The influence of performing gesture type on interpersonal musical timing, and the role of visual contact and tempo

Esther Coorevits, Pieter-Jan Maes, Joren Six, and Marc Leman

Abstract

Bodily gestures play an important role in the communication of expressive intentions between humans. Music ensemble performance, as an outstanding example of nonverbal human communication, offers an exemplary context to study and understand the gestural control and communication of these expressive intentions. An important mechanism in music ensemble performance is the anticipation and control of interpersonal timing. When performing, musicians are involved in a complex system of mutual adaptation which is not completely understood so far. In this study, we investigated the role of performers' gestures in the mediation process of interpersonal timing in a dyad performance. Therefore, we designed an experiment in which we controlled for the use of hand and arm movements in a musical task, in which dyads were asked to synchronously tap out a melody. Next to their comfortable/natural way of tapping, we instructed participants to either perform pronounced expressive hand and arm gestures in between successive taps, or to restrict from any overt body movement. In addition, we looked at effects of visual contact (yes/no) and tempo (slow: 50 beats per minute; fast: 100 beats per minute). The results show that performers' gestures improve interpersonal musical timing, in terms of the consistency and accuracy of onset asynchronies, and of the variability of produced inter-onset intervals. Interestingly, we found that the use of expressive gestures, in regard to comfortable/natural movements, add to these positive timing effects, but only when there is visual contact and at the slow tempo. In addition, we found that the type of gestures employed by musicians may modulate leader-follower dynamics. Together, these findings are explained by human anticipation mechanisms facilitated by gesturing, shedding new light on the principles underlying human communication of expressive intentions, through music.

The role of performers' gestures in joint musical communication

Music provides a powerful medium for humans to communicate, to share and explore emotions, and to create a deep sense of mutual understanding and togetherness (Lesaffre, Maes, & Leman, 2017; Miell, MacDonald, & Hargreaves, 2005). Most forms of musical communication involve direct gestural interaction, where the flow of multi-sensory information mediated by corporeal articulations enables expressive interactions (Leman, 2016). Performing music together however is a highly complex task, as it requires fine-tuned motor coordination and high-level cognitive skills. It is truly remarkable how music performers (and their audiences) manage to organize their behavior into strongly coupled activity (cf. entrainment) to create experiences of shared sense-making and feeling. From a scientific point of view, a key challenge is to better understand the fundamental mechanisms underlying the human ability to keep track of time, to produce regular temporal intervals, and to align to the timing of co-performers' actions.

Different approaches have been proposed to account for motor control mechanisms underlying (quasi-)periodic interval production. This research points out that these mechanisms are to a large degree task- and context-dependent. One of the determining factors is the availability of sensorimotor information, related to self-produced actions and to sensory information received from the external environment. Previous research puts forth that the execution of continuous bodily movements may underlie an emergent time-keeping mechanism (Cassenti, 2011; Delignières, Lemoine, & Torre, 2004; Maes, Wanderley, & Palmer, 2015; Robertson et al., 1999; Torre & Balasubramaniam, 2009). In this view, timing emerges from the control of movement dynamics, rather than being explicitly controlled by a dedicated internal clock. In addition, humans have been shown to be highly receptive to time-varying sensory patterns, which may equally support time-keeping and synchronization (Coull & Droit-Volet, 2018; Maes, Giacofci, & Leman, 2015; Motanis, Seay, & Buonomano, 2018; Schwartze, Tavano, Schröger, & Kotz, 2012). In short, spatiotemporal patterns emerging from human sensorimotor interactions with the world may provide important information about the dynamical properties in the performance of musical ensembles. However, at instances where one cannot rely on bodily gestures and/or sensory information, internal neural mechanisms are expected to be deployed instead for keeping track of time. At current, there is a general agreement that there is no single, dedicated internal “clock” residing within the brain, but that time is encoded in dynamical patterns of neural activity within a distributed network, including cortical areas, basal ganglia, and cerebellum (Allman, Teki, Griffiths, & Meck, 2014; Buhusi & Meck, 2005; Karmarkar & Buonomano, 2007; Kononowicz, van Rijn, & Meck, 2016; Remington, Egger, Narain, Wang, & Jazayeri, 2018).

Based on this theoretical framework, it is to be expected that bodily movements and available sensory information play a functional role in musical timing. Indeed, previous research has indicated that continuous body movements and auditory (tone) feedback may support (individual) timing production in musical tasks, in particular in conditions of heightened cognitive load (Maes, Giacofci, et al., 2015; Maes, Wanderley, et al., 2015). In the context of music ensemble coordination, it has been shown that head nods, body swaying and gaze patterns play an important role in establishing and maintaining interpersonal synchrony and leader-follower relations (Keller, 2014). Here, body movements associated with music performance can provide visual cues, which in combination with the timing cues from the auditory domain regulate the coordination and communication between ensemble performers. Keller and Appel (2010) for instance found that the coordination of body sway was related with the synchronization between participants. According to Goebl and Palmer (2009), pianists’ head movements became more synchronized when auditory feedback was reduced in the performance. They also noticed that pianists who were designated as the leader raised their fingers higher and preceded the other pianist in timing. This research suggests that bodily gestures can help in communicating and controlling timing in ensemble performance. This is parallel to language communication, where hand gestures have been shown to convey (semantic) information and foster mutual understanding (Krauss, Chen, & Chawla, 1996).

However, most research on ensemble coordination consist of manipulating the conditions in which sensory exchange is possible and/or manipulating the leader-follower roles, but not the bodily movements themselves. This is likely due to the fact that sensory conditions and leader-follower roles can be more easily manipulated. More critical is to manipulate the performers' body movements, as these make out a functional aspect of a musical performance. Yet research indicates that, though each musician has his/her own typical movement repertoire and peculiarities, musicians may vary gestural patterns according to differences in performance style (Caruso, Coorevits, Nijs, & Leman, 2016; Davidson, 2012; Demos, Chaffin, & Logan, 2018; Teixeira, Loureiro, Wanderley, & Yehia,

2015) (Davidson, 2012; Teixeira, Loureiro, Wanderley, & Yehia, 2015). In addition, experimental control of body movements in playing a musical piece has been applied previously, in particular to investigate effects on the perception of expressive and emotional intentions (Dahl & Friberg, 2004; Davidson, 1993; Juchniewicz, 2008) (Davidson, 1993; Dahl & Friberg, 2003; Juchniewicz, 2008). This raises the question how deliberately manipulating bodily gestures might affect the entrainment dynamics of ensemble performance in terms of timing control, synchronization, and leader-follower relationships. For that purpose, we conceived of a simple musical task, namely two musicians tapping a beat together. Tapping a beat together can be considered as the necessary basis of coupled musical interactions, on top of which more elaborated musical structures and expressive nuances can be build. Therefore, it is of interest to study deeper how gestures may regulate this basic coupling behavior. We controlled thereby for the use of hand and arm movements; next to their comfortable/natural way of tapping, we instructed participants to either perform pronounced (but naturally feeling) expressive hand and arm gestures in between successive taps, or to restrict from any overt body movement. We expected that the onset asynchronies between performers would be smaller when they perform expressive movements in between the taps, and vice versa, be larger when no movement in between onsets is allowed. We hypothesized that expressive gestures will contribute to interpersonal timing, as these gestures have time-varying properties that can support the communication of intentions. In addition, we are interested in studying how expressive gestures may modulate leader-follower dynamics in a dyad’s interaction. As expressive gestures may contain temporal information, it was expected that the participant who was attributed the expressive (continuous) gesture type, could communicate timing information more effectively, and in turn, would be inclined taking the leader role with respect to the participant who could not perform movements (and communicate timing information) in between tone onsets. The direction of asynchronies would then show who anticipates whom in terms of timing. We expect that gestures will help to convey musical time and anticipate the expressive timing patterns of a co-performer.

Further, we wanted to assess the roles of two additional factors, namely visual exchange of information and tempo, as these were expected to modulate the effects of hand/arm movement manipulations.

Concerning the exchange of visual information, previous research has shown that the observation of performers’ gestural behavior may contribute to listeners’ perception of musical parameters, such as pitch (Thompson, Russo, & Livingstone, 2010) and tone duration

(Schutz & Lipscomb, 2007). Also, studies have demonstrated that gestural information may affect the perception of musical expressiveness and emotion (Dahl & Friberg, 2007; Davidson, 1993, 2012; Krahé, Hahn, & Whitney, 2015; Thompson, Graham, & Russo, 2005; Vines, Krumhansl, Wanderley, Dalca, & Levitin, 2011; Vuoskoski, Thompson, Clarke, & Spence, 2014) and may substantially influence an audience's appreciation and evaluation of a music performance (Nusseck & Wanderley, 2009; Platz & Kopiez, 2012; Tsay, 2013, 2014). Interestingly, also in the context of ensemble music performance, research have indicated that gestures may be useful in signaling musical intentions (both on an expressive-emotional and music-structural level), and hence support joint musical coordination. In particular, gestures have been proven useful when auditory information or feedback is not reliable, ambiguous or absent (Bishop & Goebl, 2015; Demos, Carter, Wanderley, & Palmer, 2017; Goebl & Palmer, 2009). In these cases, gestures, and their visual exchange, may have a compensatory role in support of a reliable exchange of musical intentions. Based on this earlier research, we hypothesized in the present study that effects of movement type on interpersonal timing will be more pronounced in a performance context where musicians can visually communicate, compared to when musicians' cannot see each other's movements. Taking into account this factor is of particular interest as it allowed to dissociate between effects of movement execution and of the visual exchange of gestural information.

In addition, we investigated effects of the tempo at which musical tones needed to be produced (slow: 50 beats per minute, inter-onset interval (IOI)=1200 ms; fast: 100 beats per minute, IOI=600 ms). Previous research has indicated that there are differences in the mechanisms underlying the production of supra-second and sub-second temporal intervals. Early accounts pointed towards largely independent brain mechanisms; *cognitively-controlled* timing mechanisms – involving prefrontal and parietal cortical functions – for supra-second interval production, and *automatic* timing mechanisms – involving the motor system and cerebellum – for sub-second interval mechanisms (Lewis & Miall, 2003, 2006). However, more recent accounts have argued for a more nuanced and unified view by pointing out a timing mechanism common to both sub-second and supra-second temporal intervals, based on differentiated interactions within a distributed cerebellar-thalamic-striatal-cortical neural network (Koch, Oliveri, & Caltagirone, 2009; Petter, Lusk, Hesslow, & Meck, 2016). In line with this research, it is to be expected that the production of longer (supra-second) intervals relies more on an explicit and conscious representation of these intervals, while the production of shorter (sub-second) intervals relies more on implicit and automatic timing mechanisms. Hence, it is to be expected that the use of continuous expressive gestures may be most beneficial for interpersonal timing at slower tempi, as there is first of all more time to encode temporal cues within the performed gestures and second of all, as the perceived temporal cues can be processed more explicitly and consciously.

To summarize, the following research questions were investigated in the present study. (i) The central research question was whether the type of performed (hand and arm) movements (baseline/expressive gesture/no movement) had an effect on interpersonal timing, synchronization and leader-follower relationships. We expected that the ability to make expressive gestures would contribute positively to timing and synchronization, and would stimulate performers to take the leader role. (ii) In addition, we wanted to investigate whether the visual observation of gestures (yes/no), and tempo (slow/fast) are factors that could further modulate the effects of hand/arm movement manipulations. We expected that the positive effects of expressive gestures on the performance measures would be more pronounced when visual observation is possible, and at the slower tempo. In the following, we will present in detail the experiment that was conducted to test these hypotheses.

Method

Participants

The study was approved by the Ethical Committee of the Faculty of Arts and Literature of Ghent University, Belgium. Participants' written informed consent was obtained prior to participation.

In total, 14 dyads (28 participants) of musically trained people with at least six years of formal musical training and ensemble experience were tested. Two dyads were removed from further analysis due to technical problems during the recordings. The remaining participants were between 18 and 50 years old (M= 28), 12 of them were female, 12 of them were male and all of them were right-handed.

Materials

We provided each participant with a single tapping pad, by means of which they could trigger tones, audible through headphones. Successive taps on a pad triggered the successive notes of a melody, so there was a one-to-one relationship of taps to note onsets. The melody they produced was an excerpt from the Pachelbel Canon, plus a final note to end the sequence (see Fig. *1*). All bars contained four quarter notes, meaning that for each trial, both participants produced 33 tones (with the final tone included). Each participant got one voice (one got the upper voice, the other one the lower voice). As the two voices in this canon are of equal importance and have an equal note rate, no hierarchy of leader or follower was induced. To distinguish between the tones produced by their own taps and those of their co-performer, left-right panning of the voices was applied. The taps of the participants were recorded with a strain gauge-based pressure sensor under a tapping pad (in a Wheatstone bridge configuration), at a sample rate of 1000 Hz. These sensors were able to measure deformation of material caused by taps and quickly translate this deformation into voltage changes (amplifier' chip type INA 125). Once the voltage exceeded a predefined threshold, then a ‘drum-hit’ was recognized by an Axoloti device (a low-latency micro-controller suitable for digital audio production, http://www.axoloti.com/). The Axoloti then quickly reacted by providing a sound; that is, the next note in the musical stimulus (see Fig. *1*) to a participant.
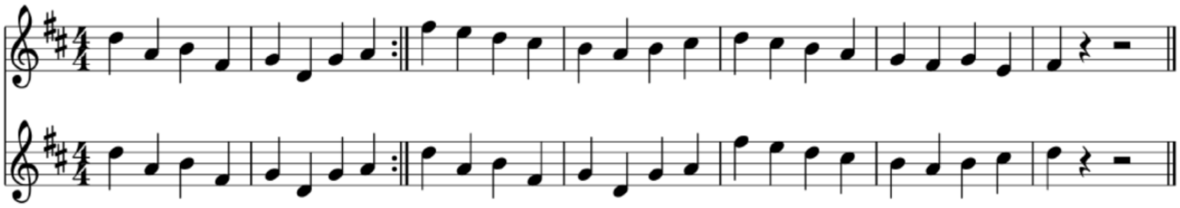
Figure 1. Musical excerpt used for the experiment. One participant triggers the notes of the upper voice, the other one the notes of the lower voice.

Design

The goal of the experiment was to investigate the role of three factors in interpersonal timing within the dyads’ performances. The first factor is Movement type. Individuals were instructed to tap (i) in a way they would do spontaneously without any further instruction, so they didn’t had to think about “how to tap” (baseline; comfortable/natural), (ii) with an additional expressive gesture of the (dominant) arm and hand in between taps, which resulted in a continuous movement trajectory between the taps (expressive gesture), or (iii) using the least movement possible by pushing on the tapping pad without lifting the finger (no movement). This leads to five combinations of performance strategies (P1(i)-P2(i), P1(ii)-P2(ii), P1(ii)-P2(iii), P1(iii)-P2(ii), P1(iii)-P2(iii))). Note that the baseline Movement type was not combined with the other movement strategies. These conditions were randomized between the participating dyads, though the combined baseline Movement type was always the first one (see Fig. *2*). The second factor was Visual contact. In one part of the experiment, participants could see each other's arms and hands, but not each other's faces (visual contact), while in the other part they could not see each other, at all (no visual contact). A third factor was Tempo. All the combinations of Movement type and Visual contact were repeated in a fast (100 beats per minute, bpm, IOI = 600 ms) and a slow tempo (50 bpm, IOI=1200 ms). In total, this sums up to 2 × 2 × 5 conditions. In each condition, participants performed three trials, making 60 trials in total.

Procedure and task

Before coming to the lab, participants were screened on musical background and dominant hand. On arriving, they got clear instructions about the experiment, and signed an informed consent form. Before starting the actual experiment, participants received enough time to learn how to trigger the melody on the tapping pad in the different performing strategies (see above (i), (ii), (iii)) and in the different tempi. Next, the actual experiment started. Each trial was initiated with a metronome that indicated the tempo with two bars of four beats. After these two bars, the metronome stopped and the participants started tapping. The main instruction they received beforehand was to play together as good as possible while performing the Pachelbel Canon, within the tempo indicated by the metronome at the beginning of the trial (in order to avoid that participants would deliberately speed up or slow down throughout a trial). As indicated in Fig. *2*, half of the dyads started with the condition where they could see each other's movements, the other half started without visual contact. When all the conditions and trials in one modality were performed, the experiment was repeated in the other modality and in the two different tempi, which were also randomized. After the participants had gone through the whole procedure, they had to fill out a questionnaire in which they were asked about their experiences during the experiment and their opinion about the task. It took approximately 2.5 hours to complete the whole experiment. Afterwards, all participants received a cd-voucher with which they could buy CDs for an amount of €15.

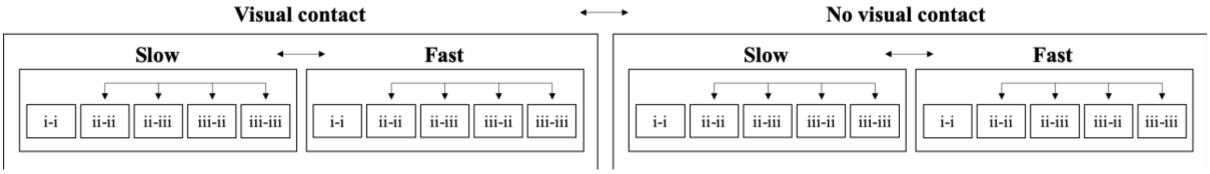

Figure 2. Schematic representation of the experimental design, and the ordering of the different factors involved: Movement type (i-i, ii-ii, ii-iii, iii-ii), Tempo (slow/fast), and Visual contact (yes/no). Arrows indicate randomizations.

Dependent variables

Interpersonal synchronization (consistency R, and asynchrony Φ). Interpersonal synchronization of dyads was assessed by looking at the phase relationship of tapping onsets throughout time. The phase of each tap of participant 2 (lower voice) relative to the closest tap of participant 1 (upper voice) was expressed as an angle between -180° and +180°; with 0° meaning that the onset of participant 2 occurred simultaneously with the onset of participant 1, a negative angle meaning that the onset of participant 2 anticipated the onset of participant 1, and a positive angle meaning that the onset of participant 2 delayed the onset of participant 1. Hence, for each performance, a distribution of 33 phase angles was obtained, which represented the relative phase differences between all corresponding onsets of the dyad. By calculating the average of the sine and cosine coordinates of all phase angles, we obtained the mean resultant vector, which has a specific length R and angle Φ (Fisher, 1995).

The resultant vector length R is related to the (circular) variance of phase angles and ranges from 0 to 1; with 0 meaning that phase angles are randomly distributed between -180° and +180°, and 1 meaning that there is a constant relative phase between a dyad’s onsets. Therefore, resultant vector length R is taken as measure for a dyad’s synchronization *consistency* (0 = minimum consistency, and 1 = maximum consistency).

The resultant vector angle Φ, ranging from -180° to +180°, is related to the average relative phase between the onsets of the dyad, and was taken as measure for a dyad’s synchronization *asynchrony*. For the analysis, two versions of this measure were considered, namely a *signed* Φ (-180° to +180°) and an *absolute* Φ (0° to 180°). The signed Φ allowed to assess leader-follower relationships (anticipation and delay) in dyads in function of the experimental conditions, while the absolute Φ provided a global measure of synchronization *accuracy*, independent of leader-follower relationships.

Timing (coefficient of variation (CV%) of inter-onset intervals). We calculated the coefficient of variation – as the standard deviation of a participant's IOIs across a trial divided by the mean IOI multiplied by 100 – to obtain a tempo-independent measure of the stability/variability of the performed inter-onset intervals (0 = no variability, to higher positive values = increased variability). As no significant differences were found between the CVs of a dyad within trials ($p > .05$), we took the average CV per trial as measure.

Analysis

Two general types of conditions can be distinguished in the experiment; a first type pertains to conditions where the members of a dyad applied the *same* Movement type (baseline, expressive gesture, or no movement), a second type to conditions where they applied a *different* Movement type (P1$_{\text{[expressive gesture]}}$–P2$_{\text{[no movement]}}$, and P1$_{\text{[no movement]}}$–P2$_{\text{[expressive gesture]}}$). The data set was split accordingly, and different analyses were performed on the two resulting data subsets (for a detailed description of the different features, see section Dependent variables). For the *same* Movement type subset, the focus was on the synchronization features consistency R and absolute asynchrony Φ, and on the timing feature CV. For each of these features, a 3×2×2 repeated measures ANOVA was applied with Movement type (baseline/expressive gesture/no movement), Visual contact (yes/no), and Tempo (fast/slow) as within-subjects factors. For the *different* Movement type subset, the focus was on leader-follower relationships, which were quantified in the synchronization feature relative asynchrony Φ. For this feature, a 2×2×2 repeated measures ANOVA was applied with Movement type (P1$_{\text{[expressive gesture]}}$–P2$_{\text{[no movement]}}$/ P1$_{\text{[no movement]}}$–P2$_{\text{[expressive gesture]}}$)), Visual contact (yes/no), and Tempo (fast/slow) as within-subjects factors. For each subject, we calculated the average value of each dependent variable across the three trials they performed in each condition. These values were then used to perform the respective repeated measures ANOVAs.

For all tests, a significance level of .05 was applied. For the repeated measures ANOVAs, Mauchly's tests of sphericity were used to check the assumption of sphericity. In the case of non-sphericity, effects were Greenhouse-Geisser corrected. Post hoc tests to follow up on main and interaction effects were conducted as *t*-tests, with significance levels corrected for multiple comparisons using the Bonferroni method.

Results

Same Movement type

Interpersonal synchronization: Absolute asynchrony Φ (Figure *3*)**.** The 3×2×2 repeated measures ANOVA revealed a main effect of Movement type, *F*(1.098,12)=6.083, *p*<.05, and Visual contact, *F*(1,12)=8.38, *p*<.05, while no interaction effects were observed. Post hoc pairwise comparisons showed that the no movement Movement type (M=11.12, SE=1.46) led to significantly higher absolute asynchronies compared to the baseline Movement type (M=6.69, SE=0.80), *t*(11)=3.43, *p*<.05, *g*=0.32, and to the expressive gesture Movement type (M=5.76, SE=0.73), *t*(11)=3.85, *p*<.05, *g*=0.39. In addition, we found that the absolute asynchrony was significantly higher when people could not see each other (M=9.28, SE=1.53) compared to when they could see each other (M=6.43, SE=1.04), *t*(11) = 2.90, *p*<.05, *g*=0.18.

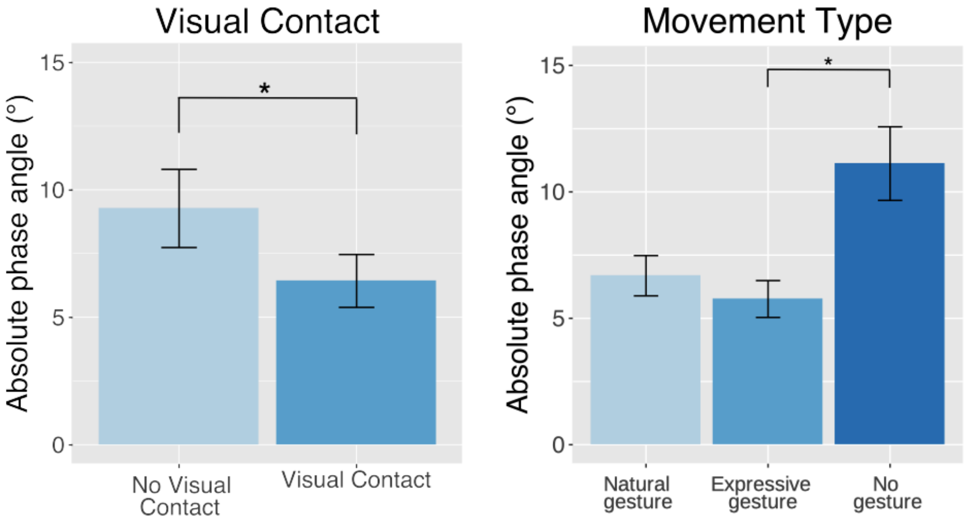

Figure 3. Mean absolute asynchrony Φ (in degrees) by: (left) Visual contact (yes/no) and (right) Movement type (baseline/expressive gesture/no movement). Error bars represent the standard error of the mean (* p<.05).

Interpersonal synchronization: Consistency R (Figure *4*)**.** The 3×2×2 repeated measures ANOVA yielded a main effect of Movement type, *F*(1.04,12)=8.65, *p*<.05, Visual contact, *F*(1,11) = 7.98, *p*<.05, and Tempo, *F*(1,12) = 125.55, *p*<.001. Also, we found a significant interaction between Movement type and Tempo, $F(1.487,12)=5.583$, $p<.05$. Post hoc tests indicated a general decrease in synchronization consistency when no movements could be performed, and at the slower tempo. Interestingly, post hoc tests indicated an increase in synchronization consistency when expressive gestures are applied (M=0.9626, SE=0.0023) compared to the baseline Movement type condition (M=0.9554, SE=0.0033), but only for the slow Tempo, $t(11)=3.72$, $p<.05$, $g=0.36$.

Although no significant interaction effect was found between Movement type and Visual contact ($F(1.227,12)=2.571$, $p=0.125$), a similar pattern can be found; post hoc pairwise comparisons revealed an increase in synchronization consistency when expressive gestures are applied (M=0.9734, SE=0.0013) compared to the baseline Movement type condition (M=0.9633, SE=0.0030), but only when there is visual contact, $t(11)= 3.72$, $p<.05$, $g=0.36$.

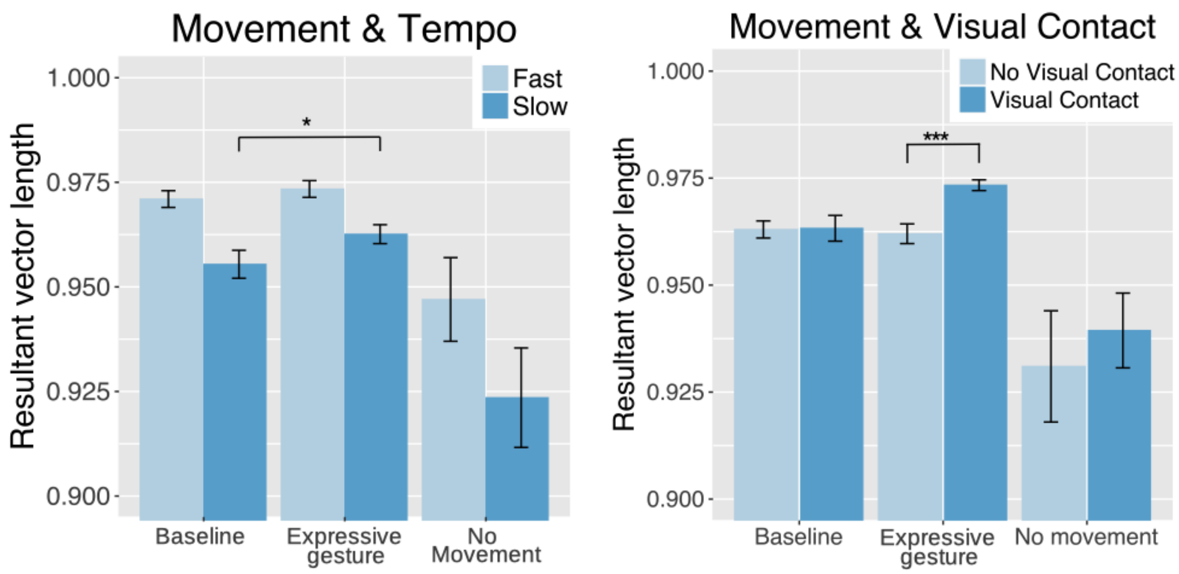

Figure 4. Mean synchronization consistency R by: (left) Movement type (baseline/expressive gesture/no movement) and Tempo (slow/fast), and (right) Movement type (baseline/expressive gesture/no movement) and Visual contact (yes/no). Error bars represent the standard error of the mean. Only significant posthoc results that are responsible for the interaction effects are indicated by asterisks (* p<.05 and *** p<.001).

Timing: Coefficient of variation (CV%) (Figure 5)**.** The 3×2×2 repeated measures ANOVA yielded a significant main effect of Movement type, $F(1.612,12)=15.662$, $p<.001$, and a significant interaction effect between Movement type and Visual contact, $F(1.89,12)=8.60$, $p<.05$. In general, it was found that the CV% was higher with the no movement Movement type, compared to the baseline Movement type, and the expressive gesture Movement type. Concerning the interaction effect between Movement type and Visual contact, post hoc tests indicated that it was driven by the finding that the expressive movement Movement type (M=3.987, SE=0.158) lowered the CV compared to the baseline Movement type (M=4.737, SE=0.205), but only when there was visual contact, $t(11)= 6.37$, $p<.001$, $g=0.53$.

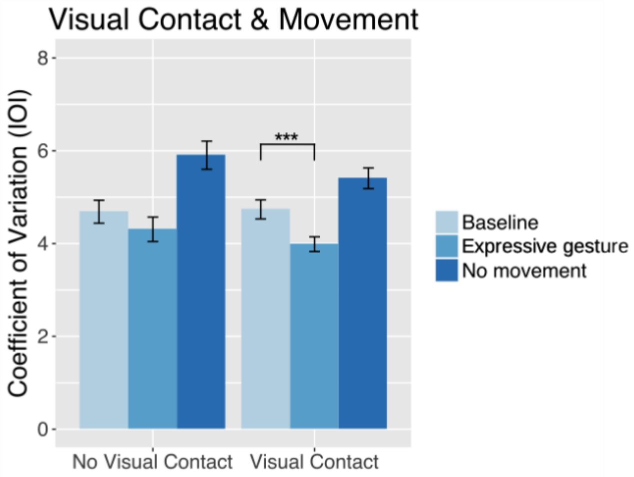

Figure 5. Mean coefficient of variation of the IOIs (CV%) by Movement type (baseline/expressive gesture/no movement) and Visual contact (yes/no). Error bars represent the standard error of the mean. Only significant posthoc results that are responsible for the interaction effects are indicated by asterisks (*** p<.001).

Different Movement type

Leader-follower relationship: Signed asynchrony Φ (Figure 6**).** The 2×2×2 repeated measures ANOVA revealed a main effect of Movement type combination, $F(1,12)=29.47$, $p<.001$, as well as significant interaction effects between Movement type and Tempo, $F(1,12)=8.31$, $p<.05$, Movement type and Visual contact, $F(1,12)=7.22$, $p<.05$, and between Movement type, Tempo and Visual contact, $F(1,12)=10.13$, $p<0.01$. The main effect of Movement type indicated that within dyads, participants that performed the expressive gesture Movement type were, on average, ahead of (cf. leading) participants that performed the no movement Movement type. Additional post hoc tests that were conducted to better understand the three-way interaction pointed out that this effect was more pronounced at the slow Tempo and when Visual contact was allowed (M=-10.41, SE=3.18 versus M=10.13, SE=2.99, t(11), p<.001, g=0.55) compared to the other conditions (slow Tempo/no Visual contact, fast Tempo/Visual contact, and fast Tempo/Visual contact), where effects were significant only at the .05 level.

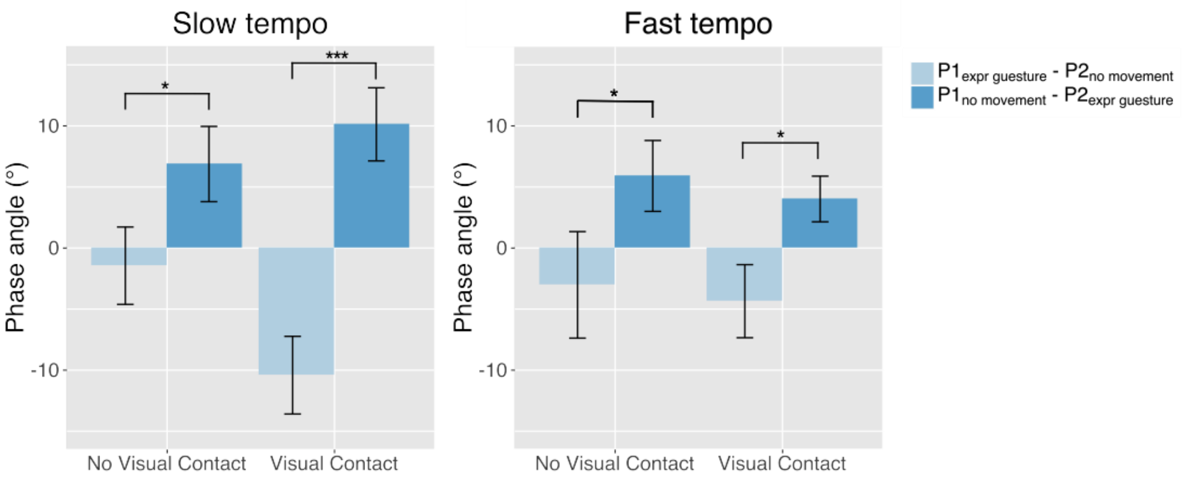

Figure 6. Mean signed asynchrony Φ (in degrees) by Movement type combination ($P1_{\text{[expressive gesture]}}$–$P2_{\text{[no movement]}}$/$P1_{\text{[no movement]}}$–$P2_{\text{[expressive gesture]}}$), Tempo (slow/fast), and Visual contact (yes/no). A negative Φ means that P1 is ahead of P2 while, reversely, a positive Φ means that P1 is lagging behind P2. Error bars represent the standard error of the mean. Only significant posthoc results that are responsible for the interaction effects are indicated by asterisks (* p<.05 and *** p<.001).

Discussion

In this study, the primary aim was to investigate the role of body movements in musical communication, in terms of interpersonal synchronization, timing control, and leader-follower relationships. We experimentally controlled for the use of hand and arm movements in a musical task, in which dyads were asked to synchronously tap out a melody. Next to their comfortable/natural way of tapping, we instructed participants to either perform pronounced expressive hand and arm gestures in between successive taps, or to restrict from any overt body movement. In general, our results show that these different movement types can indeed influence the musical communication process. In addition to the study of movement type, we investigated the role of tempo and visual contact in the communication process.

Compared to when musicians tap in a spontaneous natural way, we found that a restriction of body movements resulted in a general lower accuracy (= higher absolute asynchrony Φ) and consistency (= smaller R) of interpersonal synchronization, and in a significant increase of the variability of produced inter-onset interval durations. Effects related to the use of expressive gestures were more nuanced, as they were modulated by the factors tempo and/or visual contact. A first finding was that tapping using expressive hand and arm gestures did not change synchronization accuracy (= asynchrony) with respect to natural tapping. However, we found that interpersonal synchronization was in general more consistent (that is, tone onsets had a more consistent phase relationship) when dyads performed expressive gestures in between tapping onsets, but only when there was an exchange of visual information, and at slower tempi. A related observation was made for the variability/stability of produced inter-onset interval durations. Again, it was found that tapping using expressive gestures could improve stability of performed inter-onset interval durations, but only under the condition of an exchange of visual information.

These results suggest that the use of expressive gestures of musicians may have beneficial effects on interpersonal coordination, but that it requires sufficient time to actually execute the gestures, and that there is an exchange of visual information. The explanation of this finding is supposedly linked to the time-varying perceptual properties of both the external visual input of the partner's body movements, and of the internal somatosensory and proprioceptive feedback of one's own movements. Previous research has indicated that humans are highly receptive to these time-varying sensorimotor properties for (emergent) time-keeping purposes (Coull & Droit-Volet, 2018; Maes, Giacofci, et al., 2015; Motanis et al., 2018; Schwartze et al., 2012). This may allow humans to accurately anticipate the occurrence of specific sensory events, and hence support interpersonal synchronization and timing. A critical question though is about the principles and mechanisms underlying this anticipatory behavior. Currently, a central debate in cognitive (neuro)science and cybernetics revolves around the concepts of 'weak anticipation' and 'strong anticipation' (Dubois, 2003; Stepp & Turvey, 2010), referring respectively to a computational/inferential and behavioral dynamics/ecological approach to human motor control (Gallagher & Allen, 2018; Nasuto & Hayashi, 2016, 2019; Warren, 2006). Weak anticipatory behavior involves the construction of predictions of the environment's future states based on internal models of the environment (cf. forward models). Based on predictions and error-correction mechanisms, agents can plan their motor actions and adapt flexibly to changing environmental conditions. According to this model-based approach, the connection between the agent and its environment is not direct, but mediated by a mental model. Alternatively, strong anticipatory behavior is fundamentally rooted in the direct (action-perception) coupling of an agent with its environment, forming a dynamical system that is driven towards stable relationships between its components (as for example described by the relative phase in periodic behaviors). In this regard, anticipatory behavior in joint motor coordination may be the outcome of a moment-to-moment alignment of actions to perceptual information (Kelso, 1995; Wilson & Golonka, 2013), of the introduction of small time-delayed self-feedback (Demos, Layeghi, Wanderley, & Palmer, 2019; Roman, Washburn, Large, Chafe, & Fujioka, 2019; Stepp & Turvey, 2010; Washburn, Kallen, Coey, Shockley, & Richardson, 2015), or of long-term complexity matching of joint behavior (Fine, Likens, Amazeen, & Amazeen, 2015; Marmelat & Delignières, 2012). These studies provide valuable mathematical methods for the analysis and modeling of complexity and fluctuation structures in joint behavior, with a central focus on nonlinear, time-varying characteristics of this behavior. These methods typically rely on fairly large amounts of data, which were not available in the present study. However, our experimental paradigm, as well as similar paradigms in the domain of music interaction, are valuable as they provide excellent scenarios for studying the mechanisms of weak and strong anticipation. For that purpose, scenarios of music interaction lend themselves ideally in the way they allow to control, manipulate, and perturb variables that relate with personal background, sensory coupling, musical properties such as tempo, and so forth. The combination of versatile musical environments, and nonlinear time-series analysis methods holds great value in the further study of embodied human interaction and its underlying control mechanisms.

The main contribution of our study pertains to the role of body movement – in particular of expressive, ancillary gestures (Cadoz & Wanderley, 2000) – in musical communication. Earlier research demonstrated that visual information of performers' gestures may influence listener's perception of musical parameters (Schutz & Lipscomb, 2007; Thompson et al., 2010), as well as of musical emotion and expressiveness (Dahl & Friberg, 2007; Davidson, 1993, 2012; Krahé et al., 2015; Thompson et al., 2005; Vines et al., 2011; Vuoskoski et al., 2014). Interestingly, research showed that performers' (ancillary) gestures may equally support the communication of musical intentions to co-performers, both on the expressive-emotional and music-structural level. In particular, in cases where auditory information is absent or not reliable, performers may rely on visual cues to support musical timing (Bishop & Goebl, 2015; Demos et al., 2017; Goebl & Palmer, 2009). The current study supports this line of research by showing improved joint timing when expressive gestures are allowed. In addition to the exchange of visual information, beneficial effects on timing have also been related to the execution itself of ancillary gestures. In studies on regular interval timing, a distinction is typically made between two possible timing mechanisms, namely an event-based timing mechanism, when discrete movements are produced, and an emergent timing mechanism, when smooth continuous movements are produced (Delignières et al., 2004; Robertson et al., 1999; Torre & Balasubramaniam, 2009). In the former, timing is regulated by internal timekeeping mechanisms employing cognitive resources, while in the latter, time and timekeeping are emergent (inherent) properties of the movement and the related feedback itself. Research has shown that emergent timing may be more robust in highly cognitive demanding situations, such as music performance, as timekeeping is “outsourced” to the motor system (Maes, Giacofci, et al., 2015; Maes, Wanderley, et al., 2015). Also in this regard, the use of ancillary gestures may provide a compensatory strategy to optimize joint timing in music performance. To distinguish between the effects of visual exchange of gestural information and movement execution itself, we controlled for visual contact between musicians. The fact that the use of natural/comfortable and expressive gestures without visual contact led to a more consistent interpersonal timing and a lower IOI variability, compared to when no movement was allowed, suggests that movement execution itself – and underlying emerging timing mechanisms – positively influence interpersonal timing. Although there was no difference between the use of natural/comfortable gestures and expressive gestures when visual contact was absent, we did found a higher consistency and lower variability for expressive gestures when visual contact was allowed. This finding indicates that the exchange of visual information inherent to expressive gestures may have an additional beneficial effect on interpersonal timing, on top of the effect of movement execution itself. It must be noted that our experiment focused on musicians (> six years of formal music training). It would be of interest to investigate further whether these results are generalizable to broader populations of non-musicians. We would expect that effects in musicians are more pronounced as musicians have presumably developed specific strategies for using multimodal information about their own and/or other’s gestures for timing control through musical practice and experience.

An important additional aim of the current study was to investigate leader-follower relationships in musical dyads and the way these depend on the produced movement types. In previous research, it had been shown that assigned leader-follower roles are reflected in musicians’ body movements and their coordination (Goebl & Palmer, 2009; Keller & Appel, 2010). For instance, Keller and Appel (2010) demonstrated that sound synchrony of duetting pianists increased when the body sway of the leader (primo player) preceded the body sway of the follower (secondo player). In line with this finding, Goebl and Palmer (2009) had found that head movements of pianists that were assigned the leader role preceded those of the follower. Interestingly, it was also found that leader pianists raised fingers higher than the follower pianists. Typical for these studies was that leader-follower roles are assigned beforehand, and bodily coordination was studied as a result. In the current study, this paradigm was reversed as we assigned specific body movements to our participants beforehand (baseline, no movement, expressive gesture), and studied the effect on the leader-follower relationship within the dyads. An important finding of the study was that assigned body movements indeed had an effect. In general, the produced onsets of musicians that were asked to perform expressive gestures in between onsets were preceding their partner who could not produce any movement in between produced onsets. This supports the idea that musicians who produce expressive gestures are inclined to take the leader role and vice versa. Similar to the interpersonal timing measures consistency and variability, this effect was modulated by the factors visual contact and tempo. It was shown that the effect was most pronounced when musicians’ could see each other, and when they had enough time to effectively produce expressive gestures (slow tempo). Again, this could be explained by the presence of temporal cues in the expressive gestures. By allowing performers to gesture expressively, they are empowered to communicate temporal information, and hence, to function as a temporal reference for interpersonal coordination. These results are in line with other research studying the role of gestural communication in leader-follower dynamics. For instance, Gerpott and colleagues (2018)found that emergent leaders (in initially leaderless groups) exhibit more active body language (in particular of arms and shoulders) and less passive facial expressions. Also, Talley and Temple (2015) point towards the role of the type of hand gesture of leaders in establishing a certain relationship quality (i.e., positive hand gestures create immediacy and attraction between leaders and followers). The results of the present study contribute to this line of research within the domain of music interaction. They are promising and would benefit from follow-up studies that investigate into more detail the relationship between bodily gestures and leader-follower roles, and the influence thereon of person-, context-, and task-related factors and constraints. Particularly interesting would be to investigate this relationship in “real-life” musical contexts, such as in (jazz) ensemble improvisations, which are often characterized by a rhythm section (bass, drums) that supports a more expressive lead soloist, but where there is equally room for shifting roles through solo improvisations of any instrument.

Conclusion

The present study investigated the role of body movement (natural movement/no movement/expressive movement), visual contact (yes/no) and tempo (slow/fast) in joint musical performances (duet, melody tapping task), in terms of timing control, synchronization and leader-follower dynamics. The results show that the restriction of body movements has detrimental effects on interpersonal timing and synchronization (lower synchronization accuracy and consistency, and higher inter-onset variability). In contrast, the use of expressive gestures led to a higher synchronization consistency, compared to natural gestures, but only when there was visual contact and at the slower tempo. The same finding was observed for the inter-onset variability. Finally, results indicated that the type of body movements performed by the members of a dyad can modulate leader-follower dynamics. In general, it was found that people that performed expressive body movements tended to take the leader role in the interaction. These results suggest that expressive body movements contain time-varying cues (internal and external), that facilitate anticipation mechanisms and, in turn, may improve interpersonal timing.

Acknowledgements

This study was conducted with the support of the FWO-project ‘Foundations of expressive timing control in music’. Special thanks to Ivan Schepers for his technical support during the experiments.

References

Allman, M. J., Teki, S., Griffiths, T. D., & Meck, W. H. (2014). Properties of the internal clock: first- and second-order principles of subjective time. *Annual Review of Psychology, 65*, 743–771.

Bishop, L., & Goebl, W. (2015). When they listen and when they watch: Pianists' use of nonverbal audio and visual cues during duet performance. *Musicae Scientiae, 19*(1), 84–110.

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

Caruso, G., Coorevits, E., Nijs, L., & Leman, M. (2016). Gestures in contemporary music performance: a method to assist the performer's artistic process. *Contemporary Music Review*, *35*(4–5), 402–422.

Cassenti, D. N. (2011). The intrinsic link between motor behavior and temporal cognition. *New Ideas in Psychology*, *29*, 72–79.

Coull, J. T., & Droit-Volet, S. (2018). Explicit understanding of duration develops implicitly through action. *Trends in Cognitive Sciences*, *22*(10), 923–937.

Dahl, S., & Friberg, A. (2004). Expressiveness of musician's body movements in performances on marimba. In *Lecture Notes in Computer Science* (pp. 479–486). Berlin, Heidelberg: Springer-Verlag.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception: An Interdisciplinary Journal*, *24*(5), 433–454.

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

Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychology of Music*, *40*(5), 595–633.

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

Demos, A. P., Carter, D. J., Wanderley, M. M., & Palmer, C. (2017). The unresponsive partner: Roles of social status, auditory feedback, and animacy in coordination of joint music performance. *Frontiers in Psychology*, *8*(149).

Demos, A. P., Chaffin, R., & Logan, T. (2018). Musicians body sway embodies musical structure and expression: A recurrence-based approach. *Musicae Scientiae*, *22*(2), 244–263.

Demos, A. P., Layeghi, H., Wanderley, M. M., & Palmer, C. (2019). Staying together: A bidirectional delay–coupled approach to joint action. *Cognitive Science*, *43*(8), e12766.

Dubois, D. M. (2003). Mathematical foundations of discrete and functional systems with strong and weak anticipations. In M. V. Butz, O. Sigaud, & P. Gérard (Eds.), *Anticipatory behavior in adaptive learning systems* (pp. 110–132). Berlin, Germany: Springer.

Fine, J. M., Likens, A. D., Amazeen, E. L., & Amazeen, P. G. (2015). Emergent complexity matching in interpersonal coordination: Local dynamics and global variability. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 723–737.

Fisher, N. I. (1995). *Statistical analysis of circular data*. New York, NY: Cambridge University Press.

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

Gerpott, F. H., Lehmann-Willenbrock, N., Silvis, J. D., & Van Vugt, M. (2018). In the eye of the beholder? An eye-tracking experiment on emergent leadership in team interactions. *The Leadership Quarterly*, *29*(4), 523–532.

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception: An Interdisciplinary Journal*, *26*(5), 427–438.

Juchniewicz, J. (2008). The influence of physical movement on the perception of musical performance. *Psychology of Music*, *36*(4), 417–427.

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron*, *53*(3), 427–438.

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, UK: Oxford University Press.

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception: An Interdisciplinary Journal*, *28*(1), 27–46.

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

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 B: Biological Sciences*, *364*(1525), 1907–1918.

Kononowicz, T. W., van Rijn, H., & Meck, W. H. (2016). Timing and time perception: A critical review of neural timing signatures before, during, and after the to-be-timed Interval. In J. Wixted & J. Serences (Eds.), *Stevens Handbook of Experimental Psychology and Cognitive Neuroscience (4th ed.)* (pp. 1–35). New York, NY: Wiley.

Krahé, C., Hahn, U., & Whitney, K. (2015). Is seeing (musical) believing? The eye versus the ear in emotional responses to music. *Psychology of Music*, *43*(1), 140–148.

Krauss, R., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: what do conversational hand gestures tell us?, *28*, 389–450.

Leman, M. (2016). *The expressive moment: how interaction (with music) shapes human empowerment*. Cambridge, MA: MIT Press.

Lesaffre, M., Maes, P.-J., & Leman, M. (2017). *The Routledge companion to embodied music interaction*. New York, NY: Routledge.

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Current Opinion in Neurobiology*, *13*(2), 250–255.

Lewis, P. A., & Miall, R. C. (2006). Remembering the time: a continuous clock. *Trends in Cognitive Sciences*, *10*(9), 401–406. https://doi.org/10.1016/j.tics.2006.07.006

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. https://doi.org/10.1007/s00221-014-4108-5

Marmelat, V., & Delignières, D. (2012). Strong anticipation: Complexity matching in interpersonal coordination. *Experimental Brain Research*, *222*, 137–148.

Miell, D., MacDonald, R., & Hargreaves, D. J. (2005). *Musical communication*. (Oxford University Press, Ed.). Oxford, UK.

Motanis, H., Seay, M. J., & Buonomano, D. V. (2018). Short-term synatic plasticity as a mechanism for sensory timing. *Trends in Neurosciences*, *41*(10), 701–711.

Nasuto, S. J., & Hayashi, Y. (2016). Anticipation: Beyond synthetic biology and cognitive robotics. *Biosystems*, *148*, 22–31.

Nasuto, S. J., & Hayashi, Y. (2019). Anticipation in neurocybernetics. In R. Poli (Ed.), *Handbook of anticipation: Theoretical and applied aspects of the use of future in decision making* (pp. 1–36). Berlin, Germany: Springer.

Nusseck, M., & Wanderley, M. M. (2009). Music and motion: How music-related ancillary body movements contribute to the experience of music. *Music Perception: An Interdisciplinary Journal*, *26*(4), 335–353.

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.

Platz, F., & Kopiez, R. (2012). When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception: An Interdisciplinary Journal*, *30*(1), 71–83.

Remington, E. D., Egger, S. W., Narain, D., Wang, J., & Jazayeri, M. (2018). A dynamical systems perspective on flexible motor timing. *Trends in Cognitive Sciences*, *22*(10), 938–952.

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.

Roman, I. R., Washburn, A., Large, E. W., Chafe, C., & Fujioka, T. (2019). Delayed feedback embedded in perception- action coordination cycles results in anticipation behavior during synchronized rhythmic action: a dynamical systems approach. *PLoS ONE-Computational Biology*, *15*(10), e1007371.

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*(6), 888–897.

Schwartze, M., Tavano, A., Schröger, E., & Kotz, S. A. (2012). Temporal aspects of prediction in audition: Cortical and subcortical neural mechanisms. *International Journal of Psychophysiology*, *83*(2), 200–207.

Stepp, N., & Turvey, M. T. (2010). On strong anticipation. *Cognitive Systems Research*, *11*(2), 148–164.

Talley, L., & Temple, S. (2015). How leaders influence followers through the use of nonverbal communication. *Leadership & Organization Development Journal*, *36*(1), 69–80.

Teixeira, E. C. F., Loureiro, M. A., Wanderley, M. M., & Yehia, H. C. (2015). Motion analysis of clarinet performers. *Journal of New Music Research*, *44*(2), 97–111.

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, *156*(1–4), 203–227.

Thompson, W. F., Russo, F. A., & Livingstone, S. R. (2010). Facial expressions of singers influence perceived pitch relations. *Psychonomic Bulletin & Review*, *17*(3), 317–322.

Torre, K., & Balasubramaniam, R. (2009). Two different processes for sensorimotor synchronization in continuous and discontinuous rhythmic movements. *Experimental Brain Research*, *199*(2), 157–166.

Tsay, C. J. (2013). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences*, *110*(36), 14580–14585.

Tsay, C. J. (2014). The vision heuristic: Judging music ensembles by sight alone. *Organizational Behavior and Human Decision Processes*, *124*(1), 24–33.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*(2), 157–170.

Vuoskoski, J. K., Thompson, M. R., Clarke, E. F., & Spence, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception, & Psychophysics*, *76*(2), 591–604.

Warren, W. H. (2006). The dynamics of perception and action. *Psychological Review*, *113*(2), 358–389.

Washburn, A., Kallen, R. W., Coey, C. A., Shockley, K., & Richardson, M. J. (2015). Harmony from chaos? Perceptual-motor delays enhance behavioral anticipation in social interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4), 1166–1177.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, *4*(58).