separable such as on string instruments where the two hands play different roles: the left hand is mainly modifying the sound (choosing the pitch) while the right hand is carrying out the excitation, yet there are features of the bowing movement such as speed, pressure, acceleration, and tilt, that can be used to modify the sound (see Chapter 8).

It is important to keep in mind that a gesture typology, like all other typologies, is not to create an absolute classification system, but rather to point out some of the different functions of concepts. Furthermore, all performance movements can be considered a type of communication, but we find it useful to have a separate category for movements that are primarily intended to be *communicative*. These may be *performer–performer* and *performer–perceiver* types of communication, and range from communication in a linguistic sense (*emblems*), to a more abstract form of communication.

The term *sound-facilitating gesture is* used to cover different types of musical gestures that are not directly involved in sound production, but still play an important part in shaping the resultant sound. For example, hitting a piano key involves not only the

active finger, but also the hand, arm, and upper body. Such movements are *support* movements of the sound-producing actions. In fact, it is the preparatory movements of this complex multi-joint system that determine the trajectory and velocity of the finger before and after it hits the key. Thus, such support movements play an important role in supporting the sound-producing actions, and they may even have audible components, as shown by Wanderley (1999) in a study of clarinet performance. Here the performer's ancillary movements were seen in the movement of the clarinet bell, and this clarinet movement was shown to have an audible component due to the changing sound diffusion pattern of the instrument.

A different type of sound-facilitating gestures are the so-called *phrasing* gestures, since they are closely connected to musical phrasing. Wanderley (2002) has shown that the ancillary movements of clarinetists are an integral part of the instrumentalists' performance and are stable and reproducible even after long periods of time. Many of these repeatable movement patterns seem to be closely connected to the phrases in the music being performed, and are often related to movement of the clarinet bell (Campbell et al. 2005; Quek et al. 2006). However, it should be noted that these movements might also have a communicative function in enhancing the perceivers' experience of the phrasing of the sound.

The multi-functionality of a movement, or a gesture is well illustrated in so-called *entrained* movements, like tapping a foot, nodding the head or moving the whole upper body in synchrony with the music. These movements may help the musician to keep track of the tempo, and serve as a signal to other performers, dancers or the perceiver. Although such movements vary considerably between performers and performance styles, they may be thought of as important for the timing in a performance. It is important to notice that entrained movements can be a generator of rhythm and timing, in the same way as the rhythm and timing in music can be a generator of movements (Clarke 1999). In the clarinet experiments by Wanderley et al. (2005), the performers continued to move, albeit less, when asked to play "immobilized". The function of ancillary movements is to help the performer's needs, and could be separated from movements that are intended mainly for communication with the other performers and the audience. As such, ancillary movements are usually not carried out with a specific intention other than being the basis for, or result of, the sound-producing gestures.

Examples of support and phrasing gestures are shown in Figure 8 of a clarinetist performing the beginning of the *Allegro appassionato* from the Clarinet Sonata Op. 120, No. 1 in F minor by Brahms (1894). This *motiongram* (Jensenius, 2006) was made from a video recorded for the *Workshop on Motion Capture for Music Performance* at McGill University in October 2006,[2] and Figure 7 shows a snapshot from the video. The motiongram shows the vertical movements of the clarinetist over time (time running left to right), and facilitates studying the vertical movement of the

[2] http://www.music.mcgill.ca/musictech/idmil/mocap/mocap.html

clarinet bell, and the movement of the centre of gravity of the performer. Note how the motiongram visualizes the rhythmic sway and weight changes of the clarinetist, which seem to correspond to the breathing patterns and musical phrasings as discussed by Wanderley et al. (2005).

Figure 7: Mark Bradley performing the beginning of the *Allegro appassionato* from Clarinet Sonata Op.120, No.1 in F minor by Brahms at the Input Devices and Music Interaction Laboratory, McGill University.

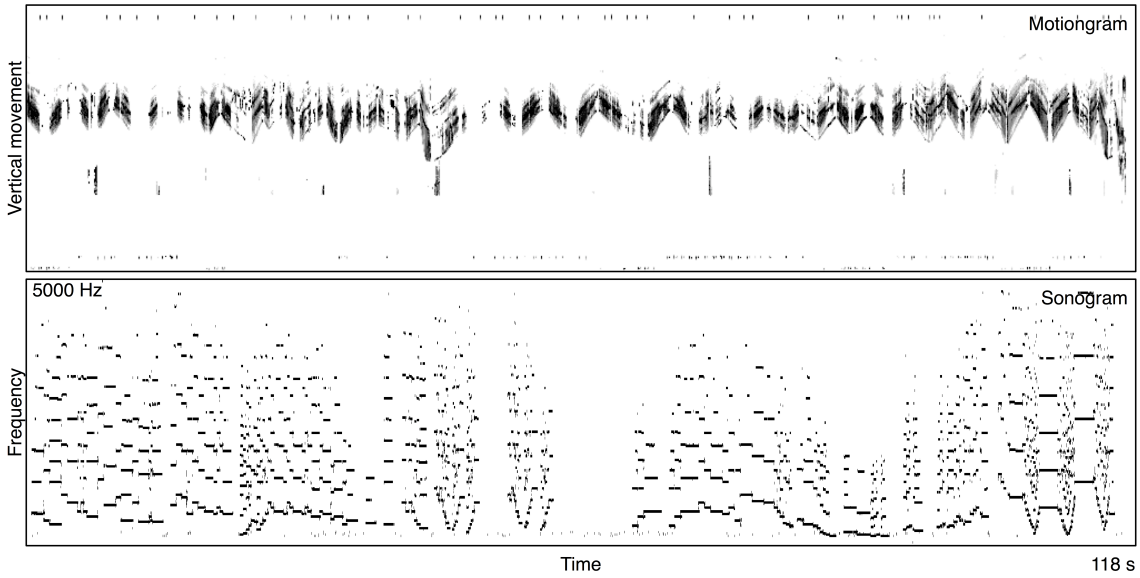

Figure 8: Motiongram made from a video recording of clarinet performance (as shown in Figure 7). The thin dotted line in the top part of the motiongram represents the changing light from an infrared motion capture camera that was also active during recording. The thin dotted line at the bottom of the display is the rhythmic tapping of the clarinetist's toes. The few vertical lines below the movement of the clarinet, shows where the clarinetist shifted his weight during performance.

The sound-facilitating gestures can be distinguished from *sound-accompanying gestures*. These are neither part of, nor ancillary to, the sound-production, but rather

intended to *follow* features in the sound. Dancing to music is perhaps the most common type of sound-accompanying movements, and something which was explored in an observation study of how dancers' movements were following qualities in the music (Casciato et al., 2005; Jensenius, 2007; Haga, 2008). However, sound-accompanying gestures may also involve entrainment. The Brazilian Samba provides an interesting example (Naveda and Leman, 2008a; in press), with strong evidence that dancing imposes a corporeal metrical grid onto musical structures that are inherently ambiguous. Thus, it is through sound-accompanying gestures, or dancing, that structure is given to the music. This complies with other recent findings that motion may influence perception (Philips-Silver and Trainor, 2008). Another type of sound-accompanying movements is what could be called *sound-tracing*, meaning tracing by hand in the air or on a surface some prominent features or features of sound. This was studied in (Godoy et al. 2006b, Leman et al., in press), where people were asked to trace sounds with a digital pen on a graphical tablet, and with a telescopic stick. Yet another example of such sound-accompanying gestures is that of *air instrument performance*, meaning imitating sound-producing gestures in the air. This is something which has received increased interest with the annual air guitar world championship, and been studied in (Godoy et al. 2006a) (See also chapters 5, 6, and 7).

5. Methods for studying musical gestures

Musical gestures are characterized by a multi-functional nature through which multiple meanings are generated. Consequently, the study of musical gesture requires an inter-disciplinary approach, with contributions from a diverse set of disciplines, including physics of musical instruments and acoustics, biomechanics and human motor control, auditory and visual perception, musical performance and dance, music theory, music technology, robotics and HCI, aesthetics and various social sciences, including the study of emotions. Given the nature of musical gestures, progress will be obtained by combining different scientific methodologies, most particularly, the methodologies from the natural sciences and the human sciences (Leman, 2008). Indeed, it is a major challenge to combine overt, observable, and measurable information with more subjective, descriptive, or even effable sensations. Being an eminently inter-disciplinary enterprise, and in particular an enterprise that combines traditionally purportedly conflicting approaches in the human sciences epitomized in 'qualitative' versus 'quantitative', it is instructive to have a brief look at some of the main perspectives in the study of musical gestures:

First, consider observation and/or introspection. The challenge is in bridging the gap between what we can spontaneously see or sense of musical gestures, be that in performance, in various sound-related movement, or just in listening, and that which we can somehow document in motion and sound data. The long-term goal would be that of discerning and give explicit representations of as many as possible features of musical

gestures, and in our work towards such a goal, there will probably be a need for alternating between observation and introspection.

Second, consider qualitative and/or quantitative methods. There seems to be converging evidence that we are quite good at perceiving movement qualities such as fast, slow, agitated, calm, tense, relaxed, in both musical gestures and in sound. One important challenge here will then be that of substantiating or documenting the basis in movement and sound for these sensations. This means to move from initial qualitative sensations to more quantitative representations by mapping out as many as possible pertinent features of movement and sound. This presents us with many methodological and technological issues of motion capture, processing, and representation of motion and sound data.

Third, consider motion capture. This is about trying to find out as much as possible of the motion in musical gestures. The technologies for this include video-based computer vision techniques, infrared, electromagnetic, ultra-sound, mechanical and inertial motion capture systems, and many other techniques. Such technologies can to a varying degree of precision and accuracy send out streams of data indicating the position and/or relative motion of various points on the body. This will in most cases result in large data sets, presenting important challenges of processing and representation in order to be useful for further analyses.

Fourth, consider processing and representation. Most kinds of motion capture data need to be filtered and transformed in order to be useful. Also, the way this data is displayed presents substantial challenges, such as how to visualize the simultaneous movements at different speeds of different points on the body in different dimensions. Synchronizing and comparing such motion capture data with other related data (e.g. MIDI) and media (e.g. audio and video) is yet another challenge, not to forget trying to find some criteria for judging similarities and differences between variant types of movements and sounds.

Next consider simulations and/or animations. Trying to simulate the kinematics and dynamics of various musical gestures may be very instructive for understanding how these gestures work. This approach implies work that is more in line with the so-called analysis-by-synthesis strategy practiced for several decades in the domain of digital sound synthesis. This means synthesizing incrementally different variants in order to discover what are the perceptually most salient features. Examples of such analysis and synthesis strategies are presented in (Bouënard, Gibet and Wanderley 2008a, b, Naveda and Leman, 2008b).

Finally, consider annotation and interpretation. Apart from the increasingly sophisticated methods and technologies for motion capture, 'naked eye' observations of musical gestures will always be needed, similar to 'naked ear' observations of gesture-related cues in the musical sound. Various schemes for making annotations have

already been developed, but there is clearly a need to continue this work in parallel with various motion capture technologies. This also goes for the largely unexplored field of social identities in musical gestures, such as studying how different cultures or cultural sub-groups have developed specific features of musical gestures and how they seem to work in various social contexts.

Needless to say, the study of musical gestures is a vast area comprising several different insights and skills.

6 Conclusion

Up to now, there is no single well-defined definition of the notion of gesture, although most authors seem to agree that gestures involve both body movement and meaning. Gestures connect well with recent approaches in embodiment music cognition, and they can be considered vehicles of human musical communication. However, this approach to the notion of gesture remain vague to some extent. In many instances of the use of the notion of gesture it will therefore be necessary to differentiate in more detail. This chapter shows that it is indeed possible to develop a proper conceptual apparatus for differentiating various functions that may be attached to musical gestures. Figure 9 gives a summary of some of the main functions that can be attributed to performers in their environment. In addition to these, also it was briefly mentioned that sound-accompanying gestures can be observed in the movements of perceivers (e.g. dancers). In all this, it is important to remember the multi-functionality of gestures, that one single gesture may have multiple functions and significations, but where in most (if not all) cases, these different functions and/or significations are but different facets of our rich musical experiences. Lastly, we should also keep in mind that musical gestures also have multiple significations ranging from the more physical to the more metaphorical, hence that we also should welcome a multiplicity of approaches to the study of musical gestures.

References

Birnbaum, D., Fiebrink, R., Malloch, J. & Wanderley, M. M. (2005), Towards a dimension space for musical artifacts, in 'NIME '05: Proceedings of the 2005 International Conference on New Interfaces for Musical Expression', Vancouver: University of British Columbia, Vancouver, BC, pp. 192–195.

Bouënard, A., Gibet, S. & Wanderley, M. M. (2008a). Enhancing the visualization of percussion gestures by virtual character animation. In 'NIME '08: Proceedings of the 2008 International Conference on New Interfaces for Musical Expression (NIME08)', Genoa, Italy, pp. 38–43.

Bouënard, A., Gibet, S. & Wanderley, M. M. (2008b). Analysis of percussion grip for physically-based character animation, in 'ENACTIVE '08: Proceedings of the 5th International Conference on Enactive Interfaces', Pisa, Italy, pp xxx.

Brahms, J. (1894), 'Clarinet sonata op.120, no.1 in f minor'.

Buxton, W. (1986), Chunking and phrasing and the design of human-computer dialogues, in 'Information Processing', Elsevier Science Publishers.

Cadoz, C. (1988), Instrumental gesture and musical composition, in 'Proceedings of the 1998International Computer Music Conference', Den Haag, Netherlands, pp. 60–73.

Cadoz, C. & Wanderley, M. M. (2000), Gesture – Music, in M. M. Wanderley & M. Battier, eds, 'Trends in Gestural Control of Music [CD-ROM]', IRCAM – Centre Pompidou, Paris, pp. 71–94.

Campbell, L., Chagnon, M.-J. & Wanderley, M. M. (2005), On the use of Laban-Bartenieff techniques to describe ancillary gestures of clarinetists, Research report, Input Devices and Music Interaction Laboratory, McGill University. URL: http://www.music.mcgill.ca/musictech/clarinet/LBMF Final% Report.pdf

Camurri, A., De Poli, G., Leman, M. & Volpe, G. (2001), A multi-layered conceptual framework for expressive gesture applications, in 'Proceedings of the International MOSART Workshop, November 2001', Barcelona, Spain.

A. Camurri, G. Volpe, G. De Poli, M. Lema (2005). ``Communicating expressiveness and affect in multimodal interactive systems''. IEEE Multimedia, 12 (1), pp. 43-53.

Casciato, C., Jensenius, A. R. & Wanderley, M. M. (2005), Studying free dance movement to music, in 'Proceedings of ESCOM 2005 Performance Matters! Conference', Porto, Portugal.

Clarke, E. (1993), 'Generativity, mimesis and the human body in music performance', Contemporary Music Review 9(1), 207–219.

Clarke, E. (1999), Rhythm and timing in music, in D. Deutsch, ed., 'The Psychology of Music', 2nd edn, Academic Press, San Diego, pp. 473–500.

Coker, W. (1972), Music & Meaning: A Theoretical Introduction to Musical Aesthetics, Free Press, New York.

Davidson, J. (1993), 'Visual perception and performance manner in the movements of solo musicians', Psychology of Music 21, 103–113.

Davidson, J. (1994), Expressive movements in musical performance, in 'Proceedings of the Third International Conference on Music Cognition (ESCOM)', Liege, Belgium, pp. 327–329.

Delalande, F. (1988), La gestique de gould: Elements pour une sémiologie du geste musical, in G. Guertin, ed., 'Glenn Gould Pluriel', Québec: Louise Courteau, pp. 85–111.

Ekman, P. & Friesen, W. (1969), 'The repertoire of nonverbal behavioral categories', Semiotica 1, 49–98.

Feyereisen, P. & de Lannoy, J.-D. (1991), Gestures and Speech: Psychological Investigations, Cambridge University Press, Cambridge.

Gabrielsson, A. (1985), 'Interplay between analysis and synthesis in the studies of music performance and music experience', Music Perception 3(1), 59–86.

Gibet, S. (1987), Codage, representation et traitement du geste instrumental: Application a la synthese de sons musicaux par simulation de mecanismes instrumentaux, PhD thesis, L'Institut National Polytechnique de Grenoble.

Godøy, R. I. (2008), Coarticulated gestural-sonorous objects in music, in 'Music and Gesture 2'.

Godøy, R. I., Haga, E. & Jensenius, A. R. (2006a), Exploring music-related gestures by sound-tracing. -a preliminary study, in '2nd ConGAS International Symposium on Gesture Interfaces for Multimedia Systems, May 9-10 2006', Leeds, UK.

Godøy, R. I., Haga, E. & Jensenius, A. R. (2006b), Playing air instruments: Mimicry of sound-producing gestures by novices and experts, in S. Gibet, N. Courty & 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', Vol. 3881/2006, Berlin Heidelberg: Springer-Verlag, pp. 256–267.

Graham, T. C. N., Watts, L. A., Calvary, G., Coutaz, J., Dubois, E. & Nigay, L. (2000), A dimension space for the design of interactive systems within their physical environments, in 'DIS '00: Proceedings of the conference on Designing interactive systems', New York, ACM Press, pp. 406–416.

Gritten, A. & King, E., eds (2006), Music and Gesture, Hampshire: Ashgate.

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

Hatten, R. S. (2003), 'Musical gesture: Theory and interpretation. course notes, indiana university. http://www.indiana.edu/~deanfac/blfal03/mus/mus t561 9824.html (accessed 1 june 2007)'.

Hatten, R. S. (2004), Interpreting musical gestures, topics, and tropes : Mozart, Beethoven, Schubert, Bloomington, IN: Indiana University Press.

Hodgins, P. (1992), Relationships Between Score and Choreography in Twentieth Century Dance: Music, Movement and Metaphor, Lewiston, NY: Edwin Mellen Press.

Jensenius, A. R. (2006), Using motiongrams in the study of musical gestures, in 'Proceedings of the 2006 International Computer Music Conference', San Francisco: ICMA, New Orleans, LA, pp. 499–502.

Jensenius, A. R. (2007), Action–Sound : Developing Methods and Tools to Study Music-Related Bodily Movement, PhD thesis, University of Oslo.

Kendon, A. (1972), Some relationships between body motion and speech, in A. Siegman & E. Pope, eds, 'Studies in dyadic communication', Pergamon Press, New York, pp. 177–210.

Kendon, A. (1980), Gesticulation and speech: Two aspects of the process of utterance, in 'The Relationship between Verbal and Nonverbal Communication', Mouton, The Hague, pp. 207–227.

Kendon, A. (1982), 'The study of gesture: some remarks on its history', Recherches Sémiotiques/Semiotic Inquiry 2, 45–62.

Kendon, A. (2004), Gesture: Visible Action as Utterance, Cambridge: Cambridge University Press.

Kurtenbach, G. & Hulteen, E. A. (1990), The art of human-computer interface design, in B. Laurel, ed., 'Gestures in Human-Computer Communication', Reading, PA: Addison Wesley, pp. 309–317.

Kvifte, T. (1989), Instruments and the Electronic Age. Towards a Terminology for a Unified Description of Playing Techniques, Solum Forlag, Oslo. URL: http://www.hf.uio.no/imv/om-instituttet/nfs/iaea/instrume.html

Laban, R. v. (1963), Modern Educational Dance, London: Macdonald & Evans Ltd.

Leman, M. (2007), Embodied Music Cognition and Mediation Technology, The MIT Press, Cambridge, MA.

Leman, M, Desmet, F, Styns, F., van Noorden, L., and Moelants, D. (in press). Sharing musical expression through embodied listening – a case study based on Chinese guqin music. Music Perception.

M. Leman (2008), ``Systematic musicology at the crossroads of modern music research", Hamburger Jahrbuch für Systematische Musikwissenschaft. In Schneider, A. (Ed.), Systematic and Comparative Musicology: Concepts, Methods, Findings. Hamburger Jahrbuch für Musikwissenschaft, 24. Frankfurt am Main: Peter Lang, pp. 89-115.

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

McNeill, D. (1992), Hand and Mind: What Gestures Reveal About Thought, University of Chicago Press, Chicago, IL.

McNeill, D. (2005), Gesture and Thought, University of Chicago Press, Chicago, IL.

McNeill, D., ed. (2000), Language and Gesture, Cambridge University Press, Cambridge.

Métois, E. (1997), Musical Sound Information: Musical Gestures and Embedding Synthesis, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Middleton, R. (1993), 'Popular music analysis and musicology: Bridging the gap', Popular Music 12(2), 177–190.

Miranda, E. R. & Wanderley, M. M. (2006), New Digital Musical Instruments: Control and Interaction Beyond the Keyboard, Vol. 21 of The Computer Music and Digital Audio Series, A-R Editions, Inc., Middleton, WI.

Naveda, L. and Leman, M. (in press). ``A cross-modal heuristic for periodic pattern analysis of Samba music and dance"

Naveda, L. and Leman, M. (2008a). Representation of Samba dance gestures, using a multi-modal analysis approach. Proceedings of the 5th International Conference on Enactive Interfaces, october 2008, Pisa.

Naveda, L. and Leman, M. (2008b). Sonification of Samba dance using periodic pattern analysis. Proceedings of the Artech Conference, Porto.

Noë, A. (2004). Action in Perception. Cambridge, MA: The MIT Press.

Pierce, A. & Pierce, R. (1989), Expressive Movement: Posture and Action in Daily Life, Sports, and the Performing Arts, Perseus Publishing, Cambridge, MA.

Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. Brain and Cognition, 67(1), 94-102.

Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.-F., Kirbas, C., McCullough, K. E. & Ansari, R. (2002), 'Multimodal human discourse: gesture and speech', ACM Transactions on Computer-Human Interaction (TOCHI) 9(3), 171–193.

Quek, O., Verfaille, V. & Wanderley, M. M. (2006), Sonification of musician's ancillary gesutres, in 'Proceedings of the Proceedings of the 12th International Conference on Auditory Display', London, UK.

Ramstein, C. (1991), Analyse, représentation et traitement du geste instrumental, Th`ese de docteur ingénieur spécialité informatique, Institut National Polytechnique de Grenoble, France.

Sacks, H. & Schegloff, E. (2002), 'Home position', Gesture (2), 133–146.

Shaffer, L. H. (1980), Analyzing piano performance: a study of concert pianists, in G. E. Stelmach & P. A. Vroon, eds, 'Tutorial in Motor Behaviour', North-Holland, Amsterdam, pp. 443–456.

Todd, N. P. M. (1995), 'The kinematics of musical expression', The Journal of the Acoustical Society of America 97(3), 1940–1949.

Varela, F., Thompson, E., and E. Rosch (1991). The Embodied Mind. Cambridge, MA: The MIT Press.

Wanderley, M. M. (1999), Non-obvious performer gestures in instrumental music, in A. Braffort, R. Gherbi, S. Gibet, J. Richardson & D. Teil, eds, 'Gesture-Based Communication in Human-Computer Interaction: International Gesture Workshop, GW'99, Gif-sur-Yvette, France, March 1999. Proceedings', Berlin Heidelberg: Springer-Verlag, pp. 37–48.

Wanderley, M. M. (2001), Performer-Instrument Interaction: Applications to Gestural Control of Sound Synthesis, Phd-thesis, Université Pierre et Marie Curie, Paris VI, Paris, France.

Wanderley, M. M. (2002), Quantitative analysis of non-obvious performer gestures, in I. Wachsmuth & T. Sowa, eds, 'Gesture and Sign Language in Human-Computer Interaction: International Gesture Workshop, GW 2001, London, UK, April 18-20, 2001. Revised Papers', Vol. LNAI 2298, Berlin Heidelberg: Springer-Verlag, pp. 241–253.

Wanderley, M. M. & Depalle, P. (2004), 'Gestural control of sound synthesis', Proceedings of the IEEE 92(4), 632–644.

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C. & Hatch, W. (2005), 'The musical significance of clarinetists' ancillary gestures: An exploration of the field', Journal of New Music Research 34(1), 97–113.

Zhao, L. (2001), Synthesis and Acquisition of Laban Movement Analysis Qualitative Parameters for Communicative Gestures, PhD thesis, CIS, University of Pennsylvania, Philadelphia, PA.