Psychological Research (2011) 75:243–258
DOI 10.1007/s00426-010-0302-7

ORIGINAL ARTICLE

Learning of timing patterns and the development of temporal expectations

Barbara Tillmann · Catherine Stevens · Peter E. Keller

Received: 26 April 2010 / Accepted: 13 July 2010 / Published online: 4 August 2010

Abstract The present study investigated the learning of a culturally unfamiliar musical rhythm, leading to the development of temporal expectations, and it explored the potential for generalization across tempi and tasks. With that aim, we adapted the serial reaction time task to examine the learning of temporal structures by an indirect method. The temporal pattern employed was based on a complex interval ratio (2:3) and compared to one based on a simple interval ratio (1:2). In the exposure phase, non-musician participants performed a two-choice speeded discrimination task that required responding by key press to each event of the simple or complex auditory pattern. Participants were not informed about the temporal regularities; their task solely concerned the discrimination task. During exposure (Experiments 1–3), response times decreased over time for both temporal patterns, but particularly for the events following the longer interval of the more complex 2:3 pattern. Exposure further influenced performance in subsequent testing phases, notably the precision of tap timing in a production task (Experiment 2) and temporal expectations in a perception task (Experiment 3). Our findings promote the new paradigm introduced here as a method to investigate the learning of temporal structures.

B. Tillmann
University of Lyon, CNRS UMR 5020, Lyon, France

B. Tillmann · C. Stevens
School of Psychology and MARCS Auditory Laboratories, University of Western Sydney, Sydney, Australia

P. E. Keller
Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

B. Tillmann (✉)
Université Claude Bernard, Lyon I, CNRS UMR 5020
Neurosciences Comportement Cognition, 50 Av. Tony Garnier, 69366 Lyon Cedex 07, France
e-mail: barbara.tillmann@olfac.univ-lyon1.fr

Introduction

Events and actions fundamental to human behavior and culture unfold in time and have onsets, rhythms, tempos, and endings. Temporal processing is required for speech perception and production, motor learning, social interaction, sport, music, and dance. One way that humans learn to anticipate events is through exposure and the development of sensitivity to statistical regularities. More specifically, infants and adults show a capacity to attune to pitch and temporal regularities—prosody, rhythm—of spoken language (e.g., Patel & Daniele, 2003; Thiessen & Saffran, 2003), and to develop musical expectations from the pitch dimension in tonal music and temporal expectations from musical rhythms and meters (Jones & Boltz, 1989; Tillmann, Bharucha, & Bigand 2000). In the laboratory, the learning of such regularities has been investigated with reduced and controlled versions of the structures in these complex, real-world patterns. The serial reaction time (SRT) task is one paradigm employed to investigate incidental learning of sequential relations between events (e.g., Nissen & Bullemer, 1987). However, little research attention has been devoted to the learning of temporal structures, and there have been no systematic investigations of learning temporal structures relevant to music using this paradigm. Suitably reduced temporal patterns characteristic of music from different cultures can provide a window into the adult capacity to process temporal structures and develop temporal expectations. The use of unfamiliar structures ensures that listeners have little prior knowledge of the material and allows investigation of the acquisition of these temporal structures under controlled conditions in the laboratory.

The vast majority of Western tonal music from classical concert and pop music repertoires (which are most often the focus of music cognition research) contain patterns with underlying beats that are separated by intervals of the same duration and equally spaced metric accents (relatively strong beats). However, musical cultures from parts of the Balkan Peninsula, West Africa, the Middle East, Asia and Latin America use rhythmic structures, in which both binary and ternary units can occur at a single metric level (Kauffman, 1980; London, 1995; Pressing, 1983). Listeners exposed predominantly to Western music develop relatively simple temporal expectations. While they can reproduce rhythms with a 2:1 ratio quite accurately, they find asymmetric patterns (e.g., based on a 3:2 ratio) difficult to produce, perceive and memorize accurately (Povel & Essens, 1985; Repp, London, & Keller, 2005; Semjen & Ivry, 2001; Summers, Bell, & Burns, 1989; Summers, Hawkins & Mayers, 1986). They tend to simplify these rhythms (Collier & Wright, 1995; Desain & Honing, 2003; Fraisse, 1956; Povel, 1981; Sternberg, Knoll, & Zukovsky, 1982); for example, stimuli in which the longer interval is less than double the duration of the shorter interval tend to be reproduced with a long/short ratio of 2:1. Even individuals who have completed long-term musical training in the Western tradition (i.e., skilled pianists or percussionists) are less accurate when coordinating movements with irregular (uneven) rhythms in comparison to regular rhythms (e.g., Fitch & Rosenfeld, 2007; Keller & Repp, 2005; Patel, Iversen, Chen & Repp, 2005).

The influence of cultural background on temporal processing has been demonstrated recently (Hannon & Trehub, 2005a, b): American adults succeeded in processing binary (even) rhythmic structures (i.e., 222), while they showed difficulties in processing culturally unfamiliar (uneven) rhythmic structures based on either 223- or 322-patterns (with 2 and 3 referring to duple and triple multiples of a basic temporal unit, 1). However, Macedonian and Bulgarian adults whose cultures also include uneven rhythms, succeeded equally well for the 223- and 322-patterns as for the 222-patterns. Hannon and Trehub (2005a) have also investigated the learning of culturally unfamiliar patterns in American listeners: brief exposure to Balkan folk music between two test sessions allowed 12-month-old infants, but not adults, to perceive rhythmic variations of 223- or 322-patterns. These findings have been interpreted as indicating a sensitive period early in life for acquiring culturally relevant rhythmic structures. Our present study aimed to reassess the learning abilities of Western (here Australian) adult listeners for unfamiliar 223-patterns by adapting a paradigm previously used in structural sequence learning (the SRT paradigm) and to explore the underlying mechanism by testing acquired temporal pattern knowledge additionally in subsequent test phases.

Providing evidence for structure learning with the SRT paradigm

Experiments using the SRT paradigm have provided evidence that participants acquire knowledge about structured sequences and that they are able to use this knowledge to predict *what* kind of event will come next, leading to faster response times (RTs). The SRT paradigm has been used with sequences based on visual events (e.g., lights), spatial locations and, less frequently, auditory events. Our study applied this paradigm to the learning of temporally structured patterns, to investigate whether participants acquire knowledge allowing them to predict *when* the next event will occur.

In the classical SRT paradigm, participants make a simple response to each element of sequentially structured sequences in the context of a choice reaction time task. They press a key that corresponds to the currently presented element, usually a stimulus light that appears at a given location. Unknown to participants, the sequence of successive events follows a repeating pattern or is governed by rules of permissible transitions (e.g., Nissen & Bullemer, 1987). Over the experimental blocks, RTs to events that respect the repeating pattern become faster. To separate motor- and task-related learning components, sequence learning has been measured with various tests. For example, participants presented with structured material produce faster RTs than those presented with random material, suggesting that the former can better prepare responses as a result of acquired knowledge of the structured pattern (e.g., Cleeremans & McClelland, 1991). Alternatively, after a series of experimental blocks respecting the sequence structure, a block with a new sequence is introduced and RTs slow down with this change (e.g., Destrebecqz & Cleeremans, 2001). Another testing method consists of presenting—after the sequenced experimental blocks—various tasks assessing recall, recognition or generation using structured or unstructured material (often with the additional aim of disentangling explicit and implicit knowledge components; e.g., Buchner & Steffens, 2001; Reed & Johnson, 1994).

When applied to the auditory modality, thus using tones instead of lights (Buchner, Steffens, Erdfelder, & Rothkegel, 1997; Perruchet, Bigand, & Benoit-Gonin, 1979, Exp. 3), "temporal" characteristics have concerned sequential patterns with different elements (i.e., successive chaining), but rarely the processing of timing, durations or intervals.

Buchner and Steffens (2001) combined tones varying in pitch with the learning of temporal intervals between-participants' responses and the presentation of the next tone (i.e., response–signal interval, RSI). These temporal sequences linked to RSI were dependent on the delay of the responses after the tones and, as response latencies varied, the temporal regularities between the tones were variable. Buchner and Steffens (2001) showed that temporal structures were only learned when they correlated with the pitch structures. This finding suggests that participants were unable to learn temporal structures (i.e., to develop temporal expectations about *when* the next event occurs) independently of event structures (expectations about *what* kind of event occurs next). Similarly, Shin and Ivry (2002) reported that temporal structures were learned solely when correlated with spatial structures, and this finding was observed not only for temporal structures based on RSI, but also for temporal structures based on intervals between the signals' onsets (stimulus onset asynchronies, SOAs). Interestingly, the authors discussed that learning of temporal patterns might be easier when the patterns form hierarchical patterns with beat structures.

The question remains whether non-musician participants can learn regularities based on the time dimension only. Specifically, it remains to be shown whether: (1) participants can learn temporal patterns independently of the regularities of the event types (e.g., location) marking them; and (2) whether the learned temporal patterns can be independent of the absolute timing, which can be tested by their presentation at a different tempo. Generalization to a different tempo would suggest a more abstract coding. The present study aimed to investigate these issues with a modified version of the SRT paradigm and by using musically relevant timing ratios (based on SOAs). Particularly, we investigated whether short-term exposure improves temporal processing of culturally unfamiliar patterns, such as 223-patterns, which cannot be successfully assimilated into familiar (culturally relevant) metrical frameworks. To that end, our SRT paradigm employed a unique combination of features including musical (i.e., non-arbitrary) rhythms, fixed SOAs, and overt motor responses.

Some evidence for the learning of temporal patterns without concurrent, correlated event structures (e.g., based on locations or tones) has been provided by Salidis (2001). Testing for temporal learning with structures based on RSIs, she used an adaptation of the SRT task using tones (the same tone presented repeatedly). Instead of asking participants to press buttons associated with a pre-defined sequence of spatial locations (one button associated with one spatial location), Salidis created a syncopation task with participants pressing a single button after each tone of the temporal pattern. The findings showed temporal structure learning, with greater decreases in RTs recorded from participants working on the temporally structured RSIs in comparison with participants working on random RSIs, though the effect was most strongly observed in response to the shortest interval. In Salidis' (2001) study, the temporal regularities of the RSIs were uncharacteristic of most music because of their symmetrical structure and arbitrary time intervals [e.g., the pattern 121323 based on intervals of 180 ms (1), 450 ms (2) and 1,125 ms (3)]. A pilot test that we conducted used a fixed pattern of SOAs (instead of a fixed pattern of RSIs) and revealed Salidis' adaptation of the SRT task to be challenging when shorter and more rhythmical musical intervals were used; it was difficult to wait rather than synchronize with the tones. Synchronization (responding at the same time as the stimulus) is easier than syncopation or interpolation when the temporal pattern is predictable (Fraisse, 1982; Keller & Repp, 2004), but synchronization necessarily causes failure in an SRT task with complex time patterns. The present study introduces a new paradigm that elicits the necessary syncopated (i.e., respond after the tone), rather than synchronized response.

A new SRT paradigm with rhythmic structures

In keeping with the incidental (non-intentional) nature of the task as employed in previous SRT research (e.g., Bremner, Mareschal, Destrebecqz, & Cleeremans, 2006), we developed a paradigm that does not require participants to respond to the temporal feature of the sequence, but rather to a feature other than event timing. To achieve this goal, the sequences were constructed from two syllables, chained randomly and whose onsets followed a particular temporal pattern. The task was to respond as quickly and as accurately as possible to each event according to whether the syllable was "ta" or "pa". For each event, participants pressed one of two response keys. This two-alternative forced choice (2AFC) task simply provided a medium for the presentation of the temporal patterns and to make participants respond to a feature other than a temporal one. No reference was made to the rules that govern the temporal structure of the sequence. As in the classical SRT paradigm, participants were not informed about learning and structure; their task was solely to make speeded responses. We selected syllables for this indirect task as syllables provided a non-musical cover story with syllable discrimination being an easy task for non-musician participants and not requiring participants' attention to the temporal dimension of the sequence. This represented an advantage over the use of tones with different pitches that create melodic lines with contour and rhythms, which might be processed in an integrated way (e.g., Jones & Boltz, 1989).

Using relatively simple temporal patterns, the aim was to investigate the incidental learning of temporal structures, and the influence of these structures on temporal processing and expectancy formation, in non-musician adult listeners. The rhythmic patterns investigated here were SOA-patterns of 223 versus 224, with 2, 3 and 4 referring to duple, triple and quadruple multiples of a basic temporal unit (1). 224 represents a simple pattern based on the simple integer ratio 1:2, while 223 is more complex (2:3 or 1:1.5) and an uneven pattern, uncharacteristic of Western tonal music. The 223-pattern was the experimental pattern that was the focus of the current investigation of sequence learning. The 224-pattern served as a comparison (or control) condition for the 223-pattern: first, participants performed the syllable discrimination task in an exposure phase, thus providing a baseline for exposure- and task-related influences in the second phase of the experiment (see Experiments 2 and 3); second, the 224-pattern also contained a short–short–long interval structure as does the 223-pattern. The two short–short–long interval structures thus have the same perceptual groupings based on the longer interval. However, the 224-pattern has an even meter (i.e., reinforcing the cultural bias of processing more simple, binary meter patterns) rather than an uneven meter as does the 223-pattern. If participants perceive not only the grouping structures, but also the exact rhythmic structure (i.e., the exact duration of the third, longer interval) and its underlying meter (e.g., Handel, 1992; Hébert & Cuddy, 2002), then performance differences should be observed between the two exposure patterns, and the exposure patterns should differentially influence performance on the subsequent tasks (see below).

In three experiments, two groups of participants received exposure to either the experimental 223-pattern or the control 224-pattern, both presented in the form of the syllable discrimination task. If structure learning occurs, then listeners will come to anticipate the temporal occurrence of the next syllable and respond more quickly as exposure progresses. Experiment 1 introduced this new paradigm and tested this basic prediction (see below for details). Experiments 2 and 3 tested whether knowledge acquired during the exposure phase generalizes to more accurate processing of temporal patterns in a second phase of the experiment. Experiment 2 examined the production of even and uneven patterns in synchronization and continuation tasks. Experiment 3 tested for temporal expectations in a temporal priming paradigm by presenting target syllables at the end of a short, structured sequence intended to provide a temporal context. To investigate whether temporal sequence structures and interval ratios had been learned (rather than interval durations only), patterns in these test phases were played at tempi that differed from that of the exposure phase. It was hypothesized that 223-exposure might help to decrease the cognitive cost of 223-pattern processing. More generally, the exposure phase might lead to a non-specific advantage of temporal processing, and particularly for a simple, even test pattern (i.e., 222). This advantage should be observed for the 223-exposure group, but also (or even more strongly) for the 224-exposure group (i.e., serving as a control group with exposure). For this reason, an additional control group who received no exposure was included in Experiments 2 and 3 (i.e., No-Exposure group); its performance presumably reflected the general, cultural bias of temporal processing in Western listeners.

Experiment 1

In Experiment 1, participants listened to syllable sequences (composed of the syllables "pa" and "ta"), with syllable onsets following a temporal pattern of either 223 or 224. Participants responded with a key press to each event, indicating which of the two syllables had been presented. Consequently, participants performed motor responses whose timing reflected the temporal pattern of the sequences (even though some irregularities might be added by the variability of response latencies). The hypothesis was that RTs in the syllable discrimination task should decrease over time. If this improvement reflects not only motor and task learning, but also increasing sensitivity to the underlying temporal pattern, which should lead to more precise temporal expectations, differences in improvement should occur for the 2-, 3- and 4-intervals. As longer intervals are perceived and produced less reliably than shorter intervals (within certain time ranges; see Eisler, Eisler, & Helström, 2008), we expected the longer intervals to benefit to a greater degree from the increased exposure (e.g., shorter RTs). In addition, if RTs are not determined solely by the interval lengths, we expected differences between 2-intervals depending on the position within each pattern: notably, a 2-interval following another 2-interval should benefit from the repetition of the same interval in comparison to a 2-interval following a longer interval (i.e., 3- or 4-interval) (see Schupp & Schlier, 1972).

Method

Participants

Forty-one students from the University of Western Sydney participated in Experiment 1 for partial course credit: 21 and 20 were allocated to 223- and 224-Exposure groups, respectively. The groups were comparable in their musical background, as measured by years of musical training on an instrument or voice with respective averages of 0.71 years (±1.01) and 0.73 years (±1.48) and a median of 0 for both groups. Participants had self-reported normal hearing.

Material

Two syllables ("pa", "ta") were synthesized with a male voice using publicly available Text-To-Speech software (http://www.cslu.ogi.edu/tts/demos/index.html).[1] The duration of each syllable was 470 ms. Three sequences of 246 syllables each (i.e., 123 instances of "pa" and "ta", respectively) were created for each timing pattern, with a random order of the two syllables. The SOA was 700 ms for the 2-interval, 1,050 ms for the 3-interval and 1,400 ms for the 4-interval. The chaining followed the pattern of 223 or 224, which cycled continuously without pauses or other cues. For the short and the long intervals in each pattern, 50% of the syllables were "pa" and 50% were "ta". The order of the three sequences was counterbalanced across participants in each group.

Equipment

The experiment was presented to participants on an iMac running Psyscope X (Cohen, MacWhinney, Flatt, & Provost 1993) connected to a USB soundcard (Edirol UA-25) and Sennheiser HD-25 headphones. Responses were made using an ioLab Response Box.

Procedure

Written consent was obtained from all participants as per the University of Western Sydney Human Research Ethics Committee approval (HREC 07/006). Participants were informed that they would hear sequences of the syllables "pa" and "ta". Their task was to identify as quickly and as accurately as possible whether the presented syllable was "pa" or "ta", by pressing one of two adjacent keys on a computer keyboard (using right index and middle fingers, respectively). They were encouraged to respond while the syllable was still being pronounced and before the next syllable was presented. No error feedback was given.

Data analyses

Correct RTs (averaged over syllables) of participants with overall performance superior to 55% (i.e., 3 participants were omitted in each exposure group) were analyzed by a 3 × 3 × 2 ANOVA with Block (1/2/3) and Position (1/2/3) as within-participant factors and Exposure (223/224) as a between-participants factor. For the position factor, Position 1 referred to the syllable following the longer interval (i.e., duration of 3 or 4), while Position 2 and Position 3 referred to the syllables following the shorter interval (i.e., 2). This is illustrated in Fig. 1.

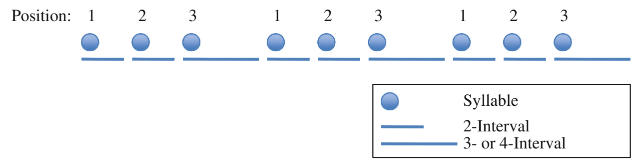

Fig. 1 Schematic representation of the temporal pattern used as stimuli (223- or 224-pattern) and the three positions inside the pattern (with Position 1 following the longer interval). The syllables are the events instantiating the temporal structure and serving for the syllable discrimination task used in the here introduced SRT paradigm

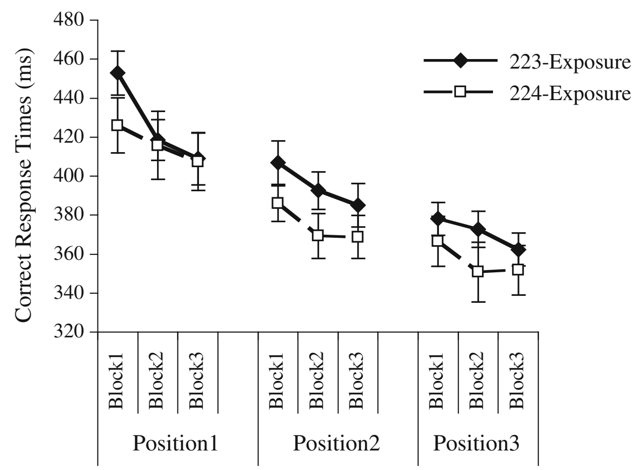

Fig. 2 Average correct RTs for syllable discrimination observed in Experiment 1, presented as a function of Exposure Group (223 or 224), Position (Position 1 referring to the syllable following the longer time interval, 3 or 4; Positions 2 and 3 referring to syllables following the standard interval of 2), and Block (1, 2 or 3). *Error bars* represent standard errors

Results

For the samples (i.e., $n = 18$ and 17 for 223- and 224-Exposure groups, respectively), mean accuracy was at 84.17% (ranging from 74 to 95) and 84.78% (ranging from 68 to 97) for 223- and 224-Exposure groups, respectively.[2]

As hypothesized, correct RTs (Fig. 2) became faster over time, as shown by a significant main effect of Block, $F(2,66) = 8.66$, MSE = 1,568.01, $p < 0.0001$. The main effect of Position, $F(2,66) = 80.27$, MSE = 1,122.06,

[1] These syllables might differ slightly in their perceptual center, which can be influenced by the initial consonant. However, the influence of these differences should be rather minimal because the syllables had a duration of 470 ms, the minimal SOA was 700 ms and the syllables were presented in random order in the exposure blocks.

[2] The observation that syllable identification was not at ceiling was certainly due to the speeded response requirement, the sequential chaining and the continuous responding. The syllables were clearly distinguishable, as evidenced by a short perceptual test: the syllables were presented (out of context with 5 repetitions for each syllable, presented in random order) to 14 participants; identification scores were 100% for each syllable and participant.

$p < 0.0001$, was also significant: orthogonal contrasts showed that RTs were slower for syllables following the longer interval (i.e., 3 or 4) than for those following the short intervals (i.e., 2), $F(1,33) = 108.42$, $p < 0.0001$, and that RTs for the first 2-interval (i.e., Position 2) were slower than for the second 2-interval (Position 3), $F(1,33) = 29.20$, $p < 0.0001$. There was an interaction between Block and Position, $F(4,132) = 2.43$, MSE = 264.73, $p = 0.05$, and these two factors interacted with Group, as reflected in a significant three-way interaction, $F(4,132) = 2.88$, MSE = 264.73, $p = 0.03$. For the 223-Exposure, the decrease in RTs over blocks was more pronounced for Position 1 (i.e., following the longer time interval), than for Positions 2 and 3, $F(1, 33) = 9.33$, $p = 0.004$ (but not for the 224-Exposure, $p = 0.76$). The RT difference between the two exposure groups for Positions 2 and 3 (i.e., the 2-intervals) was not significant, $p = 0.21$. The result for Position 1 remained significant with Bonferroni correction applied to the contrasts addressing this interaction ($p < 0.017$). Neither the main effect of Group nor its interaction with Block or Position were significant ($ps > 0.30$).

Discussion

Experiment 1 used a new sequence-learning paradigm that combined a repeating temporal sequence with a syllable discrimination task. Participants were asked to discriminate two syllables without their attention being drawn explicitly to the timing pattern of the sequence. The results of Experiment 1 show that (1) this temporally structured, continuous syllable discrimination task, which required speeded 2AFC responses, was achievable with reasonable accuracy despite the time constraints, (2) RTs decreased over time (i.e., across the three blocks) and (3) this decrease was more pronounced for RTs following the longer intervals (Position 1) for the 223-Exposure group. For Position 1 in the last two blocks, the 223-Exposure group attained the same speed as the 224-Exposure group; even though the 3-interval does not represent an integral multiple of the short interval (2). In addition, for both exposure groups, the RTs for syllables after the 2-intervals were faster than after the long intervals, and they were fastest for the 2-interval that followed the other 2-interval. The 2-interval represents the standard timing interval in the temporal patterns (i.e., the most frequent interval, which was also the shortest), thus facilitating the anticipation of the next event respecting this time delay, and particularly when this delay was directly repeated. These differences in RT cannot be accounted for by a simple foreperiod effect, which would have predicted faster RTs for longer foreperiod durations (i.e., the time between the previous event and the subsequent target presentation) (e.g., Niemi & Näätänen, 1981). Schupp and Schlier (1972) have reported similar results for short and long intervals (i.e., in their experimental sessions consisting of only two types of intervals): RTs after two short intervals were faster than after a short interval that was preceded by a long interval. Based on findings related to additional experimental manipulations, they concluded that participants become sensitive to the distribution of inter-stimulus intervals within an experimental session, leading to expectations for event n, $n + 1$ and partly even $n + 2$.

In our experiment, the observed differences in RT patterns for different intervals and their changes over the experimental blocks suggest that participants learned the temporal patterns in the exposure phase. This learning allowed them to anticipate more precisely *when* the next syllable will occur, thus leading to faster syllable discrimination.

The observed differences between positions and groups suggest that the improvement observed over exposure does not solely reflect motor learning and/or task learning, but also temporal structure learning. To test for structure learning (beyond motor and/or task learning), previous SRT studies have adopted two methods: (1) they introduced a test or transfer block based on a new pattern structure (i.e., leading to increased RTs) or (2) they used additional tasks in subsequent test phases, such as recognition, generation or prediction tasks (e.g., Buchner & Steffens, 2001; Reed & Johnson, 1994). In a recent study using an adaptation of our new paradigm, we have used the first method and provided evidence for temporal structure learning with a test/transfer block: after several blocks of exposure to a temporal pattern, the introduction of a different temporal pattern slowed RTs (Brandon, Tillmann, Stevens & Terry, in preparation). In the present study, we adopted the second approach aiming to confirm that learning in the exposure phase included temporal learning and not only perceptual-motor learning: in Experiments 2 and 3, the exposure phase (using the paradigm introduced in Experiment 1) was followed by two different, subsequent testing phases. These tasks investigated whether the benefits of the temporal learning in the exposure phase extend to the production and perception of temporal patterns with the same tempo and with a different tempo. In particular, we predicted that the exposure to a 223-pattern should benefit the production of uneven patterns based on a 2:3 ratio, which had been shown to be difficult to process for Western listeners (e.g., Fitch & Rosenfeld, 2007; Hannon & Trehub, 2005a, b).

Experiment 2

Experiment 2 combined the exposure RT task with a subsequent rhythm production task. Experiment 2 thus had two purposes: first, to replicate the findings of the exposure phase observed in Experiment 1 for the newly introduced paradigm; second, to test for benefits of the exposure phase on production performance. For this second purpose, the two exposure groups responding to 223- or 224-patterns (with the 224-Exposure serving as a control group for the 223-exposure[3]) were compared to an additional control group that did not receive any exposure. Performance of this No-Exposure group was expected to reflect the general cultural bias of the Australian participants in favor of even rhythms, such as 224, allowing us to assess the influence of exposure phases on rhythm production. The production task contained the complex, uneven 223-pattern and an isochronous 222-pattern, which was used to gauge the effects of the exposure RT task on basic temporal processing (i.e., of a regular, isochronous pattern with short intervals). As our main goal was to test for improved 3-interval processing after 223-exposure (with 224-exposure serving as a control), we did not add a 224-pattern in the test phase. In the production task, participants were required to tap along with the 223- or 222-patterns (synchronization phase) and to continue tapping without a model (continuation phase). To strengthen the test for generalization, the 222- and 223-patterns in the production task were presented at either the same tempo (i.e., with the same interval durations) as during exposure or a faster tempo (doubling the speed).

If the exposure phase has an influence on participants' temporal processing, production performance should be more precise for the exposure groups than the No-Exposure group, and particularly for the 223-pattern in the 223-Exposure. Based on the findings of studies of uneven rhythm production (e.g., Repp et al., 2005), we expected that, for the 223-pattern, produced durations (as measured by inter-tap-intervals, ITIs) would be too short for 2-intervals and too long for 3-intervals, and that tap timing variability (as measured by standard deviations and coefficients of variation, CV) would be higher for 3-intervals. Importantly, we hypothesized that these inaccuracies would be less evident in the 223- relative to the 224-Exposure. If the influence of the exposure phase extends beyond the learning of a given temporal pattern (i.e., its absolute timing), notably to a more abstract processing of the interval ratio and the underlying meter (due to hierarchical processing or to multiple internal oscillators becoming entrained to the signal; Large, 2008; Large & Velasco, 2009), improved production performance should be observed also at the faster tempo.

Method

Participants

Sixty-three students of the University of Western Sydney participated in Experiment 2 for partial course credit; none had participated in Experiment 1. They were divided equally between the three experimental conditions ($n = 21$). 223-Exposure, 224-Exposure and No-Exposure groups were comparable in their musical background, as measured by years of musical training on an instrument or voice, with respective averages of 0.76 years (±1.30), 0.48 years (±0.83) and 1.1 years (±2.72) as well as a median of 0 for all groups. Participants had self-reported normal hearing.

Material

The exposure phase was as described in Experiment 1. In the production task, the 222- and 223-patterns were presented at either a slow or a fast tempo. At the slow tempo, the SOAs were 700 ms for the 2-interval and 1,050 ms for the 3-interval. At the fast tempo, the SOAs were 350 ms for the 2-interval and 525 ms for the 3-interval. The trials were created with MAX/MSP, exported from EXS24 and using the built-in sampler of LogicPro. Three types of tone of 120 ms duration were used: one played with a marimba timbre at C4 for the introductory presentation cycles; another played with a piano timbre at C3 for the synchronization phase; and a single tone played with a piano timbre at C7 indicating the end of the continuation phase. In the production task, the marimba timbre was used for 2 cycles (each cycle being one 222- or one 223-pattern), directly followed by 10 cycles using the C3 piano tone, followed by 10 cycles without sound and then the C7 piano tone, indicating the end of the trial.

For the experimental session, the trials were presented to participants with Max/MSP. To ensure no delay in the recorded response, participants tapped on the table (with their finger), with a microphone (AKG condenser C391B) placed close to the tapping location. Each trial was recorded into one stereo file with tapping on one channel and the presented stimulus pattern on the other channel.

Procedure

The exposure phase was as described in Experiment 1 and was followed by the production task. The No-Exposure group encountered only the production task. For each trial,

[3] The 224-Exposure served as a control group for the 223-Exposure in two respects: (1) it controlled for exposure to the syllable sequences with a short–short–long grouping, which reinforced binary interval structure with the 224-pattern, and (2) it allowed us to estimate whether the switch in tasks (from the exposure RT task to the production task in Experiment 2 and to the detection task in Experiment 3) might incur a general switch cost (e.g., Schneider & Logan, 2006).

participants were instructed to listen carefully to the patterns played by the marimba and, using one finger, to start tapping with the cycles played by the piano (synchronization phase). Once the piano stopped, they were told to continue tapping the same pattern up to the sounding of the high-pitch piano tone, indicating the trial's end (i.e., continuation phase). One experimental block contained four trials (the 222- and 223-patterns at slow and fast tempo), presented in counterbalanced orders over participants and blocks. Participants started with one practice block and continued with six experimental blocks.

Data analyses

For the exposure phase, RTs of correct responses were analyzed as described in Experiment 1; three participants and one participant were omitted, respectively, in 223- and 224-Exposure groups because of low accuracy. For the production task, the recorded tapping performance was analyzed with Matlab, which determined tap onsets by locating peak amplitudes (generated when the tapping finger impacted upon the table) against the ambient background noise. The first two cycles of synchronization and continuation phases were omitted from the analyses (i.e., warm-up cycles). For the remaining eight cycles, the following parameters were analyzed separately for the two patterns and the two tempi using ANOVAs with Group (No-Exposure, 223-Exposure, 224-Exposure) as the between-participants factor: average Inter-Tap-Intervals (ITIs) and their within-trial variability (standard deviation) for synchronization and continuation; average asynchrony (tapping earlier or later than the presented tone, leading to negative and positive asynchronies, respectively) and its standard deviation for the synchronization phase. For the 223-pattern, these parameters were calculated separately for the 3-intervals and 2-intervals. To investigate whether effects of variability were related to changes in ITI, we performed additional analyses using the coefficients of variation (CV) defined as the standard deviation of ITIs divided by the mean ITI.

Results

Exposure phase

For the samples (i.e., $n = 18$ and 20 for 223- and 224-Exposure groups, respectively), correct responses were 82.75% (ranging from 68 to 94) and 83.69% (ranging from 68 to 94) for 223- and 224-Exposure groups, respectively.

As in Experiment 1, correct RTs became faster over time (Fig. 3), as shown by a significant main effect of Block, $F(2,72) = 8.56$, MSE = 1,503.89, $p < 0.001$. The main effect of Position was also significant: $F(2,72) =$ 64.69, MSE = 1,261.16, $p < 0.0001$: orthogonal contrasts showed that RTs were slower for syllables following the longer interval (i.e., 3 or 4) than for those following the short intervals (i.e., 2), $F(1, 36) = 70.94$, $p < 0.0001$, and that RTs for the first 2-interval (i.e., Position 2) were slower than for the second one (Position 3), $F(1, 36) = 39.76$, $p < 0.0001$. In addition, the interaction between Block and Position was significant, $F(4,144) = 2.54$, MSE = 292.08, $p = 0.04$. The decrease in RTs over blocks was more pronounced for Position 1 (following the longer time interval) than for Positions 2 and 3 (i.e., the 2-interval), $F(1, 36) = 15.29$, $p < 0.001$. The main effect of Group and its interactions with Block and Position were not significant, $p\text{s} > 0.10$. Note that while this differed from Experiment 1, an analysis combining Experiments 1 and 2 (with Experiment as an additional between-participants factor) confirmed the Group × Block × Position interaction, $p = 0.02$, which did not interact with Experiment, $p = 0.59$.

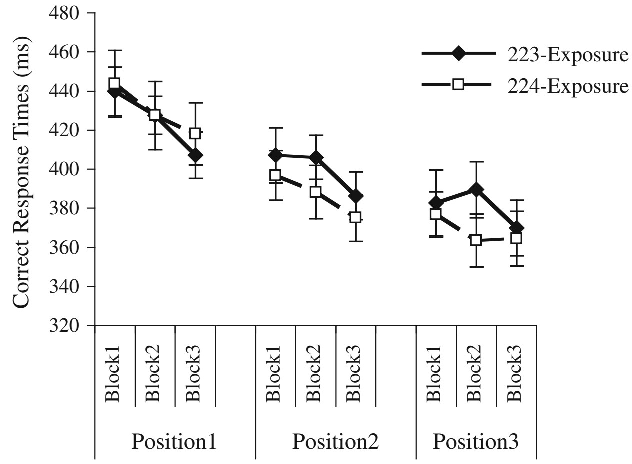

Fig. 3 Average correct RTs for syllable discrimination observed in Experiment 2, presented as a function of Exposure Group (223 or 224), Position (Position 1 referring to the syllable following the longer time interval, 3 or 4; Positions 2 and 3 referring to syllables following the standard interval of 2), and Block (1, 2 or 3). *Error bars* represent standard errors

Production phase

Because of weak production performance (e.g., missing taps or additional taps leading to difficulties aligning the production with the presented temporal pattern), two participants were omitted from the analyses for each the 223-Exposure and the No-Exposure groups, and three participants were omitted from the 224-Exposure group. For the remaining participants (19, 18, 19, respectively, in the 223-Exposure, 224-Exposure and No-Exposure groups), data analyses were performed as described above (Table 1).

Table 1 Results of the Production Task in Experiment 2

	223-Exposure		224-Exposure		No-Exposure	
	Mean	SD	Mean	SD	Mean	SD
222-pattern/slow tempo						
Synchronization ITI	699	31	700	29	699	31
Asynchrony	−60	32	−43	29	−55	31
Continuation ITI	657	30	672	27	668	28
222-pattern/fast tempo						
Synchronization ITI	349	18	349	17	350	18
Asynchrony	−18	20	−19	18	−21	20
Continuation ITI	341	17	342	16	346	20
223-pattern/slow tempo/2-intervals						
Synchronization ITI	647	53	640	49	649	51
Asynchrony	−78	34	−65	38	−69	39
Continuation ITI	589	29	615	27	610	30
223-pattern/slow tempo/3-interval						
Synchronization ITI	1,153	50	1,165	53	1,147	58
Asynchrony	19	49	41	53	21	57
Continuation ITI	1,076	54	1,140	52	1,082	61
223-pattern/fast tempo/2-intervals						
Synchronization ITI	312	*21*	314	*22*	323	*27*
Asynchrony	−15	33	−13	30	−8	34
Continuation ITI	*319*	*16*	*325*	*17*	*336*	*24*
223-pattern/fast tempo/3-interval						
Synchronization ITI	595	32	591	35	577	33
Asynchrony	44	33	43	36	33	38
Continuation ITI	604	*25*	601	*34*	588	*35*

Means and Standard Deviations (SD) of Inter-Tap-Interval (ITI) and Asynchrony in the Synchronization phase and means and SD of ITI in the Continuation phase, presented as a function of Group (223-Exposure, 224-Exposure and No-Exposure) and presented patterns (222, 223). At the slow tempo, the 2-interval and 3-interval lasted for 700 ms and 1,050 ms, respectively. At the fast tempo, the 2-interval and 3-interval lasted for 350 ms and 525 ms, respectively. For the 223-pattern, data are presented separately for responses corresponding to events following 2- or 3-intervals. Significant group effects are indicated in italics (see text)

At the slow tempo (i.e., the same tempo as exposure) for 222- and 223-patterns, no significant effects of Group were observed. For the 222-patterns, participants were close to the target tempo during synchronization ($m = 699$ ms) and their taps anticipated the tones (−53 ms); during continuation, they tapped slightly faster than the target tempo (666 ms). For the 223-patterns, during synchronization and continuation, the 2-intervals were shortened (646 and 605 ms) and the 3-intervals lengthened (1,144 and 1,099 ms) relative to the target values (see Table 1). During synchronization, participants anticipated the tones after the 2-intervals (−71 ms), but not after the 3-interval (+27 ms).

At the fast tempo, significant effects of Group were observed only for the 223-pattern (see below), but not for the 222-pattern. For the 222-pattern, participants adhered to the tempo during synchronization (349 ms) and continuation (343 ms); they also anticipated the tones (−19 ms) during synchronization. For the 223-patterns, during synchronization and continuation, the 2-interval was shortened (317 ms and 326 ms) and the 3-interval lengthened (587 ms and 598 ms) relative to the target values (see Table 1). Participants also anticipated the tones for the 2-intervals (−12 ms) and tapped with a delay after the 3-interval (+40 ms). For the following parameters, significant effects of Group were observed in the ANOVAs, and contrast analyses were run to test for an overall benefit of exposure over No-Exposure (comparing the 223- and 224-Exposures to the No-Exposure group) and to compare the 223-Exposure to the two control groups (the 224-Exposure and the No-Exposure group). During synchronization and continuation (see Table 1), the main effect of Group was significant for standard deviations of ITIs including 2-intervals, $F(2,53) = 3.27$, $MSE = 58.25$, $p = 0.046$; and $F(2,53) = 5.57$, $MSE = 66.28$, $p = 0.006$, respectively: the two exposure groups showed less variability than the No-Exposure group, $F(1,53) = 6.40$, $p = 0.02$, and $F(1,53) = 10.97$, $p < 0.01$, but there was no specific advantage for the 223-Exposure. During continuation, the main effect of Group for standard deviations of ITIs including 3-intervals was significant, $F(2,53) = 3.17$, $MSE = 155.65$, $p = 0.05$. The variability data for the 3-intervals did not reflect an overall benefit of exposure, but a specific benefit of 223-exposure: the 223-Exposure showed less variability than the control groups, $F(1,53) = 6.25$, $p = 0.02$. Finally, during the continuation phase, the main effect of Group for the ITIs corresponding to the 2-intervals was significant, $F(2, 53) = 3.18$, $MSE = 466.60$, $p = 0.05$: all groups tapped too fast, but the two exposure groups tapped faster than the No-Exposure group, $F(1, 53) = 5.49$, $p = 0.02$.

To investigate whether the effects on variability observed at the fast tempo can be explained solely by differences in ITI durations, we calculated CV and ran the appropriate contrast analyses testing whether the significant effects observed for variability hold when the effects of ITI differences are partialled out. These analyses confirmed the patterns reported above for variability: during synchronization, mean CVs for 2-intervals were smaller in 223- and 224-Exposure groups (0.07; 0.07) than in the No-Exposure group (0.08), $F(1,53) = 5.44$, $p = 0.02$. During continuation, mean CVs for 2-intervals were smaller in 223- and 224-Exposure groups (0.05; 0.05) than in the No-Exposure group (0.07), $F(1,53) = 10.28$, $p = 0.002$. During continuation, mean CVs for 3-intervals in the 223-Exposure group (0.04) were smaller than in the 224-Exposure group (0.06) and the No-Exposure group (0.06), $F(1,53) = 5.91$, $p = 0.02$, thus supporting the hypothesis that the exposure phase including the 3-interval reduces inaccuracies for this uneven rhythm.

To test whether participants showed a stronger influence of the exposure phase at the beginning of the production phase, we analyzed the results separately for the first three production blocks. This analysis revealed a significant Group effect for the 223-patterns at the slow tempo that had not been observed in the analyses on all blocks: standard deviations of asynchrony corresponding to the 3-interval were the largest for the No-Exposure group, followed by the 224-Exposure group and then the 223-Exposure group (63, 54 and 48, respectively, $F(2,53) = 3.48$, MSE = 315.06, $p = 0.04$).

Finally, we tested whether the amount of learning in the exposure phase correlated with production performance for Position 1 of the 223-patterns (the event following the 3-interval, thus the interval creating the complex ratio). For each participant (for whom performance had been analyzed in both exposure and production phases), we calculated the difference between block 1 and block 3 in the exposure phase for RTs at Position 1 (the syllable following the long interval, either 3 or 4) and correlated that difference with production performance. Significant correlations were observed only for the 224-Exposure ($n = 17$) in the synchronization phase for standard deviation of ITIs [$r(15) = 0.60$, $p < 0.05$] and standard deviation of asynchronies [$r(15) = 0.60$, $p < 0.05$]: the more the 224-Exposure group became faster for the 4-interval over exposure, the more variability they showed in the production phase for a 3-interval.

Discussion

The exposure phase of Experiment 2 replicated the main findings of Experiment 1, with RTs decreasing over blocks and with differences between positions. Despite the lack of an interaction with exposure type, the differential decrease in RTs across positions supports the interpretation that participants learned the temporal pattern. In addition, the production task revealed some benefits of the exposure phase for tapping performance; a general benefit of exposure over no exposure and a specific benefit of 223-exposure for the production of the 3-intervals in the uneven rhythm. Results for the 223-pattern at the fast tempo revealed that participants were more precise in their tapping (i.e., tapping with lower variability) after exposure. For the 2-intervals, the benefit was shown in synchronization and continuation phases for both exposure groups in comparison to the No-Exposure group. This observation suggests that the advantage was due to the exposure to a temporal pattern and probably also to having made a motor response to each syllable in the discrimination task. For the production of the 3-intervals, more specific influences were observed as a function of exposure type, notably a benefit for the 223-Exposure group and a cost for the 224-Exposure group: the 223-Exposure group showed decreased variability for 3-intervals in comparison to 224-Exposure and No-Exposure groups. However, for the 224-Exposure group, the correlation analyses between exposure and production performance suggest a cost in producing the 3-interval due to learning of the 4-interval timing in the exposure phase. In sum, Experiment 2 showed general and specific influences of the exposure RT task on production performance. Of particular interest is the benefit for the 3-interval: the 223-exposure phase facilitated the processing of the uneven rhythm, which was similar to those that have been shown to be processed less accurately in previous studies (e.g., Repp et al., 2005; Fitch & Rosenfeld, 2007).

It is worth noting that this advantage of the 223-Exposure was observed without differences in the tempo of production. For the 2-interval, the reduced variability observed for the two exposure groups (in comparison to the No-Exposure group) might be, at least partially, related to the increased tempo in the continuation phase. Indeed, previous studies have reported that variability is linked with ITI, notably variability decreases linearly with decreasing ITI in accordance with a generalized form of Weber's law (e.g., Fraisse, 1966; Peters, 1989; Repp, 1997; Repp et al., 2005). It might be argued that participants after exposure are more comfortable with the uneven rhythms, leading them to produce these rhythms at a faster tempo. However, this link to tempo cannot explain the entire influence of exposure on variability, as the tempo related differences were also observed in analyses of coefficients of variation (CV).

The benefits of the exposure phase on tapping were only present for the production task using the 223-pattern, but not the 222-pattern. To further investigate the benefit of the exposure phase on the production of long intervals more generally, future adaptations of the present paradigm should use both 224- and 223-patterns in the production phase, thus testing the general and specific benefits of 224-exposure on the production of longer, but even intervals.

Finally, for all groups, the results for the 223-pattern in synchronization and continuation showed that accurate and precise production in the tapping task was difficult for the non-musician participants. All groups showed a contrast effect similar to those previously observed in studies using rhythms with 2:3 ratios (e.g., Repp et al., 2005; Snyder, Hannon, Large, & Christiansen, 2006), leading to shortening of the shorter interval and lengthening of the longer interval to yield an average ratio of 2:3.6 rather than 2:3. Despite the difficulty of the uneven rhythm production for non-musicians, their performance showed previously observed characteristics, such as the contrast effect and anticipation of 2-intervals (Aschersleben, 2002).

Experiment 3

RTs during exposure phases in Experiments 1 and 2 suggest that participants became sensitive to the temporal pattern with which the syllables had been presented. The hypothesis is that exposure allows participants to develop more precise temporal expectations for the next upcoming syllable, thus being more prepared to respond and showing faster RTs. Experiment 3 further investigated the possible generalization of the learned temporal patterns to a subsequent testing phase using 222- and 223-patterns presented at a different tempo. In this test phase, participants listened to 7-syllable sequences and were instructed to make speeded discrimination judgments on the last syllable (i.e., deciding whether it was "pa" or "ta"). The first six syllables were presented via a repeated 222- or a 223-pattern; the target syllable either adhered to this temporal pattern or occurred earlier or later (with deviations of ±20%; Jones, Johnston & Puente, 2006). The RTs should be shortest for the expected time delay (i.e., on-time targets) and increased for earlier and later targets (or at least for earlier targets, Schmuckler & Boltz, 1994; Tillmann & Lebrun-Guillaud, 2006; but see Penel & Jones, 2005 for a different pattern of results). If the exposure phase allows for more precise temporal expectations, we predicted that, in the test phase, faster RTs and a more pronounced influence of expectancy violation would be observed for the time patterns that had been processed during the exposure phase (e.g., the 3-intervals for the 223-Exposure group). The test patterns were presented at a different tempo than the tempo used during exposure in order to test for generalization of the pattern learned in the exposure phase; notably to test for learning of the more abstract (relative) rhythmic pattern in contrast to the potential benefits being based on the specific interval durations (the absolute temporal patterns).

Method

Participants

Sixty-one students of the University of Western Sydney participated in Experiment 3 for partial course credit; none had participated in Experiments 1 or 2. They were divided between the three experimental conditions, with 20 each in the 223-Exposure group and No-Exposure group and 21 in the 224-Exposure group. The groups were comparable in their musical background, as measured by years of musical training on an instrument or voice, with averages of 0.88 years (±1.86), 0.63 years (±1.59) and 0.7 years (±0.98) for 223-, 224-Exposure and No-Exposure groups, as well as a median of 0 for all groups. Participants had self-reported normal hearing.

Materials

The exposure phase was as described in Experiment 1. For the discrimination task of the test phase, the syllables "pa" and "ta" of the exposure phase were used, together with the syllable "ga", constructed with the same text-to-speech software and method as described for Experiment 1. With these syllables, 7-syllable sequences were created, with the first six syllables being "ga" and the last one (i.e., the target) either "pa" or "ta". Syllables were chained following either the 222- or the 223-pattern: the SOAs were 600 ms for the 2-intervals and 900 ms for the 3-intervals. The occurrence of the target either adhered to the SOA of the 222- or 223-patterns (target on-time) or occurred earlier (SOAs of 480 and 720 ms for 2- and 3-intervals, respectively) or later (SOAs of 720 and 1,080 ms for 2- and 3-intervals, respectively). The 12 sequence types (2 patterns × 3 target timings × 2 target syllables) were constructed with Audacity digital sound software and the experiment was run with Psyscope X (Cohen et al., 1993) on an iMac.

Procedure

After the exposure phase, which was as described in Experiment 1, participants performed the discrimination task. They were instructed that they should listen to 7-syllable sequences, that the first six sounds were the syllable "ga", and that they had to decide as fast and as accurately as possible whether the seventh syllable was "pa" or "ta". Response keys and finger assignments were as in the exposure phase. The next trial started when participants pressed a third key (using the left hand). To encourage participants to answer as quickly and accurately as possible, a time-out of 1,500 ms was applied, an incorrect response was accompanied by an alerting feedback signal, and a correct response stopped the sounding of the target, allowing participants to continue with the next trial. Participants performed first a practice block with six sequences (covering the two patterns with all target timings, presented with either "pa" or "ta") and then two blocks of 60 trials, separated by a short break. In total, participants performed 20 trials for each experimental condition, presented in random order [2 patterns (222/223) × 3 target timings (on-time/earlier/later)], with 50% of trials containing each target syllable.

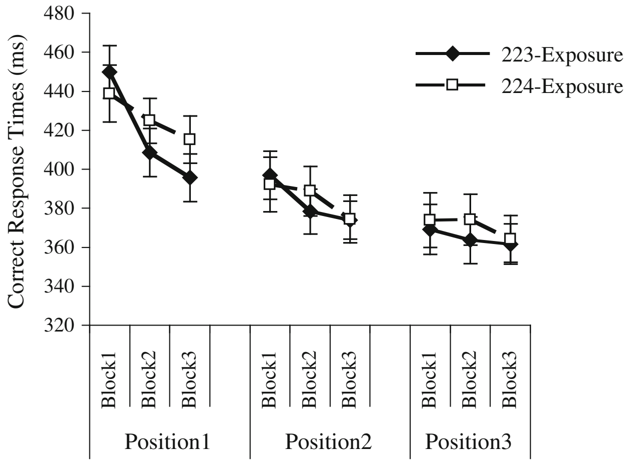

Fig. 4 Average correct RTs for syllable discrimination observed in Experiment 3, presented as a function of Exposure Group (223 or 224), Position (Position 1 referring to the syllable following the longer time interval, 3 or 4; Positions 2 and 3 referring to syllables following the standard interval of 2), and Block (1, 2 or 3). *Error bars* represent standard errors

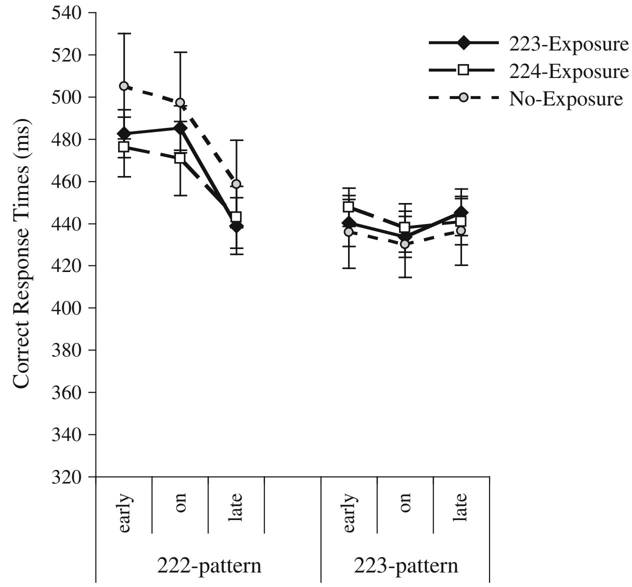

Fig. 5 Average correct RTs for the discrimination task in Experiment 3, presented as a function of Group (223- and 224-Exposure groups, No-Exposure group), Time pattern (222, 223) and Target timing (early, on-time, late). *Error bars* represent standard errors

Data analyses

For the exposure phase, RTs for correct responses were analyzed as described in Experiment 1; one participant was omitted in the 224-Exposure group because of low accuracy. For the discrimination task, accuracy and correct RTs were analyzed, averaged over the 20 trials for each of the 6 within-participant experimental conditions.

Results

Exposure phase

For the samples (i.e., 20 in each group), correct responses were 82.95% (ranging from 63 to 95) and 85.83% (ranging from 67 to 96) for 223- and 224-Exposure groups, respectively. RTs became faster over time (Fig. 4), as shown by a significant main effect of Block, $F(2,76) = 8.93$, $MSE = 1{,}742$, $p < 0.001$. The main effect of Position was also significant, $F(2,76) = 94.71$, $MSE = 986.80$, $p < 0.0001$): orthogonal contrasts showed that RTs were slower for syllables following the longer interval (i.e., 3 or 4) than for those following the short intervals (i.e., 2), $F(1, 38) = 112.32$, $p < 0.0001$, and that RTs for the first 2-interval (i.e., Position 2) were slower than for the second one (Position 3), $F(1, 38) = 35.43$, $p < 0.0001$. The interaction between Block and Position was significant, $F(4, 152) = 8.56$, $MSE = 315.20$, $p < 0.0001$, and, as in Experiment 1, the interaction between Group × Block × Position was significant, $F(4, 152) = 2.43$, $MSE = 315.20$, $p = 0.05$: the stronger decrease over blocks for Position 1 (vs. Positions 2 and 3) was significant for the 223-Exposure Group, $F(1, 38) = 19.19$, $p < 0.0001$, but not the 224-Exposure Group, $p = 0.27$. This effect remained significant with Bonferroni correction ($p < 0.025$). Neither the main effect of Group nor its interaction with Block or Position was significant (ps > 0.31).

Test phase (discrimination task in the 7-syllable sequences)

Participants achieved high accuracy in the syllable discrimination task (98.47%). For correct RTs (Fig. 5), one participant was omitted in each of the groups because of average RTs 2 SDs slower than the group mean. For the remaining participants (i.e., 19, 20, and 19 for 223-, 224-Exposure, and No-Exposure groups, respectively), correct RTs were analyzed with a 2 × 3 × 2 ANOVA, with Time patterns (222/223) and Target timings (on-time/earlier/later) as within-participant factors and Group (223-, 224-Exposure and No-Exposure) as between-participants factor. The main effects of Time pattern and Target timing were significant, $F(1,56) = 40.92$, $MSE = 2{,}554.83$, $p < 0.0001$, and $F(2,112) = 26.30$, $MSE = 521.60$, $p < 0.0001$, respectively, as was their interaction, $F(2, 112) = 38.36$, $MSE = 470.57$, $p < 0.0001$. This interaction showed that Target timing influenced RTs only for the 222-pattern, and notably RTs were slower for early and on-time targets than for late targets. In addition, the interaction between Group and Time pattern just failed to reach significance, $F(2,56) = 3.06$, $MSE = 2{,}554.83$, $p = 0.055$. The difference between the participant groups was observed mostly for the 222-pattern and not for the 223-pattern, for which all groups responded faster overall. No other effects were significant. Finally, there were no significant correlations between the indicator of learning in the exposure phase (see Experiment 2) and RTs involving the 3-interval.

Discussion

RTs in the exposure phase replicated the improvement over time, with stronger decrease for the longer intervals and particularly the 3-interval for the 223-Exposure group (as in Experiment 1).[4] This finding buttresses the interpretation in terms of learning the temporal pattern and suggests that participants develop increasingly precise temporal expectations during the exposure phase (i.e., allowing them to anticipate the next event and thus leading to faster RTs).

In the test phase, the discrimination data for the 222-pattern showed some influence of the exposure phase: for all target timings, the No-Exposure group responded slower than the 223-Exposure group, while this group responded slower than the 224-Exposure group. Both exposure groups seemed to have taken advantage of the exposure phase, particularly the 224-Exposure group who had been exposed solely to binary intervals. However, exposure did not influence RTs for the 223-pattern, for which participants responded faster overall. These faster RTs might be due to the 3-interval allowing participants to group the six events preceding the target syllable into two groups. This grouping (based on an acoustic cue, the longer silence) was not possible with the isochronous 222-pattern, thus rendering the anticipation of the target more difficult. Future experiments could address this by introducing an additional accent pattern to the 222-sequences, thereby helping to structure the sequence.

In contrast to our hypothesis and to most results of previous temporal expectancy studies (e.g., Schmuckler & Boltz, 1994), the present data set did not show slower RTs for the early-occurring targets and for late targets relative to on-time targets. The faster RTs for the late targets might have occurred because participants expected the target to appear on-time and were thus ready to respond to a late-appearing target (see Tillmann & Lebrun-Guillaud, 2006, for a similar discussion). The late targets might also benefit from effects related to the decreasing uncertainty of the foreperiod duration (e.g., Niemi & Näätänen, 1981). In contrast to expectation for early and on-time targets, for which there were, respectively, three and two possibilities, there was no uncertainty about the target's possible onset for late targets (i.e., only one remaining onset option) and participants could prepare to respond (though without knowing yet which response to make). Two other reasons for faster RTs with late targets might be that (1) the late target was preceded by a short silence that gave an additional cue that the target event follows; this cue should be particularly helpful for the 222-pattern, which does not benefit from grouping; and (2) the temporal interval created by the late target occurred also in the 223-patterns with early-occurring targets (i.e., 720 ms). Furthermore, as we did not observe slower RTs for early targets than for on-time targets, the data of the 223- and 224-Exposure groups have to be compared to the data of the No-Exposure group. This comparison revealed some advantage of the exposure phase (particularly of the exposure to 2-intervals) on performance for the 222-pattern.

General discussion

Inspired by previous research using SRT tasks, the present study introduced a new paradigm to investigate temporal pattern learning. We were interested in testing whether participants can benefit from short-term exposure in the laboratory, specifically for the processing of a culturally less familiar musical rhythm, the 223-pattern. Over the exposure phase, RTs of participants became faster not only for the 224-pattern, but also for the 223-pattern (Experiments 1–3). Experiments 2 and 3 revealed that the exposure phase influenced performance in subsequent testing phases to some extent. These influences were observed for test materials presented at different tempi than the tempo of the exposure phase materials. These findings suggest that listeners learn the relative timing patterns rather than the absolute temporal patterns. Experiment 2 showed that the exposure phase influenced production performance, particularly by decreasing tap timing variability. For the production of the more difficult 223-pattern, the performance of the 223-Exposure group showed a benefit of exposure for the 3-interval, while the 224-Exposure group showed (for the production of the 3-interval) a cost of having been exposed to the 4-interval (reflected in the correlation between amount of learning in exposure and variability in production). Experiment 3 aimed to directly test the processing advantage linked to more precise temporal expectations. RTs showed some influence of exposure on temporal expectation for the 222-pattern, while the 223-pattern may have benefited from grouping, which led to overall faster RTs.

The observation of temporal pattern learning, as reflected in faster RTs and subsequent test performance, extends previous findings by Salidis (2001) who reported faster RTs over increasing exposure to a RSI-based cycling temporal sequence. Using post-tests (i.e., interview, generation and

[4] A pooled analysis on the exposure phase of Experiments 1–3 (with Experiment as a between-participants factor) confirmed the interaction between Group, Block and Position ($p = 0.001$), which did not interact with Experiment ($p = 0.69$).

prediction tasks), Salidis showed that this temporal learning occurred without awareness. She concluded that complex temporal patterns might be more easily processed implicitly rather than explicitly, thus pointing out the power of implicit cognition and its relevance for our understanding of temporal cognition. Because of the subsequent testing phases, our study did not assess the implicit/explicit nature of the acquired knowledge. This needs to be aimed for in future studies (see Brandon et al., for a first attempt), keeping in mind that measuring awareness is difficult and the distinction of implicit versus explicit learning depends on criteria and tasks (e.g., Berry, 2002; Perruchet, 2008; Shanks & St. John, 1994). Whether being restricted entirely to implicit processing or including some explicit knowledge, the main contribution of the present study was to show that non-musician adults can learn complex, uneven temporal patterns that are culturally unfamiliar to them. This contrasts with conclusions by Hannon and Trehub (2005a) and suggests that it would be useful to run further studies testing learning with different materials and tasks (e.g., similarity judgments as used in Hannon & Trehub).

In contrast to Salidis (2001) who used RSI-based temporal sequences, our study showed learning with temporal sequences whose events were not under the control of the participants (i.e., timing patterns were based on SOAs instead of RSIs). This aspect also differs from other studies implementing the temporal sequences with RSIs. Such studies either failed to show temporal pattern learning (Buchner & Steffens, 2001; Shin & Ivry, 2002) or showed learning mostly for one of the intervals used (Salidis, 2001). The difficulty to show temporal learning in these earlier studies may be due to the use of RSI-based patterns with participants' delays in responding influencing the timing pattern between events. Comparison with the present findings suggests that participants might learn temporal patterns between event onsets (e.g., between tones) more easily than temporal patterns between an action and an event (as implemented in RSIs), which are further disturbed by variable temporal patterns between event onsets.

In addition, two of these previous studies tested for the conjoint learning of regularities on two dimensions. In Buchner and Steffens (2001), structured tone sequences were presented together with an RSI pattern that was either in a fixed relation or an ambiguous relation to the tones. Evidence for learning of the RSI pattern was only found when the RSI pattern was uniquely related to the tone patterns. Similarly, Shin and Ivry (2002) reported that temporal patterns (i.e., in RSI patterns) cannot be learned independently from a concurrent spatial pattern. The necessity to combine temporal and spatial information into an integrated representation has also been observed for temporal sequences that are based on fixed SOAs, thus independently of participants' response latencies (Shin & Ivry, 2002, Experiment 2). However, in their experiment, participants were tested for the learning of two sequences; with one based on a spatial pattern and one on a temporal pattern. As the spatial pattern was directly associated with the motor response for the task (i.e., pressing spatially arranged keys), the learning of the spatial patterns might have benefitted from this arrangement, while the temporal pattern was less relevant for the task.

In contrast to Buchner and Steffens (2001) and Shin and Ivry (2002), participants in our study were (1) presented only with one regular sequence (on the time dimension), with syllables chained at random, and (2) the SOAs were fixed and independent of participants' response delays. Our syllable discrimination task provided the medium to make participants press a key following the occurrence of an event (i.e., whose onsets defined the temporal pattern). Consequently, there was only one pattern to be learned (the timing). In addition, participants' motor responses associated with the task might have facilitated the learning of the temporal sequences. Indeed, SRT studies testing for perceptual learning (i.e., by removing the contribution of motor learning) reported that pure perceptual learning occurs only for simple deterministic sequence structure and that the motor component of the responses was necessary for the acquisition of more complex sequences (Deroost & Soetens, 2006; see also Remillard, 2003). Similarly, in our study the motor component might have boosted temporal learning. In view of previous research suggesting that learning is difficult on the temporal dimension (e.g., Shin & Ivry, 2002), we employed an SRT task that included the motor component. The present results indicating temporal learning (see also Brandon et al.) provide now a basis for future studies refining this observation, notably to compare the cognitive capacity of temporal sequence learning for perception with action (as implemented here) and for perception only (without action).

During the exposure phase, participants became sensitive to the presented temporal patterns, and this temporal learning influenced some aspects of production and perception performance in the subsequent testing phases. As our focus was on short-term exposure effects, the present findings might be interpreted in the theoretical framework of neural oscillator models linked to dynamic attention (e.g., Large, 2008; Large & Jones, 1999). According to such models, a pulse that is induced in response to periodicity in rhythmic structure (in the exposure phase) facilitates synchronization and the anticipation of events in time. Following Large and Jones (1999), internal oscillators allow perceivers to focus attention towards expected points in time, enabling anticipation and efficient processing of upcoming events. In our experiment, relevant neural oscillators may have become tuned to the temporal patterns during exposure, thus getting more precise over time (with improvement being greatest for the 3-interval, which is not a multiple of the more frequent 2-interval) (Large & Velasco, 2009). In the testing phase, the effect of exposure for the processing of patterns at different tempi might be mediated by coupling between multiple oscillators or networks of oscillators or time-keepers, allowing for hierarchical processing and entrainment to the signal at various time scales, and thus resistant to changes in tempo.

In sum, the present study showed learning of temporal patterns via exposure by using an experimental task that focused participants' attention on another dimension of the material. The findings of exposure and test phases showed that with exposure to a temporal pattern, but without being told to learn the temporal structures, non-musician participants acquire some temporal pattern knowledge and become sensitive to timing features of the exposure patterns. Even if the patterns used here remain rather simple time patterns (e.g., with 2- and 3-intervals being chained in cyclically repeating sequences), the findings are encouraging in showing that adults can learn new, unfamiliar temporal patterns by exposure. The paradigm provides a method to further our investigation of temporal cognition and to examine the learning of more complex timing patterns such as those found in real music. In addition, future research should compare our present paradigm with a paradigm using explicit learning instructions (i.e., explicitly drawing participants' attention to the temporal patterns and asking them to tap along or reproduce them). As has been previously shown for spatial sequence learning (e.g., Fletcher, Zafiris & Frith, 2005), the SRT-based, incidental learning approach might be more powerful than explicit learning approaches for complex temporal patterns, which Western listeners find challenging in perception, production and memory (e.g., Povel & Essens, 1985).

Acknowledgments This research was supported by the Eminent Visiting Researcher scheme of the University of Western Sydney and the Australian Research Council Discovery Project (DP0986394) scheme. We thank Catherine Best for discussions on the present project as well as Clare Howell, Julien Tardieu, and Tim Byron for assistance with data collection and analyses.

References

Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain and Cognition, 48*, 66–79.

Berry, D. C. (2002). How implicit is implicit learning? In G. Underwood (Ed.), *Implicit Cognition* (2nd ed., pp. 203–226). Oxford: OUP.

Brandon, M., Tillmann, B., Stevens, C., & Terry, J. (in preparation). Implicit learning of temporal structures with weak and strong meter.

Bremner, A., Mareschal, D., Destrebecqz, A., & Cleeremans, A. (2007). Cognitive control of sequential knowledge at 2 years of age: Evidence from an incidental sequence learning and generation task. *Psychological Science, 18*(3), 261–266.

Buchner, A., & Steffens, M. C. (2001). Simultaneous learning of different regularities in sequence learning tasks: limits and characteristics. *Psychological Research, 65*, 71–80.

Buchner, A., Steffens, M. C., Erdfelder, E., & Rothkegel, R. (1997). A multinomial model to assess fluency and recollection in a sequence learning task. *Quarterly Journal of Experimental Psychology, 50*, 631–663.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120*, 235–253.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments and Computers, 25*, 257–271.

Collier, G. L., & Wright, C. E. (1995). Temporal rescaling of simple and complex ratios in rhythmic tapping. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 602–627.

Deroost, N., & Soetens, E. (2006). Perceptual or motor learning in SRT tasks with complex sequence structures. *Psychological Research, 70*, 88–102.

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception, 32*, 341–365.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin and Review, 8*(2), 343–350.

Eisler, H., Eisler, A. D., & Hellström, Å. (2008). Psychophysical issues in the study of time perception. In S. Grondin (Ed.), *Psychology of time* (pp. 75–109). Bingley, UK: Emerald.

Fitch, W. T., & Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms. *Music Perception, 25*, 43–58.

Fletcher, P. C., Zafiris, O., Frith, C. D., et al. (2005). On the benefits of not trying: Brain activity and connectivity reflecting the interactions of explicit and implicit sequence learning. *Cerebral Cortex, 15*, 1002–1015.

Fraisse, P. (1956). *Les structures rhythmiques*. Louvain: Publication Universitaires de Louvain.

Fraisse, P. (1966). L'anticipation de stimulus rythmiques: Vitesse d'établissement et précision de la synchronisation. [Anticipation of rhythmic stimuli: Speed of establishment and precision of synchronization.] *L'Annee Psychologique, 66,* 15–36.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). NY: Academic Press.

Handel, S. (1992). The differentiation of rhythmic structure. *Perception and Psychophysics, 52*, 497–507.

Hannon, E. E., & Trehub, S. E. (2005a). Tuning into musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences (USA), 102*, 12639–12643.

Hannon, E. E., & Trehub, S. E. (2005b). Metrical categories in infancy and adulthood. *Psychological Science, 16*, 48–55.

Hébert, S., & Cuddy, L. L. (2002). Detection of metric structure in auditory figural patterns. *Perception and Psychophysics, 64*, 909–918.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459–491.

Jones, M. R., Johnston, H., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology, 53*, 59–96.

Kauffman, R. (1980). African rhythm: a reassessment. *Ethnomusicology, 24*, 393–415.

Keller, P. E., & Repp, B. H. (2004). When two limbs are weaker than one: Sensorimotor syncopation with alternating hands. *The Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology, 57A*, 1085–1101.

Keller, P. E., & Repp, B. H. (2005). Staying offbeat: Sensorimotor syncopation with structured and unstructured auditory sequences. *Psychological Research, 69*, 292–309.

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In Simon Grondin, (Ed.) *The psychology of time.* Emerald.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119–159.

Large, E. W. & Velasco, M. G. (2009). Modeling of pulse and meter as neural oscillation. Presentation at the Conference of the Society for Music Perception and Cognition, 3–6 August 2009 Indiana University–Purdue University, Indianapolis Indianapolis, Indiana, USA.

London, J. (1995). Some examples of complex meters and their implications for models of metric perception. *Music Perception, 13*, 59–77.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin, 89*, 133–162.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cognitive Psychology, 19*, 1–32.

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87*, B35–B45.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research, 163*, 226–238.

Penel, A., & Jones, M. R. (2005). Speeded detection of a tone embedded in a quasi-isochronous sequence: Effects of a task-irrelevant temporal irregularity. *Music Perception, 22*, 371–388.

Perruchet, P. (2008). Implicit learning. In H.L. Roediger, III (Ed.), *Cognitive psychology of memory. Vol. 2 of Learning and memory: A comprehensive reference,* 4 vols. (J. Byrne, Editor). Oxford: Elsevier (p. 597–621).

Perruchet, P., Bigand, E., & Benoit-Gonin, F. (1997). The emergence of explicit knowledge during the early phase of learning in serial reaction time tasks. *Psychological Research, 60*, 4–13.

Peters, M. (1989). The relationship between variability of intertap intervals and interval duration. *Psychological Research, 51*, 38–42.

Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 3–18.

Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2*, 411–440.

Pressing, J. (1983). Cognitive isomorphisms between pitch and rhythm in world musics: West Africa, the Balkans and Western tonality. *Studies in Music, 17*, 38–61.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 585–594.

Remillard, G. (2003). Pure perceptual-based sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 581–597.

Repp, B. H. (1997). Variability of timing in expressive piano performance increases with interval duration. *Psychonomic Bulletin and Review, 4*, 530–534.

Repp, B. H., London, J., & Keller, P. E. (2005). Production and synchronization of uneven rhythms at fast tempi. *Music Perception, 23*, 61–78.

Salidis, J. (2001). Nonconscious temporal cognition: Learning rhythms implicitly. *Memory and Cognition, 29*, 1111–1119.

Schmuckler, M. A., & Boltz, M. G. (1994). Harmonic and rhythmic influences on musical expectancy. *Perception and Psychophysics, 56*, 313–325.

Schneider, D. W., & Logan, G. D. (2006). Hierarchical control of cognitive processes: Switching tasks in sequences. *Journal of Experimental Psychology: General, 135*, 623–640.

Schupp, W., & Schlier, C. (1972). The dependence of simple reaction time on temporal patterns of stimuli. *Kybernetik, 11*, 105–111.

Semjen, A., & Ivry, R. B. (2001). The coupled oscillator model of between-hand coordination in alternate-hand tapping: A reappraisal. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 251–265.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*, 367–447.

Shin, J. C., & Ivry, R. B. (2002). Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology Learning Memory and Cognition, 28*, 445–457.

Snyder, J. S., Hannon, E. E., Large, E. W., & Christiansen, M. H. (2006). Synchronization and continuation tapping to complex meters. *Music Perception, 24*, 135–146.

Sternberg, S., Knoll, R. L., & Zukofsky, P. (1982). Timing by skilled musicians. In D. Deutsch (Ed.), *The psychology of music* (pp. 181–239). Orlando, FL: Academic.

Summers, J. J., Bell, R., & Burns, B. D. (1989). Perceptual and motor factors in the imitation of simple temporal patterns. *Psychological Research, 50*, 23–27.

Summers, J. J., Hawkins, S. R., & Mayers, H. (1986). Imitation and production of interval ratios. *Perception and Psychophysics, 39*, 437–444.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*, 706–716.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychological Review, 107*, 885–913.

Tillmann, B., & Lebrun-Guillaud, G. (2006). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research, 70*, 345–358.