20

Expressive Timing in Music and Dance Interactions

A Dynamic Perspective

Pieter-Jan Maes and Marc Leman

20.1 Introduction

In music and dance performance, timing involves the alignment of body movements with musical rhythms and temporal structures, such as tempo, metre, and phrasing. Timing may also imply *co-regulation* of this alignment among performers, based on adaptations of individual body movements to one another. In order for a performance to be successful, timing is characterized by fine-grained temporal nuances and variations that result in compelling dance grooves, typically embedded in configurations that display synchronization and counterpoint between dance and music, and exhibiting profound expressive qualities and emotional excitement. In music and dance, bodily coordination and the co-regulation of timing means walking the thin line between order and variation, between prediction and surprise, and between tension and release.

In recent decades, our scientific understanding of timing—in terms of sensorimotor coordination and co-regulation—has been growing steadily. The integration of novel technologies for recording and analysing sound and body movement within empirical research methods has greatly improved our understanding of timing, offering a fine-grained picture of specific relationships between body movement and musical signals. For example, several studies have used a frequency–time analysis of music–dance relationships; these are typically based on Fourier or wavelet analysis of movement velocity and involve subsequent correlation analysis and dimensionality reductions on frequency and/or time representations of body parts in relation to music (Amelynck et al., 2014; Hartmann et al., 2019; Toiviainen et al., 2010). However, despite overall progress in the field, two types of problems deserve our attention. The first type of problems relates to methodology and technology—when optical occlusion occurs when a body part is hidden from the view of a motion caption camera; when synchronization issues occur in contexts of multimodal data playback and recording; or simply when technology drives us to focus on kinematic parameters so that we tend to neglect other aspects of human movement, such as

Pieter-Jan Maes and Marc Leman, *Expressive Timing in Music and Dance Interactions* In: *Performing Time*. Edited by: Clemens Wöllner and Justin London, Oxford University Press. © Oxford University Press 2023. DOI: 10.1093/oso/9780192896254.003.0021

kinetic energy and muscle tension. It can be assumed that, for these problems, better measurement will move the field a step further, and issues in analysis will probably improve by applying and refining advanced analysis methods. The second type of problems, often closely related to the first type, relates to epistemology. These problems involve the gap between the artistic and the scientific approach, also known as the gap between insider and outsider knowledge, or the gap between first-person and third-person perspective (Leman, 2010). Typically, a dancer would use the term 'expressive timing' in view of an intended dance narrative, and understand timing in terms of corporeal articulations, gesturing, phrasing, communication of emotion, and storylines in relation to music. In contrast, scientists face the problem of measuring 'expressive timing', using theory to understand what measured signals mean in their context. Up to now, the scientific understanding of expressive timing is based on knowledge of short time frame sensorimotor mechanisms rather than in terms of artistic intentions. Closing this epistemological gap requires theoretical perspectives and conceptualizations from a humanities point of view, on top of the engineering challenges of measurement and analysis.

This chapter aims to provide some ingredients for understanding timing in the context of the above-mentioned epistemological problem. We show how expressive timing in music and dance can be understood in terms of ongoing sensorimotor processes and interactions within our environment. While that understanding currently focuses on short time frames, we aim to expand our understanding of expressive timing in larg(er) time frames. We suggest that the concept of embodied hierarchies and dynamical theory might help our scientific understanding of expressive timing to progress beyond the current state of the art.

At this point, most researchers tend to agree that timing is a critical issue in music and dance (Keller, 2014). Timing is typically a stabilizing factor and constancy in timing is needed to allow performers to predict the future so that collaboration and co-regulated alignments between dance and music become possible. Accordingly, the timing should be such that it entrains others to respond and that, through mutual exchanges of timing, time patterns (endowed with expression) emerge and become a new level for excitement and appeal to engagement (Leman, 2016). Key concepts in our understanding of timing are entrainment and emergence (see Madison, this volume). Entrainment occurs when patterns evolve over time towards particularly stable pattern configurations, such as when dance and music rhythms gradually reach in-phase or anti-phase synchronization. Emergence occurs when interactions between micro-patterns generate macro-pattern configurations, such as when a metre emerges from polyrhythms (both in music and body swaying).

Apart from these considerations about entrainment and emergence, we argue that a combination of two major theoretical frameworks, known as embodiment and predictive coding, can shed light on the underlying principles of timing in music and dance interactions. Embodiment theory relates to bodily realizations of timing, and to the idea that timing in humans is grounded in action repertoires shared among conspecifics (Leman, 2007; Leman et al., 2017). Predictive coding theory (Koelsch et al., 2019; Vuust & Witek, 2014) involves predictions of timing and the idea that participants in an ensemble adapt their timings based on individual assumptions about joint timing constancy. While these theories support our understanding of expressive timing in short time frames, it is of interest to investigate their potential for understanding expressive timing in large time frames such as dance narratives. A narrative occurs when entities of expressive timing (such as gestures) are sequentially ordered in particular ways, such that their contrasts and tensions tell a story. Such a (non-verbal) story offers a powerful way to create affective and emotional responses in those who perceive the narrative, even if the expressive nature of the narrative—or the responses to it—cannot be accurately described in verbal terms. At this point, it is of interest to note that expressive timing, given its non-verbal character, can be guided by a linguistic storyline such as the libretto in opera, or the lyrics of a song. For example, rather than counting a 13/12 metre, it may be easier to articulate a linguistic mantra and/or incorporate the metre in particular choreographies, which then can be cognitively *outsourced* (Maes, Giacofci, & Leman, 2015; Maes, Wanderley, & Palmer, 2015). But we should also consider facilitation of expressive timing through images that support particular postures, attitudes, and movement concepts (see also Godøy & Leman, 2010). For example, in Tchaikovsky's *Swan Lake*, the behaviour of swans (i.e. how swans move in their natural environment) provides an inspiration for a narrative of subsequent gestures endowed with expressive timing, revealing intentions related to seduction and enchantment. In short, to better understand concepts like 'expressive timing' and 'narrative of expressive timing' in terms of measured quantities, we envision an expansion of the embodied predictive coding theory in the direction of larger time frames. Given the state of the field, such an ambition cannot be anything but embryonic. Nevertheless, some ingredients are worth considering.

20.2 Outsourcing Timing to Sensorimotor Processes: The Body and Sound as Time-Keeper

At present, it is generally agreed that there is no single dedicated locus within the human brain that is solely responsible for time processing, like the central master clock operating in a computer. In contrast, there is ample support for the idea that time-keeping and -production are mediated by the intrinsic activation dynamics of a distributed network of cortical and subcortical brain regions (Buhusi & Meck, 2005; Paton & Buonomano, 2018, see Doelling et al., this volume). The specific brain regions that comprise a timing network may vary and depend on factors such as the duration of the timed intervals and the sensorimotor modalities involved (Grahn, 2012; Grondin, 2010). A better understanding of the contribution of sensorimotor mechanisms in timing is important, as timing is foundational to both music and dance.

In an earlier series of studies, we investigated in more detail how sensorimotor mechanisms, related to action-perception processing, may support timing in music performance (Maes, Giacofci, & Leman, 2015; Maes, Wanderley, & Pamer, 2015). The core idea was that time is inherently embedded in the performance of bodily gestures and in the sensory information perceived during live musical interactions (the sounds, the visual information about other's gestures, the tactile vibrations, etc.). To understand the extent to which people can rely on self-performed actions and perceived sensory patterns to regulate individual and joint time-keeping in musical performances, we designed experiments in which we studied participants' time-keeping abilities in rhythmical interval production tasks. A comparison between conditions in which, on the one hand, sensory and/or motor information was continuously available, and, on the other, conditions in which no (or only minimal) sensorimotor information was available thereby revealed underlying timing control mechanisms. For example, we hypothesized that when no continuous sensorimotor information was available, participants would need to rely more on their cognitive resources to keep track of time. This idea was grounded in earlier timing research positing a distinction between an emergent timing mechanism and an event-based timing mechanism underlying the control of rhythmic movements (Delignières et al., 2004; Robertson et al., 1999; Torre & Balasubramaniam, 2009; Zelaznik, 2005). The emergent timing mechanism is based on the idea that time-keeping is rooted in the dynamically unfolding action-perception processes going on in live embodied musical interactions. For instance, the performance of a bodily gesture, even the simplest one, necessarily implies a demarcation of a temporal interval, from the beginning of the gesture until its ending. Time thus becomes a property that is embedded in—and emerges from—the control of movement dynamics, and the body and motor system may function jointly as a timekeeper in support of rhythmical behaviour. This emergent timing mechanism assumes the presence of a continuous movement, typically periodic and oscillatory in nature, that helps to keep track of time. In many instances, however, rhythmical behaviour is performed by discrete movements, characterized by short, salient movement events that demarcate (quasi-)periodic temporal intervals in which no movement occurs. Keeping track of time during these intervals is regulated by what is called an event-based timing mechanism. It is often assumed that event-based timing relies on an explicit mental representation of time, requiring cognitive resources such as attention and working memory, such as in counting.

As a matter of fact, music and dance provide relevant contexts to study these distinct timing mechanisms. In a study on samba dance, Naveda and Leman (2010) suggested that continuous gestures mark the time points of the metre in space, with an event-based representation offering a concise representation of the dynamic continuous movement. In another study by Maes, Wanderley, and Palmer (2015), cellists were asked to perform simple melodies using two distinct types of bowing gestures, namely legato articulation and staccato articulation. Specific for legato articulation is that bowing strokes are smoothly tied to one another (cf. continuous rhythmic movement), while staccato articulation is characterized by short tone onsets separated by interspersed pauses (cf. discrete rhythmic movement). Hence, in accordance with the existing theory on emergent versus event-based timing, we hypothesized that timing of staccato articulation would rely more on cognitive resources compared to legato articulation. To test this hypothesis, we integrated a dual-task paradigm in which we crossed the primary musical timing task (production of cello tones at regular temporal intervals) with a secondary 'cognitive load' task relying on attention and working memory. Accordingly, if both tasks tapped into the same resources, it was to be expected that the tasks would mutually impair one another. In that way, it became possible to assess the relative roles of sensorimotor and cognitive resources in different musical timing contexts. The results of this experiment showed that only the timing of staccato articulation was impaired under conditions of heightened cognitive load, in particular at slower performance tempi. These findings indicated that, in line with the existing theory on emergent versus event-based timing, the temporal control of rhythmic movements may, indeed, recruit distinct neural networks depending on the performance context and modalities involved (Koch et al., 2009; Petter et al., 2016). It is also of practical interest to know that, in conditions of heightened cognitive load—frequent in music and dance contexts—body movements may support the regulation of timing.

In a follow-up study, the focus was further directed towards the possible role of sounds in support of an emergent timing mechanism (Maes, Giacofci, & Leman, 2015). As it does for bodily gestures, time may be similarly embedded in, and emerge from, the unfolding of acoustical energy and/or musical patterns. Through repeated experiences, musicians and dancers may establish an association between perceived auditory regularities in music and the duration of temporal intervals. In turn, in coupling action to perception, body movements may be aligned in time to perceived auditory regularities in order to obtain successful performance timing. Sensorimotor processing is at the core of this emergent timing mechanism. Keeping track of time becomes a matter of coupling actions to perception and of underlying sensorimotor learning processes, rather than of the purely mental act of keeping track of time. Consequently, it is to be expected that this emergent sensorimotor timing mechanism is more robust in situations of heightened cognitive load, as it relies less on cognitive resources for keeping track of time.

This idea was tested in the study by Maes, Giacofci, and Leman (2015). In this study, participants were asked to tap out a simple melody as regularly as possible by repeatedly pressing a keyboard key. This timing task was integrated into the synchronization-continuation paradigm, whereby participants were first asked to synchronize their taps to an auditory metronome (indicating the target temporal interval to be produced) and to maintain that regular pace after the metronome had stopped. The crucial aspect of the experiment resided in the manipulation of the auditory tones that were produced, in particular their amplitude envelope that captured the amplitude change of a tone over time (attack, decay, sustain, and release). Tapping the key either produced a short piano tone (discrete tone) or a long piano tone of which the amplitude envelope exactly fitted the temporal interval that needed to be produced (continuous tone). In synchronizing taps to the initial metronome ticks, participants were expected to implicitly learn the relationship between the amplitude envelope of the continuous tone, and the target temporal interval that needed to be produced; in other words, the onset of each next metronome tick (the point where a tap should be performed) occurred systematically at the point where the previous piano tone ceased to be heard. Results of the experiment showed that continuous tones contributed to regular timing production, suggesting that participants could, indeed, rely on the perceived temporal characteristics of the self-produced piano tones to regulate their timing. This conclusion was further supported by an additional temporal manipulation of the amplitude envelope of the continuous tone. While participants were tapping at a regular pace (without a metronome), we gradually shortened the duration of the continuous tone, from a duration that equalled the target temporal interval of 1,100 ms to a duration of 867 ms. Consequently, we observed that participants adapted their temporal tapping interval correspondingly in order to maintain the alignment of each tap with the ending of the piano tone produced by the previous tap. Again, this indicated that the amplitude envelope of piano tones was taken as temporal reference for participants' control of their own timing.

The studies above demonstrated that both bodily gestures and sounds can support an emergent timing mechanism, rooted in sensorimotor processing. This points to an interesting mechanism that can be of practical benefit to musicians and dancers. When relying on the alternative *event-based* timing mechanism, keeping track of time requires cognitive resources, such as attention and memory, meaning that these resources—at least to some extent—cannot be allocated anymore to other tasks in music and dance interaction that require them. In that regard, the use of sensorimotor processing skills underlying the emergent timing mechanism may be a valuable and welcome strategy to control timing in music and dance interactions in an alternative way, freeing up cognitive resources for executing additional tasks. In earlier work (Coorevits et al., 2020; Maes, Giacofci, & Leman, 2015) we have referred to this strategy as an *outsourcing* strategy. With this concept, we point to the idea that timing can be transferred, at least partly, from cognitive resources to sensorimotor processing. This idea fits within a more general tendency to use physical actions—as well as phenomena, objects, and technologies within their environment—to reduce the cognitive demands of tasks such as mental timekeeping.

In one study, we investigated how this emergent timing mechanism may not only be of relevance for regulating individual timekeeping, but also for interpersonal co-regulation in music and dance performance. In Coorevits et al. (2020), we applied the framework of emergent versus event-based timing to a context of social music interaction. In particular, we were interested whether so-called ancillary movements in music performance could contribute to joint regular interval production. Unlike sound-producing movements, ancillary movements do not contribute directly to the production of musical sounds (Cadoz & Wanderley, 2000). However, they do support embodied expression and communication of musical intentions (Davidson, 1993; Wanderley, 2002). Because of the inherent spatiotemporal nature of ancillary movements, we were interested in how they might support temporal co-regulation in social music performance. In our experiment, we manipulated the type of ancillary gestures that participants were allowed to perform in a joint tapping task. The goal of this task was to perform the same musical melody together as regularly as possible by tapping their finger on tapping pads. Between conditions, we controlled for the type of tapping gestures performed. Participants were either instructed to perform prominent, continuous up/downward movements in between successive taps (continuous gestures), or were restricted to perform any movements between successive taps (discrete gestures). In line with the findings on individual timing, results showed that when participants of a dyad could see each other, continuous gestures (performed by both) led to better joint timing in terms of consistency and accuracy of synchronized tapping. In addition, we found that the type of gestures employed by musicians modulated leader–follower dynamics, in the sense that participants who performed continuous bodily gestures tended to take the leader role in the interaction.

20.3 Embodied Predictive Coding and Expressive Narratives

Until recently, sensorimotor coupling was typically studied from the viewpoint of rhythmic tapping and timekeeping (Repp & Su, 2013). For understanding the basics of music and dance interactions, this paradigm is still very relevant. Currently, however, the tapping studies are expanded in several directions, including neuroscience (Elst et al., 2021) and musicology, where limbs, full body movement, and more realistic (or artistic) movements are studied (see previous section for examples). Overall, our understanding of timekeeping and of simple (discrete, continuous) sensorimotor-based timing (both in individual and social contexts), is based on prediction, as captured by the predictive coding theory, a brain theory which holds that the brain is a prediction generating engine (Koelsch et al., 2019). In the context of timing, the engine would constantly adapt a prediction about timing based on prediction error, which is the difference between the predicted timing and the real timing (Vuust & Witek, 2014). In the context of co-regulated timing, Leman (2021) has implemented this principle in an algorithm called *BListener*, an application that solely focuses on perception (not involving any embodiment) but that clearly explicates how Bayesian inferencing about timing works. BListener could become part of an action-perception coupling mechanism in a virtual musician that could—with some further programming effort—play along with human musicians and interact with dancers. BListener perceives the duration between two consecutive onsets (so-called inter-onset intervals (IOIs)) and estimates the constancy of subsequent fluctuating IOIs. The onsets come from musical onsets of notes, hand clapping, or from dance onsets, such as feet touching the ground. The

Bayesian inference component in this approach relates to the updating mechanism for predicting timing. The Bayesian view involves a likelihood (i.e. how likely it is that a new temporal event stems from the assumed timing constancy) and a prior (the expected timing constancy in absence of observations), from which a posterior (expected timing constancy, given the new temporal event) is inferred. In view of a subsequent onset, the prior is updated by replacing the old prior with the posterior. This Bayesian inference is embedded in a system dynamic that updates the priors every sample (e.g. 100 times per second). In music and dance, the rhythms can often be conceived as a concatenation of IOIs taken from a limited set of IOI values having a simple relationship among each other, such as binary or ternary. These binary and ternary components together mark the metre. Constancy in timing is therefore typically a constancy in the fluctuation of subsequent durations that establish the metre. Repetitive dances, for example, can thus be captured in terms of metre constancy, which may even may be a marker of a qualitative performance, attracting interest and leading to excitement or empowerment.

While this theory provides a solid basis for timing (see also Vuust & Witek, 2014), it nevertheless does not incorporate the embodiment perspective, which we believe to be essential—not only for expressive timing being in line with the timing predictions, but also for generating narratives of expressive timing. Given its relation to the unfolding of a gesture over time, expression implies careful planning ahead of time, and hence careful prediction of timing. The Bayesian brain could be conceived as a controller mechanism for expressive timing necessarily connected with corporal articulations that unfold in a pattern over time (see the previous paragraphs). Looking back at finger tapping studies and extensions of the experimental paradigm of sensorimotor control in music and dance interaction studies, it can be noticed that sensorimotor mechanisms and predictions work well in the millisecond range. However, these mechanisms can typically have no awareness of larger-scale developments such as we know them in artistic performances. At that level, we often speak about patterns at larger time frames, such as expressive arcs (as in the unfolding of a gesture that starts smoothly, going to a maximum in movement speed and articulation, and then ends again smoothly) and tension bows (as in a gesture that starts in sync with the music, becomes highly de-synchronized, and ends again in sync) that may reach far beyond the 10–20 s range. Such expressive arcs and tension bows are assumed to influence the timings at the local level. Hence, for understanding this aspect, it is necessary to call upon a more global level, which is typically associated with the narrative and its associated intentionality, or goal-driven character. There is a considerable gap between sensorimotor studies, on the one hand (their underpinning by predictive coding theories and micro-embodiment), and studies about artistic expression, narratives, and intentions, on the other, and this issue is currently not well understood.

Therefore, the idea is that the outsourcing of sensorimotor processes could be conceived within a broader perspective on 'the narrative of expressive timing', or, alternatively called, 'the expressive narrative'. This narrative, for example, can be conceived as a storyline expressed in music and through dance, as in Tchaikovsky's *Swan Lake*, or it can be an abstract story line, such as the dance sequence in Baroque group dances, which is meant to change partnership among a group of male and female dancers. The narrative element can also be implied by a ritual activity aimed at reaching the divine, such as in the Zekir ritual of the Soefi (Rouget, 1985); or it can be an inherited story line in oral tradition about cattle, providing an ingredient for forward and backward moving in African circular dances (Phyfferoen et al., 2017). The narrative can be anything, as long as it involves sequences of expressive timing. Often, these sequences aim at generating tension arcs in pursuit of an artistic effect. Tension arcs indeed capture the attention of audiences.

In the examples mentioned, we assume that the overall narrative component acts as a guiding principle for intentionality, or goal-directed music–dance action, that creates tension, surprise, and emotional effects through contrasts, both in global timing (the storyline as intentional guide) and local timing (the outsourcing of time through movement). Huron's theory of expectation (2006) is based on the contrast between fast reactive responses and slow appraisal responses. Here we assume that expectations (of timing) are embodied, through the outsourcing of timing in gestures, and that musicians/dancers engage in recruiting these gestures to enact the narrative. As such, the outsourced embodied timing supports the expressive narrative. However, one could also argue that the expressive narrative guides the outsourced embodied timing.

Tensions between biological bias in timing (such as entrainment to beats) and cultural habits in timing (such as the codifications of how beats are accommodated through dance) can be a rich source for artistic expression. This view perhaps challenges a too narrow view on predictive coding as local prediction error minimalization engine. Being guided and misled by musicians or dancers, so that audiences become surprised and novel musical patterns emerge, is likely to be an essential part of expressive and creative dance–music interactions.

Given the current state of the art, we propose two pathways towards a better understanding of expressive timing in short time frames (the sensorimotor perspective) and timing in large time frames (the narrative perspective). The first pathway consists of a gradual expansion of the timescale at which action-perception coupling is studied. The second pathway consists in the development of a theoretical perspective that incorporates sensorimotor prediction mechanisms with the expressive narrative in a perspective that is driven by a dynamic theory.

20.3.1 Embodied Hierarchies

The concept of embodied hierarchies encompasses the idea that dance and music draw upon gestures as a mid-level concept between low-level sensorimotor action and high-level narrative (Godøy & Leman, 2010). Previous work on gesture suggested that gestures can be conceived as a concatenation of basic gestures that compose a gesture narrative, or as a gesture narrative that can be decomposed in units subsuming basic gestures (Leman & Camurri, 2006; Leman & Naveda, 2010; Naveda & Leman, 2010, see also Jensenius et al., 2010). Hilt and colleagues (2019) propose a multilayer approach to group coordination, suggesting that the co-regulation of group behaviour is based on the exchange of information across several layers, each of them tuned to carry specific coordinative signals. Multilayer sensorimotor communication in which several timescales are embedded through several sensory channels (auditory, visual, spatial) may be the key concept for understanding how musicians and, more generally humans, communicate with each other (see also Eerola et al., 2018).

20.3.2 Dynamic Theory

A dynamic theory might help us to understand how expressive timing at the sensorimotor level gets embedded in expressive narratives. Interactions between the two levels are assumed to draw upon balancing mechanisms, creating states of equilibrium (or homeostasis) among involved processes. For example, the tempo of a joint performance can be seen as a state of equilibrium reached through co-regulated expressive timing. This interaction state emerges from a dynamic interaction among performers. Small differences in the timing of one musician or dancer may affect overall tempo-balance at the group level. Several small differences from different musicians or dancers thereby appear as small fluctuations that are compensated for (negative feedback) in view of keeping the balance (i.e. the timing constancy). However, during a performance, one tempo may suddenly have to change to another. That sudden change in tempo would imply another type of regulation (called positive feedback) in the sense that all processes adapt to the newly intended equilibrium state for co-regulated timing. The music and dance ensemble's timing is an example of a system capable of building and maintaining a series of homeostatic co-regulated states that are intrinsically sense-giving, both during the act of creating music and dance and during the act of perceiving both. The latter can be understood as the motivational drive for acting and engaging in music and dance interactions. The narrative can be conceived as an overall structuring element for expressive timing. It can globally define the tempo (e.g. fast or slow), the articulation (e.g. soft or firm), and it can guide the tension of an intended story line (e.g. from fast to slow; from soft to firm). The narrative thus appears as an intentional co-regulating factor, a hyper-parameter that co-steer the parameters that define the semi-automated and automated low-level mechanisms at longer terms. This intentional co-regulation can be understood from the viewpoint of predictive coding as it can probably set expectations for expectations, itself being steered by the possibilities and limitations of the human body.

20.4 Conclusion

Some problems in current dance and music research are concerned with intrinsically different types of knowledge about expressive timing. We reflected on the idea that our scientific knowledge in terms of (subliminal) sensorimotor control and our artistic knowledge in terms of (intended) narrative, can, to some extent, be filled. Thanks to novel technologies and methodologies, artistic intuitions about timing tend to push scientific research further in directions that go beyond classical research topics in sensorimotor timing, such as finger tapping and time-keeping. By looking at expressive timing in music and dance at large time frames, we tend to walk the thin line between order and variation, between prediction and surprise, between tension and release, and between sensorimotor and narrative perspectives. Research that originated in (systematic, cognitive, empirical) musicology has, indeed, contributed many exciting new insights and much understanding of our intuitions about music and dance interactions. Expressive timing has been fully incorporated in modern cognitive (neuro)science. Yet, many challenges remain to be addressed in future research. In this chapter, we suggested that the concept of embodied hierarchies could be helpful as a starting point for understanding multiple time frames of expressive timing. In parallel with this concept, we suggested that expressive timing can be understood from the viewpoint of homeostatic (co-)regulation, which is a dynamic principle.

Clearly, more insights about the relationship and interdependencies of timing on these different temporal scales is needed. Therefore, the further exploration and application of (non-)linear time-series analysis tools that may unveil interdependencies across different temporal scales is important. Consequently, we would gain deeper insights into the biological and cultural principles regulating (interpersonal) timing behaviour, and contribute to the refinement of existing theoretical frameworks. Theory formation will need to integrate the various levels and mechanisms of timing control and co-regulation, ranging from low-level spontaneous coordination based on dynamical principles (Kelso, 1995, 2021; Leman, 2021), to higher-level learning, to predictive processing, and active inference (Gallagher & Allen, 2018; Koban et al., 2019; Sebanz & Knoblich, 2009). In addition, further research is required on the role of variability, deviation, and surprise in human engagement with music and dance. More than other daily activities, for which optimal functioning in terms of behaviour and decision-taking is often crucial, music and art in general offer room for exploring the dynamics of prediction and surprise, of recurrence and variability (Schiavio et al., 2021). Far from being undesirable, surprise and deviation may reveal a diversity of emerging novel patterns in both sound and bodily behaviour, leading to highly rewarding and novel subjective experiences.

New technologies for recording and analysing sound, bodily behaviour, and (neuro)physiological responses have dramatically improved this research on timing in music and dance in the past decades. Recently, new technologies in the domain of extended reality (virtual, augmented, and mixed reality) are being incorporated into the empirical study of embodied music interaction (Turchet et al., 2021; Van Kerrebroeck et al., 2021). They offer radically new possibilities for multisensory, immersive stimulus creation in experimental research on narratives of expressive timing. Advances in (neuro)cognitive research into the fundamental principles underlying the embodied nature of expressive timing, along with emergent new technologies, provide a vibrant context in which research on timing may flourish in the coming years.

References

Amelynck, D., Maes, P.-J., Martens, J. P., & Leman, M. (2014). Expressive body movement responses to music are coherent, consistent, and low dimensional. *IEEE Transactions on Cybernetics*, *44*(12), 2288–2301.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, *6*(10), 755–765.

Cadoz, C., & Wanderley, M. M. (2000). Gesture—music. In M. M. Wanderley & M. Battier (Eds.), *Trends in gestural control of music* (pp. 71–93). Ircam-Centre Pompidou.

Coorevits, E., Maes, P.-J., Six, J., & Leman, M. (2020). The influence of performing gesture type on interpersonal musical timing, and the role of visual contact and tempo. *Acta Psychologica*, *210*, 103166.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*(3), 103–113.

Delignières, D., Lemoine, L., & Torre, K. (2004). Time intervals production in tapping and oscillatory motion. *Human Movement Science*, *23*(2), 87–103.

Eerola, T., Jakubowski, K., Moran, N., Keller, P. E., & Clayton, M. (2018). Shared periodic performer movements coordinate interactions in duo improvisations. *Royal Society Open Science*, *5*(2), 171520.

Elst, O. F., Vuust, P., Kringelbach, M. L., & Foster, N. H. (2021). *The neuroscience of dance: A systematic review of the present state of research and suggestions for future work*. PsyArXiv. https://doi.org/10.31234/osf.io/kfpcx

Gallagher, S., & Allen, M. (2018). Active inference, enactivism and the hermeneutics of social cognition. *Synthese*, *196*(6), 2627–2648.

Godøy, R. I., & Leman, M. (2010). *Musical gestures: Sound, movement, and meaning*. Routledge.

Grahn, J. (2012). Neural mechanisms of rhythm perception: Current findings and future perspectives. *Topics in Cognitive Science*, *4*(4), 585–606.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception & Psychophysics*, *72*(3), 561–582.

Hartmann, M., Mavrolampados, A., Allingham, E., Carlson, E., Burger, B., & Toiviainen, P. (2019). Kinematics of perceived dyadic coordination in dance. *Scientific Reports*, *9*(1), 1–14.

Hilt, P. M., Badino, L., D'Ausilio, A., Volpe, G., Tokay, S., Fadiga, L., & Camurri, A. (2019). Multi-layer adaptation of group coordination in musical ensembles. *Scientific Reports*, *9*(5854), 1–10.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT Press.

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). Routledge.

Keller, P. E. (2014). Ensemble performance: Interpersonal alignment of musical expression. In R. Timmers & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (pp. 260–282). Oxford University Press.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. A Bradford Book.

Kelso, J. A. S. (2021). Unifying large-and small-scale theories of coordination. *Entropy*, *23*(5), 537.

Koban, L., Ramamoorthy, A., & Konvalinka, I. (2019). Why do we fall into sync with others? Interpersonal synchronization and the brain's optimization principle. *Social Neuroscience*, *14*(1), 1–9.

Koch, G. K., Oliveri, M., & Caltagirone, C. (2009). Neural networks engaged in milliseconds and seconds time processing: Evidence from transcranial magnetic stimulation and patients with cortical or subcortical dysfunction. *Philosophical Transactions of the Royal Society. Series B, Biological Sciences*, *364*(1525), 1907–1918.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, *23*(1), 63–77.

Leman, M. (2007). *Embodied music cognition and mediation technology*. MIT Press.

Leman, M. (2010). Music, gesture and the formation of embodied meaning. In R. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 126–153). Routledge.
Leman, M. (2016). *The expressive moment: How interaction (with music) shapes human empowerment*. MIT Press.
Leman, M. (2021). Co-regulated timing in music ensembles: A Bayesian listener perspective. *Journal of New Music Research*, *50*(2), 121–132.
Leman, M., & Camurri, A. (2006). Understanding musical expressiveness using interactive multimedia platforms. *Musicae Scientiae*, *10*(1 Suppl), 209–233.
Leman, M., Lesaffre, M., & Maes, P.-J. (2017). What is embodied music interaction? In M. Lesaffre, P.-J. Maes, & M. Leman (Eds.), *Routledge companion to embodied music interaction* (pp. 1–10). Routledge.
Leman, M., & Naveda, L. (2010). Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in Samba and Charleston. *Music Perception*, *28*(1), 71–91.
Maes, P.-J., Giacofci, M., & Leman, M. (2015). Auditory and motor contributions to the timing of melodies under cognitive load. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1336–1352.
Maes, P.-J., Wanderley, M. M., & Palmer, C. (2015). The role of working memory in the temporal control of discrete and continuous movements. *Experimental Brain Research*, *233*(1), 263–273.
Naveda, L., & Leman, M. (2010). The spatiotemporal representation of dance and music gestures using topological gesture analysis (TGA). *Music Perception*, *28*(1), 93–111.
Paton, J. J., & Buonomano, D. V. (2018). The neural basis of timing: Distributed mechanisms for diverse functions. *Neuron*, *98*(4), 687–705.
Petter, E. A., Lusk, N. A., Hesslow, G., & Meck, W. H. (2016). Interactive roles of the cerebellum and striatum in sub-second and supra-second timing: Support for an initiation, continuation, adjustment, and termination (ICAT) model of temporal processing. *Neuroscience & Biobehavioral Reviews*, *71*, 739–755.
Phyfferoen, D., Stroeken, K., & Leman, M. (2017). The Hiplife zone: Cultural transformation processes in African music seen from the angle of embodied music interactions. In M. Lesaffre, P.-J. Maes, & M. Leman (Eds.), *The Routledge companion to embodied music interaction* (pp. 232–240). Routledge.
Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, *20*(3), 403–452.
Robertson, S. D., Zelaznik, H. N., Lantero, D. A., Bojczyk, K. G., Spencer, R. M., Doffin, J. G., & Schneidt, T. (1999). Correlations for timing consistency among tapping and drawing tasks: Evidence against a single timing process for motor control. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1316–1330.
Rouget, G. (1985). *Music and trance: A theory of the relations between music and possession*. University of Chicago Press.
Schiavio, A., Maes, P.-J., & van der Schyff, D. (2021). The dynamics of musical participation. *Musicae Scientiae*, *26*(3), 604–626.
Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*(2), 353–367.
Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception*, *28*(1), 59–70.
Torre, K., & Balasubramaniam, R. (2009). Two different processes for sensorimotor synchronization in continuous and discontinuous rhythmic movements. *Experimental Brain Research*, *199*(2), 157–166.
Turchet, L., Hamilton, R., & Çamci, A. (2021). Music in extended realities. *IEEE Access*, *9*, 15810–15832.
Van Kerrebroeck, B., Caruso, G., & Maes, P.-J. (2021). A methodological framework for assessing social presence in music interactions in virtual reality. *Frontiers in Psychology*, *12*, 663725.
Vuust, P., & Witek, M. A. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*, 1111.
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. GW 2001*. (Lecture Notes in Computer Science) (Vol. 2298, pp. 241–253). Springer.
Zelaznik, H. N. (2005). Timing variability in circle drawing and tapping: Probing the relationship between event and emergent timing. *Journal of Motor Behavior*, *37*(5), 395–403.