

Motor Constraints Shaping Musical Experience

Rolf Inge Godøy

KEYWORDS: Sound-producing body motion, constraints, motor control, biomechanics, coarticulation

ABSTRACT: In recent decades, we have seen a surge in published work on embodied music cognition, and it is now broadly accepted that musical experience is intimately linked with experiences of body motion. It is also clear that music performance is not something abstract and without restrictions, but something traditionally (i.e., before the advent of electronic music) constrained by our possibilities for body motion. The focus of this paper is on these various constraints of sound-producing body motion that shape the emergent perceptual features of musical sound, as well as on how these constraints may enhance our understanding of agency in music perception.

Received April 2018

[PDF text](#) | [PDF examples](#)

Volume 24, Number 3, September 2018
Copyright © 2018 Society for Music Theory

1. Introduction

[1.1] With the exception of mechanical instrument sounds, before the advent of electronic music musical sound was always produced by some kind of body motion, such as by blowing, stroking, bowing, shaking, hitting, kicking, squeezing, etc. on some kind of physical instrument or by motion of the human vocal apparatus. Therefore, there has traditionally been an imprint of human effort and of human artifacts, in short, of human agency, on music as we know it. It may be argued that this imprint has persisted even after the advent of various means for electronic sound production, by a projection of traditional mental schemas of body motion onto musical sound that draws its energy from an electric power supply.

[1.2] All musical sound is somehow related to body motion trajectories: think of the hand moving from an initial position to an impact with the drum membrane and back again, or the inhaling before the onset of a long trumpet tone and the subsequent blowing through the mouthpiece. And musical sound is also often related to posture shapes: for instance, the vocal tract shaped for some vowel in singing, or the hand on the keyboard shaped for some chord in piano performance. There usually is continuous body motion between distinct onsets of sonic events, for the simple reason that all body motion takes time.

[1.3] As suggested by research on music and body motion ([Godøy and Leman 2010](#)), we seem to ascribe images of body motion to music that we hear (e.g., images of an energetic hand or mallet motion when hearing a ferocious drum passage, or images of a slow protracted bow motion when hearing soft and slow string music), because of previous experience of sound-producing body motion. This means that musical experience not only consists of a series of sonic events, but also of an extended choreography of body motion interacting with instruments, interactions that carry salient sensations of human agency in music.

[1.4] Body motion is continuous in music; there is no instantaneous displacement of so-called *effectors* (e.g., fingers, hands, arms, lips, tongues). As such, it is a fundamental constraint on music making and affects performers' interaction with musical instruments, including the human voice. From such interactions, we see the emergence of so-called *idioms* in music making, 'idioms' here defined as sets of sounds and/or sound passages that are easy to produce and well-sounding.

There are a number of possible sound-producing actions in music performance, as well as a number of difficult or conceivably impossible ones. The distinction derives from what can be denoted *motor constraints* in music. This paper will explore these constraints and reflect on how they are embedded in our sensations of agency in musical experience.

[1.5] Claiming that there are a number of motor constraints embedded in music should in no way limit or diminish the value of music. On the contrary, detecting these constraints should enhance our understanding of music as an art that is intimately linked with human behavior, by making us recognize that human agency in music is clearly manifested in the motor constraints of music making.

[1.6] To substantiate this claim of human agency by motor constraints in music, I shall in the following sections give a step-by-step presentation of the various elements involved here, starting with what may be called a *motor theory perspective* on musical experience. The main point of this perspective is to understand all musical experience as somehow related to sensations of sound-producing body motion. But this perspective will also necessitate discussions of *timescales*, as there are distinct qualitative features of both sound and motion at the different timescales involved in musical experience. Furthermore, there are a number of *instrument constraints* and *body motion constraints* that together limit possibilities of musical expression, and we need to assess how these constraints help shape *musical expression*—in particular, shape the possibilities for what could be called *musical translations*, the transfer of musical ideas from one set of motor constraints to another, as is done in orchestrating a piano score. Towards the end of the paper, this will lead to some *ontological reflections*, reflections on ‘what-is-what’ in musical experience, such as the relationships between notation, sound, and body motion in our concepts of music. The main contention will then be that the perception and/or imagery of human body motion is a salient reflection of human agency in music.

2. Motor theory perspectives

[2.1] Obviously, there are many close links between sound and body motion in music, and this has been the focus of a number of publications for the past couple of decades (see [Godøy and Leman 2010](#) for an overview). In parallel, an understanding of body motion has emerged as an essential element of human perception and cognition in general ([Gallese and Metzinger 2003](#)), such as in social cognition ([Wilson and Knöblich 2005](#)), and even in more abstract thought ([Gallese and Lakoff 2005](#)). In our present context, we should note that the crucial role of body motion in human perception and cognition has permeated some of the leading research on human agency, with the contention that we perceive the world, including the behavior and utterances of fellow humans, by way of mental simulation of the body motion related to whatever it is that we are perceiving. Also, it has been argued that agency, in the sense of mental simulation of motion, is indeed embedded in our basic neurocognitive faculties (see [Gallese 2000](#) for details).

[2.2] Unmistakably, in the past couple of decades we have seen an ‘embodied turn’ in several human sciences, and it is now not controversial to suggest that sensations of body motion are involved in most, if not perhaps all, domains of human perception and cognition. This may be seen as valid for different modalities, such as in vision, with the principle of active tracing of whatever it is that we are seeing ([Berthoz 1997](#)), and in audition with the so-called *motor theory* of auditory perception ([Liberman and Mattingly 1985](#)). Originating in linguistics, the motor theory (or rather *motor theories*, as there are variants here) has since been extended to other domains ([Galantucci, Fowler, and Turvey 2006](#)). It was initially met with a great deal of skepticism; however brain observation studies in recent decades have lent much support to this theory (e.g., [Haueisen and Knösche 2001](#), [Kohler et al. 2002](#), [Bangert and Altenmüller 2003](#)).

[2.3] Regarding sound perception, the gist of motor theory can be expressed as the tendency to covertly (but sometimes also overtly) simulate the body motion that we believe is at the source of whatever it is that we are hearing. In the case of language, this means mentally simulating the vocal apparatus motion of some language—its so-called *phonological gestures*. In the case of music, it means mentally simulating instrumental or vocal sound-producing motion. One crucial, and sometimes not so well understood, feature of motor theory, is that of approximate renderings, or approximate imitation. For instance, there are many languages that I can neither speak nor understand (e.g., Arabic and Chinese), yet I believe that I am able to distinguish between these two by way of rather vague and inexact mental renderings of the two languages’ different sets of phonological gestures. Likewise, phenomena such as scat singing and beatboxing in music attest to the possibility of approximate imitations of non-vocal sounds by the human vocal apparatus, imitations sufficiently similar to the original sounds so as to enable recognition.

[2.4] As for the sound-producing body motions that may be perceived and remembered, and that

subsequently form the basis for motor imagery in music perception, they may be differentiated as follows (see [Godøy 2010](#) for details):

- *Excitatory* motion, denoting an energy transfer from the body to the instrument or the vocal apparatus, including both kinematic (i.e. visible motion trajectory) images, and, by empathy or deduction from the kinematics, also effort sensations (e.g., *heavy, light, tense, relaxed*).
- *Modulatory* motion that modifies the sound (e.g., the left hand changing pitch on a string instrument, or the opening and closing of a mute on a brass instrument)
- *Ancillary* motion, a collective term for motion used by musicians to help in the expressive shaping of the music, and also to avoid fatigue and strain injury, hence motion not strictly excitatory nor modulatory, yet it seems, indispensable for most musicians.
- *Communicative* motion, which includes cues between the performers in an ensemble, as well as theatrical motions that can make an impression on the audience.

[2.5] Given most peoples' exposure to such body motion, it is reasonable to propose a *motormimetic* element in music perception, meaning that listeners (variably so) mentally imitate sound-producing motions when listening to music ([Godøy 2001, 2003, 2004](#)). In order to find out more about this, we carried out some studies on *air instrument performance* ([Godøy, Haga, and Jensenius 2006](#)), exploring the imitative capabilities of different listeners, ranging from experts to people with little or no musical training. The results suggest that people have a high degree of knowledge of sound-producing body motion in music, despite the variations in accuracy and detail related to levels of musical training.

[2.6] There is also music-related motion that does not reflect the details of sound-production—what we collectively call *sound-accompanying* body motion, such as in dancing, walking, or gesticulating to musical sound. Music induces images of motion that reflect salient features of musical sound, such as pitch contours, dynamics, various rhythmical patterns, and motion and effort, all of which may vary in motion trajectory (i.e., people may move differently to the same musical sound). Such variations in sound-accompanying body motion can be ascribed to listeners focusing on different aspects of the musical sound texture (foreground melody, accompaniment patterns, percussion patterns, etc.) and can be understood as interpretations of the music's multiple *gestural affordances* ([Godøy 2010](#)).

[2.7] Spontaneous body motion to music can be observed in a multitude of everyday situations, most clearly in what is called *entrainment*, meaning that listeners are able to synchronize with what they detect as the pulse of the music (see [Clayton, Dueck, and Leante 2013](#) for extensive discussions of this). In some music, the pulse may be very salient, and in other cases not ([Lartillot et al. 2008](#)), suggesting that listeners have an ability to extract pulse information from rather complex sound textures—an ability related to what could be called *auditory gist perception* ([Harding, Cooke, and König 2007](#)), an under-researched topic. In summary, music may evoke a wide range of motion sensations in listening, ranging from details of sound-producing motion to more global sensations of energy and pulse. In most cases, these sensations are quite closely linked to one or more salient features of the sound.

3. Timescales of music-related sound and body motion

[3.1] To better understand how sensations of body motion are evoked by musical sound, we will briefly look at the different timescales possible here. To begin, we have quasi-stationary features in the range of audible vibrations extending from approximately 20 to 20,000Hz, including perceptually salient features such as pitch and the stationary components of timbre and loudness. Next, below the 20Hz region, we find some other very salient features of music: in the fastest range of this region are various transients and fluctuations in pitch, timbre, and dynamics of the sound, including tremolos and trills; in a slower range of this region are the dynamic, timbral, and pitch-related envelopes of individual tones or sound events; and in a still slower range of this region are constellations of tones or sounds in succession to form various rhythmical, textural, melodic, and timbral patterns. In our work on music-related body motion, we have made a tentative distinction between three main timescales that we believe also are relevant for our considerations here of motor constraints in musical experience:

- *Micro timescale*, which encompasses continuous motion (such as in continuous bowing on a string instrument) or stationary postures (such as in sustained tones on a wind instruments or by the human voice), resulting in continuous sound. This timescale also includes very rapid repeated motion such as vibrato, trills, or tremolos, resulting in micro-textural features that appear to be continuous or quasi-stationary. Listeners are very quick to recognize such features of timbre and micro-textures (often referred to as "sound" in popular music research), sometimes in fragments as short as 250 milliseconds ([Gjerdingen and Perrott 2008](#)).
- *Meso timescale*, approximately in the 0.5 to 5 seconds duration range, typically encompass singular sonic

events or chunks of fused sonic events, such as rhythmical, textural, and melodic patterns. Meso timescale chunks are holistically perceived and conceived as units, and include stylistic, expressive, and affective elements readily perceived by listeners (Godøy 2013).

- *Macro timescale*, which includes several meso timescale chunks in succession, and may typically encompass sections and even whole works of music. At this timescale, we can speak of more global sensations of motion, such as *calm*, *agitated*, *jerky*, or *smooth*, and their corresponding kinematic and effort-related features, features that are detectable with motion capture and video processing methods.

[3.2] In other words, these different timescales encompass not only different sonic features, but also rather different motion features. We can see a further parallelism between sound and body-motion features in the following scheme for classifying sonic objects, which is adopted from Pierre Schaeffer's so-called *typology* (Schaeffer 1966, 429–59):

- *Sustained*, meaning continuous sound and a continuous effort and transfer of energy from the body to the instrument, such as bowing, blowing, and singing.
- *Impulsive*, denoting a short burst of sound resulting from a discontinuous peak of effort, such as hitting or plucking.
- *Iterative*, meaning the rapid repetition of a sound or group of sounds, such as in a tremolo or a trill, which typically corresponds to a rapid back-and-forth motion of the hands or fingers (see an example of this in section 7 below).

[3.3] There are qualitative thresholds between these sound-motion categories, depending on duration and rate of events, by the phenomenon of so-called *phase-transitions* in dynamical theory (Haken, Kelso, and Bunz 1985). For instance, if an impulsive sound gradually accelerates, or increases in rate, it will sooner or later turn into an iterative sound (e.g., a tremolo); conversely, if an iterative sound is gradually slowed down, it will sooner or later turn into an impulsive sound.

[3.4] As for the quasi-stationary micro timescale of continuous musical sound and motion, it is useful to adopt Schaeffer's term *morphology*, as this will cover several salient pitch-related and timbre-related features. The central concept here is *mass*, denoting the spectral spread of sound, which in our perspective is related to body postures in sound-producing body motion (vocal tract shape or hand/fingers shape in relation to instrument). 'Mass' is then a general concept for quasi-stationary spectral distribution with the main classes *tonic* (clearly pitched), *complex* (inharmonic or noise type), and *variable* (changing sensation of pitch). Additionally, there are also two main classes of fluctuations in the quasi-stationary mass, related to body motion in performance:

- *Grain*, denoting the fast fluctuations within the sound—sometimes the 'natural' outcome of the sound-producing motion (e.g., the 'brrrrr' sound of a deep double bass sound or the stroking of a maraca), and sometimes the performer's back-and-forth or rotational body motion in tremolos or trills.
- *Gait*, denoting fluctuations in the sound at the slower pace of body motion, as in walking, dancing, or gesticulating.

[3.5] In musical practice, we often see simultaneous layers at different timescales, i.e., composite textures, typically with a sustained background, or accompaniment, and a slow foreground, or melody, often with a faster, grain-type texture added to the sustained basic harmonic scheme. Different sound and body-motion features at different timescales may be simultaneously at work in music, and these different timescales converge in meso timescale chunks that reflect human agency in music.

4. Instrument constraints

[4.1] As musical sound is produced by body motion, we can identify which features come from which part of this interaction, starting with the response of musical instruments to body motion.

[4.2] Traditional musical instruments involve an energy transfer from the human body to the instrument, resulting in sound output from the instrument. Acoustics explains this transfer of energy and further propagation within the instrument, as well as the subsequent radiation of sound out from the instrument. This process results in patterns of overall energy dissipation over time (i.e., in envelopes of dynamics, pitch, and timbre, as well as in initial, or attack, transients and various subsequent fluctuations in the course of the sound). These are the elements that contribute to our subjective perception of the instrument's sound.

[4.3] As for excitation, the various classes of sound-producing body motion (blowing, bowing, scraping, shaking, hitting, kicking, etc.), as well as the typological categories mentioned earlier (*sustained*, *impulsive*, *iterative*), may variably apply here, but there are a number of constraints in the responses of any instrument: an impulsive excited instrument (drum, vibraphone, marimba, piano) will typically have a decaying envelope of variable length, whereas a sustained excited instrument (bowed, blown) may have a continuous, steady envelope with a duration dependent on the bow

speed or amount of breath. The physical makeup of an instrument constrains its possible sound output and some of the micro-textural features of the sound, such as its grain (e.g., the mentioned ‘brrrrr’-type grain texture of a deep double bass tone that is produced by a continuous, smooth bowing motion).

[4.4] Importantly, the design of the instrument also strongly constrains its ergonomics, determining the limits of the possible and the impossible sound-producing body motion, although expert instrumentalists and singers may dedicate the major part of their lives to extending the limits of what is possible on various instruments or with the human voice. The abovementioned idioms are situated at the intersection of the physical makeup of an instrument and performance motion. Historically, there has been an interactive process of instrument, ergonomics, and sound design, leading to the emergence of repertoires of highly successful idioms and to the exploitation of these idioms for maximal sound effect, at minimal effort cost or difficulty for the performer. The orchestrations of Nikolai Rimsky-Korsakov in his mature period are prime examples of this optimal adaptation of instrument constraints to body motion constraints.

[4.5] As a simple yet effective example of this, consider the excerpt of the cello part from the second movement of Rimsky-Korsakov's *Capriccio Espagnol* in **Example 1**. This movement starts out in the key of F major and modulates to C major, making an ordinary modulation to the dominant. This allows for the cello figure, shown in Example 1, an energetic back-and-forth bowing motion with a *forte* sonic output, easy to

Example 1. Rimsky-Korsakov, *Capriccio Espagnol*, II, m. 70, cello fragment



(click to enlarge)

perform because of the two lower open strings. This passage would be rather uncomfortable to play, and not sound as good in keys that would not employ open strings. Actually, the whole movement, and the entire *Capriccio Espagnol* as such, is clearly conceived with a view towards exploiting instrumental idioms in combination with optimal acoustic dispositions for a robust and well-sounding output. The orchestration reaps the combined benefits of the best possible individual roles for the musicians with the best possible (i.e., subjectively well-sounding) timbral results.

[4.6] In general, the basic mental schema of a division between instrument and performer seems to have become so ingrained that it has in some cases also been projected onto new technologies for making musical sound:

- In the early days of digital music, software developers introduced a structure of “instrument” and “score” modules, hence, a control structure that resembled the traditional division of instrument and performance, to facilitate music generation.
- Also, with the possibility of producing any sound in theory, whether previously heard or unheard, the difficulty comes in precisely controlling the desired sound features in sound synthesis. One solution is to simulate the sound generation of real physical instruments by various, and often simplified, models, where it should be possible to predict features of the output in response to the input of the model. For example, increased force in the excitation of a so-called *physical model* piano string should also produce changes in the timbre of the output sound, and not only an increase in the amplitude of the sound.
- The desire to control expressivity has spurred an extensive effort in re-introducing the human touch in digital music by developing so-called *new interfaces for musical expression* (NIME).

[4.7] Additionally, amplification technology may dramatically increase the physical energy level of musical sound (e.g., the PA-system of a rock concert at a sports stadium), yet such technology will still transmit what is basically small-scale sound-producing body motion (e.g., fingers plucking guitar strings)—in effect, conserving some features of the original motor image across large-scale amplification of the sound.

5. Body motion constraints

[5.1] In spite of the great flexibility of humans in developing body-motion skills, the possibilities of human body motion are also limited: there are types of body motion that are very difficult and possible only with substantial practice, and there are types that are easy, requiring little or no training. A brief overview of some relevant body motion constraints is discussed here.

[5.2] The overall constraint that all human motion takes time means that there are speed limitations on music-related body motion. There are also limits on maximal force, and there is the need for rests and change of posture to avoid fatigue or strain injury. Added to such basic biomechanical constraints are also those of motor control, with constraints of reaction speed, manifest in the so-

called *psychological refractory period* (Klapp and Jagacinski 2011). This in turn entails the need for anticipation, i.e. of moving an effector in place before the sound onset motion can take place, the need for grouping motion into chunks and into action hierarchies, and the need for automatization in order to be fast enough. In short, there are several constraints on sound-producing motion that are well known in musical performance, and in our context, we also have the following constraints that strongly contribute to shaping musical performance:

- The most obvious is the earlier mentioned *phase-transition*, which is dependent on speed and density of events and results in groupings of body motion.
- Related to phase-transition is the phenomenon of *coarticulation*, meaning the fusion of otherwise separate motion units and sounds into coherent chunks that are perceived (and conceived) holistically. Coarticulation is studied extensively in linguistics and, to a certain extent in robotics and other movement-related sciences, but less so in music. The main elements of coarticulation in music is that the motion of the effectors are embedded in contexts of past and future motion, effectively resulting in a contextual smearing of both the body motion and the resultant sound.

[5.3] The main features of coarticulation in music may be summarized as follows (see Godøy 2014 for details):

- *Temporal coarticulation*: an effector's present position or posture (e.g., hand shape and position on the keyboard) is determined by the immediate past and immediate future actions (fingers on the last keys and keys to be pressed). For this reason, we may speak of *carry over effects* (i.e., the context of recently past events) and *anticipatory effects* (i.e., the context of coming events), in coarticulation.
- *Spatial coarticulation*: an effector may, depending on required force and on context, variably recruit neighboring effectors (e.g., the finger may recruit hand, arm, elbow, shoulder and even torso motion in hitting a key).

[5.4] In short, coarticulation means contextual smearing of both the sound-producing body motion and of the resultant sound, leading to the formation of qualitatively new units of musical sound and motion at the meso timescale. Another way of putting this is that groupings, or the emergence of chunks in music, is conditioned by body-motion constraints, hence bearing the imprint of human agency in music.

6. Constraint-based musical expression

[6.1] If we recognize the existence of the combined instrumental or vocal human body-motion constraints mentioned above, it would make sense to suggest that musical expression is piggy-backing on these constraints, and that listeners, through extensive experience of musical performances, have come to not only accept these constraints, but also expect them when listening to music. It seems that listeners are sensitive to features of 'mechanical' renderings of musical sound in cases of machine-generated performances, such as the playback of music by sequencer-controlled MIDI instruments, and distinguish these performances from more 'natural' ones by living musicians.

[6.2] We have seen several research projects aimed at finding principles of expressivity in musicians' performances (for an overview, see Goebel et al. 2008) and at developing commercial software that re-introduces 'human touch,' by implementing some imperfections in the otherwise strictly quantized note-on data of MIDI files. One potential source of such 'human touch' here could again be that of coarticulation, based on the fact that musicians will usually have some amount of contextual smearing in their performances, and that listeners probably also expect coarticulation-based contextual smearing in music (Godøy 2014).

[6.3] What is at stake here is the relationship between real, human generated sound-producing body motion and its resultant sound or, put differently, musical sound that we perceive as bearing the imprint of human agency, versus some more abstract notion of music often propagated by our Western notational system and its associated concepts. If we accept that coarticulation is an integral part of human motion in musical performance, with consequences for the shaping of the resultant musical sound, the next step will be to recognize that performance is not just a matter of 'interpretation,' but more precisely that *performance is a transformation of the score to a series of articulated human motion chunks and sonic objects*. This recognition should also spur us to make some critical reflections on the status of musical works as such, starting with the question: Can a musical work survive a transfer from one instrumental or vocal setting to another, and somehow retain its identity across a re-orchestration or, what could be called, a 'musical translation'?

7. Musical translation

[7.1] Given the large number of musical translations in various orchestrations and the huge number of arrangements found in Western music, the paucity of research that focuses on issues of musical

translation is quite remarkable. Although we do have a number of manuals on how to orchestrate or how to make arrangements for various ensembles, we still lack more systematic and in-depth studies of the ergonomics and associated body-motion constraints involved. However, Nikolai Rimsky-Korsakov does address musical idioms in his orchestration treatise ([Rimsky-Korsakov 1964](#)), and above all, systematically exploits them in works such as the abovementioned *Capriccio Espagnol*. It should also be added that some orchestration books ([Jacob 1940](#); [Piston 1991](#)) do indeed treat the topic of, what we might call, *idiom translation* from a pragmatic perspective, and we shall investigate a small example of the challenges and possible solutions that such a translation presents.

[7.2] It is well known that in language, local idioms are not easy to translate. A literal translation of an idiom may often lead to rather absurd results, whereas the overall content and intention of a text may be more easily translated. In the case of music, we have a number of constraint-based instrument-specific idioms that, if just transcribed note-by-note from one setting to another, might have some very unfortunate consequences. In such instances, a non-literal and effect-oriented translation might be more successful. To produce a successful orchestration, we may have to deviate from the note-by-note score of the original and, within reasonable limits, transform the musical material in order to make the new orchestral version loyal to, what we believe are, the overall aesthetic intentions or ‘spirit’ of the original.

[7.3] As a small thought exercise here, consider translating the excerpt in **Example 2** from the last movement of Beethoven’s Piano Sonata in A-flat major, op. 26, for a string ensemble. This excerpt is a typical piano idiomatic passage, not particularly difficult to play on the piano at a fast tempo, if using a tilting motion of the wrist, making what is called an *iterative* sound-producing motion.

Example 2. Beethoven, Piano Sonata in A \flat Major, op. 26, “Allegro,” IV, mm. 3.2–5.1



(click to enlarge)

[7.4] When translating this passage for a string orchestra, it would not be a good idea to make a ‘literal’ note-by-note transcription, because that would be unduly difficult (however, transposing this down a semitone might make it considerably easier to play by using the open D-string on the violins). To translate this for string orchestra, we could try more transformative versions, such as the version shown in **Example 3** or **Example 4**.

Example 3. String orchestra version of the Beethoven excerpt in Example 2



(click to enlarge)

Example 4. Alternate string orchestra version of the Beethoven excerpt in Example 2



(click to enlarge)

[7.5] In rewriting a fast *back-and-forth* tilting type figure to a *down-and-up repeated note* figure, the latter is much easier to play on bowed string instruments by an iterative tremolo-like down-up bow motion; however this string version of the Beethoven fragment now would be uncomfortable or difficult to play on the piano. What we need to evaluate, then, is whether this string ensemble translation is too drastic, compared to the note features of the original, or whether this translation is good because it is adapted to instrument-performance constraints.

[7.6] There are a number of musical translations that have made a substantial impact on our Western music culture (see, for instance, Stokowski’s orchestral transcriptions of various works by J. S. Bach). The main question in our context is to what extent the sense of agency survives such translations. In other words: What is the motor image that we have of one version of a musical excerpt versus another version? Something similar could be asked for cases going in the opposite

direction: translating a texturally highly complex work to a simpler version, such as an orchestral work to a piano reduction. To what extent is the overall sense of body motion of the original preserved in such a piano reduction? Think of the many cases where a crescendo in the sustained tone instruments (woodwinds, brass, and bowed strings) of the orchestral original has been translated into a tremolo figure in a piano reduction, making the crescendo possible, at the cost of introducing a new textural element not found in the original (the tremolos with rapid repeated onsets).

[7.7] Musical translations present us with intriguing questions that could be studied more systematically in music perception, and that could also be linked with the general phenomenon of so-called *motor equivalence* (Rosenbaum 2009), meaning our ability to use alternative effectors in performance of musical ideas. For example, in simple cases, this would consist of the ability to play a tune with the left hand rather than the right hand, and in more complex cases, to reproduce musical ideas across different instrument-specific idioms, as illustrated in the Beethoven examples above (Examples 3 and 4). Another way of understanding the phenomenon of motor equivalence is that humans seem to have a more general, adaptable, and goal-directed capacity for motor planning and control, so that when needed, the same goals may be achieved with alternative effectors and motion trajectories, suggesting that some aspects of human motion are robust across different detail instantiations.

8. Ontological reflections

[8.1] Considering the constraints involved in music performance and musical translation, we can reflect on *the ontological status of musical works* in Western musical thought. The term *ontological* is used here to denote the existence of musical works in, and across, different guises, such as scores, performances, recordings, and memory traces, as well as the existence of the multitude of features typically manifest in musical works, such as tunings, intervals, chords, modes, timbres, textures, rhythmical patterns, articulations, body motion, expressivity, sense of energy and affect. Needless to say, these are all very extensive topics, and the aim here is limited to briefly assessing the role of body motion in the ontology of musical works.

[8.2] Clearly, performance is usually regarded as an essential element in 'bringing to life' otherwise inert symbols of the score in Western music, and it is also generally recognized that performance traditions dictate a number of "dos and don'ts," which supplement the scores. Yet it is not clear what a musical work is: Is it an ideal entity with n-number of more or less acceptable performance renderings, or is it an ideal entity independent of any particular performed version? And what about its features: Is a musical work dependent on having all of its features present (e.g., all instruments of the original score performing), or can it survive and retain its identity with a reduction in the number of features (e.g., a reduced ensemble), or changes in detail features (e.g., in various translations as briefly discussed above), or even in various low-resolution or distorted versions (e.g., in listening to a historical wax roll recording)?

[8.3] From a motor theory perspective, an answer could be that musical works are strongly associated with motion sensations, tentatively to the point of being scripts of music-related body motion, motion that then would be re-created in our minds when listening or merely imagining by so-called *musical imagery* (Godøy 2001). This capacity for mental simulation of music-related motion could be considered a very basic feature of human musical perception and cognition, in line with the paramount importance attached to motor cognition in the various motor theory research briefly mentioned earlier (e.g. Gallese and Metzinger 2003), constituting what we could see as a *motor ontology* at the very base of music as we know it. In other words, this would suggest that we experience, remember, and imagine music by way of not only sound, but by motion scripts that may run parallel with images of sound.

[8.4] Western music notation has been very successful in conserving vast amounts of music from past centuries, as well as in enabling coordinated performances of complex musical works in large ensembles. However, what has been missing is a versatile scheme for representing salient features of unfolding musical sound and music-related body motion. Fortunately, we see the emergence of methods for signal-based sound analysis that enable us to represent more perceptually salient features of musical sound, such as timbre, texture, dynamics, and also various expressive nuances. These methods help close the gaps between the symbolic representations of Western music notation and the continuous, sub-symbolic features of musical sound (see Castellengo 2015 for an excellent overview of signal-based representations). Likewise, we now have methods and technologies for capturing and studying music-related body motion, and we are also able to explore in detail the correlations between features of sound and body motion in musical experience (Godøy and Leman 2010).

[8.5] Such studies of correlations between sound and body motion in music serve to demonstrate that music is multimodal, involving the main modalities of sound and motion, with motion in turn generally considered as a composite that comprises not only kinematics and effort, but also touch, proprioception, balance, and possibly other components as well. Needless to say, there remain substantial challenges for the study of musical experiences based on continuous sound and continuous body motion, yet it is a worthwhile and necessary undertaking, in view of the crucial ontological status of continuous sound and body motion in music from our motor theory perspective.

[8.6] As a point of method, a common denominator for all of the multidimensional features of sound and body motion is that they can be represented as *shapes*, meaning fundamentally *non-abstract* and *concrete* features, to use the terminology of Schaeffer (1966). Thinking of musical features as shapes should help us to think of music as continuous unfolding motion, rather than as discrete symbols, and encourage us to think of musical performance as a fusion of individual motion and sound events that form coarticulated motion-sound chunks, thus enhancing our capacity for recognizing and exploring human agency in musical experience.

9. Summary

[9.1] One main result of research on music-related body motion during the last couple of decades is the recognition of human body motion as a fundamental element in music perception and cognition, even to the extent that sound and body motion may be seen as inseparable from our experiences of music. A consequence of this view will be the recognition of the imprint of human agency everywhere in music. And with the use of available methods and technologies, we can in historically unprecedented ways zoom in on and represent through graphs, animations, and pictures what we believe are typical reflections of human body motion in music.

[9.2] This is relevant for the large domain that is often labelled *musical expressivity*. However, this domain should not only include the affective features of musical performance, but also the typical human body-motion features of musical performances (as opposed to machine-based performances), on the meso timescale of tones, rhythms, grooves, and phrases, as well as on the micro timescale of within-tone and between-tone expressive nuances. This means that we need to recognize human motor constraints on sound-producing motion as integral to musical sound.

[Return to beginning](#)

Rolf Inge Godøy
 Department of Musicology
 University of Oslo
 P.B. 1017 Blindern, N-0315 Oslo, Norway
r.i.godoy@imv.uio.no

[Return to beginning](#)

Works Cited

- Bangert, Marc and Eckart O. Altenmüller. 2003. "Mapping Perception to Action in Piano Practice: A Longitudinal DC-EEG study." *BMC Neuroscience* 4:26. doi.org/10.1186/1471-2202-4-26
- Berthoz, Alain. 1997. *Le sens du mouvement*. Odile Jacob.
- Clayton, Martin, Byron Dueck, and Laura Leante, eds. 2013. *Experience and Meaning in Music Performance*. Oxford University Press.
- Castellengo, Michèle, 2015. *Écoute musicale et acoustique*. Éditions Eyrolles.
- Galantucci, Bruno, Carol A. Fowler, and Michael T. Turvey. 2006. "The Motor Theory of Speech Perception Reviewed." *Psychonomic Bulletin & Review* 13 (3): 361–77.
- Gallese, Vittorio. 2000. "The Inner Sense of Action. Agency and Motor Representations." *Journal of Consciousness Studies* 7 (10): 23–40.
- Gallese, Vittorio, and George Lakoff. 2005. "The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge." *Cognitive Neuropsychology* 22 (3/4): 455–79.
- Gallese, Vittorio and Thomas Metzinger. 2003. "Motor ontology: The Representational Reality of Goals, Actions and Selves." *Philosophical Psychology* 16(3): 365–88.

- Gjerdingen, Robert O. and David Perrott. 2008. "Scanning the Dial: The Rapid Recognition of Music Genres." *Journal of New Music Research* 37 (2): 93–100.
- Godøy, Rolf Inge. 2001. "Imagined Action, Excitation, and Resonance." In *Musical Imagery*, ed. Rolf Inge Godøy and Harald Jorgensen, 239–52. Lisse: Swets and Zeitlinger.
- . 2003. "Motor-Mimetic Music Cognition." *Leonardo* 36 (4): 317–19.
- . 2004. "Gestural Imagery in the Service of Musical Imagery." In *Gesture-Based Communication in Human-Computer Interaction: 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15–17, 2003, Selected Revised Papers, LNAI 2915*, eds. Antonio Camurri and Gulatiero Volpe, 55–62. Springer.
- . 2010. "Gestural Affordances of Musical Sound." In *Musical Gestures: Sound, Movement, and Meaning*, ed. Rolf Inge Godøy and Marc Leman, 103–25. Routledge.
- . 2013. "Quantal Elements in Musical Experience." In *Sound, Perception, Performance. Current Research in Systematic Musicology, Vol. 1*, ed. Rolf Bader, 113–28. Springer.
- . 2014. "Understanding Coarticulation in Musical Experience." In *Sound, Music, and Motion. Lecture Notes in Computer Science*, ed. Mitsuko Aramaki, Olivier Derrien, Richard Kronland-Martinet, and Sølvi Ystad, 535–47. Springer.
- Godøy, Rolf Inge, Haga, Egil, and Jensenius, Alexander Refsum. 2006. "Playing 'Air Instruments': Mimicry of Sound-producing Gestures by Novices and Experts." In *GW2005, LNAI3881*, ed. Sylvie Gibet, Nicolas Courty, and Jean-Francois Kamp, 256–67. Springer.
- Godøy, Rolf Inge and Leman, Marc, eds. 2010. *Musical Gestures: Sound, Movement, and Meaning*. Routledge.
- Goebel, Werner, Simon Dixon, Giovanni De Poli, Anders Friberg, Roberto Bresin, and Gerhard Widmer. 2008. "Sense in Expressive Music Performance: Data Acquisition, Computational Studies, and Models." In *Sound to Sense, Sense to Sound: A State of the Art in Sound and Music Computing*, ed. Pietro Polotti and Davide Rocchesso, 195–242. Logos.
- Haken, Hermann, Scott Kelso, and Herbert Bunz. 1985. "A Theoretical Model of Phase Transitions in Human Hand Movements." *Biological Cybernetics* 51 (5): 347–56.
- Harding, Sue, Martin Cooke, and Peter König. 2007. "Auditory Gist Perception: An Alternative to Attentional Selection of Auditory Streams?" In *WAPCV 2007, LNAI 4840*, ed. Lucas Paletta and Erich Rome, 399–416. Springer.
- Hauelsen, Jens and Thomas R. Knösche. 2001. "Involuntary Motor Activity in Pianists Evoked by Music Perception." *Journal of Cognitive Neuroscience*, 13 (6): 786–92.
- Jacob, Gordon. 1940. *Orchestral Technique*. Oxford University Press.
- Clapp, Stuart T., and Richard J. Jagacinski. 2011. "Gestalt Principles in the Control of Motor Action." *Psychological Bulletin* 137 (3): 443–62.
- Kohler, Evelyne, Christian Keysers, M. Alessandra Umiltà, Leonardo Fogassi, Vittorio Gallese, and Giacomo Rizzolatti. 2002. "Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons." *Science* 297 (5582): 846–48.
- Lartillot, Olivier, Tuomas Eerola, Petri Toiviainen, and Jose Fornari. 2008. "Multi-feature Modeling of Pulse Clarity: Design, Validation, and Optimization." In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08), Helsinki University of Technology, Espoo, Finland, September 1–4, 2008*, 305–8.
- Lieberman, Alvin M. and Mattingly, Ignatus G. 1985. "The motor theory of speech perception revised." *Cognition* 21: 1–36.
- Piston, Walter. 1991. *Orchestration*. Victor Gollancz Ltd.
- Rimsky-Korsakov, Nikolai. 1964. *Principles of Orchestration*. Dover Publications.
- Rosenbaum, David, 2009. *Human Motor Control*. 2nd edition. Elsevier.
- Schaeffer, Pierre, 1966. *Traité des objets musicaux*. Éditions du Seuil.

Wilson, Margaret and Günther Knöblich. 2005. "The Case for Motor Involvement in Perceiving Conspecifics." *Psychological Bulletin* 131 (3): 460–73.

[Return to beginning](#)

Copyright Statement

Copyright © 2018 by the Society for Music Theory. All rights reserved.

[1] Copyrights for individual items published in *Music Theory Online (MTO)* are held by their authors. Items appearing in *MTO* may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may *not* be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of *MTO*.

[2] Any redistributed form of items published in *MTO* must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of *MTO* in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of *MTO*, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may be copied and/or distributed for research purposes only.

[Return to beginning](#)

Prepared by Sam Reenan, Editorial Assistant



Number of visits: 837