

Body movement enhances the extraction of temporal structures in auditory sequences

Yi-Huang Su · Ernst Pöppel

Received: 14 January 2011 / Accepted: 23 May 2011 / Published online: 22 June 2011
© Springer-Verlag 2011

Abstract Auditory and motor systems interact in processing auditory rhythms. This study investigated the effect of intuitive body movement, such as head nodding or foot tapping, on listeners' ability to entrain to the pulse of an auditory sequence. A pulse-finding task was employed using an isochronous sequence of tones in which tones were omitted at pseudorandom positions. Musicians and non-musicians identified their subjectively fitting pulse either using periodic body movement or through listening only. The identified pulse was measured subsequently by finger tapping. Movement appeared to assist pulse extraction especially for non-musicians. The chosen pulse tempi tended to be faster with movement. Additionally, movement led to higher synchronization stabilities of the produced pulse along the sequence, regardless of musical training. These findings demonstrated the facilitatory role of body movement in entraining to auditory rhythms and its interaction with musical training.

Introduction

Imagine sitting in a live jazz concert: the listeners feel immersed in the music and start intuitively to move along with different parts of their body. The same is seen in the musicians: they automatically move their head or tap their feet in a periodic manner while playing some rhythmically engaging passages. Even in a classical concert where excessive gestures are not encouraged by convention, we often observe musicians rhythmically moving some parts of their body that are not engaged in playing the instruments, along with the music. Moving one's body periodically to the music, be it foot tapping or head nodding, is a frequent manifestation in listeners as well as in performing musicians. It presents a common example of audio–motor crosstalk in experiencing musical rhythms, and poses an interesting question about the nature of rhythm perception: do we move only because we react to the rhythm we hear, or does the movement itself contribute to the process of hearing the rhythm?

Perception and action are believed to share common representational mechanisms through which they interact (Prinz, 1997). Neurophysiological studies on rhythm perception have also concurred that processing auditory rhythms engages both auditory and motor areas of the brain (Chen, Penhune, & Zatorre, 2008b; Grahn & Brett, 2007; Bengtsson et al., 2009), and the motor system can be crucial in this act (Grahn & Brett, 2009). A direct behavioral link between 'hearing the rhythm' and 'moving the body' has been established in two studies where the interpretation of the same auditory rhythm was shaped by different patterns in which the listeners bounced their body (Phillips-Silver & Trainor, 2005, 2007). Given this audio–motor interplay, the present study further pursued the hypothesis that the use of body movement is not merely a reaction to

Electronic supplementary material The online version of this article (doi:10.1007/s00426-011-0346-3) contains supplementary material, which is available to authorized users.

Y.-H. Su (✉)
Institute of Medical Psychology,
Ludwig-Maximilians-University Munich,
Goethestr. 31, 80336 Munich, Germany
e-mail: jasmine.su@med.uni-muenchen.de

E. Pöppel
Human Science Center,
Ludwig-Maximilians-University Munich,
Munich, Germany

hearing rhythmic input, but could actively *assist* the processing of temporal structures in the auditory events.

The temporal structure of interest is the pulse.¹ In music, pulse is defined as a series of stable and undifferentiated *psychological events* arising endogenously in response to musical rhythms (Cooper & Meyer, 1960). Rather than a physical property of the stimuli, the pulse is a subjectively experienced isochrony. We chose to target this process because perceiving such isochrony is the basic principle of human entrainment to auditory stimuli (Merker, Madison, & Eckerdal, 2009). It serves as the subjective referent by which we experience complex temporal relations in musical rhythms (Large, 2008), and corresponds to the felt tempo. A relevant *sensorimotor theory* of temporal tracking has been proposed by Todd and colleagues (Todd, 1999; Todd, Lee, & O’Boyle, 2002). It incorporates importantly an internal motor representation of the body along with the sensory input and the motor output as coordinated mechanisms of tracking and synchronizing to an isochronous pulse. This theory emphasizes the relation between the embodied motor process and the percept of isochronous structure in the rhythm, which constitutes the central idea of the present study.

This idea is further supported by several complementary findings: human’s ability to perceive a regular pulse seems innate (Winkler, Háden, Ladinig, Sziller, & Honing, 2009) and is proposed to arise from an endogenous neural oscillation entraining to rhythmic stimuli (Large & Snyder, 2009). Interestingly, such oscillations are also associated with motor tasks (Fujioka, Trainor, Large, & Ross, 2009; Salenius & Hari, 2003). Consistently, the premotor activation in the brain is enhanced by listening to rhythms at one’s preferred tempo (Kornysheva, von Cramon, Jacobsen, & Schubotz, 2010), and human listeners’ preferred tempo in music (Moelants, 2002) corresponds to the preferred frequency (~ 2 Hz) in locomotion (Macdougall & Moore, 2005). These findings seem to suggest that the percept of regular pulse, which also defines the tempo, could entail a motor component: forming a pulse by means of entrainment may require a motor process, at least internally (Grahn & Rowe, 2009).

In light of the audio–motor interaction, this study investigated whether an *external motor process* such as moving one’s body to the rhythm—an intuitive behavior for many

people—could actually assist the extraction of its pulse by means of facilitated entrainment. A pulse-finding task was employed where a tone sequence of no particular metrical or accent structure was presented continuously. The structure of the sequence could be seen as underlyingly isochronous (based on the nominal stimulus tempo) with tones omitted at pseudorandomly chosen positions, similar to one of the sequence types employed in the study of Patel, Iversen, Chen, and Repp (2005, sequence type 7: I-WM). The listeners first established their subjectively fitting pulse either using preferred periodic body movement or through listening only, and then produced their identified pulse by finger tapping. Critically, in establishing the pulse, the engagement of body movement was expected to initiate an overt motor activity *while* the listener searched for the regular pulse to which to entrain. Rather than a mere manifestation of the already established pulse, the movement should be adopted from the beginning of the entrainment process in order to assist finding a stable pulse. That is, the movement could be initially out of synchrony with any pulse period, but would gradually (or quickly, as we hypothesized the presence of movement to be facilitatory for the process) synchronize to the pulse of the sequence. Without movement, such entrainment process would have to be internally generated and might require more a cognitive strategy to analyze the temporal structure of the sequence.

As such, movement was expected to assist the listeners to more easily ‘tune in’ to the temporal information and to establish their pulse at one of several possible (sub)-harmonic frequencies (Large & Snyder, 2009). The stability of one’s tapped pulse, preceded and accompanied by movement, was also expected to be higher as a result of enhanced sensorimotor integration (Chen, Penhune, & Zatorre, 2009). In addition, the movement effect was compared between musicians and non-musicians. Musicians were expected to be able to analyze the structure of the sequence (Chen, Penhune, & Zatorre, 2008a) and to generate the pulse internally (Grahn & Rowe, 2009) even in the absence of body movement, while non-musicians were expected to depend more on such overt motor activity to discover and entrain to the pulse.

Method

Participants

20 young, healthy, right-handed participants (range 20–35 years, mean age 24, SD 3.8) participated in the experiment via on-campus recruitment, and received payment in return. Ten were musically-trained (amateurs with at least 8 years training, 6 pianists and 4 violinists, 3 of whom were amateur orchestra members); the other ten had

¹ The terms pulse and beat are often used interchangeably in a musical context. However, beat implies a defined metrical organization based on the alternating strong and weak accentuation (Cooper & Meyer, 1960), which involves the perceptual grouping of pulse, e.g. groups of two or four as in a duple meter, or groups of three as in a waltz meter. Pulse itself, on the other hand, is not confined by metrical specifications; it exists as long as the isochrony is felt by the listener, and is generalizable in processing rhythms across different cultures and musical genres. Therefore, we prefer to use the term pulse here.

never received formal musical training. All reported music listening as leisure activities.

Stimuli and materials

Auditory stimuli were generated as wave files by the music software Logic 8 Express (Apple Inc. California) using a synthesized woodblock sound (as the instrument “clave”), with 42 ms tone duration. Each wave file was a ‘building block’ consisting of five isochronous time points; each point could be either occupied by a tone or not, resulting in 31 possible building blocks, excluding the block with no tone. The blocks were generated at six tempi: 60, 90, 120, 150, 180, and 210 beats per minute (BPM), corresponding to the shortest inter-onset interval (IOI) of 1000, 666.7, 500, 400, 333.3, and 285.7 ms. The experiment was carried out in Matlab[®] 2009a (Mathworks) using Psychophysics Toolbox extensions version 3 (Brainard, 1997), running on a MacBook Pro laptop computer. Participants sat in a comfortably-lit sound-proof room. The sound was delivered via headphones (Philips SBC HS900).

Design and procedure

The participants were divided into four groups based on the instructed task strategy and musical training: (1) movement, musicians, (2) movement, non-musicians, (3) no-movement, musicians; and (4) no-movement, non-musicians. There were five participants in each group.

Instruction

Pulse in the task was explained as (translated from German) ‘the successive time points with equal intervals which are subjectively fitting tactus to the ongoing tone sequence. It should be as stable as possible throughout the trial and should not alternate between different levels.’ Besides the verbal explanation, in order to ensure the same understanding of the task from both musicians and non-musicians, an instructional demonstration was carried out for each participant

prior to the experiment. The demonstration differed between movement and no-movement groups as follows.

For the movement groups, the experimenter played an example tone sequence (as would be played in a real experimental trial) and demonstrated behaviorally with foot tapping where the pulse should be temporally. The experimenter demonstrated two different possibilities of the fitting pulse, one being the subharmonic of the other (i.e. twice as slow), which exemplified the notion of ‘different pulse levels’.

For the no-movement groups, the experimenter played an example tone sequence which was accompanied by an additional sequence of low tones illustrating pulse. Two examples of such a combined sequence were played, the low tones in each example demonstrating a different (but fitting) pulse level. Crucially, there was no mention of the link between the present task and the everyday behavior such as ‘tapping one’s feet to the music’.

Experiment

In the beginning of each trial, 31 building blocks of one tempo were strung up in a randomized order, with the rule that all blocks were selected once, and the very first time point was occupied by a tone. The 31 concatenated blocks made up a long sequence that was looped within a trial. Each trial consisted of two consecutive phases: (a) pulse extraction, and (b) pulse production (Fig. 1).

Prior to the instruction, each participant in the *movement* groups had been asked to report their preferred means of body movement when they listened to music. In the extraction phase, they were requested to use their reported preferred movement (e.g. foot tapping, head nodding) from the start of the sequence to assist finding the pulse. As they started moving, their movement were usually not immediately in synchrony with any pulse of the sequence, but should be tuned to a subjectively fitting pulse level before they proceeded to the production phase. For the *no-movement* groups during the extraction phase, the participants were requested to try finding the pulse only by listening, strictly

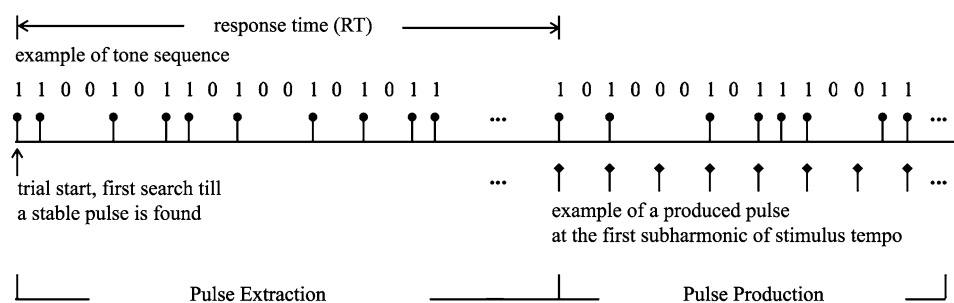


Fig. 1 Illustration of the trial procedure. The *upper panel* depicts an example of the stimulus sequence. The number 0s and 1s denote the theoretical positions of the isochronous pulse according to the

nominal stimulus tempo, where (in the sequence) 1 is occupied by a tone, and 0 is not. The *lower panel* depicts an example of a pulse identified at the 1:2 subharmonic of the stimulus tempo

without any movement, until they felt a fitting stable pulse was found. The phase of pulse extraction was not speeded. When the participants felt sure of their identified pulse, they (in all groups) were requested to start the production phase by tapping their pulse on the computer key “B” along with the sequence in a synchronized manner. Participants in the no-movement groups were instructed to restrict the movement during the production phase to only the index finger, while those in the movement groups were not particularly requested to stop movement during tapping. This was meant to maximize the contrast between movement and no-movement groups throughout the task. 16 consecutive taps were recorded per trial (inter-tap intervals representing the identified inter-pulse intervals) before the next trial commenced. The time needed for pulse extraction (henceforth referred to as response time, RT)—the time between the start of stimulus presentation and the first pulse tap—was also recorded in each trial to index the subjective task difficulty.

The stimuli were presented in 6 tempi and 30 trials each, randomly assigned to 4 blocks. The whole experiment lasted 2.5–3 h depending on the individual speed, with breaks after each block. Before starting the experimental session, each participant underwent at least five practice trials and more if they did not show enough understanding of the task. One basic sign of the participant’s understanding of the task was that, during practice, he or she did not produce taps that were simply time-locked (i.e. as a response) to the tones, but instead taps that exhibited certain degree of periodicity.

Data analyses and results

Percentages of stable and unstable pulse

For each trial, the mean inter-tap interval (ITI, in milliseconds) and the coefficient of variation (CV = within-trial standard deviation divided by mean ITI \times 100%) were calculated, excluding always the first four taps. To index the task performance, each trial was first categorized as being *stable* or *unstable* by the following criteria:

Stable trials A criterion of CV \leq 10% was first applied to identify trials with stable pulse series.² In order to

reliably identify trials in which a pulse had really been found, as opposed to trials with stable taps around a mean ITI that was irrelevant to the correct pulse period, we applied an additional criterion on the mean ITI of every stable trial regarding its identified pulse period, within which the pulse was considered to be successfully found: $(N \times \text{IOI}) \pm (N \times \text{IOI}) \times 10\%$, $N = 0.25, 0.5, 1, 2, 3, 4$, etc. N represented the chosen pulse level in each trial (i.e. the mean ITI being around N times of the shortest stimulus IOI). This criterion filtered out the stable trials with a mean ITI that exceeded 10% deviation from the correct inter-pulse interval. As such, the stable trials were further divided into two sub-types: (1) stable pulse, and (2) stable, but not considered pulse.

Unstable trials Trials produced with CV $>$ 10% were labeled as unstable trials. Each unstable trial was further categorized as reflecting one of the three behaviors which most often cause a large within-trial CV³: (1) Type 1—constantly irregular and unstable ITIs, (2) Type 2—pulse switching between different (sub)-harmonic levels, and (3) Type 3—rarely occurring missing taps or a pause within an otherwise stable tap series.

The occurrence of unstable Type 2 was generally very low (average frequency $<$ 0.1%), so we excluded it from further analyses. Of the four analyzed pulse types—stable pulse, stable no pulse, unstable Type 1, and unstable Type 3—only the first one (stable pulse) represented the successful trials. The percentages of these four types were submitted to a mixed-model ANOVA with one within-subject factor: produced pulse type (4 levels), and two between-subject factors: movement (2 levels) and musical training (2 levels). It revealed a significant pulse type \times movement \times musical training interaction, $F(3, 48) = 4.85$, $p < 0.01$, $\eta_p^2 = 0.23$ (Fig. 2). Follow-up partial ANOVAs revealed that the three-way interaction resulted from a significant interaction between musical training and pulse types in the *no-movement* groups, $F(3, 24) = 6.73$, $p < 0.01$, $\eta_p^2 = 0.46$, but not in the movement groups, $F(3, 24) = 1.91$, $p > 0.15$. Post-hoc comparison (two-sample t test) revealed that for the *no-movement* groups, the percentage of stable pulse was different between musicians and non-musicians (78 vs. 29%), $p < 0.05$, $t(8) = 2.95$, and the percentage of Type 1 unstable pulse also differed between these two groups (11 vs. 60%), $p < 0.05$, $t(8) = 2.54$. For the *movement* groups, a main effect of pulse type was significant $F(2, 16) = 9.30$, $p < 0.001$, $\eta_p^2 = 0.99$, and the post-hoc comparisons (Tukey HSD) showed that the percentage of stable pulse was significantly higher than any of the three unsuccessful types (all $ps < 0.001$), while the percentages amongst these three types did not differ (all

² A criterion of ITI stability for constituting a ‘pulse’ has not been established in the literature, as it would depend on the task condition and the given stimuli. A study on the perceptual threshold of pulse attribution (Madison & Merker, 2002) found an average 8.6% deviation of the inter-tone intervals in the sequence, beyond which the participants were unable to identify the pulse. Considering the higher difficulty in the present task as the tones in a sequence did not appear regularly, and that the pulse was measured by production, a criterion of 10% was used. This, together with the criterion on inter-pulse interval, appeared to reflect the interaction between musical training and movement well (Fig. 2).

³ For detailed classification criteria, see Figure S1 and the described procedure in the supplementary material.

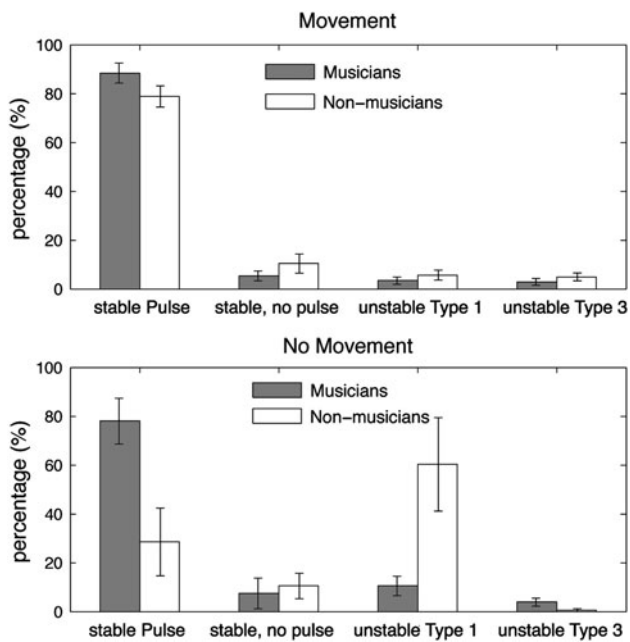


Fig. 2 Mean percentages of the four produced pulse types—stable pulse, stable but no pulse, unstable Type 1, and unstable Type 3—from each of the four participant groups. Error bars represent standard errors of the mean

$p > 0.5$). In short, non-musicians without movement produced a significantly higher percentage of unstable pulse than musicians without movement, and their unstable pulse mostly resulted from high variabilities of within-trial ITIs (Type 1). Between moving musicians and moving non-musicians, however, the distribution of produced pulse types did not differ, and they produced mostly stable pulse (88 and 79% in musicians and non-musicians.).

Similarly, the partial ANOVA between the two *musician* groups yielded no significant interaction between movement and pulse type, $F(3, 24) = 0.95$, $p > 0.4$, suggesting similar type distributions from musicians with and without movement. The partial ANOVA between the two *non-musician* groups, however, yielded a significant interaction between movement and pulse type, $F(3, 24) = 8.78$, $p < 0.001$, $\eta_p^2 = 0.52$. Post-hoc comparison (two-sample t test) revealed that non-musicians with movement produced a higher percentage of stable pulse than non-musicians without movement, $p < 0.01$, $t(8) = 3.44$ (79 vs. 29%, movement vs. no-movement).

Identified pulse tempo

The mean ITI from every stable pulse trial was transformed into the corresponding tempo (BPM) and then scaled as the *ratio* to the nominal stimulus tempo. Each resultant ratio was then logarithmically transformed before being plotted against the stimulus tempo. In this way the (sub)-harmonic relationship between the subjectively tuned-in pulse tempo

(especially at slower subharmonics such as 1:2, 1:3, and 1:4) and the given stimulus tempo can be more clearly shown. Results from each participant group were plotted together, each cross representing a single trial (Fig. 3). It shows the tendency from each participant group to select certain pulse levels under each stimulus tempo. For the exact frequency of each cluster, see Figure S2 in the supporting information for detailed histogram distributions.

As seen in the scatterplot, in establishing pulse, the movement groups showed more focused tuning to the stimulus tempo and its 1:2 subharmonic (0.5 ratio). The no-movement groups tended more to scatter and shifted toward the 1:4 subharmonic (0.25 ratio) as the tempo increased. Musicians produced better-tuned tempi especially with movement. Non-musicians using movement could tune to similar pulse tempi as musicians with movement.

Time needed for pulse extraction

RTs were submitted to a mixed-model ANOVA with one within-subject factor, tempo (6 levels), and two between-subject factors, movement (2 levels) and musical training (2 levels). A main effect was found only for tempo, $F(5, 80) = 25.25$, $p < 0.001$, $\eta_p^2 = 0.61$, with longer RTs at slower tempi (Fig. 4a). Post-hoc comparisons (Tukey HSD) found significant differences between 60 BPM and all the other tempi (all $ps < 0.001$), and between 90 BPM and all the other tempi (all $ps < 0.05$). RTs appeared to decrease with increasing tempo until 120 BPM, above which they were not significantly differentiated by tempo. Interaction of movement \times musical training was close to significant, $F(1, 16) = 3.55$, $p = 0.07$, $\eta_p^2 = 0.18$. As Fig. 4a shows, while both movement groups behaved similarly, non-musicians without movement seemed to need longer time than musicians without movement.

As an alternative, the RT data were also plotted not as the measured time but as the number of underlying pulse cycles (=RT/stimulus inter-pulse interval). The ANOVA naturally yielded the same between-group results as for RT, but the number of needed pulse cycles increased with the tempo (see Figure S3 in supplementary material).

Degree of synchronization

To measure the pulse stability by degrees of synchrony between the produced pulse and the sequence, the asynchrony was calculated between each tap and its theoretically correct position (based on the chosen pulse tempo). The variability was indexed as the within-trial standard deviation (SD) of the asynchronies—higher SD indicating lower stability—and submitted to a mixed-model ANOVA with one within-subject factor, tempo, and two between-subject factors, movement and musical training. Main

Fig. 3 Scatterplot of the produced pulse tempi as the ratio to the stimulus tempi, for each participant group separately. Each ratio was plotted as its logarithmic transformation for better viewing of pulse at slower subharmonics. *X* axis depicts each stimulus tempo condition. *Y* axis depicts the (sub)-harmonics of the stimulus tempo (1 = stimulus tempo, 2 = twice the stimulus tempo, 1/2, 1/3, and 1/4 = 0.5, 0.33, and 0.25 of the stimulus tempo). Only the tempi from *stable pulse* trials are plotted in this chart. Each *cross* represents a single trial. The *number in each chart* denotes the total percentage of *stable pulse* from this participant group

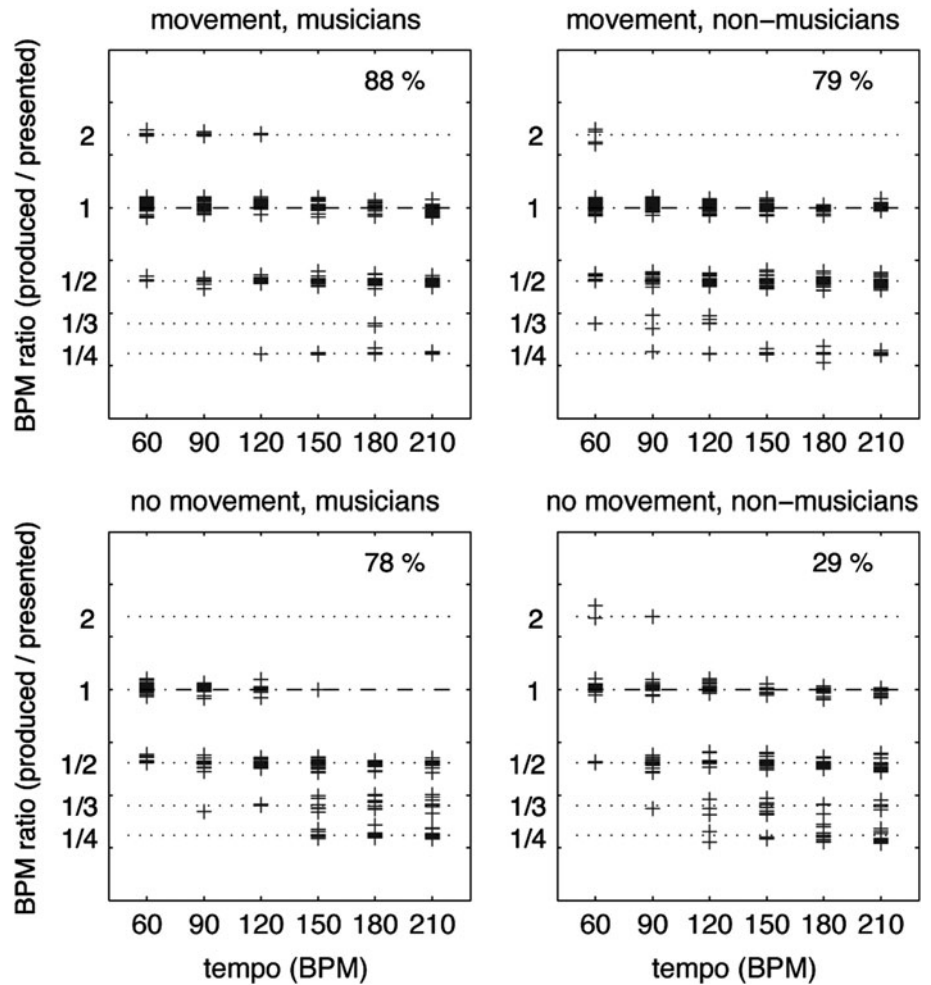
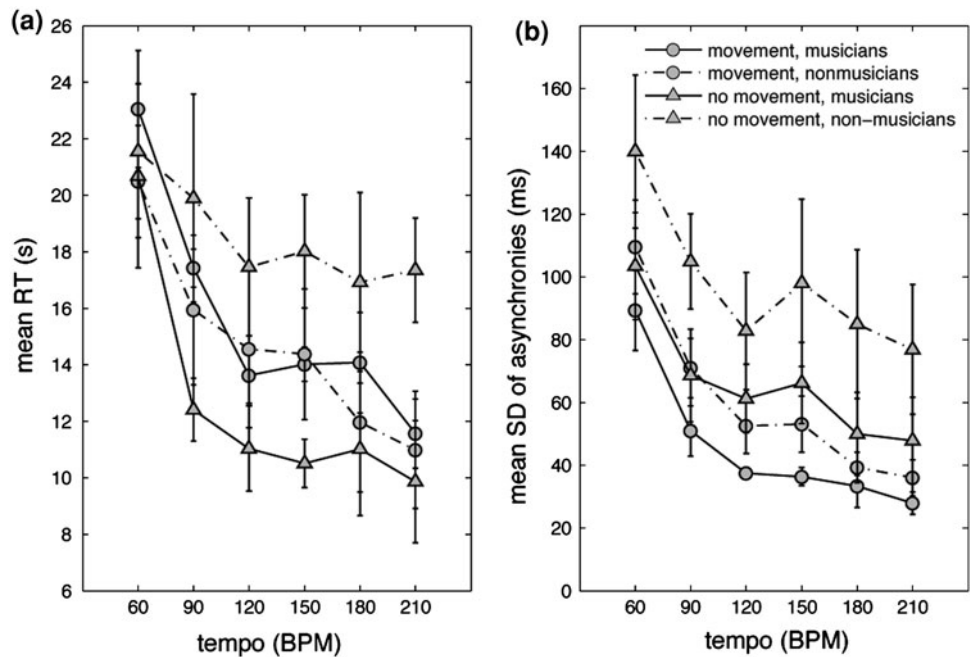


Fig. 4 **a** Mean RT as a function of the stimulus tempo, for each participant group. *Error bars* represent standard errors of the mean. **b** Mean standard deviation of asynchronies as a function of the stimulus tempo, for each participant group. *Error bars* represent standard errors of the mean



effects were found for movement $F(1, 16) = 10.11$, $p < 0.01$, $\eta_p^2 = 0.39$ (mean SD 53 vs. 97 ms, movement vs. no-movement), musical training $F(1, 16) = 7.51$, $p < 0.05$, $\eta_p^2 = 0.32$ (56 vs. 97 ms, musicians vs. non-musicians), and tempo $F(5, 80) = 44.42$, $p < 0.001$, $\eta_p^2 = 0.74$, without interactions (Fig. 4b). Post-hoc comparisons (Tukey HSD) found significant differences between the following tempi: 60 BPM versus all the other tempi; 90 versus 120, 180 and 210 BPM; 120 versus 210 BPM; 150 versus 180 and 210 BPM. The results showed that movement led to higher stability of synchronization in both musicians and non-musicians.

Discussion

Effect of movement on pulse extraction and entrainment

Our results highlight that moving one's body to an auditory sequence could indeed facilitate the extraction of the temporal structure such as the subjective pulse in a sequence.

The extent of this facilitation depended on musical training. Musicians are rhythmically trained and typically perform better in sensorimotor tasks (Chen et al., 2008a; Repp & Doggett, 2006; Franek, Mates, Radil, Beck, & Pöppel, 1994) and cross-modal timing tasks (Wöllner & Cañal-Bruland, 2010; Pecenka & Keller, 2009). It was not surprising that their training enabled them to analyze the temporal structure and establish stable pulse overall, even in the absence of movement. This proved to be much more challenging for non-musicians. With the assistance of body movement, however, non-musicians could find their pulse to a similar extent as the musicians.

What role does body movement play in this case, and what could account for its benefit? The use of body movement has been postulated as an intrinsic part of human entrainment to isochronous stimuli (Madison & Merker, 2002; Bolton 1894). Here, however, we tested the role of body movement in entraining to stimuli where the isochrony was implied but not explicitly or regularly given, and the pulse was thus more difficult to discover. Moreover, the movement we investigated was *not* a mere manifestation of the already extracted pulse such that the participants would first find the pulse internally and then start to move according to it. Instead, they started moving as soon as the sequence began, using the overt motor activities to facilitate the tuning to the pulse periodicity. In doing so, the exhibited movement for each sequence (as observed during the experimental session⁴) mostly did not

start as being immediately in synchrony with the pulse, but rather went through a bit of adjustment before tuning to one of the fitting pulse levels. An interactive dynamic might be taking place during this process: the self-initiated movement frequency, which is not tuned-in at first, could be attracted to one of the underlying periodicities of the sequence (Repp, 2006), and in doing so leads the listener to start 'hearing' the pulse at that level, forming a positive audio–motor feedback loop. In the absence of overt movement, by contrast, this tuning process must then rely on the internal motor entrainment and/or the ability to analyze the sequence. Our results show that, unlike musicians, non-musician seemed to be lacking an effective internal motor simulation that entrained to the pulse when it was not regularly present at the rhythmic surface; nor did they possess additional musical knowledge as a compensatory strategy. They thus appeared to benefit much from the external motor process in order to entrain to the structure of the rhythm. This parallels the finding of Grahn and Rowe (2009) where, compared to non-musicians, musicians more often perceived the beat when it was less explicitly presented, and this was accompanied by higher connectivities between auditory and motor cortical areas, suggesting a higher level of internal audio–motor coupling.

Notably, our task required the search for a subjective temporal referent while no particular metrical accent was given, contrary to most people's experience of music listening. Meter has been defined as 'the measurement of the number of pulses between more or less regularly occurring accents' (Cooper & Meyer, 1960). While there are many cultural differences in meter, music from most cultures is pulse-based (Large, 2008; Arom, 1989; Humble, 2002). By not giving any metrical cues, we aimed to link body movement to a temporal process that was not strongly constrained by the previously-shaped listening experiences (e.g. Iversen, Patel, & Ohgushi, 2008). Namely, one did not necessarily need to recognize a particular meter (such as 2/4 or 3/4) before identifying the pulse. Although humans exhibit a preference for culturally familiar meters (Trehub & Hannon, 2009; Soley & Hannon, 2010) and might find it more difficult to follow some 'exotic' meters, our study demonstrated an approach that relied solely on the search for a pulse, regardless of metrical preferences. This search was found to be facilitated by the accompanying body movement—a potentially useful 'hearing by moving' strategy.

Effect of movement on pulse tempo

The presence and absence of movement as a pulse-search strategy seemed to lead to different preferred pulse levels. Movement was expected to predispose the chosen pulse tempi to a range of comfortable movement frequencies

⁴ Every participant was observed for around 20–30 min during the first experimental block, and also for a shorter while in the beginning of each successive block.

(London, 2002; Macdougall & Moore, 2005), which appeared to be the case: with movement, the pulse was more often tuned to the nominal stimulus tempo, or to its 1:2 subharmonic when the stimulus tempo increased. The 1:4 subharmonic was rarely chosen, as it would have been too slow for continuous periodic movement. Without movement, they tended more to correspond to the slower subharmonics and especially more often to the 1:4 subharmonic as the stimulus tempo increased. We speculate that, in the absence of movement, the participants resorted more to a cognitive strategy to analyze the temporal structure of the sequence, especially in the case of musicians. Quite likely they would group the pulse automatically by imposing mental accents (Bolton, 1894; Repp, Iversen, & Patel, 2008), thus rendering the sequence to be heard as metrical in different ways. This would allow them to flexibly tune to different referent levels (Drake, Jones, & Baruch, 2000), though in the end they tended to opt for the slower subharmonics, as observed, because their internal pulse at a higher metrical level could be kept more stable against the irregular tones (Patel et al., 2005).

Therefore, complementary to the finding that different patterns of body bouncing can bias the metrical interpretation of a rhythm (Phillips-Silver & Trainor, 2005, 2007), the results here further demonstrate the differentiating role of the presence/absence of body movement in perceiving different pulse levels in an auditory sequence.

Time for pulse extraction

Though the time needed for pulse extraction—as measured in the experiment—may have been a function of both task difficulty and subjective readiness, it nevertheless revealed the between-group differences. That non-musicians without movement needed overall longer time than the other groups indicated the felt task difficulty, which paralleled the outcome of their pulse production. With movement, non-musicians needed similar amount of time as musicians with movement, arguing for the facilitatory effect of movement in the absence of compensating musical skills. In addition, RT decreased as the tempo increased up to about 120 BPM. The observation that RT did not decrease systematically above 120 BPM seems to reflect the relation between stimulus tempo and human's maximal pulse saliency around 80–100 BPM (London, 2002), outside of which it could be more difficult to feel the pulse.

Pulse synchronization

The degree of synchrony between the produced pulse and the sequence was also influenced by movement and musical training. Both musicians and non-musicians in the

movement groups exhibited higher stability of synchronization than those in the no-movement groups. Since movement was present in the extraction phase and as observed also often in the production phase, it would be difficult to distinguish whether the pulse stability (as measured by the variability of asynchronies) benefited from the movement in either phase alone. Though the overall facilitatory effect on pulse entrainment should have derived from the movement prior to tapping, it seems reasonable to assume that concurrent body movement during tapping might play a positive role in tapping stability. It has been found that simultaneous bimanual tapping reduces the within-hand variabilities compared to tapping with only one hand, and this advantage is accounted for by the decreased variability in the central timing process (Wing & Kristofferson, 1973; Helmuth & Ivry, 1996) or the increased sensory reafference (Drewing & Aschersleben, 2003; Prinz, 1997). In this view, our result of reduced (single-handed) tapping variability could also be attributed to the concomitant larger-scale body movement. Further investigation is needed to elucidate whether different kinds of body movement leads to the same stabilization in finger tapping, and whether the improvement can be accommodated in the same theoretical framework. Our findings provide empirical support for the idea that musicians can indeed benefit from such natural body movement while playing the instrument, in keeping up a stable tempo.

Presence versus absence of movement

In interpreting our results as demonstrating the effect of body movement and musical training on pulse finding, two questions may arise: (1) Was the poor performance of non-musicians in the no-movement group attributable to the lack of understanding of the task, because no explanation linked to movement was given? (2) Did musicians in the no-movement condition perform better because of some micro-movements they used secretly though they were not supposed to?

The first question can be dismissed because of our instruction with the auditory demonstration, showing what the pulse was and where it should be temporally in relation to the tone sequence. The participants did not have to possess specific musical knowledge to understand the temporal nature of the task. They went through practice trials, and received feedback and explanations during practice until they showed sufficient understanding of the task. This ensured that the outcome of their performance was not due to lower understanding, but rather due to the task difficulty under the appointed experimental condition.

Regarding the second question, it is possible that musicians could potentially carry out some micro-

movement, perhaps without being aware of it. If they had indeed moved secretly and constantly though they were not supposed to, the pattern of their results should have been very similar to that produced by musicians in the movement group. This was, however, not supported by our data: (1) without movement, musicians produced a rather different range of the chosen pulse tempi from musicians with movement (Fig. 3), which tended to be slower than would be naturally carried out by continuous movement; and (2) the stability of their produced pulse was also lower than that of the musicians with movement (Fig. 4b), signifying the absence of concurrent body movement to help stabilize the taps. Therefore, granted a higher tendency in musicians to carry out micro-movement in no-movement condition, our observed results suggested that this possibility either did not occur, or even if it had, its effect was both qualitatively and quantitatively different from that of the natural overt movement, and more similar to that of the no-movement condition. Two points may distinguish such potential micro-movement from the overt body movement in terms of its effect on pulse entrainment: it may have occurred not in a continuously periodic manner, or may have involved less motor activation in the brain, or both. If we were to explain the observed data from musicians without movement as a result of using secret micro-movement, it would still suggest that such micro-movement must function differently and less effectively as overt movement. As opposed to the external motor entrainment initiated by overt body movement, micro-movement might be a natural manifestation of the *internal* motor engagement. This explanation would not contradict our interpretation of facilitation by overt body movement and its interaction with musical training, but would rather point out the unique advantage of overt natural movements compared to less intuitive and much smaller-scale ones during the entrainment process.

What kind of movement?

The aim of our study was to investigate the effect of the presence and absence of movement on entraining to the pulse, and the movement of interest is the kind that a listener would naturally employ in a real-life scenario such as when listening to the music. The kinds of movement reported and performed subsequently by the participants in the task included most often head nodding (often involving the neck and the upper back) and foot tapping (often accompanied by slight head movement). In one case it was elected to be arm swiveling with foot tapping.⁵ Indeed

⁵ Often two kinds of movement were adopted together by the same participant. The frequency of each employed movement was as follows: head nodding (6), foot tapping (9), and arm swiveling (1).

finger tapping was also a movement, though a smaller one and not so commonly observed as a natural listening habit. In order to register the pulse, it had to be performed by all participants including those in the no-movement groups. However, since none of the participants reported nor chose to use finger tapping itself during pulse extraction phase, the result was not confounded with a ‘practice effect’ of finger tapping. We can, therefore, attribute the observed effects to the opted larger body movements during the pulse discovery/entrainment process, which possibly also stabilized pulse tapping. Future studies might attempt to reveal whether different scales of movement, e.g., larger body movement versus smaller one such as finger tapping, would lead to different effects of motor simulation for entraining to the rhythm.

Overall our study demonstrated that overt body movement assisted the extraction of the underlying pulse in a non-isochronous sequence. It also led to better tuning to the sequence tempo and better synchronization to the sequence. The results provide empirical evidence of body movement as being a useful strategy especially for untrained listeners to approach auditory rhythms, and when musicians intuitively move their head or tap their feet while playing an instrument, it could help them keep up a stable tempo.

Acknowledgments This work was supported by the doctoral scholarship to the first author from the Bayerische Forschungsstiftung, and the experimental expenses were supported by the Andrea von Braun Foundation. The authors thank Marc Wittmann, Dragan Rangelov, and Björn Merker for earlier discussions of the work, as well as Bruno Repp and an anonymous reviewer for very useful inputs on the manuscript.

Conflict of interest The authors declare no conflict of interest.

References

- Arom, S. (1989). Time structure in the music of central Africa: periodicity, meter, rhythm and polyrhythmics. *Leonardo*, 22, 91–99.
- Bengtsson, S. L., Ullen, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., et al. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex*, 45, 62–71.
- Bolton, T. L. (1894). Rhythm. *American Journal of Psychology*, 6, 145–238.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008a). Moving on time: brain network for auditory–motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, 20, 226–239.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008b). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, 18, 2844–2854.

- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Annals of the New York Academy of Sciences*, *1169*, 15–34.
- Cooper, G., & Meyer, A. B. (1960). *The rhythmic structure of music*. London: The University of Chicago Press.
- Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, *77*, 251–288.
- Drewing, K., & Aschersleben, G. (2003). Reduced timing variability during bimanual coupling: a role for sensory information. *The Quarterly Journal of Experimental Psychology Section A*, *56*, 329–350.
- Franek, M., Mates, J., Radil, T., Beck, K., & Pöppel, E. (1994). Sensorimotor synchronization: motor responses to pseudoregular auditory patterns. *Perception & Psychophysics*, *55*, 204–217.
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of New York Academy of Science*, *1169*, 89–92.
- Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*, 893–906.
- Grahn, J. A., & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex*, *45*, 54–61.
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, *29*, 7540–7548.
- Helmuth, L. L., & Ivry, R. B. (1996). When two hands are better than one: reduced timing variability during bimanual movements. *Journal of Experimental Psychology Human Perception and Performance*, *22*, 278–293.
- Humble, M. (2002). The development of rhythmic organization in Indian classical music. Unpublished dissertation, University of London.
- Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of Acoustic Society America*, *124*(4), 2263–2271.
- Kornysheva, K., von Cramon, D. Y., Jacobsen, T., & Schubotz, R. I. (2010). Tuning-into the beat: aesthetic appreciation of musical rhythms correlates with a premotor activity boost. *Human Brain Mapping*, *31*, 48–64.
- Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. In S. Grondin (Ed.), *The psychology of time* (pp. 189–231). United Kingdom: Emerald.
- Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of New York Academy of Science*, *1169*, 46–57.
- London, J. (2002). Cognitive constraints on metric systems: some observations and hypotheses. *Music Perception*, *19*, 529–550.
- Macdougall, H. G., & Moore, S. T. (2005). Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. *Journal of Applied Physiology*, *99*, 1164–1173.
- Madison, G., & Merker, B. (2002). On the limits of anisochrony in pulse attribution. *Psychological Research*, *66*, 201–207.
- Merker, B., Madison, G., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, *45*, 4–17.
- Moelants, D. (2002). Preferred tempo reconsidered. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Ed.), *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 580–583). Sydney.
- Patel, A., Iversen, J., Chen, Y., & Repp, B. H. (2005). The influence of metricity and modality on synchronization with a beat. *Experimental Brain Research*, *163*, 226–238.
- Pecenkova, N., & Keller, P. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of New York Academy of Science*, *1169*, 282–286.
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: movement influences infant rhythm perception. *Science*, *308*, 1430.
- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition*, *105*, 533–546.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129–154.
- Repp, B. (2006). Does an auditory distractor sequence affect self-paced tapping. *Acta Psychologica*, *121*(1), 81–107. doi: [10.1016/j.actpsy.2005.06.006](https://doi.org/10.1016/j.actpsy.2005.06.006).
- Repp, B. H., & Doggett, R. (2006). Tapping to a very slow beat: a comparison of musicians and nonmusicians. *Music Perception*, *24*, 367–376.
- Repp, B. H., Iversen, J. R., & Patel, A. D. (2008). Tracking an imposed beat within a metrical grid. *Music Perception*, *26*, 1–18.
- Salenius, S., & Hari, R. (2003). Synchronous cortical oscillatory activity during motor action. *Current Opinion in Neurobiology*, *13*, 678–684.
- Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: a cross-cultural comparison. *Developmental Psychology*, *46*, 286–292.
- Todd, N. P. M. (1999). Motion in music: a neurobiological perspective. *Music Perception*, *17*, 115–126.
- Todd, N. P. M., Lee, C. S., & O'Boyle, D. J. (2002). A sensorimotor theory of temporal tracking and beat induction. *Psychological Research*, *66*, 26–39.
- Trehub, S. E., & Hannon, E. E. (2009). Conventional rhythms enhance infants' and adults' perception of musical patterns. *Cortex*, *45*, 110–118.
- Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, *14*, 5–12.
- Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Science*, *106*, 2468–2471.
- Wöllner, C., & Cañal-Bruland, R. (2010). Keeping an eye on the violinist: motor experts show superior timing consistency in a visual perception task. *Psychological Research*, *74*, 579–585.