1244

In W. Prinz & B. Hommel (Eds.) (2002), *Common mechanisms in perception and action: Attention and Performance XIX* (pp. 245–265). Oxford, U.K.: Oxford University Press.

12 The embodiment of musical structure: effects of musical context on sensorimotor synchronization with complex timing patterns

Bruno H. Repp

Abstract. Two experiments demonstrate that musical context facilitates sensorimotor synchronization with complex timing patterns that are compatible with the musical structure. Several very different timing patterns were derived from an analysis of expressive performances of a musical excerpt. A random pattern (Exp. 1) or phase-shifted versions of the musical patterns (Exp. 2) served as comparisons, and an isochronous pattern served as practice. Musically trained participants first attempted repeatedly to synchronize their finger taps with click sequences instantiating these timing patterns. Subsequent repetitions of the click sequences were accompanied by the identically timed music, and finally, the music disappeared and was only to be imagined in synchrony with the clicks. Compared with the random or phase-shifted patterns, synchronization accuracy for the musical patterns improved as soon as the music was introduced, especially when the pattern was highly typical. This relative improvement was reduced or absent when the music was merely imagined. Nevertheless, both musical context and imagery systematically modulated the timing of finger taps in synchronization with strictly isochronous click sequences. Thus perception or imagination of musical structure can involuntarily affect the timing of concurrent action, presumably by modulating the timekeeping processes that pace the motor behavior. This study also demonstrates that radically different timing patterns are compatible with the same musical structure, as they seem to be in expert artistic performance.

12.1 Introduction

There is an intimate relationship between music and the human body (see, e.g. Clarke 1993a; Iyer 1998; Pierce and Pierce 1989; Repp 1993; Shove and Repp 1995). Music is produced by moving various extremities across musical instruments, or by engaging the mouth, lungs, and vocal tract. These moving parts of the body are attached to (or embedded in) the trunk which provides structural support and often participates by swaying along. In most cultures, listeners participate in music by dancing, clapping, tapping, or rocking in synchrony with its rhythm. Only in the Western tradition of serious art music, overt movement is proscribed for audiences in concert halls, but listeners still feel a readiness to move, or imagine themselves moving along with the music, or speak about being moved by the music. Thus there is a very close relation between music perception and action, particularly with regard to rhythm and timing.

Human music performances are distinguished from machine renditions (unless they successfully simulate human performance) by the presence of many subtle features that originate in the musicians' movements. Clynes (1983) has referred to these features as 'expressive microstructure' which conveys 'living qualities'. One of these features is expressive timing. It consists in systematic deviations from temporal regularity which signify to a listener that the music was not produced by a machine but by a thinking, feeling, and moving being. Expressive timing originates from three sources (cf. Penel and Drake 1998, 1999): (1) biomechanical constraints in technically difficult passages; (2) obligatory perceptual–motor patterns related primarily to rhythm and rhythmic grouping; and (3) intentional communication of structural or emotional aspects of the music. The present study is mainly concerned with the second of these factors, only incidentally with the third, and not at all with the first.

Recent research has produced considerable evidence that a particular musical structure is often associated with a particular expressive timing pattern. This most typical pattern corresponds to the average timing pattern of a large sample of human performances. It is representative of many individual performances (Repp 1998a) and is judged to be aesthetically pleasing (Repp 1997). When musicians are requested to play with perfectly regular timing (as specified in a musical score) or in synchrony with a metronome (Repp 1999c), or when they try to create a perceptually regular performance on a computer by adjusting successive temporal intervals (Penel 2000), they nevertheless produce small but systematic timing variations whose pattern resembles that of the typical expressive timing pattern (Behne and Wetekam 1993; Drake and Palmer 1993; Palmer 1989; Penel and Drake 1998; Repp 1999a,c). A complementary pattern of perceptual biases is observed when listeners are asked to detect local deviations from perfect regularity in a musical passage (Repp 1992b, 1998b,c,d, 1999b,c). These findings suggest that there is a level of subconscious and obligatory timing variation upon which larger intentional expressive variations are superimposed. The obligatory variations seem to be linked to the lowest level of rhythmic grouping in the music, whereas intentional expressive timing reflects several hierarchical levels of grouping (Penel and Drake, 1998) as well as possibly other factors (meter, melodic contour, harmony, etc.). The similarity of the pattern of obligatory variations to the typical expressive timing pattern may be explained by the fact that they share the lowest level of grouping, which accounts for much of the timing variation (Penel 2000).

However, intentional expressive timing does not always follow the most typical pattern. The timing patterns produced by experienced concert artists sometimes represent quite radical departures from the norm (Repp 1992a, 1998a). While such highly individual timing patterns may sound strange on first hearing, the fact that they were produced by outstanding musicians indicates that they are not arbitrary or inappropriate to the musical structure. Nevertheless, it seems that these patterns are not strongly *implied* by the musical structure, if at all. It appears that creative performers must overcome a natural tendency to produce the most typical timing pattern (Repp 2000b). Penel and Drake (1998) have argued that typical timing is a form of motor compensation for perceptual timing distortions caused by rhythmic grouping. If so, then the typical timing pattern must always be present underlyingly, even if it is overridden by different intentions. Alternatively, the typical timing pattern may be regarded as a natural strategy for representing rhythmic groups in action, a strategy that in turn causes perceptual biases via a motor–perceptual interaction (Repp 1998d; Viviani and Stucchi 1992).

Perhaps, then, the typical (obligatory) timing pattern is a consequence of carrying out grouped actions on a musical instrument. However, Repp (1999a,b,c) eliminated this factor by asking participants (including non-pianists and even non-musicians) to tap with the index finger in synchrony with piano music that was reproduced under computer control in a perfectly regular fashion. The tap-tone asynchronies and inter-tap intervals were still found to exhibit systematic deviations from regularity that tended to be positively correlated with the typical expressive timing profile. Thus, perception of musical structure exerted an influence even on a concomitant action pattern that had no structure of its own. The correlation between the obtained timing pattern and the typical expressive timing pattern was relatively small; this may have been due in part to an additional process of automatic error correction in synchronization (Mates 1994; Pressing 1998; Repp 2000a; Vorberg and Wing 1996), which counteracted the emergence of the typical timing pattern. The tentative conclusion from these results, therefore, was that a musical structure tends to induce a tendency towards the typical timing pattern in concurrent motor activity.

It may be predicted, then, that this tendency to move expressively should facilitate the synchronization of movements with music that exhibits the typical (intentional) expressive timing pattern, even though that pattern shows much larger deviations from regularity than the obligatory timing variations induced by the music, which are generally below the perceptual detection threshold. This prediction has been investigated previously by asking pianists to tap their index finger in synchrony with (1) one of their own previously recorded expressive performances, (2) a computer-generated version that exhibited the typical timing pattern (the average timing pattern of a large number of human performances), and (3) a sequence of clicks that instantiated the typical timing pattern, while participants imagined the music in synchrony with the clicks (Repp 1999a).[1] The pianists were quite successful in all three tasks (though not as accurate as in tapping to an isochronous sequence). Moreover, their synchronization was as accurate with the clicks as with the music itself, which suggested that musical imagery could effectively substitute for the musical sound. However, one shortcoming of that study was that it included no other conditions that the pianists' synchronization accuracy could be compared with. For example, it was not determined how well they could synchronize with the clicks without imagining the music, or with music having expressive timing patterns other than the most typical one, or with non-musical timing patterns of comparable average tempo and variability. Thus it was not clear whether synchronization with the most typical timing pattern in music was better than with other possible timing patterns, or indeed whether the relatively good synchronization performance had anything to do with music at all.

It was the purpose of the present study to make these additional comparisons. Two similar experiments were conducted to address five hypotheses or predictions. One hypothesis was that synchronization with music exhibiting a typical expressive timing pattern would be more accurate than synchronization with music exhibiting a less typical (but still structurally appropriate) timing pattern, because the former pattern is more strongly implied by the musical structure than the latter. To that end, several timing patterns of different typicality, derived from an extensive performance analysis (Repp 1998a), were used. Another hypothesis was that synchronization with even the less typical musical timing patterns would be more accurate than synchronization with an arbitrary or structurally inappropriate timing pattern, imposed on the same music. To test this prediction, synchronization with the musical patterns was compared to synchronization with a random pattern (Exp. 1) or with phase-shifted versions of the musical patterns (Exp. 2). A third hypothesis was that the differences just predicted would also emerge, though perhaps be smaller in magnitude, when the music was merely imagined in synchrony with a click sequence instantiating the timing patterns. (This click sequence also accompanied the music when music was present.)

A fourth hypothesis was that timing patterns derived from expressive music performance might be easier to synchronize with than arbitrary timing patterns even in the absence of real or imagined music, simply because musical patterns are more regular. Moreover, musical timing patterns may differ from each other in their degree of regularity (i.e. periodicity or predictability), and hence in how difficult they are to learn and predict in repeated presentations.[2] Therefore, synchronization accuracy was also assessed in a condition in which music was neither present nor imagined (i.e. where the timing pattern was carried only by a click sequence). This condition provided a crucial baseline for interpreting the findings in the music and imagery conditions, and it necessitates an important qualification of the first three hypotheses. Specifically, their predictions are that synchronization with musical timing patterns should be *selectively facilitated* when music is present or imagined, compared with a condition in which music is neither present nor imagined. This selective facilitation should be largest for the most typical timing pattern and smaller for the less typical musical patterns. There should be no facilitation and possibly even interference for arbitrary or structurally inappropriate timing patterns. Viewed from an ANOVA perspective, the effects of primary interest in this study thus were interactions between condition and pattern type, not main effects.

Finally, a fifth hypothesis was that synchronization accuracy would improve as a function of repeated presentation of the same timing pattern, but more so for musical patterns than for structurally inappropriate patterns (and most clearly for the most typical pattern) when music is present or imagined. Thus, an interaction between pattern type and trial number was also predicted.

To get used to the synchronization task and the three experimental conditions (clicks only, clicks plus music, clicks plus imagined music), participants first tapped in time with an isochronous pattern. This made it possible to address another interesting issue in passing, as it were. As mentioned earlier, tapping in synchrony with isochronous music leads to systematic deviations from regularity in the timing of the taps (Repp 1999a,b,c). One question was whether that finding would be replicated when the music merely accompanies an isochronous click sequence that participants try to synchronize with. Even more interesting, however, was the question of whether similar systematic deviations from regularity would be evident when the music was merely imagined in synchrony with the isochronous click sequence. A previous attempt to determine this (Repp 1999a) led to unclear results, perhaps because the instructions had not sufficiently emphasized musical imagery. If a significant effect of musical imagery were found in this very simple synchronization task, this would constitute convincing evidence of the reality of musical imagery and provide further proof of a close connection between music perception and action.

12.2 Experiment 1

12.2.1 Methods

12.2.1.1 Materials

The timing patterns were derived from an analysis of 115 expert performances of the opening (bars 1–5) of Frédéric Chopin's Etude in E major, op. 10, No. 3 (Repp 1998a). A computer-generated score of this music is shown on top of Fig. 12.1. The second half of the original bar 5 was condensed into a chord to give maximal closure to the excerpt, as heard in the experiment.

Below the musical score and vertically aligned with it, Fig. 12.1(a) shows the most typical expressive timing pattern (or timing profile) for this excerpt (T0). This is the average profile of the 115 performances whose timing was measured from digitized acoustic recordings. It is equivalent to the first unrotated principal component obtained in a principal components analysis of the performance timing profiles, a component which accounted for 61% of the variance. The graph depicts tone inter-onset intervals (IOIs) as a function of metrical (score) position, with 8 sixteenth-note subdivisions per bar. The initial upbeat IOI, corresponding to an eighth note in the score, has been excluded from all graphs and statistics; its average duration was 1122 ms. All other IOIs represent nominal sixteenth-

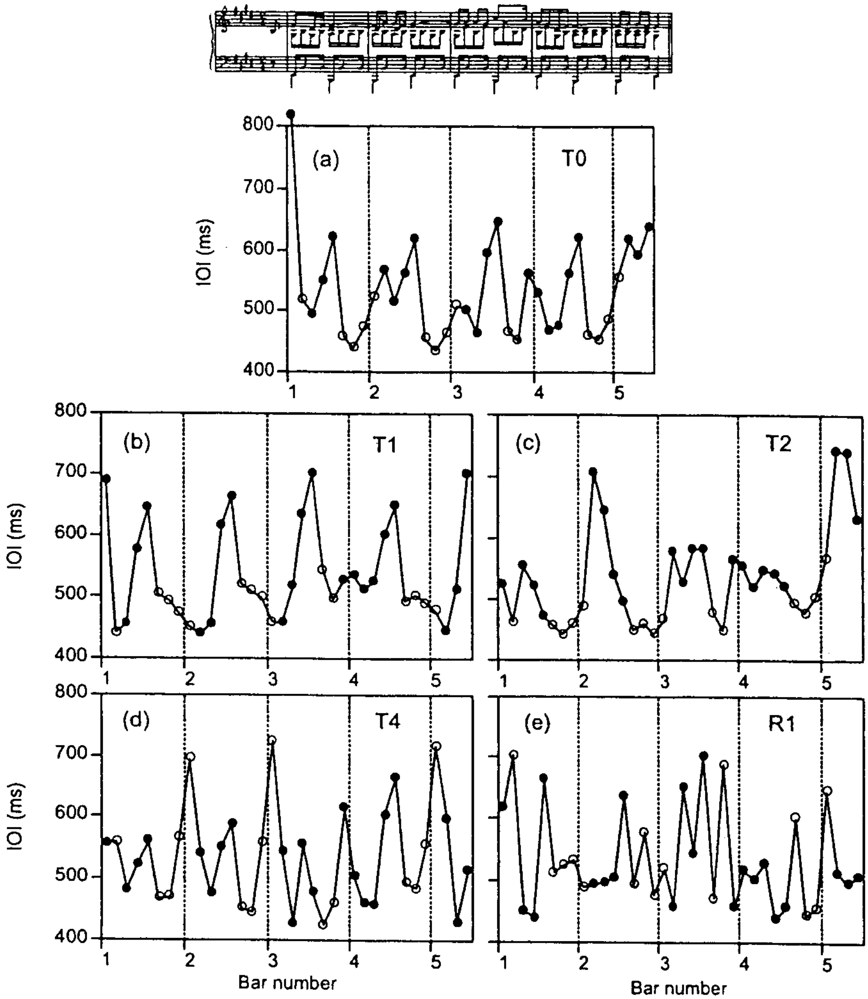

Fig. 12.1 (*top*) A computer-generated score of the opening of Etude in E major, op. 10, No. 3, by Frédéric Chopin. (a) The most typical expressive timing profile (T0) for this music. (b), (c), (d) Mutually uncorrelated timing profiles (T1, T2, T4) representing principal components of the timing patterns observed in expert performances. (e) An arbitrary timing pattern (R1) obtained by randomizing the inter-onset intervals (IOIs) of T1. Solid circles indicate IOIs initiated by melody notes, open circles those initiated by accompaniment notes only.

note intervals. IOIs initiated by melody tones (among other tones) are shown as filled circles, those initiated only by accompaniment tones as open circles. The melody, in the highest voice, is divided into six rhythmic groups (runs of filled circles in the graph), each ending with a sustained tone during which the accompaniment in the other voices continues. It can be seen that the T0 pattern includes *ritardandi* (final slowing) within each of the melodic segments, as well as a lengthening of the final IOI in bar 3 (which is the initial IOI of the longest melodic group) and sometimes of the final IOI of an accompaniment passage immediately preceding a melodic group (the initial IOIs in bars 2, 3, and 5). The T0 pattern was not used in Experiment 1 because of a concern that its correlation with the other patterns, especially T1, might lead to carry-over effects of pattern learning. However, it was used in Experiment 2.

Three additional musical timing profiles (T1, T2, T4) were used in Experiment 1 and are shown in Fig. 12.1(b), (c), and (d). They represent the first, second, and fourth Varimax-rotated principal components of the timing patterns of the 115 expert performances (Repp, 1998a) and respectively accounted for 31%, 17%, and 11% of the variance.[3] Thus, T1 was more typical of expert performance than were T2 or T4, and this was also reflected in their respective correlations with T0 (see Table 12.1), which may serve as indices of typicality. Being principal components, these three profiles were mutually uncorrelated. Originally vectors of standard scores, they were converted into IOIs by multiplying them with the average within-performance standard deviation (80 ms) and adding them to the grand average IOI duration of the 115 performances (533 ms). Thus they all had the same basic tempo and degree of timing modulation. A fourth pattern, R1, was generated by randomly scrambling the IOI durations of the T1 pattern (Fig. 12.1(e)). As can be seen in Table 12.1, the typicality of R1 was even lower than that of T4. The R1 profile correlated with the three musical profiles 0.21, −0.18, and −0.04, respectively (all n.s.). The duration of the initial upbeat IOI (not shown) was 1000 ms in all four patterns used in Experiment 1.

The four timing patterns also differed in complexity or regularity. For example, T1 is characterized by strong *ritardandi* within all melodic groups, but it lacks the other timing features seen in T0, and this results in a very clear periodicity. By contrast, T2 shows a striking *accelerando* in the melodic group of bar 2 and to a lesser degree also in bars 1 and 5, but not at all in bars 3 and 4, which makes this pattern more complex than T1. T4 shows pronounced between-group *ritardandi* that exceed the within-group *ritardandi*, as well as a lengthening of the final IOI in bar 3. It seems to be of intermediate complexity. The random pattern, of course, is the most complex pattern. To quantify these intuitions, an index of the degree of pattern periodicity was computed in the form of the lag-8 autocorrelation (ac8), which assesses the average similarity of timing from one bar to the next. A measure of relative pattern complexity was then obtained by subtracting ac8 from 1. These complexity indices are shown in Table 12.1. Furthermore, Table 12.1 includes the lag-1 autocorrelations (ac1) of the four patterns, which will be referred to later.

In addition to the four timing patterns, an isochronous sequence with constant IOIs of 500 ms (except for an initial 1000-ms IOI) was presented. Each of these five timing patterns was imposed

Table 12.1 Typicality indices (i.e. correlations with T0), complexity indices (i.e. 1 − ac8; see text for explanation), and lag-1 autocorrelations (ac1) for the four timing patterns used in Exp. 1

Pattern	*Typicality*	*Complexity*	*ac1*
T1	0.67	0.31	0.29
T2	0.46	0.73	0.59
T4	0.36	0.58	0.15
R1	0.20	0.90	−0.25

on a series of what will informally be called 'clicks'. Each click was in fact a high-pitched tone (C8, MIDI pitch 108, fundamental frequency 4,168 Hz) produced on a Roland RD-250 s digital piano, with a nominal duration of 20 ms. The tones had sharp onsets followed by a rapid decay and a longer soft ringing. Each click sequence comprised 38 identical sounds. When the click sequence was accompanied by the music, the music had exactly the same timing pattern in terms of its top-line tones (the highest tones in all sixteenth-note positions). The clicks coincided (within 1 ms) with the onsets of these top-line tones and were clearly audible above the music. The precise methods for synthesizing the music performances are described in Repp (2000b).

12.2.1.2 Participants

Twelve undergraduate students from Yale University were paid to participate. All had advanced musical training, which was a necessary requirement in a study of expressive timing and musical imagery. Three of them were pianists, and the others, several of whom also played the piano, were players of string instruments in the Yale Symphony Orchestra.

12.2.1.3 Procedure

Participants were first instructed in the use of the response device, a Fatar Studio 37 MIDI controller (a silent three-octave piano keyboard). They were instructed to hold the controller on their lap, to keep their index finger in contact with a self-chosen white key, to release the key fully before pressing it again, to start tapping with the second click in each sequence, to stay in synchrony with the clicks at all times, and not to count the clicks. The response key moved about 10 mm from its resting position to the (cushioned) bottom position, but the electronic contact occurred before the lowest position was reached, which added a small negative constant to the tap-tone asynchronies. The response key did not make any audible sound unless it was struck very hard, so that participants generally had to gauge their synchronization errors cross-modally. The keypresses were registered by a MAX patch running on a Macintosh Quadra 660AV computer, which also controlled the playback of the sequences on a Roland RD-250s digital piano ('Piano 1' sound).[4] Participants sat in front of the computer monitor, which displayed the trial number, and listened binaurally over Sennheiser HD540 II earphones.

The three conditions (clicks only, clicks with accompanying music, and clicks with imagined music—referred to in the following as 'clicks', 'music', and 'imagery') were presented in the same order to all participants, constituting three successive parts of the experimental session. Within each condition, all timing patterns were presented in the same order. Each condition started with the isochronous sequence, but the order of the other four sequences was varied across participants, according to three different 4×4 Latin squares. Each timing pattern was presented 10 times in succession, without any preceding practice trials. The participants' task was to tap in synchrony with each pattern to the best of their ability, and to try to predict the pattern with their taps from the second trial on. In the music condition, the instruction was to tap in synchrony with the clicks and not to pay any special attention to the music. In the imagery task, participants were told to imagine the music in synchrony with the clicks and to be sure not to make an extra tap at the end, since this would indicate that they had not imagined the music correctly. A copy of the musical score (Fig. 12.1, top) was in view throughout the music and imagery conditions, propped up below the computer monitor. There were 3 seconds of silence between trials, short breaks between timing patterns, and longer breaks between conditions.

12.2.1.4 Analysis

Three different measures of synchronization accuracy were used. One was the *standard deviation of the asynchronies* (sda). This measure was useful because all sequences used had the same average tempo (i.e. mean IOI duration) and the same average timing modulation (i.e. standard deviation of IOIs). If participants were able to perfectly predict a timing pattern with their taps, then the standard deviation of the asynchronies would be equal to that found with an isochronous sequence. Of course, in view of the complexity of the patterns, prediction was not expected to be perfect in any condition.

The other two measures of synchronization accuracy were correlational. One was the *lag-0 cross-correlation* ($r0$) between the inter-tap intervals (ITIs) and the click IOIs. If the taps predict the sequence timing pattern accurately, then $r0$ will be high. The other measure was in a way the converse of $r0$. Michon (1967) first demonstrated that attempts to synchronize with an auditory sequence whose temporal intervals vary unpredictably result in ITIs that echo the sequence IOIs at a lag of one. (See also Hary and Moore 1985, 1987; Schulze 1992.) This temporal tracking behavior seems to be the consequence of an automatic error-correction process that tries (unsuccessfully) to minimize the synchronization error. It results in a high *lag-1 cross-correlation* ($r1$) between ITIs and IOIs, which thus is a measure of the participant's inability to predict the temporal pattern. Thaut, Tian, and Azimi-Sadjadi (1998) found that tracking occurred even with sequences that were modulated in a regular, periodic fashion, but this may have been due to the small size of the modulations. When larger modulations of a regular, meaningful, or familiar nature are imposed on a stimulus sequence, the participant's taps will tend to predict the sequence timing, which reduces $r1$ and increases $r0$ (Michon, 1967).

However, it is problematic to rely on the raw values of $r0$ and $r1$. Each of these correlations has a theoretical lower limit that depends on the temporal structure of the sequence. In fact, it seems that both correlations have the same lower limit, namely the *lag-1 autocorrelation* (ac1) of the sequence timing pattern: when prediction ($r0$) is optimal, $r1$ will approach ac1 because the sequence of ITIs is similar to the sequence of IOIs. When tracking ($r1$) is maximal, $r0$ will approach ac1 because the ITIs echo the IOIs at a lag of 1. Therefore, a correction was applied to both $r0$ and $r1$, in order to take into account the fact that different timing patterns have different ac1 values (see Table 12.1). The *prediction index* ($r0^*$) thus was computed as $(r0-\mathrm{ac}1)/(1-\mathrm{ac}1)$, and the *tracking index* ($r1^*$) was computed as $(r1-ac1)/(1-ac1)$. Both indices had a theoretical range from near zero to 1.

12.2.2 *Results and discussion*

Because of space restrictions, only the results for one of the three indices of synchronization accuracy, the prediction index ($r0^*$), will be reported in detail. In general, the results for the sda index were similar, whereas those for the tracking index ($r1^*$) were less clear, suggesting that, despite a strong negative relationship with $r0^*$, $r1^*$ captures somewhat different aspects of synchronization behavior.

The results for $r0^*$ are shown as a function of trial number in Fig. 12.2. Rather than comparing the results among all three experimental conditions at once, three separate repeated-measures ANOVAs were conducted, each of which compared two conditions. The fixed variables in each ANOVA were condition (2), pattern (4), and trial (10). Within each ANOVA, separate comparisons were carried out between each musical pattern and the R1 pattern, which served as the baseline. Additional two-way ANOVAs were conducted on each individual condition.

The main effect of pattern was highly significant in all three two-way ANOVAs, $F(3,33)>7.8$, $p<0.0005$. Overall, performance tended to be best for the T1 pattern, followed by T4, R1, and T2. It

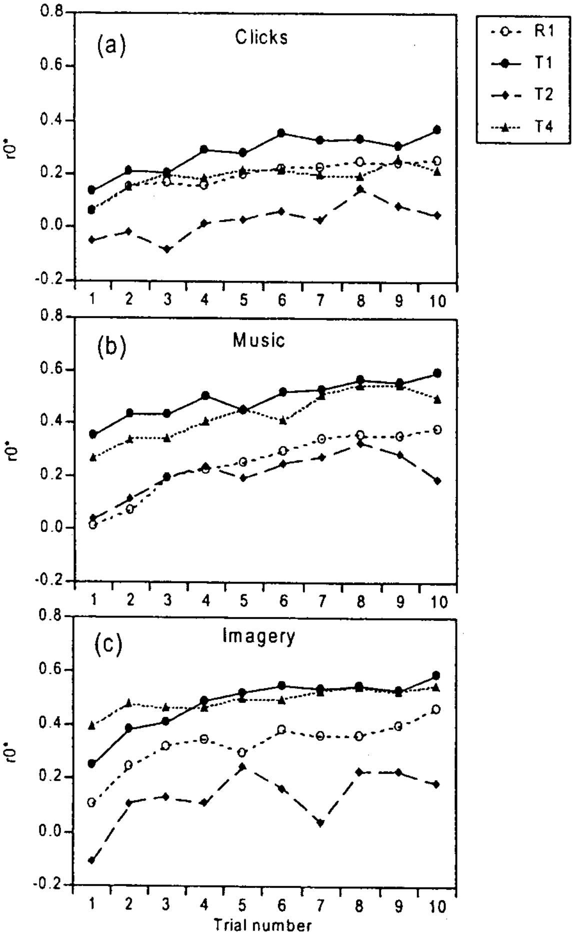

Fig. 12.2 Average prediction indices ($r0^*$) as a function of trial number for four timing patterns in the three conditions of Experiment 1. (a) Click condition. (b) Music condition. (c) Imagery condition.

was surprising that T2 yielded poorer performance than R1, but T2 happened to be the musical pattern with the highest ac1 coefficient (Table 12.1), so that its $r0$ coefficient was most affected by the correction that turned it into $r0^*$. It is possible that this correction was too extreme, as it did not take into account automatic error correction in tracking. The better performance with T1 and T4 is consistent with the lower complexity of these patterns.

The main effect of trial was also highly significant in all analyses, $F(9, 99) > 8.7$, $p < 0.0001$, due to gradual improvement within conditions. The Pattern × Trial interaction reached significance in the music condition, $F(27, 297) = 1.6$, $p < 0.03$, and in the imagery condition, $F(27, 297) = 1.9$, $p < 0.008$, but not in the click condition, $F(27, 297) = 0.6$. These interactions are difficult to interpret,

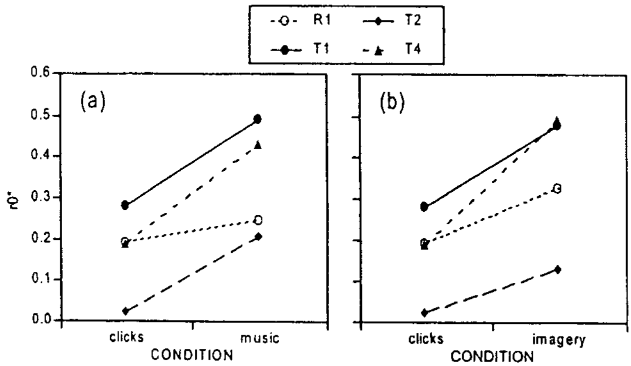

Fig. 12.3 Prediction indices ($r0^*$), averaged across trials within conditions, as a function of condition in Experiment 1. (a) Click and music conditions. (b) Click and imagery conditions.

however. More rapid improvement for the more typical patterns, as hypothesized in the Introduction, was not evident. Rather, all patterns seemed to improve at about the same rate.

The Condition × Trial interaction was significant when comparing the music and imagery conditions, $F(9, 99) = 3.0$, $p < 0.004$, and also for clicks vs. music, $F(9, 99) = 2.2$, $p < 0.04$, but not for clicks vs. imagery, $F(9, 99) = 1.7$, $p < 0.10$. These interactions were due to somewhat greater improvement for all patterns within the music condition than within the other two conditions.

The main effect of condition was highly significant in the two ANOVAs involving the click condition, $F(1, 11) > 43$, $p < 0.0001$, but not in the music vs. imagery comparison. Performance in these latter two conditions did not differ, but was substantially better than in the click condition. This suggests that pattern prediction was improved by both the presence and imagery of music, but the improvement could also have been due to general pattern learning, as observed within conditions. Therefore, the Pattern × Condition interaction was the crucial statistic. That interaction was significant in all three ANOVAs: clicks vs. music, $F(3, 33) = 3.2$, $p < 0.04$; clicks vs. imagery, $F(3, 33) = 4.1$, $p < 0.02$; and music vs. imagery, $F(3, 33) = 4.9$, $p < 0.007$. The corresponding data, averaged over trials, are shown in Fig. 12.3, which focuses on the two interactions of primary interest. Individual comparisons of each musical pattern with R1 in the clicks vs. music analysis (Fig. 12.3(a)) confirmed that the presence of music selectively improved prediction performance for T1, $F(1, 11) = 6.6$, $p < 0.03$, and for T4, $F(1, 11) = 12.3$, $p < 0.005$, but not significantly for T2, $F(1, 11) = 3.6$, $p < 0.09$. An interesting aspect of these data is that the selective advantage for the musical patterns was already present in the first trial of the music condition (see Fig. 12.2). In the comparison of the click and imagery conditions (Fig. 12.3(b)), a selective facilitation relative to R1 was evident only for T4, $F(1, 11) = 7.5$, $p < 0.02$. The significant Pattern × Condition interaction in the music vs. imagery ANOVA was mainy due to T2, for which prediction performance was worse in the imagery than in the music condition.

12.2.2.1 Average timing profiles

Figure 12.4 shows the results for the isochronous sequence, which is represented by the horizontal dotted line (IOI = 500 ms). The ITIs are shown as data points with double standard errors (roughly, 95% confidence intervals). In the click condition (Fig. 12.4(a)), the ITIs closely matched the sequence IOIs from the fourth IOI on. The initial three ITIs reflect a 'tuning in' to the sequence (see also Fraisse 1966; Repp 1999b; Semjen, Vorberg, and Schulze 1998): Despite the constant sequence tempo from trial to trial, the first tap tended to occur too late, so that the following ITIs had to be shortened to achieve synchrony; however, there were also substantial individual differences in that respect, as reflected in the large standard errors. The pattern of tap timing from the fourth ITI on did not show any significant deviation from uniformity in a one-way ANOVA with the independent variable of position (33), $F(32, 352) = 1.1$. By contrast, the tap timing profile in the music condition (Fig. 12.4(b)) did show significant variation from the fourth ITI on, $F(32, 352) = 14.7$, $p < 0.0001$, and also showed a different pattern of the initial three ITIs. Moreover, the pattern of systematic deviations from regularity was quite similar to that obtained in a previous study (Repp 1999b: Exp. 3) with the

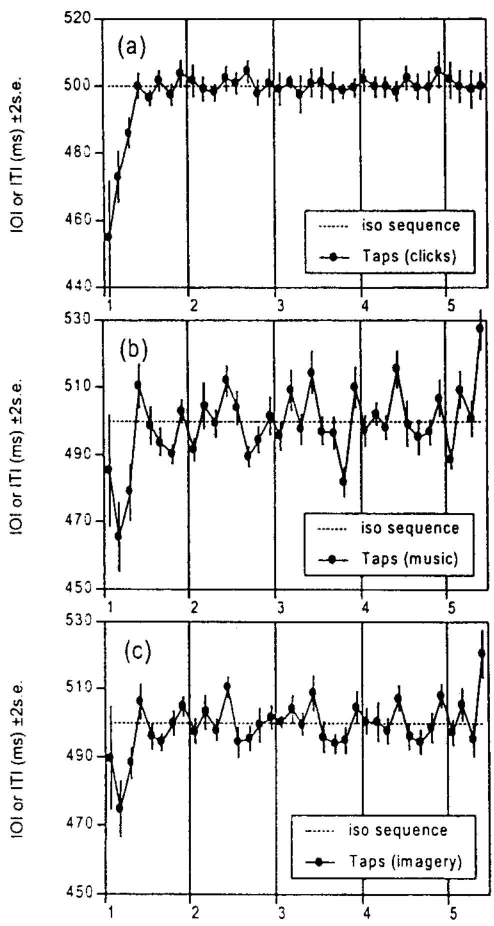

Fig. 12.4 Average inter-tap interval (ITI) profiles (with double standard errors) in the three conditions of Experiment 1. (a) Click condition. (b) Music condition. (c) Imagery condition.

same musical excerpt, but without superimposed clicks: the correlation was 0.76 ($p < 0.001$), or 0.86 if the initial three ITIs are included. Thus the earlier results for synchronization with music alone were replicated, even though the present task required synchronization with clicks that were merely accompanied by music. It appears that the effect of the musical structure on tap timing is unavoidable (see also Repp 1998b: Exp. 3). The most interesting and novel finding, however, is that this systematic tap timing pattern persisted in attenuated form in the imagery condition (Fig. 12.4(c)). Here there was again a significant deviation from uniformity from the fourth ITI on, $F(32, 352) = 9.2$, $p < 0.0001$, and the pattern correlated 0.84 with that in the music condition (Fig. 12.4(b)), or 0.91 if the initial three ITIs are included. Thus musical imagery had a significant effect on motor timing in synchronization with a perfectly isochronous click sequence.[5]

12.2.2.2 Summary

In terms of the five hypotheses outlined in the Introduction, the results may be summarized as follows. The first hypothesis was that more typical musical timing patterns would be synchronized with more accurately than less typical timing patterns when music is actually present. The predicted rank order of T1 > T2 > T4 was only partially confirmed, due to an unexpectedly (perhaps artifactually) low prediction index for T2. The second hypothesis was that all three musical patterns would be synchronized with more accurately than the R1 pattern when music was present. This was true for T1 and T4 but not for T2, for the same reason as before. The third hypothesis was that the first two predictions would also hold in the imagery condition, though perhaps less clearly. Indeed, the results in the imagery condition were similar to those in the music condition, only less pronounced.

The fourth hypothesis was that there would be significant differences among the patterns already in the click condition, due to differences in pattern complexity. Significant differences were indeed obtained, but they did not reflect differences in pattern complexity in a straightforward way. Consideration of these differences led to qualified predictions with respect to the first three hypotheses. One prediction was that synchronization with typical musical patterns should be selectively facilitated compared to less typical patterns in the music and imagery conditions. This prediction received little support. The second and most important prediction was that, in comparing the click and music conditions, synchronization with musical patterns should be selectively facilitated compared with the random pattern in the music condition. This prediction received substantial support. The third hypothesis, that the same would be true in the comparison of the click and imagery conditions, received only weak support.

Finally, the fifth hypothesis, that pattern learning would be faster for musical than for random patterns when music is present or imagined, was not supported. Instead, it appeared that music facilitated the learning of all patterns to some extent.

12.3 Experiment 2

Experiment 1 provided reasonable evidence that synchronization with complex timing patterns derived from music performance is facilitated when the appropriate music is heard or imagined, relative to a condition in which the music is neither heard nor imagined. The results were not as strong as expected, however, and this may be attributed to a methodological weakness having to do with the R1 pattern. In hindsight, it was not a good idea to employ only a single random pattern for comparison; it would have been better to use a different random pattern for each participant. Accidentally, R1 had some features in common with T1, namely long IOIs at the ends of several melodic groups (see Fig. 12.1). Thus, this pattern was not as inappropriate to the music as it could have been and may actually have received some slight facilitation from the musical context.

Experiment 2 took a different approach. Instead of constructing arbitrary timing patterns for comparison with the musical patterns, a phase-shifted version of each musical timing pattern (a method employed previously by Clarke 1993b, in an imitation study) was constructed to serve as its specific comparison. Without a musical context, the phase shift had little significance, but once the music was present or imagined, the original patterns were properly aligned with the musical structure whereas the phase-shifted patterns were not. Thus the prediction was that synchronization with each musical pattern would be selectively facilitated relative to its phase-shifted version in both the music and imagery conditions, but not in the click condition. Indeed, it was considered possible that musical context would even impair synchronization with phase-shifted patterns, relative to the click condition.

Experiment 1 provided only limited support for the hypothesis that the degree of facilitation of synchronization with musical patterns in musical contexts would be positively related to the typicality of these patterns in music performance, in the form of an advantage of T1 over T2 and T4. However, the experiment did not include the most typical musical timing pattern, T0, for which the greatest amount of facilitation should be expected. Experiment 2 included this pattern as well, at the risk of some carry-over of learning between it and the fairly similar T1 pattern (see Fig. 12.1).

Another methodological change concerned the arrangement of the three experimental conditions. In Experiment 1, all timing patterns were presented in one condition before being presented in the next one. The main advantage of this design was that participants did not hear the music until after the click condition. A possible disadvantage was the temporal separation of the music and imagery conditions, which may have weakened the strength of the musical imagery. In Experiment 2, the design was blocked by timing pattern instead. For each timing pattern, an unbroken series of trials was presented, in the course of which the three conditions followed each other in the same fixed order as previously. This design had the advantage of revealing the transitions between the three conditions more clearly, but the disadvantage that participants might feel tempted to imagine the music during the click condition, despite instructions that discouraged this strategy. The new design was motivated by the intriguing observation in Experiment 1 that the selective advantage for the musical patterns seemed to be present on the very first trial in the music condition. In Experiment 2, the immediacy of such contextual effects could be observed more directly, without any intervening breaks.

12.3.1 Methods

12.3.1.1 Materials

The materials were the same as in Experiment 1, except for the following differences. The R1 pattern was no longer employed. Instead, there were four musical patterns (T0, T1, T2, T4) and a phase-shifted version of each (T0′, T1′, T2′, T4′). The phase-shifted patterns were obtained by moving the first two IOIs (following the initial 1000-ms 'upbeat' IOI) to the end of the pattern. Thus the phase shift amounted to one-eighth note, or −90 degrees relative to the metrical cycle defined by the musical bars.[6] Table 12.2 shows that the phase-shifted versions were all atypical of expressive performance, with one (T4′) actually contradicting the most typical pattern. However, the complexity and ac1 indices were only slightly affected by the IOI manipulation. When the music accompanied

Table 12.2 Typicality indices (i.e. correlations with T0), complexity indices (i.e. 1 − ac8; see text for explanation), and lag-1 autocorrelations (ac1) for the eight timing patterns used in Exp. 2

Pattern	*Typicality*	*Complexity*	*ac1*
T0	1.00	0.53	0.19
T0′	−0.17	0.31	0.29
T1	0.67	0.31	0.29
T1′	−0.03	0.14	0.35
T2	0.46	0.73	0.59
T2′	−0.15	0.62	0.59
T4	0.36	0.58	0.15
T4′	−0.54	0.60	0.15

the clicks, it started and stopped with the click sequence and followed the same timing pattern. An isochronous pattern was also included, mainly for practice but also to replicate the intriguing effect of musical imagery on tap timing (Fig. 12.4(c)).

12.3.1.2 Participants

Twelve musically trained Yale undergraduates were paid for their participation. Nine of them were players of string instruments in the Yale Symphony Orchestra. Five of them had participated in Experiment 1, but since one year had elapsed between experiments, no carry-over of learning was expected. The remaining three participants had less advanced musical training but instead had considerable practice in synchronization tasks.

12.3.1.3 Procedure

Each temporal pattern was repeated 20 times, with 3 s of silence between repetitions. Trials 1–8 constituted the click condition; the first two of these trials were considered practice and were not analyzed. Trials 9–14 constituted the music condition, and trials 15–20 the imagery condition. Participants were urged *not* to imagine the music during the initial 8 trials; otherwise, the instructions were the same as in Experiment 1. The isochronous pattern was always presented first, and the remaining 8 patterns were presented in an order that was counterbalanced across participants according to 1.5 Latin squares, constructed so that original and phase-shifted patterns alternated and the other three patterns intervened between the original and phase-shifted versions of the same pattern. The musical score was in view throughout the experiment.

12.3.2 Results and discussion

The data were again analyzed in terms of the three indices of synchronization accuracy, but only the results for $r0^*$ are reported here. The results in terms of sda were similar, whereas $r1^*$ again yielded a less clear picture. The ANOVAs were largely analogous to those in Experiment 1 but included the variable of version (2) in addition to pattern, condition, and trial.

Figure 12.5 shows that the difference between the original and phase-shifted versions of each timing pattern increased substantially in favor of the original version in the music condition relative to the click condition, and decreased again in the imagery condition. The Condition × Version

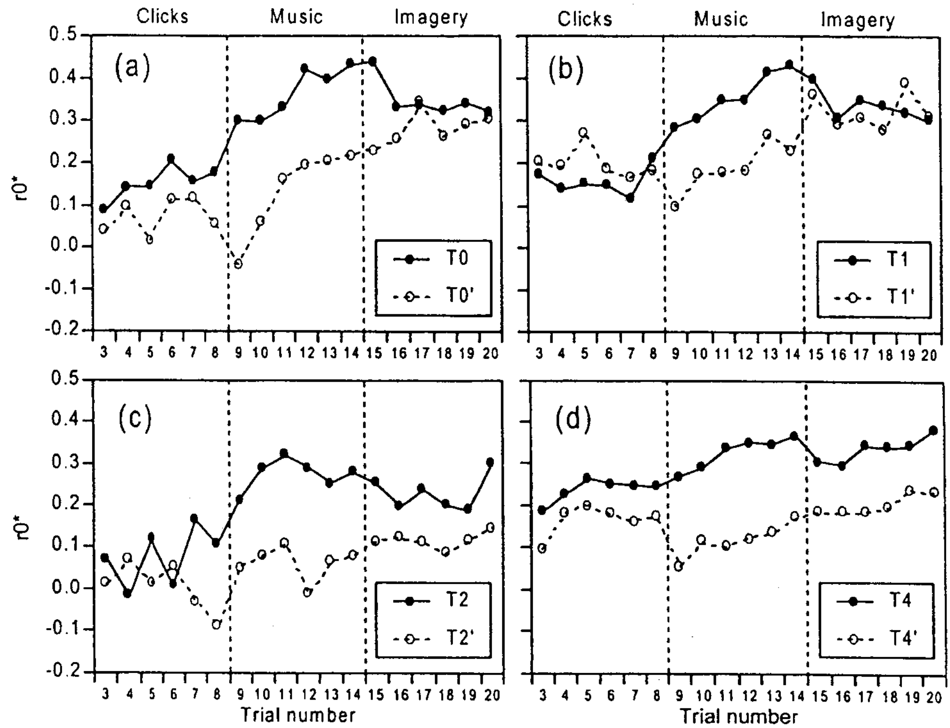

Fig. 12.5 Average prediction indices ($r0^*$) as a function of trial number for eight timing patterns in Experiment 2. (a) T0 and T0′. (b) T1 and T1′. (c) T2 and T2′. (d) T4 and T4′.

interaction was highly significant when comparing the click and music conditions, $F(1, 11) = 24.0$, $p < 0.0006$, and also when comparing the music and imagery condition, $F(1, 11) = 36.6$, $p < 0.0002$, but not when comparing the click and imagery conditions, $F(1, 11) = 1.5$, $p < 0.25$. The triple interaction with pattern was not significant, indicating that the Condition × Version interaction was similar for all four patterns. The main effect of version in favor of the original patterns was significant not only in the comparisons involving the music condition but also in the comparison of the click and imagery conditions, $F(1, 11) = 5.0$, $p < 0.05$; however, it did not change significantly between these two conditions, nor did it interact significantly with pattern in any condition.

Prediction performance increased across trials in the music condition, $F(5, 55) = 7.9$, $p < 0.0001$, but not in the click and imagery conditions, where the main effect of trials was nonsignificant. This was also reflected in a significant Condition × Trials interaction for clicks vs. music, $F(5, 55) = 2.9$, $p < 0.03$, and for music vs. imagery, $F(5, 55) = 5.8$, $p < 0.0003$, but not for clicks vs. imagery. In the music condition, there was also a Pattern × Trials interaction, $F(15, 165) = 2.4$, $p < 0.004$. The largest improvement over trials was shown by T0/T0′ and the smallest by T2/T2′. Note that original and phase-shifted versions improved at the same rate; the Version × Trials interaction was nonsignificant.

12.3.2.1 Average timing profiles

The average ITI profiles for the isochronous sequence were extremely similar to those of Experiment 1 (Fig. 12.4) and therefore are not shown separately. Apart from the initial three ITIs (which were omitted in the statistical analyses), there was no significant deviation from uniformity in the click condition, $F(32, 352) = 1.2$, $p < 0.23$, whereas there were highly significant deviations in both the music condition, $F(32, 352) = 9.9$, $p < 0.0001$, and the imagery condition, $F(32, 352) = 4.8$, $p < 0.0001$. The pattern of the deviations was highly similar in these two conditions, $r(31) = 0.81$, $p < 0.0001$, although it was less pronounced in the imagery condition, and the ITI profiles also correlated highly with those obtained in Experiment 1, $r(31) = 0.93$ and 0.80, $p < 0.0001$, for the music and imagery conditions, respectively. These results indicate that participants were imagining the music correctly in the imagery condition, even though they had heard it only 6 times at that point in the experiment.

One curious result worth noting is that the (nonsignificant) pattern of deviations from regularity in the click condition exhibited some resemblance to the (significant) patterns obtained in the music and imagery conditions, $r(31) = 0.52$ and 0.50, $p < 0.01$, respectively. This had not been the case in Experiment 1. In neither experiment had the participants yet heard the music. However, the participants in Experiment 2 knew which music they would be hearing subsequently and had the musical score in front of them. Thus it is possible that some participants imagined the music spontaneously from the notation (cf. Brodsky, Henik, Rubinstein, and Zorman 1998), especially since emphatic instructions *not* to imagine the music during the click condition were given only after the isochronous practice sequence.[7]

12.3.2.2 Summary

The results of Experiment 2 largely confirm those of Experiment 1. The first hypothesis, predicting that more typical original timing patterns would be synchronized with more accurately than less typical original timing patterns when music was actually present, received some support in that performance for T0 and T1 was better than for T2 and T4. The second hypothesis, that all original patterns would be synchronized with more accurately than their phase-shifted versions when music was present, was strongly supported. The third hypothesis, that the first two predictions would also hold in the imagery condition, was supported in the case of the second prediction only.

The fourth hypothesis, that there would be significant differences among the patterns in the click condition, was supported, but not for the reason originally envisioned: differences in pattern complexity did not seem to play an important role. The qualified predictions of the first two hypotheses, which take differences among patterns in the click condition into account, were strongly confirmed in that more typical patterns benefited more from musical context than less typical (original) patterns, and especially in that original patterns benefited more than phase-shifted patterns. However, the qualified third hypothesis, concerning imagery, was not supported by the results of Experiment 2, which did not demonstrate a selective benefit of musical imagery for musically appropriate patterns. The fifth hypothesis, that pattern learning would be faster for original than for phase-shifted patterns when music is present or imagined, was not supported. Instead, in agreement with Experiment 1, the results suggested that audible music facilitates the learning of timing patterns regardless of their appropriateness.

12.4 General discussion

The present study investigated the ability of musically trained participants to synchronize a simple motor activity with complex timing patterns derived from expressive timing in music performance.

These patterns were of a kind not previously investigated in pattern learning or synchronization tasks: they were neither isochronous nor rhythmic nor random (except in one case), but are best described as semi-regular or quasi-periodic in various degrees. Their regularities derived from their original association with a musical structure.

In the click condition, especially in Experiment 1 (where the participants were unaware that music would be introduced later) but also in Experiment 2 (to the extent that the participants followed instructions to refrain from musical imagery), the question of interest was whether the regularities inherent in the timing patterns would help participants to learn and predict the timing variations to some extent. The participants' success in this task was limited, which is not surprising in view of the small number of trials (10 in Exp. 1, 8 in Exp. 2). Their synchronization performance was characterized primarily by tracking, which, as Michon (1967) and others have shown, is the characteristic response to unpredictable temporal patterns in a synchronization task. Only in Experiment 1 was there evidence for improvement across trials within the click condition. However, this improvement did not differ among timing patterns, which suggests a general learning effect that was independent of pattern structure. Nevertheless, there were significant differences among patterns from the very beginning in the click condition. For example, the T1 and T4 patterns exhibited larger $r0^*$ indices than T2 in Experiment 1 (Fig. 12.2(a)), and T4 was more predictable than all other patterns in Experiment 2 (Fig. 12.5). The reasons for these differences are not well understood at present. Differences in pattern complexity, defined here as the degree of periodicity, did not seem to be the only cause.

The music condition was the primary focus of interest here. The main hypothesis was that complex timing patterns derived from expressive music performance, which are quite meaningless when carried by a click sequence, would suddenly gain meaning and structural support when they are appropriately instantiated by the accompanying music, and that this would automatically facilitate pattern prediction in synchronization. Synchronization with random or phase-shifted timing patterns, by contrast, was not expected to benefit from the musical context. These predictions received strong confirmation in both experiments. In Experiment 2, all four original patterns were shown to benefit much more from the musical context than their phase-shifted versions. In fact, three of the phase-shifted patterns seemed to suffer interference from the music, at least on the first music trial (see Fig. 12.5).

These effects evidently derive from the relative compatibility of the timing patterns with the musical structure, particularly with the rhythmic grouping in the melody (cf. Clarke 1985, 1993b). Auditory perception of musical structure primes certain action patterns that are expressive of that structure, and timing is the most important characteristic of these action patterns. Shin and Ivry (1999), in a recent study, proposed a similar explanation for their finding that incidental learning of arbitrary temporal patterns occurred only when these patterns were systematically associated with a constant action pattern, in their case spatial hand movements in response to visual stimuli. (One of their manipulations also involved a phase shift of the temporal pattern relative to the spatial pattern.) Timing is a property of actions or events. In the case of music, appropriate actions are implied by the sound structure which defines compatibility with regard to timing.

Previous research has demonstrated that the most typical timing pattern, T0, has a privileged relation to the musical structure: it is representative of many individual performances (Repp 1998a); it is observed when pianists try to play in strict time (Repp 1999a,c); it is aesthetically appealing (Repp 1997); it biases imitation of expressive timing (Repp 2000b); and perception of timing in music exhibits strong biases whose pattern closely resembles T0 (Repp 1998b,c,d) and which have been attributed to basic auditory and/or motor grouping processes (Penel and Drake 1998; Repp 1998d). Therefore, it was of special interest to see whether timing patterns other than T0 would be selectively facilitated by the music in the present task. There was clear evidence for facilitation of T1 in both experiments, but that pattern is moderately correlated with T0 and hence fairly typical as well. By contrast, the T2 and T4 patterns are of low typicality, although they do resemble the expressive timing patterns of some outstanding pianists. Nevertheless, selective facilitation of these patterns by the music did occur. This seems to be the first demonstration, other than by the original pianists' performances themselves, that radically different timing patterns can be compatible with the same musical structure, as hypothesized by Repp (1998a).

The differences between the original timing patterns and their phase-shifted versions can also be viewed in terms of relative typicality. The seemingly equal benefit bestowed by musical context on the four original patterns relative to their phase-shifted versions may be a consequence of the fact that the lower relative typicality of the phase-shifted versions varied in parallel with the higher relative typicality of the original patterns (see Table 12.2).

An interesting and somewhat unexpected finding was that synchronization performance improved more during the music condition than during the preceding click condition or the following imagery condition, and that this improvement occurred regardless of pattern typicality. It appears that the musical context provided a structural framework that facilitated pattern learning, regardless of the appropriateness of the pattern. In other words, the temporal pattern could be 'pegged to' the musical structure, which served as a memory aid. This process presumably also accounts for musicians' ability to reproduce structurally inappropriate timing patterns reasonably well in an imitation task (Clarke 1993b; Repp 2000b).

The present study also addressed the question of whether musical imagery can have behavioral effects similar to those of music actually heard. In the present context, musical imagery refers to the generation of auditory and/or motor images from a memory representation of recently heard music. Basically, this amounts to an 'inner singing' of the melody, perhaps with accompaniment notes filled in where there are no melody note onsets. How vivid or detailed the participants' imagery was is not known. What is clear from the results, however, is that imagery was not an effective substitute for hearing the music. Evidence for a benefit due to imagined music was weak in Experiment 1 and effectively absent in Experiment 2. This is a somewhat disappointing result, but it may simply indicate that the participants' imagery was not strong enough. In Repp's (1999a) study, skilled pianists who had played the Chopin Etude excerpt earlier in the same experimental session were capable of equally accurate synchronization performance in music and imagery conditions; their synchronization performance was also much better overall as that of the present participants. Thus, more experienced or more practiced individuals may well show a clearer benefit of musical imagery.

In Experiment 2, a tendency of some participants to imagine the music during the click condition may have worked against finding a relative benefit in the imagery condition. Indeed, 45% of the trials in the click condition did not exhibit an extra tap at the end, which indicates that the end of the sequence had not come as a surprise. However, strategies other than outright imagery (e.g. counting, grouping, or memory for local temporal pattern features near the end) could also have been responsible. By contrast, 96% of the imagery trials ended without an extra (or missing) tap, which suggests that the music was imagined correctly, though perhaps not vividly enough, in synchrony with the clicks.

Evidence that musical imagery occurred also comes from the results with isochronous sequences. Here imagery induced systematic deviations from regularity in the finger taps, similar to those that are observed in tapping to isochronous music (Repp 1999a,b) or, as also demonstrated here, to isochronous clicks accompanied by isochronous music. The deviations induced by imagery were smaller than those induced by real music, which again shows that imagery was less effective than hearing the actual sound. However, the finding that musical imagery can have involuntary effects on motor timing is theoretically interesting. It suggests a close connection between musical imagery and movement timing, just as there is a close connection between music perception and movement timing. The pattern of systematic deviations from regularity in tapping may represent a combination of expressive tendencies and automatic error correction, which is required to maintain synchronization. This issue is in need of further research, however.

Automatic error correction is also responsible for the tracking tendency which dominated synchronization performance, especially in Experiment 2. Tracking is the consequence of unsuccessful synchronization, where each large asynchrony is partially corrected on the next tap while simultaneously a new large asynchrony may arise from the unpredicted time of occurrence of the next tone. The underlying mechanism is likely to be phase correction (Mates 1994; Pressing 1998; Vorberg and Wing 1996), which is an obligatory process that commonly occurs without awareness (Repp 2000a). A second error-correction mechanism hypothesized to underlie synchronization performance, timekeeper period correction (Mates 1994), probably does not play any important role in tracking as long as the average tempo of the sequence is constant, as it was in the present experiments. However, the period correction mechanism may well be responsible for the prediction of a learned pattern. In other words, remembered aspects of timing patterns as well as perceived or imagined musical structure may influence tap timing via intentional or unintentional modulations of the timekeeper period. Period correction may in part be a top-down mechanism which mediates temporal expectations and governs intentional temporal control, whereas phase correction is largely bottom-up and input-driven (Repp, 2001). If this interpretation is correct, then the error correction mechanisms that have been identified in simple synchronization tasks may have broader implications for temporal pattern learning, motor control, and perception. Indeed, Large and Jones (1999) have proposed a perceptual model of beat tracking that incorporates analogous mechanisms. The possible parallels between error correction processes in perception and production of timing warrant further study.

Acknowledgments

This research was supported by NIH grant MH-51230. I am grateful to Paul Buechler and Steve Garrett for assistance. Address correspondence to Bruno H. Repp, Haskins Laboratories, 270 Crown Street, New Haven, CT 06511–6695 (e-mail: *repp@haskins.yale.edu*).

Notes

1. Throughout this article, the term 'clicks' is used to refer to what was actually a series of very high-pitched digital piano tones (see Methods).
2. The term prediction rather than anticipation is used to avoid confusion with the anticipation tendency (i.e. taps precede sequence events) commonly observed in synchronization tasks (see, e.g. Aschersleben and Prinz 1995). As the terms are used here, prediction is pattern-specific whereas anticipation is not.
3. The third principal component was of little interest because it mainly consisted of a greatly lengthened IOI following the initial downbeat (cf. Fig. 12.1(a)).

4. A MAX patch is a program written in the graphic MAX programming environment. Due to a peculiarity of that software, real-time durations were 2.4% shorter than specified or recorded, and than reported here.
5. Figures comparing the IOIs and average ITIs for the four modulated timing patterns may be found in an electronic appendix to this article on the author's web page: <www.haskins.yale.edu/haskins/STAFF/repp.html>
6. A graphic example of T0 and T0′ is included in the electronic appendix.
7. For a figure illustrating the average ITI profiles for the T0 and T0′ timing patterns, see the electronic appendix.

References

Aschersleben, G. and Prinz, W. (1995). Synchronizing actions with events: The role of sensory information. *Perception and Psychophysics, 57*, 305–317.

Behne, K.-E. and Wetekam, B. (1993). Musikpsychologische Interpretationsforschung: Individualität und Intention. *Musikpsychologie, 10*, 24–37.

Brodsky, W., Henik, A., Rubinstein, B., and Zorman, M. (1998). Demonstrating inner hearing among highly-trained expert musicians. In S.W. Yi (Ed.), *Proceedings of the Fifth International Conference on Music Perception and Cognition*, pp. 237–242. Seoul, Korea: Western Music Research Institute.

Clarke, E.F. (1985). Structure and expression in rhythmic performance. In P. Howell, I. Cross, and R. West (Eds.), *Musical structure and cognition*, pp. 209–236. London: Academic Press.

Clarke, E.F. (1993a). Generativity, mimesis and the human body in music performance. *Contemporary Music Review, 9*, 207–219.

Clarke, E.F. (1993b). Imitating and evaluating real and transformed musical performances. *Music Perception, 10*, 317–342.

Clynes, M. (1983). Expressive microstructure in music, linked to living qualities. In J. Sundberg (Ed.), *Studies of music performance*, pp. 76–181. Stockholm: Royal Swedish Academy of Music.

Drake, C. and Palmer, C. (1993). Accent structures in music performance. *Music Perception, 10*, 343–378.

Fraisse, P. (1966). L'anticipation de stimulus rythmiques: Vitesse d'établissement et précision de la synchronisation. *L'Année Psychologique, 66*, 15–36.

Hary, D. and Moore, G.P. (1985). Temporal tracking and synchronization strategies. *Human Neurobiology, 4*, 73–77.

Hary, D. and Moore, G.P. (1987). Synchronizing human movement with an external clock source. *Biological Cybernetics, 56*, 305–311.

Iyer, V.S. (1998). Microstructures of feel, macrostructures of sound: Embodied cognition in West African and African-American musics. Unpublished doctoral dissertation, University of California, Berkeley.

Large, E.W. and Jones, M.R. (1999). The dynamics of attending: How we track time-varying events. *Psychological Review, 106*, 119–159.

Mates, J. (1994). A model of synchronization of motor acts to a stimulus sequence. I. Timing and error corrections. *Biological Cybernetics, 70*, 463–473.

Michon, J.A. (1967). *Timing in temporal tracking*. Assen, NL: van Gorcum.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 331–346.

Penel, A. (2000). Variations temporelles dans l'interpretation musicale: Processus perceptifs et cognitifs. Unpublished Ph.D. dissertation, University of Paris 6.

Penel, A. and Drake, C. (1998). Sources of timing variations in music performance: A psychological segmentation model. *Psychological Research, 61*, 12–32.

Penel, A. and Drake, C. (1999). Seeking 'one' explanation for expressive timing. In S.W. Yi (Ed.), *Music, mind, and science*, pp. 271–297. Seoul, Korea: Seoul National University Press.

Pierce, A. and Pierce, R. (1989). *Expressive movement: Posture and action in daily life, sports, and the performing arts*. New York: Plenum Press.

Pressing, J. (1998). Error correction processes in temporal pattern production. *Journal of Mathematical Psychology, 42*, 63–101.

Repp, B.H. (1992a). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's 'Träumerei'. *Journal of the Acoustical Society of America, 92*, 2546–2568.

Repp, B.H. (1992b). Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations. *Cognition, 44*, 241–281.

Repp, B.H. (1993). Music as motion: A synopsis of Alexander Truslit's 'Gestaltung und Bewegung in der Musik' (1938). *Psychology of Music, 21*, 48–72.

Repp, B.H. (1997). The aesthetic quality of a quantitatively average music performance: Two preliminary experiments. *Music Perception, 14*, 419–444.

Repp, B.H. (1998a). A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 104*, 1085–1100.

Repp, B.H. (1998b). Obligatory 'expectations' of expressive timing induced by perception of musical structure. *Psychological Research, 61*, 33–43.

Repp, B.H. (1998c). The detectability of local deviations from a typical expressive timing pattern. *Music Perception. 15*, 265–290.

Repp, B.H. (1998d). Variations on a theme by Chopin: Relations between perception and production of deviations from isochrony in music. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 791–811.

Repp, B.H. (1999a). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior, 31*, 145–164.

Repp, B.H. (1999b). Detecting deviations from metronomic timing in music: Effects of perceptual structure on the mental timekeeper. *Perception and Psychophysics, 61*, 529–548.

Repp, B.H. (1999c). Relationships between performance timing, perception of timing perturbations, and perceptual–motor synchronization in two Chopin preludes. *Australian Journal of Psychology, 51*, 188–203.

Repp, B.H. (2000a). Compensation for subliminal timing perturbations in perceptual–motor synchronization. *Psychological Research, 63*, 106–128.

Repp, B.H. (2000b). Pattern typicality and dimensional interactions in pianists' imitation of expressive timing and dynamics. *Music Perception, 18*, 173–211.

Repp, B.H. (2001). Processes underlying adaptation to tempo changes in sensorimotor synchronization. *Human Movement Science, 20*, 277–312.

Schulze, H.-H. (1992). The error correction model for the tracking of a random metronome: Statistical properties and an empirical test. In F. Macar, V. Pouthas, and W.J. Friedman (Eds.), *Time, action, and cognition*, pp. 275–286. Dordrecht, The Netherlands: Kluwer.

Semjen, A., Vorberg, D., and Schulze, H.-H. (1998). Getting synchronized with the metronome: Comparisons between phase and period correction. *Psychological Research, 61*, 44–55.

Shin, J.C. and Ivry, R.B. (1999). Concurrent temporal and spatial learning in a serial reaction time task. Poster presented at the Annual Meeting of the Psychonomic Society, Los Angeles, CA.

Shove, P. and Repp, B.H. (1995). Musical motion and performance: Theoretical and empirical perspectives. In J. Rink (Ed.), *The practice of performance*, pp. 55–83. Cambridge, UK: Cambridge University Press.

Thaut, M.H., Tian, B., and Azimi-Sadjadi, M.R. (1998). Rhythmic finger tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment. *Human Movement Science, 17*, 839–863.

Viviani, P. and Stucchi, N. (1992). Motor–perceptual interactions. In G.E. Stelmach and J. Requin (Eds.), *Tutorials in motor behavior II*, pp. 229–248. Amsterdam: Elsevier.

Vorberg, D. and Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer and S. W. Keele (Eds.), *Handbook of perception and action*, Vol. 2, pp. 181–262. London: Academic Press.