Motormimetic Features in Musical Experience
Rolf Inge Godøy

1. Introduction
There can be little doubt that we experience links between different sense modalities in music: at a concert, we typically hear sound and see the sound-producing body motion of the musicians. At a dance performance, we may likewise perceive sound, motion, and probably also sense the effort made by the dancers. And for both sound and motion, we often encounter visual terms such as rough, smooth, narrow, broad, etc. in music-related discourse. In short, it should not be controversial to regard music as a multimodal form of art, since it involves other sensations in addition to sound. But precisely how the different sense modalities are activated and interact in musical experience, does indeed present us with a great number of unanswered questions.

We have in our own research been pursuing what we call motormimetic cognition (Godøy, 2001; 2003), denoting an incessant mental simulation of sound-related body motion in musical experience, primarily of assumed sound-producing body motion (e.g. hitting, kicking, stroking, bowing, blowing), but also of various kinds of sound-accompanying body motion (e.g. dancing, walking, gesticulating). We regard such mental simulation of body motion as applicable to most features of music, including sound onsets, dynamics, timbre, texture, rhythm, melody, harmony, articulation, and expressivity. Furthermore, we believe motormimetic cognition can create links between the different modalities, e.g. that mental simulation of sound-producing body motion can also link sound to vision and/or sound to touch. In short, we believe motormimetic cognition can be the basis for a coherent theory of feature mapping between different modalities in music, based on studying both musicians' and listeners' body motion and posture shapes in music-related contexts.

In this chapter we shall present what we see as motormimetic features in musical experience, starting with a review of some basic ideas of motormimetic cognition and associated concepts in musical thought, in what we call ontological reflections, preparing the ground for a closer look at the main elements of music-related body motion and salient features of musical sound. This leads on to what we call amodal shape cognition, a general cognitive faculty that may bridge gaps between modalities involved in musical experience, in turn leading to the topic of sound-motion objects, entities that saliently manifest motormimetic features in musical experience. We also believe that studying motormimetic features in musical experience can provide us with a conceptual apparatus and practical tools (in combination with new technologies) for exploring aesthetic and affective images of sound and motion in music, contributing to bridging gaps between quantitative and qualitative approaches in musical research.

2. Motormimetic cognition
The background for our focus on motormimetic cognition was an interest in musical imagery, defined as "our mental capacity for imagining musical sound in the absence of a directly audible sound source, meaning that we can recall and re-experience or even invent new musical sound through our 'inner ear'" (Godøy & Jørgensen, 2001, p ix). We were particularly interested in the triggers of mental images of musical sound, i.e. in how such images are produced. In line with the so-called ecological approach to mental imagery (Shepard, 1984), we envisaged that mental images of music would also include components of sound-production, both of the instruments (their material as well as resonant features) and the sound-producing body motion associated with any particular sound. This meant that music perception would also include the mental
imitation of such assumed sound-producing motion, hence the idea of motormimetic cognition in music (Godøy, 2001; 2003). For instance, listening to, or just imagining, the sound of ferocious drumming, would also engender sensations of energetic mallet and hand motion in our minds, and listening to, or imagining, soft string music, would lead to sensations of slow, protracted bowing motion.

Searching the literature for research on mental simulation of sound generation in perception and imagery, we became aware of research in linguistics that in the 1960s had suggested that speech perception also includes the mental simulation of the sound-producing motion of the vocal apparatus believed to be involved in the production of speech sound. Called the motor theory of perception (Liberman and Mattingly 1985), this idea was first met with a lot of opposition, however, in the past couple of decades it has become quite well supported by both behavioural and neurocognitive research.

Moreover, it has been suggested that the motor theory (with variants) could be a more general theory of perception, applicable to audition in general and to other sense modalities (Galantucci, Fowler, and Turvey 2006). Applications include vision, with the suggestion that perceiving visual patterns may also involve a mental tracing of the patterns (Berthoz 1997), and various other areas of human behaviour (Jeannerod, 2006), including social behaviour (Wilson & Knoblich, 2005), leading to the idea of a general motor ontology, i.e. a general theory of understanding the world in terms of actions and action goals (Gallese & Metzinger, 2003). Also, there have been studies suggesting close relationships between the perception of sound in general and motor-related brain areas (Kohler, Keysers, Umiltà, Fogassi, Gallese, & Rizzolatti, 2002). And as for suggestions of 'hard wired' connections, there is the so-called McGurk effect which demonstrates that images of sound production may even override the information provided by the sound (McGurk & MacDonald, 1976). Similar close relationships between sound and motion in the human brain have been documented in musical contexts (Haueisen & Knösche, 2001; Zatorre & Halpern, 2005; see Godøy, 2010a for an overview).

Curious about listeners' spontaneous associations of musical sound with body motion, we carried out some studies of so-called air instrument performance such as air guitar, air drums, and air piano, based on suspicions that even people with little or no musical training could be able to associate sound-producing body motion with musical excerpts that they heard (Godøy, Haga, & Jensenius, 2006). We were also interested in the spontaneous body motion that listeners would make to musical sound in what we called free dance (Haga, 2008), and more specifically in how listeners would spontaneously trace the subjectively perceived sound fragments they heard with hand motions in what we called sound-tracing (Nymoen, Godøy, Jensenius, & Torresen, 2013). In all these cases, we could observe a fair level of similarity in the listeners' motion to the musical sound, but there were of course also significant differences between listeners with different levels of musical training, i.e. trained listeners tended to render more details than untrained listeners.

One crucial feature of motormimetic cognition is the capacity for approximate, or 'imperfect' kinds of imitation, as we typically found in our own studies and as may be readily observed in other cases of imitation, such as in so-called scat singing and beatboxing, i.e. of trying to reproduce non-vocal sounds with the vocal apparatus. In terms of perception, we thus see what we could call a vague-to-salient continuum of similarity in imitation and recognition. Consider as an example someone like me who cannot speak or understand either Finnish or Arabic, but yet I believe I am able to
recognize the difference between these languages because the sound-producing vocal motion (so-called *phonological gestures*) is so different.

Interestingly, similar 'low-resolution' images seem also to work in vision where for instance the motion of just a very limited set of reflective markers, i.e. light reflecting points in so-called *motion capture* recordings, and no other information, are sufficient to effectively convey extremely salient features of body motion such as speed (slow, fast), body type (small, large), gender (male, female), and mode of gait (calm, agitated), as well as affective states (sad, happy, etc.). It seems then that images of music-related motion are both readily perceived by most people, and that they may also be rather sparse ('low-resolution') and still work quite well as carriers of salient information. And in sound perception, we may have similar low-resolution sensory material (e.g. highly compressed audio, poor signal to noise ratios, various distortions, etc.), yet still seemingly sufficient to afford meaningful musical experiences. This raises questions of what on the one hand are inherent (and physical) features of music, and what on the other are our internal (mental simulation based) factors shaping musical experience, questions that should encourage us to make some reflections on the ontology of musical experience.

3. Ontological reflections
When we delve deeper into questions of music-related body motion, we invariably have to deal with some basic issues of 'what is what' in music: granted that we often tend to think of music as a sonic art, what criteria do we have for sonic components in musical experience? Or: what is in the sound, and what is in our perceptions and our processing of sound (so as to make the mentioned, even quite poor acoustic renderings of musical sound somehow meaningful)? And what about all the other elements that go into making and perceiving musical sound? For instance, a rhythmical pattern played on a standard drum set: is this pattern a sonic pattern of drum and cymbal sounds? Or is it just as much a pattern of the motion of the mallets/hands/arms/ in relation to the instruments? Or could it be a combination that we shall later call *sound-motion objects*? And furthermore, as musical sound is ephemeral, present in a moment 'here-and-now' and then vanished, what then is a rhythmical pattern? Is it something that may be represented by notation for the performer, or by signal-based representation (e.g. a spectrogram), or by some other kind of visual pattern? And what about our memory image of such a rhythmical pattern: does it contain salient mental images of the various sounds? Or: does it rather contain a rather vague sensation of the body motion effort of drumming? Or does it again contain multimodal images with sound, vision, haptics, and sense of effort? These and similar questions call for what could be called *ontological reflections*, and far from being tedious exercises of categorization, are in fact necessary differentiations in our studies of motormimetic cognition in music.

For a start, it could be useful to take timescales into consideration, because a differentiation of timescales can help us better see the different components at work in the production and perception of musical sound and music-related motion. We have found it useful in our own research to consider both sound and body motion according to these three main categories of timescales:

- **Micro timescale** of continuous sound and motion including sound features such as pitch, dynamics, stationary timbre, and corresponding continuous motion features such as in bowing and blowing. It may be obvious, but most motion and sound is continuous within some minimum duration range, i.e. it is not mainly chaotic, but may have some chaotic components in addition to the predominantly non-chaotic
components, i.e. noise mixed into periodic sound, such as a certain amount of white noise in addition to harmonic components in bowed string instrument sound.

- **Meso timescale**, typically in the 0.5 to 5 seconds duration range, containing salient features such as overall dynamical, timbral, and pitch-related envelopes of sounds. The meso timescale undoubtedly contains the most salient musical features such as melodic, harmonic, rhythmical, and textural patterns. Notably, this is also the duration range of the so-called *sonic objects* in the theory of Pierre Schaeffer and co-workers (Schaeffer, 1966; Chion, 1983, Godøy, 2006) and of what we call *sound-motion objects*.

- **Macro timescale**, with several meso timescale chunks in concatenation, typically the timescale of sections, songs, movements, and whole works, and often with more superordinate dramaturgical and/or narrative effects.

The privileged role of the meso timescale has also been supported by findings in other domains of the cognitive sciences, in short, designating the meso timescale as optimal both in terms of attention, memory, action planning and control (Godøy, 2013). And as noted by Schaeffer and co-workers more than half a century ago, focusing on sonic objects on the meso timescale allows us to go beyond the limitations of Western common practice notation and to consider emergent holistic features of these sonic objects.

Undoubtedly, Western common practice notation has had a strong influence on the ontology of music in our Western culture: notation has enabled the conservation and ensuing reproduction of at times highly complex music in performances, but the information in this notation has been limited, requiring familiarity with various so-called *performance traditions* in parallel with the notation representations. The basis for Western common practice notation is *categorization* of pitch and duration, and to a limited extent, also of tempi, dynamics, timbre (by denoting instruments and/or mode of performance) and some rudimentary indications of expression. This has come at the cost of not having more explicit representations for the many and salient features of musical sound, e.g. nuances in pitch that may be found in various Indian musical cultures. As for ontology, this means that we need to clarify what we are referring to: when we speak of say a melody, are we referring to the notation, to a sequence of sounds, or to some more abstract mental entity of pitches in succession? And what about the timbral and expressive features? Are they considered secondary, just a kind of incidental ‘wrapping’ of the main content of melody?

Schaeffer made the distinction *abstract-concrete* in his music theory, denoting that which can be represented as symbols, i.e. such as pitch and duration in Western music theory, as abstract, and the until then mostly un-named dynamic, pitch-related, timbral, textural, etc. nuance features as concrete. The concrete sound features were typically characterized by *shape*, potentially applicable to all perceptual and physical features of musical sound, and also to music-related body motion (Godøy, 2006).

The basic idea of sonic objects was that of having a holistic perception of a fragment of musical sound from any source (traditional instruments, electronic instruments, any environmental sound), typically in the duration range of very approximately 0.5 to 5 seconds. Research on sonic objects started with depicting the overall dynamic and pitch-related shapes of these sounds (e.g. of a violin tone, of a snare drum roll, of a squeaking door, etc.) and continued with differentiating various sub-features of these sonic objects. For one thing, this meant that the entire sonic object must be kept in memory as a whole, as it was found that the overall shape of the sonic object and the succession of sub-features together influenced the emergent
perceptual qualities of the sonic object as a whole. In terms of ontology, this means that the sonic object is a mental entity, relying on memory and imagery, but according to Schaeffer, having a correlate in the physical sound signal (Schaeffer, 1966).

This holistic perception of objects from a continuous stream of sensations seems to fit well with basic ideas in phenomenology: as suggested by Husserl, perception will by necessity have to proceed by a series of more discontinuous points in time, by what he called 'now-points', i.e. intermittent points where we may accumulate the continuous stream of sensations into somehow meaningful chunks (Husserl, 1991; Godøy, 2010c). It can be argued that such a discontinuous element of chunk-by-chunk perception is also well in accordance with more recent findings in human motor control, i.e. with the need for anticipatory plans in the form of action gestals (Klapp & Jagacinski 2011; Godøy, 2011; 2013), and hence that we can apply a similar top-down, object-focused approach to both sound and motion in our research.

4. Music-related body motion
Finding music-related body motion in so many different contexts, through an international research effort lasting several years, we have tried to develop a conceptual apparatus for handling different kinds of music-related motion. According to Godøy & Leman (2010), we may think of the following two main categories of music-related motion:

- **Sound-producing**, meaning motion that transfers energy from the human body to the instrument (or the vocal apparatus in the case of singing), first of all in so-called excitatory motion, e.g. hitting, stroking, bowing, blowing, etc. but which may also include motion that modifies the sound in so-called modulatory motion, e.g. the vibrato motion of the left hand on a violin, the opening and closing of a mute on a trumpet, etc. In this main category, we also find various kinds of so-called ancillary motion, motion that is not strictly speaking producing sound, yet often necessary for facilitating the sound-producing motion, for avoiding fatigue and/or strain injury, and for helping in the expressive shaping of the music, such as sweeping elbow, shoulder, and even torso motion in piano performance. Various kinds of communicative motion may also be included here, motion directed towards fellow musicians and/or towards the audience.

- **Sound-accompanying**, meaning all kinds of motion somehow related to the sound features of the music, typically in synchrony with the perceived beat such as in dance, walking, and other kinds of entrained motion, or more free kinds of gesticulating to some prominent features of the music such as a melodic contour or a particular percussion pattern. There may be similarities between the energy envelopes of the sound and the motion, e.g. smooth motion to calm, protracted sound, more jerky motion to heavily accented sound. But in some cases, sound-accompanying motion may be more independent of the music, and add some new contrapuntal element to the music.

Notably, the boundaries between these categories may not always be so clear, e.g. we may find sound-accompanying motion in dance or in other contexts resembling sound-producing motion. A famous example of this is in the Barber Scene in Charlie Chaplin's *The Great Dictator* where the barber (Chaplin) carries out various shaving motions as sound-accompanying motion to Brahms' *Hungarian Dance Number 5*, and where there is a striking similarity between the shapes of the sound-producing motion and the barber's motion (Godøy, 2010b). Also, motion may be multifunctional, such as when a pianist raises elbows and shoulders in preparation for the following
accented downbeat chord, this motion may also serve as an upbeat cue to fellow musicians and even as a theatrical gesture of vitality towards the audience.

With motion capture technologies (and also with video) and suitable processing tools, it is now possible to extract several salient motion features such as the overall quantity of motion, all the motion trajectories, both local and global (individual body parts and whole body), and importantly, the derivatives, i.e. velocity, acceleration, and jerk, features that seem to be important for our subjective perception of mode of motion and affect. It seems that we may be able to establish correlations between more classical motion categories as presented in the works of Rudolf Laban (Laban, 1980) and the concrete physical motion features derived from the motion capture data.

Moreover, although musicians and dancers may exhibit great motion skills, there are still a number of constraints on human motion, both in sound-producing motion and in sound-accompanying motion, constraints of speed, endurance, force, coupling of body parts, etc., and notably also constraints of motor control, such as reaction times and need for anticipatory programming (cf. previous comment on chunking in a phenomenological perspective), constraints that together contribute to what we perceive as typically human (as opposed to robotic) motion. These constraints are manifest as so-called phase-transitions, i.e. qualitative changes due to various parameter changes (more on this below), coarticulation, i.e. the fusion of small-scale motion into more large-scale motion chunks, and chunking due to anticipatory control, more specifically the so-called psychological refractory period that limits human reaction time (Klapp & Jagacinski, 2011). That said, it should be emphasized that these constraints should not be seen as some kind of impediment to musical and corporeal expression, but rather that what we experience as human and expressive motion is piggybacking on constraints, i.e. that we in a sense expect to observe such constraints when humans are involved.

Furthermore, in most kinds of music-related body motion, we see an interaction of continuous motion trajectories and more stationary postures. Postures may be seen as points of orientation for motion control (Rosenbaum, Cohen, Weiss, & van der Wel, 2007), something that we in our own studies of music-related motion have called key-postures, typically to be found at downbeats and other salient points in the music, and we see continuous motion between these key-postures (Godøy, 2013; 2014; in press).

We also believe that there are some basic music-related motion shapes, shapes based on body motion constraints, and interestingly, also associated with similarly distinct sound features:

- **Sustained**, meaning a continuous transfer of energy from the body to the instrument (or the vocal apparatus) such as in bowing or blowing
- **Impulsive**, with a discontinuous burst of effort such as in hitting or stroking
- **Iterative**, in the sense of rapid and repeated motion such as in a tremolo or in a trill

These three motion shapes are mutually exclusive i.e. there are categorical boundaries between them: if a sustained motion is shortened beyond a certain duration threshold, it will become an impulsive motion, and conversely, an impulsive motion lengthened beyond a certain duration threshold will become a sustained motion. Similarly, if an iterative motion is slowed down, it becomes a series of impulsive motions, and conversely, if a series of impulsive motions is speeded up, it may become an iterative motion. Such categorical transitions can be characterized as phase-transitions, often referred to in dynamical theory (Haken, Kelso, & Bunz, 1985), and we believe, also a constraint factor in music-related motion. Furthermore, these three shape categories may equally well be applied to sound. In fact, we have adapted these categories from
Schaeffer's theory of sonic objects, and they are prime examples of what we shall later in this chapter call sound-motion objects.

Combinations of such motion shapes may be found both in single performer and in ensemble motion, resulting in various textures. In music, composite textures are very common, both by a single performer (e.g. a pianist) and in the division of labour amongst the members of an ensemble, e.g. with a sustained background harmony in some instruments, a more pulsating, impulsive sound type accompaniment in others, and a free moving foreground melodic line in a soloist, often with iterative motion added to these textures as embellishments in other instruments. This means that there are often close relationships between motion and sound features, making it difficult to separate these two.

5. Musical sound features

From a motormimetic perspective, musical sound may be perceived as the result of some kind of sound-producing choreography. This goes for all features included in the earlier mentioned Western common practice and notation domain, i.e. typically, tone onsets, duration, dynamics, and timbre, as well as some expressive features. However, the more extensive differentiation and representation of sound features such as various nuances of pitch, dynamics, timbre and timing will need representations as shapes, as was one of the main ideas of Schaeffer and co-workers.

Their work on the perception and classification of sonic objects started with the overall dynamic and pitch-related shapes of the sonic objects in what was called **typology**, and continued downwards into the various details of the sonic object in what was called **morphology**. The typology contained three basic dynamical shapes, shapes that interestingly correspond to three equally distinct motion types presented in the previous section:

- **Sustained**: protracted, relatively stable sounds
- **Impulsive**: sounds with an abrupt (percussive or plucked) onset followed by decay
- **Iterative**: sounds exhibiting fast fluctuations as in tremolos and trills

and the pitch-related categories:

- **Pitched**, meaning relatively unambiguous and stable pitch
- **Varied**, with a changing pitch, such as in a glissando tone
- **Non-pitched**, meaning either heavily inharmonic or noise-dominated sounds

This produced a 3 x 3 matrix for the initial and rather coarse classification of sonic objects. In the next step, the morphology would lead to a much more extensive differentiation and qualification of internal features of the sonic objects, here just two will be mentioned:

- **Grain**, meaning the fast fluctuations within a sound, such as by tremolos and trills, as well as by inherent fluctuations such as the snaring sound of a deep double bass tone, and the differentiation of rate, amplitude, shape, etc. of these fluctuations.
- **Gait**, referring to slower fluctuations in the sound, typically in the timescale of what we would call a pulse or beat, often also organized into various metrical patterns such as of different dance forms, e.g. waltz, tango, samba, etc. This is now related to the large and fast growing domain of so-called **entrainment** in music, with documentation showing that music has a particular capability to induce motion in listeners in different cultures (Clayton, Dueck, & Leante, 2013).

In most Western and non-Western music, we will often have combinations of different features into more composite textures, often also with a division of roles amongst instruments in ensembles as mentioned in the previous section. What is all the more intriguing then is how listeners, independently of musical training, seem to
be able to extract salient motion features from complex sound textures, often referred to as beat tracking, which is an all the more impressive feat considering the difficulties of trying to make machines do such beat tracking. Thus one possible way to understand human beat tracking is that of motormimetic cognition, i.e. of projecting the overall shape of assumed sound-producing body motion onto the sound, hence, to correlate the subjectively perceived energy shapes of musical sound with motion shapes.

6. Amodal shape cognition
The widespread use of shape metaphors in music-related discourse could suggest that there is a special bond between music and the visual domain, and we have seen a number of publications that document such relationships, see e.g. (Eitan, 2013) for an updated overview. However, with our focus on motormimetic cognition in this chapter, we need a kind 'translator' between the modalities of sound and motion, and it seems the phenomenon of amodal shape cognition may serve this purpose here.

We may encounter a number of different terms concerned with the relationships between the modalities such as cross-modal, multimodal, inter-modal, and lastly, amodal, and a useful discussion is to be found in (Macpherson, 2011). We take the expression 'amodal shape cognition' here to denote the mental scheme of geometric forms independent of any particular sense modality, and hence as something that may be transposed from one modality to another, e.g. as was the case of the typology shapes for motion and sound presented earlier.

Amodal shape cognition seems to be a universal faculty of the human mind, as has been suggested by so-called morphodynamic theory (Thom, 1983; Petitot, 1985). This theory sees shape as the fundamental mode of perceiving and understanding the world in general. This fits well with what we now know of features of musical sound as non-abstract and shape-related, both subjectively, as suggested by Schaeffer, and more signal-based, as now recognized in acoustics (shapes of envelopes, spectrum, various transients, etc.), and with what we know of motion trajectories and posture shapes, as well as the shapes of significant derivatives (velocity, acceleration, and jerk) based on various motion capture data.

Amodal shape cognition is thus the unifying and translation-enabling element for mapping motion to sound and sound to motion. Most notably, this may be done both with high-end technologies, i.e. motion capture, various sensors, and sophisticated processing software for both motion and sound, or it may be done with the more modest means of eyes, ears, paper, and pencil, where the crucial element is to ceaselessly shift our attention between the sound shapes and motion shapes. In epistemological terms, this means actually practicing a heuristic of exploring features of sound and motion, even a kind of so-called hermeneutic circle of 'understanding' the sound by motion and 'understanding' the motion by the sound.

However, when mapping shapes from one domain to another, care must be taken to conserve the overall perceptual features, i.e. to avoid making abstractions from the data that effectively destroys the shape-nature of whatever it is that we are mapping. The topic of mapping has been central to developments of new musical instruments, given the fact that any input data can in principle be used to control any output features of a sound synthesis model, however not all mappings are suitable in terms of perceived shape features. We have seen different schemes for such mappings (Hunt, Wanderley, & Paradis, 2003), and this problematic extends to the entire domain of new musical instruments as well as to sonifications in general (Jensenius & Godøy,
the core challenge here is somehow to enable multidimensional control of the multidimensional nature of musical sound and musical expression.

7. Sound-motion objects

From a motormimetic perspective, sound and motion may be seen as two sides of the one and same phenomenon. Again, consider the mentioned example of a rhythmical pattern performed on a drum set: it is both sound and motion, and we are hard put to separate these two unless we focus on one aspect at a time.

One crucial feature here is that of focusing on time-limited fragments of sound-motion on the meso timescale, on what we like to call sound-motion objects. Besides emphasizing the dual, i.e. multimodal, nature of what we are dealing with here, it also emphasizes that we focus on a delimited chunk as a coherent entity.

From a production point of view, meso timescale sound-motion objects seem to be optimal. Research on human motor control points to the goal-directed and hierarchical nature of motion control (Grafton & Hamilton, 2007), and more specifically, to posture-centered motion as optimal for planning and control (Rosenbaum et al. 2007). In line with this, discontinuous, i.e. intermittent, motion control seems more efficient than continuous control (Loram, van De Kamp, Lakie, Gollee, & Gawthrop, 2014), and chunking motion into action gestalts seems optimal for demanding tasks (Klapp & Jagacinski, 2011).

From a perceptual point of view, there are a number of publications that converge in suggesting the meso timescale as optimal for sensing salient, meaningful features in various domains (Godøy, 2013). In sound perception too, it is well known that the entire sound needs to be somehow kept in memory because the different stages of the unfolding all contribute to the perceptual result, and we assume the same to be the case for motion, i.e. that motion chunks need to have a minimum duration in order to make sense. Recent work on auditory objects points to the interaction of low-level features with more high-level schema-based factors in our ability to discern sonic objects in listening (Bizley & Cohen, 2013), as well as to the multimodality involved in auditory object perception (Griffiths & Warren, 2004). We thus seem to have several findings in recent research converging on suggesting that sound-motion objects play a crucial role in perception.

However, we should keep in mind that research on sound-motion relationships is not new, see e.g. (Schneider, 2001) for an interesting overview of past research, and the work of one of the pioneers in the field, Alexander Truslit, has recently become available with supporting audio-visual material (Truslit, 2015). But needless to say, in the last couple of decades we have had important advances in methods and technologies for studying sound-motion objects. Most importantly, we now have technologies for high resolution motion recording, first of all motion capture with infrared camera systems, but also with other sensors such as inertial and pressure sensors, and various sensors built into musical instruments, and not to forget, video recordings. We also have sophisticated software tools for extracting perceptually salient information from both motion and sound data, as well as for making various correlations between sound and motion data, e.g. various statistical and machine learning tools. And it is now possible to generate high quality sound and motion material to be used in so-called analysis-by-synthesis and match-mismatch experiments, as well as in various artistic applications. The last point is particularly significant as a contribution to the field of artistic research in enabling the direct generation of high-level aesthetic and affective sound-motion material, effectively combining qualitative and quantitative approaches to the study of musical experience.
8. Conclusion
Over the past couple of decades, we have seen a number of publications arguing in favour of music as a multimodal form of art (Eitan, 2013). This may be obvious in everyday circumstances with the common experience of seeing musicians perform, seeing their body motion, and seeing people dance or make all kinds of body motion to music in various contexts. Yet the more systematic study of the links between motion features and sound features seems to be only in its beginning. Given available methods and technologies, this should indeed be a promising field of research in the coming years.

We see the aims here as two-fold: the aim of enhancing our basic understanding of how music works and affects us, and the aim of developing new means of artistic expression. We believe awareness of motormimetic elements in music could in particular be beneficial for more structured artistic practice research. This is based on the belief that the motormimetic rendering of musical sound is a method for becoming more familiar with the sound, actually also a kind of hermeneutic method of tracing musical sound by body motion in order to delve deeper into musical experience.

References


