
Tapping in Time With Mechanically and Expressively Performed Music

CAROLYN DRAKE & AMANDINE PENEL

Université René Descartes, Paris

EMMANUEL BIGAND

Université de Bourgogne, Dijon

We investigate how the presence of performance microstructure (small variations in timing, intensity, and articulation) influences listeners' perception of musical excerpts, by measuring the way in which listeners synchronize with the excerpts. Musicians and nonmusicians tapped on a drum in synchrony with six musical excerpts, each presented in three versions: mechanical (synthesized from the score, without microstructure), accented (mechanical, with intensity accents), and expressive (performed by a concert pianist, with all types of microstructure). Participants' synchronizations with these excerpts were characterized in terms of three processes described in Mari Riess Jones's Dynamic Attending Theory: attunement (ease of synchronization), use of a referent level (spontaneous synchronization rate), and focal attending (range of synchronization levels). As predicted by beat induction models, synchronization was better with the temporally regular mechanical and accented versions than with the expressive versions. However, synchronization with expressive versions occurred at higher (slower) levels, within a narrower range of synchronization levels, and corresponded more frequently to the theoretically correct metrical hierarchy. We conclude that performance microstructure transmits a particular metrical interpretation to the listener and enables the perceptual organization of events over longer time spans. Compared with nonmusicians, musicians synchronized more accurately (heightened attunement), tapped more slowly (slower referent level), and used a wider range of hierarchical levels when instructed (enhanced focal attending), more often corresponding to the theoretically correct metrical hierarchy. We conclude that musicians perceptually organize events over longer time spans and have a more complete hierarchical representation of the music than do nonmusicians.

Received June 5, 1998; accepted for publication April 22, 2000.

Address correspondence to Carolyn Drake, Laboratoire de Psychologie Expérimentale, CNRS UMR 8581, Université René Descartes, Institut de Psychologie, Centre Universitaire de Boulogne, 71, avenue Edouard Vaillant, 92774 Boulogne-Billancourt Cedex, France. (e-mail: drake@idf.ext.jussieu.fr)

ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

WATCHING an audience listen to a piece of music with a clear metric structure (e.g., a march or waltz) can provide interesting insights into the ways listeners perceive the music. Listeners often sway or tap their foot in a quasi-regular fashion in time with the music. This regular motor activity is likely to reflect the perceived metric structure, a cognitive organization that allows nonadjacent events to be situated in relation to each other. These motor actions occur at an intermediate rate, usually not as fast as the fastest notes in the melody, but not too slow either. More surprisingly, not all the listeners move at the same rate: some tap at the beat level, others at the measure level, still others at the two-measure level, and still others at less easily defined rates. This study provides a systematic investigation of the way in which musicians and nonmusicians synchronize their motor activities with music. Contrary to other studies, we are not interested in how precisely listeners' taps correspond to acoustic events and so do not measure synchronization accuracy by mean deviation or standard deviation from the model. Rather, we are interested in the *way* in which listeners synchronize, that is, the structural aspects of the sequence with which listeners synchronize their taps. We therefore adopt a generous synchronization criterion: a tap is considered as synchronized with the acoustic event if it falls within a 10% window (see methods section for more details).

Influence of Expressive Performance Variations on Synchronization

The main question addressed in this article concerns the influence of expressive performance variations (microvariations in timing, intensity, articulation, and pedaling) on the way in which listeners apprehend the metrical structure. Previous research on beat induction has usually involved synchronized tapping with very simple sequences (in a musical sense) (e.g., Duke, 1989, 1994; Duke, Geringer, & Madsen, 1991; Parncutt, 1994). One exception is a study by Repp (1999) that involved tapping along with expressive music. However, Repp was not interested in the way in which listeners synchronized with the music, as he instructed his participants to tap at the lowest level of the metrical hierarchy. The few studies that have used musical stimuli have adopted expressionless, mechanical, computer-generated music, containing only the information indicated in the score, without added expressive variations (Vos, van Dijk, & Schomaker, 1994). To our knowledge, ours is the first study to investigate unconstrained synchronization with expressive performances of musical excerpts.

Similarly, very little theoretical attention has been aimed at this issue. Although they do not address directly the question of how people synchronize with musical sequences, two sets of approaches are relevant to our

a clear metric
interesting insights
may or tap their
regular motor
cognitive organi-
zation to each
usually not as fast
More surpris-
the beat level,
level, and still
systematic investi-
synchronize their
are not inter-
events and so
on or standard
way in which
sequence with
generous syn-
with the acous-
tication for more

Synchronization

the influence of
ng, intensity,
pprehend the
ually involved
al sense) (e.g.,
ncutt, 1994).
ng along with
way in which
participants to
ies that have
l, computer-
in the score,
aker, 1994).
strained syn-
ts.
at this issue.
ple synchro-
evant to our

concerns. First, *beat and meter induction models* were developed to explain how listeners extract a beat or meter from a musical sequence. Second, *the structural communication hypothesis* was proposed to explain how expressive variations may aid the performer to transmit a particular musical interpretation to the listener. Although these approaches were not developed to address the same issues, they lead to contrasting predictions concerning the role of expressive microvariations in our synchronization task.

Beat and meter induction models (Desain, 1992; Longuet-Higgins & Lee, 1982; Parncutt, 1994; Povel & Essens, 1985; Rosenthal, 1992; Steedman, 1977) suggest that the more temporally regular a musical excerpt is, the easier it should be to extract the underlying beat, which is temporally regular by definition. These models therefore lead to the prediction that it should be easier to synchronize with mechanical performances synthesized by computer (maximum temporal predictability) than with real performances, which contain multiple expressive variations (low temporal predictability).

The structural communication hypothesis leads to different predictions. Research guided by this perspective has demonstrated the intimate relation between performance variations and aspects of the musical structure (e.g., Clarke, 1988; Palmer, 1989; Shaffer, Clarke, & Todd, 1985). The idea is that performers construct a mental representation of the musical structure they wish to communicate to listeners and that this structure is conveyed to listeners by the performance microstructure. The performance microstructure should aid listeners in the creation of an appropriate mental representation. The intended metric structure may be conveyed either directly or indirectly to the listener by the use of expressive variations. Direct communication of the metric structure may be provided by specific, systematic expressive variations related to the metric structure: events in strong metric locations are often delayed in time and played louder, more legato, and longer than events in weaker metric locations (Drake & Palmer, 1993; Gabrielsson, Bengtsson, & Gabrielsson, 1983; Penel & Drake, 1998; Sloboda, 1983, 1985). Listeners may use these cues (in addition to those in the musical structure itself) to focus on one particular metric level (Sloboda, 1983, 1985). Indirect communication of the metric structure may be provided by expressive variations related to other aspects of the musical structure (such as grouping, harmonic, and voice structures). These other expressive variations may enhance the creation of a mental representation for the music and thus help listeners code, in a more efficient way, other aspects of the musical structure such as the metric structure, even without specific metric expressive variations. Both the direct and indirect structural communication hypotheses suggest that synchronization should be easier with real performances (containing multiple expressive variations that high-

light the musical structure) than with mechanical performances (containing no such cues).

In order to evaluate the role of expressive microstructure in synchronization performance, musicians and nonmusicians were required to tap in time with three versions of six musical excerpts: mechanical versions (synthesized by computer according to the score), accented versions (mechanical versions containing intensity accents on the first beat of each measure), and expressive versions (performed by an expert pianist, thus containing many types of expressive variation). The main focus was on the differences between synchronization with the mechanical and expressive versions (contrasting completely expressionless with completely expressive performances). The accented versions were added to examine the influence of intensity accents on the first beats of measures, supposedly major cues in conveying and perceiving meter (Sloboda, 1985).

Psychological Processes Involved in Synchronizing With Music

What psychological mechanisms may be involved in synchronization with musical excerpts? The Dynamic Attending Theory proposed by Mari Riess Jones (Jones, 1976, 1987, 1990; Jones & Boltz, 1989) suggests one possible set of processes. We will use this theoretical framework in evaluating how well, and in what way, listeners are able to synchronize with our musical excerpts.

Jones considers that each individual has an internal rhythm or rate at which event processing is optimal (the referent period). When listening to a complex sequence such as music (composed of multiple hierarchical metric levels), listeners *attune* their internal rhythms to one of those in the sequence. Attunement usually occurs with events at the hierarchical level closest to a listener's referent period, called the *referent level* (see Figure 1).

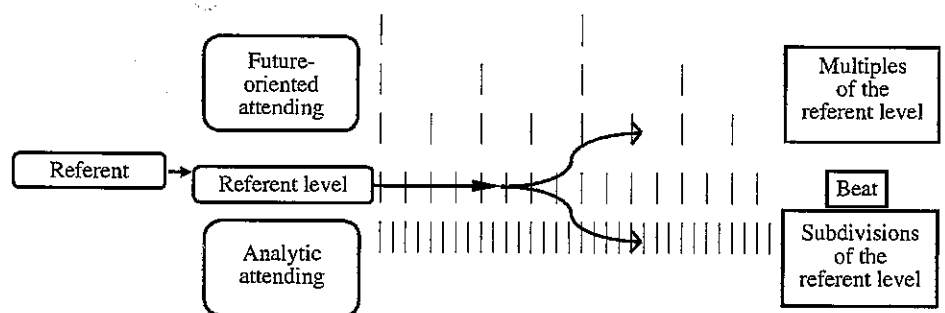


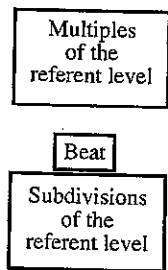
Fig. 1. Summary of some aspects of the Dynamic Attending Theory of Jones. See text for details.

ances (containing
 e in synchroni-
 quired to tap in
 l versions (syn-
 ions (mechanical
 each measure),
 thus containing
 the differences
 e versions (con-
 performances).
 ce of intensity
 es in conveying

With Music

onization with
 by Mari Riess
 suggests one pos-
 k in evaluating
 with our mu-

hm or rate at
 n listening to a
 rchical metric
 ose in the se-
 rchical level
 (see Figure 1).



nes. See text for

Once listeners have “latched on to” the music at this intermediate rate, they are able to direct their attention towards other hierarchical levels by a process of *focal attending*. They can direct their attention down the metric hierarchy toward events occurring at faster rates (*analytic attending*) or up the metric hierarchy toward events occurring at slower rates (Jones calls this *future-oriented attending*). These focal listening strategies do not exclude the possibility of adopting a global listening strategy whereby the multiple metrical levels are perceived simultaneously. Here, we examine the functioning of three processes when synchronizing with musical excerpts: (1) Attunement abilities are measured by the degree to which participants are able to successfully synchronize with musical excerpts. (2) The referent level is measured by the hierarchical metrical level corresponding to the rate at which participants spontaneously synchronize. (3) Focal attending is measured by asking listeners to synchronize with music at rates faster (or slower) than their referent level, and then faster (or slower) still, until they can no longer accomplish the task. This method provides an indication of the ease with which individuals are able to move away from their referent level and the range of hierarchical levels accessible to them.

Specific predictions can be derived from the two theoretical positions of beat induction and structural communication about the influence of expressive variations on the three processes derived from Jones’ theory: (1) Attunement: Beat induction models predict that it should be easier to synchronize with mechanical performances than with expressive ones because of the temporal regularity of the mechanical performances. The structural communication hypothesis predicts the opposite (expressive easier than mechanical) because the metric structure is highlighted directly or indirectly by the performance microstructure. (2) Referent level: No predictions can be drawn from beat induction models about the influence of performance microstructure on the rate of spontaneous synchronization. The structural communication hypothesis predicts slower spontaneous synchronization rates for expressive than for mechanical performances: Performance microstructure should facilitate the creation of an appropriate mental representation for the music, thus enabling listeners to perceptually organize events over longer time spans and to perceive longer units. (3) Focal attending: Beat induction models predict a wider range of accessible hierarchical levels for mechanical than for expressive performances because of the greater temporal regularity of the mechanical performances. Two opposing predictions can be derived from the structural communication hypothesis. A wider range of accessible hierarchical levels would be predicted if performance microstructure communicates (either directly or indirectly) the metrical hierarchy, whereas fewer hierarchical levels would be accessible if uncertainty about the appropriate metrical level is reduced by the performance microstructure.

Role of Musical Training

We investigate these abilities in both musicians and nonmusicians because previous research has suggested that musical training results in a change in the way listeners perceive rhythm (Bamberger, 1986; Drake, 1993a, 1993b; Palmer & Krumhansl, 1990). Concerning the role of expressive microstructure, both beat induction models and the structural communication approach suggest that musicians should be better synchronizers than nonmusicians are, in one case because of enhanced temporal control, in the other case because of enhanced skills in "using" the expressive variations. We make three specific predictions about changes with musical expertise. First, musicians should be better able to attune their internal rhythms to those of the musical excerpts, and so they should be better able to synchronize their taps with them. Second, previous work has suggested that children who have received musical training tend to use slower referent levels than people who have not received such training (Drake, Jones & Baruch, in press). We consider that this reflects musicians' greater ability to perceptually organize events over long time spans, perhaps resulting in the perception of longer units. Third, musicians may have enhanced focal attending abilities, being more able to extract the metrical hierarchy from music (Bigand, Lerdahl, & Pineau, 1994; Palmer & Krumhansl, 1990) and thus having a more complete hierarchical representation of it. If this is the case, musicians should be able to synchronize with more hierarchical levels than nonmusicians are able to. We therefore predicted a wider range of accessible hierarchical levels for musicians.

Relationship Between Perception and Action

An assumption underlying the present study is that spontaneous tapping in relation to music (a motor behavior) allows the investigation of the above-mentioned processes (attuning ones' internal rhythms to those in the environment, spontaneously focusing attention on events at one particular hierarchical level, and directing attention toward multiple hierarchical levels). We make this assumption because we consider that both beat perception and beat synchronization access similar central timing activities (Frasse, 1963; Ivry & Hazeltine, 1995; Keele, Pokorny, Corcos, & Ivry, 1985; see also Drake et al., in press). Previous studies have made the same assumption. For instance, Duke (1989) examined the way tempo affects the perceived beat of isochronous sequences by asking nonmusicians to tap along. Similarly, Parncutt (1994) investigated the way rhythmic structure influences the perceived beat of metronomic, cyclic, rhythmic sequences, and

Vos et al. (1994) investigated the perception of meter by asking participants to tap at the beginning of each perceived measure.

Method

PARTICIPANTS

Thirty-six participants took part in the experiment, all of whom reported normal hearing. Eighteen were nonmusicians and 18 were musicians. Groups were matched in age (nonmusicians: mean = 23.6 years, range = 20 to 28 years, SD = 2.4 years; musicians: mean = 24.4 years, range = 20 to 35 years, SD = 3.2 years). Nonmusicians had never received any formal music training or played any instrument, whereas musicians had a mean of 7 years of musical training (range = 2 to 12 years), 14 years of practice (range = 5 to 27 years), and 8 hours of practice per week (range = 2 to 25 hours). Thus, participants regarded as musicians covered a wide range of training and practice. Consequently, differences obtained between our two groups of participants reflect robust effects that might have been even stronger with a more homogeneous group of musicians.

MATERIALS

Six pieces of music for piano were used, covering a wide range of styles encountered in Western tonal music: the Minuet I from Partita No.1 by J. S. Bach, the Sonata in B (K.261-L.148) by D. Scarlatti, the first movement of an easy sonata (op. 49, no. 2) by L. van Beethoven, a mazurka (op. 7, no. 1) by F. Chopin, the "Menuet sur le nom de Haydn" by M. Ravel, and a Romanian folk dance ("The Dance with the Staff") by B. Bartók. The average tempi used by the pianist for the pieces corresponded to the following quarter-note durations: Bach (3/4) 508 ms (118 bpm), Scarlatti (2/4) 667 ms (90 bpm), Beethoven (4/4) 400 ms (150 bpm), Chopin (3/4) 417 ms (144 bpm), Ravel (3/4) 600 ms (100 bpm), Bartók (2/4) 667 ms (90 bpm).

None of the musicians had ever performed any of the pieces, but some musicians knew some of them. We do not consider this a methodological problem for two reasons. First, we think that musical training is necessarily confounded with familiarity with music and in particular with the different musical styles of the excerpts chosen: A trained musician is necessarily familiar with them and becomes increasingly familiar with training. Second, we believe that familiarity with the musical style is more important than familiarity with a particular excerpt in determining musicians' behavior compared with nonmusicians' behavior.

Three versions of each of the pieces were created. The *expressive* version was a real performance by a professional pianist (concert pianist and piano teacher). It thus contained many performance variations in timing, intensity, articulation and pedaling. The *mechanical* version was synthesized by computer. All expressive variations were absent in this version, that is, the acoustic result corresponded exactly to the written score (without what may be considered to be implied by a musical score, for instance louder melody notes; we consider such pattern to be part of expression, expression being of course partly implied by the score). The *mechanical* versions were played at the mean tempi used by the pianist, and for each piece, a mean intensity representing the arithmetic mean of the MIDI velocities of the notes in the expressive version was applied to all notes. Articulation was defined for each note by its duration in the score (hence, everything was played legato except for rests). No pedal was used. The *accented* version was a synthesized mechanical version in which metrical accents were added by increasing by 10 dB the intensity of notes situated on the first beat of each measure (10 dB seemed appropriate perceptually: not too loud and not too soft). Thus, the mechanical versions were also regular in both time and intensity, the ac-

cented versions were regular in time and intensity (with 2 intensity levels and intensity accents occurring regularly), but the expressive versions were regular in neither time nor intensity.

APPARATUS

The performances of the expressive versions were recorded on a Yamaha Disklavier II monitored by a computer using the sequencer Master Tracks Pro4. The pianist was instructed to perform the pieces in a natural, musical fashion. In the same silent room, all versions were played back on the Disklavier using Master Tracks Pro4, line out. Participants tapped in time with the music on a Roland pad PD7 connected to a sound synthesizer Roland TD5 and to the computer for the recording of the taps in Master Tracks Pro4, line in. Participants received auditory feedback from their taps, corresponding to a drum sound.

PROCEDURE

Each participant heard all six pieces, each in only one of the three versions, so that each type of version occurred twice (two pieces in the expressive version, two in the mechanical version, and two in the accented version). The two pieces heard in each version varied systematically between participants, and the order of the versions and of the pieces was counterbalanced in a mixed order.

For each piece, the participants first listened to the music to familiarize themselves with it. They could listen to it several times or interrupt the listening phase before the end of the piece. Then, during the production phase, the piece was started again and they were asked to tap on the pad with their dominant hand in the way they found the easiest (some with the tips of their fingers, others with their knuckles). They were instructed to tap regularly and in synchrony with the music, as soon as they were able to. In a first phase, they were asked to tap at their spontaneous speed, that is, at the rate that seemed most appropriate to the musical excerpt. For half of the participants, in a second phase, they were asked to tap, if they could, faster than their spontaneous speed, but still staying in time with the music, and then if they could faster still, and so on. In a third phase, they were asked to tap, if they could, slower than their spontaneous speed, but still staying in time with the music, and then if they could slower still, and so on. The other half of the participants did the opposite order (slower and then faster). There was no pause between phases, but a small break between pieces. The experimental session lasted about half an hour.

DEPENDENT MEASURE

A participant was considered to have successfully synchronized with a piece of music if the taps coincided with a particular metrical level of the *acoustic sequence* (with a tolerance window of $\pm 10\%$ of interval duration). In the following, the ability to synchronize will implicitly refer to the ability to synchronize *successfully*, not to the ability to tap anything on the pad. We compared the timing of the produced taps with the acoustic sequence rather than with a theoretical isochronous sequence in order to prevent rejecting as incorrect taps in the expressive versions that were not regularly timed but fluctuated in the same way as the performers' tempo. We then established the metrical level at which the participant synchronized and calculated the corresponding average interonset interval (IOI) between taps (in milliseconds).

Results

In general, the results described here were valid for all six excerpts. In a few cases, a different pattern was observed for one excerpt, but the particu-

lar one varied for each dependent measure. The data were therefore collapsed across excerpts for clarity.

ATTUNEMENT (ABILITY TO SYNCHRONIZE SPONTANEOUSLY)

General Synchronization

A rough indication of the ability to synchronize with the musical excerpts was obtained by counting the number of times participants were able to synchronize with at least part (10 successive taps) of the excerpt. Synchronization was generally very good (93%). An analysis of variance (ANOVA) on the number of successful synchronizations by group (musicians and nonmusicians) and version (mechanical, accented, expressive) revealed that musicians (98%) were better than nonmusicians (88%), $F(1, 34) = 5.3$; $p < .03$, and that there were no significant differences among the three versions (mechanical = 94%, accented = 96%, expressive = 89%).

Initial Synchronization

This more sensitive score provided an indication of how quickly participants were able to start synchronizing (they were required to start synchronizing as soon as possible). Figure 2 shows the average number of measures out of the first 10 measures (measures 2-11) with which participants syn-

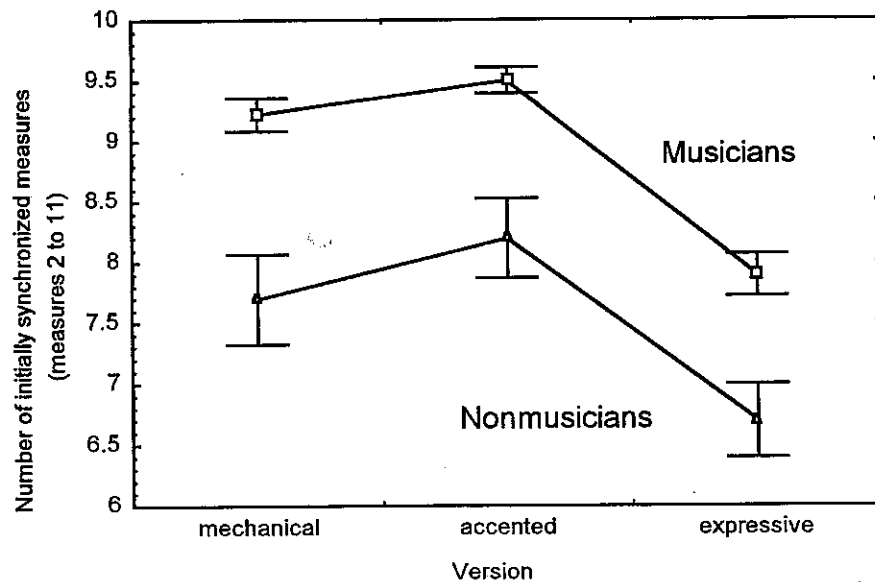


Fig. 2. Number of successfully synchronized measures (measures 2-11) for the three versions (mechanical, accented and expressive) and two groups of participants (musicians and nonmusicians).

chronized successfully. Initial synchronization was good (8.2 successfully synchronized measures). An ANOVA on the number of initially synchronized measures by group (musicians and nonmusicians) and version (mechanical, accented, and expressive) revealed that musicians (8.9 measures) did better than nonmusicians (7.5 measures), $F(1, 34) = 20.5$; $p < .001$. Initial synchronization also varied with version, $F(2, 68) = 7.15$; $p < .01$: It was significantly more difficult with expressive (7.3 measures) than with mechanical (8.5 measures) and accented (8.8 measures) versions, $F(1, 34) = 13.3$; $p < .0001$, which did not differ significantly. Thus, temporal regularity facilitated initial synchronization, whereas the additional presence of regularly occurring intensity accents did not contribute further. This effect of version was the same for both groups; there was no significant interaction between group and version.

Optimal Synchronization

To find out whether the same pattern of results characterized synchronization later in the excerpts, we identified, for each participant and excerpt, which measures exhibited successful synchronization and identified the 10 successive measures for which the number of correctly synchronized measures was the highest. Not surprisingly, optimal synchronization was better than initial synchronization (9.3 measures compared with 8.2 measures for initial synchronization, $t(34) = 11.17$; $p < .0001$). A ceiling effect was observed for musicians. In all other respects, an ANOVA on the number of optimally synchronized measures by group (musicians and nonmusicians) and version (mechanical, accented, expressive) confirmed the results obtained with global and initial synchronizations. Optimal synchronization was better for musicians (9.8 measures) than for nonmusicians (8.8 measures), $F(1, 34) = 5.76$; $p < .02$. It also varied with version, $F(2, 68) = 2.68$; $p < .04$: optimal synchronization was significantly worse with expressive (8.8 measures) than with mechanical (9.4 measures) and accented (9.7 measures) versions, $F(1, 34) = 5.38$; $p < .03$, which did not differ significantly. No significant interaction was observed between group and version, indicating that the effect of version was the same for both groups.

In sum, three measures of the ability to spontaneously synchronize provide converging results. Musicians were better than nonmusicians at synchronizing with the musical excerpts, demonstrating heightened attunement abilities. The expressive versions were somewhat more difficult to synchronize with than the temporally regular mechanical and accented versions, whatever the level of musical training. The regularly occurring intensity accents of the accented versions did not lead to better performance compared with the entirely mechanical versions.

REFERENT LEVEL (SPONTANEOUS SYNCHRONIZATION RATE)

An indication of the referent level was obtained by establishing how fast participants tapped when they spontaneously synchronized successfully. All subsequent analyses on spontaneous synchronizations are based on the 10 measures that exhibited optimal synchronization, as defined earlier.

Spontaneous Synchronization IOI

Figure 3 presents the mean tapping rates (IOI in milliseconds) for spontaneous synchronizations. Note that these means were computed over tapping rates corresponding to discrete levels in the metrical hierarchy. The results were analyzed in a new ANOVA on the spontaneous synchronization IOI with the variables of group and version. On average, musicians tapped more slowly than nonmusicians tapped (musicians: IOI = 1042 ms, nonmusicians: IOI = 858 ms), $F(1, 34) = 5.7$; $p < .03$. Synchronization IOIs also varied with version, $F(2, 68) = 6.0$; $p < .01$: they were slower with accented and expressive versions than with mechanical versions (mechanical: IOI = 798 ms, accented: IOI = 1028 ms, expressive: IOI = 1024 ms), $F(1, 34) = 8.5$; $p < .01$. A significant interaction between group and version, $F(2, 68) = 3.5$; $p < .04$, indicates that different effects of version were observed for the two groups. Planned comparisons showed that musicians tapped significantly more slowly for expressive than for accented and me-

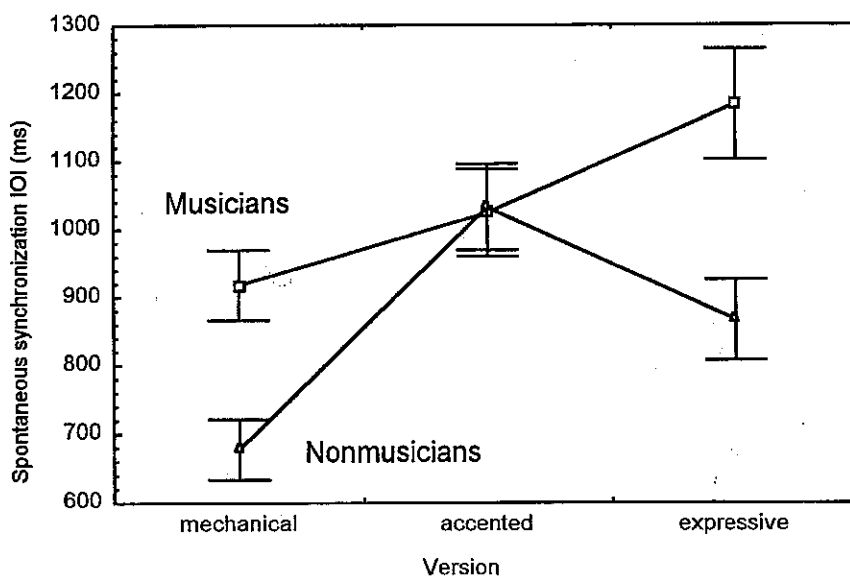


Fig. 3. Mean interonset interval (IOI) between taps (in milliseconds) for spontaneous synchronization, for the three versions and two groups of participants.

chanical versions, $F(1, 34) = 3.8$; $p < .05$, which did not differ significantly. Nonmusicians tapped significantly faster for mechanical than for accented and expressive versions, $F(1, 34) = 5.7$; $p < .02$, which did not differ significantly. Thus, both groups tapped the fastest with the mechanical versions. The added intensity metrical accents in the accented versions and the performance microstructure in the expressive versions led to slower tapping rates compared with the perfectly regular mechanical versions, at least with the nonmusicians. The general trend for musicians to tap more slowly than nonmusicians was observed for the mechanical and expressive versions, $F(1, 34) = 4.1$; $p < .05$; $F(1, 34) = 6.1$; $p < .01$, respectively, but not for the accented versions. This difference suggests that the regularly occurring intensity accents in the accented versions led to the metric structure of the pieces being perceived in the same way by both groups of participants.

Metrical Level

As we have seen, tapping was slower in some conditions than in others, but because the tempo and meter varied for the six excerpts (there were two pieces in 2/4, one piece in 4/4, and three pieces in 3/4), this slower tapping may or may not have corresponded to higher metrical levels. In order to investigate this issue, we considered five categories of metrical level: below the quarter-note level, the quarter-note level, between the quarter-note level and the measure level, the measure level, and above the measure level. Figure 4 presents histograms of spontaneous tapping levels for the two groups of participants for the three versions. Most synchronizations occurred either at the quarter-note or measure levels, with a two-peaked distribution for both groups in all versions, with the exception of nonmusicians in the mechanical versions. In the mechanical versions, the nonmusicians tapped mainly at the quarter-note level, whereas the musicians tapped either at the quarter-note or measure level. When regularly occurring intensity accents were added at the measure level (accented versions), more nonmusicians tapped at this level, whereas the musician's behavior was not significantly modified. The presence of performance microstructure in the expressive versions lead to a greater proportion of synchronizations at other levels (below, in between, above).

Musicians tapped more frequently at the measure level and above than nonmusicians did (musicians, 46%; nonmusicians, 32%), $\chi^2(1, N = 201) = 3.96$; $p < .05$, and inversely, nonmusicians tapped more frequently at the quarter-note level and below than musicians did (musicians, 44%; nonmusicians, 58%), $\chi^2(1, N = 201) = 4.21$; $p < .05$). In these χ^2 analyses, participant-excerpt combinations were considered to be independent observations. Thus, the slower tapping of musicians compared with nonmusicians derived from tapping at a higher metrical level. Synchronization level also varied for the three versions, with a different pattern for the

and
 offer significantly.
 than for accented
 not differ signifi-
 canical versions.
 ons and the per-
 slower tapping
 ons, at least with
 more slowly than
 ressive versions,
 , but not for the
 ly occurring in-
 structure of the
 participants.

s than in others,
 epts (there were
 /4), this slower
 etrical levels. In
 ries of metrical
 between the quar-
 above the mea-
 apping levels for
 st synchroniza-
 ls, with a two-
 he exception of
 al versions, the
 ereas the musi-
 When regularly
 l (accented ver-
 usician's behav-
 rmance micro-
 proportion of

and above than
 32%), χ^2 (1,
 pped more fre-
 did (musicians,
 5). In these χ^2
 to be independ-
 compared with
 l. Synchroniza-
 pattern for the

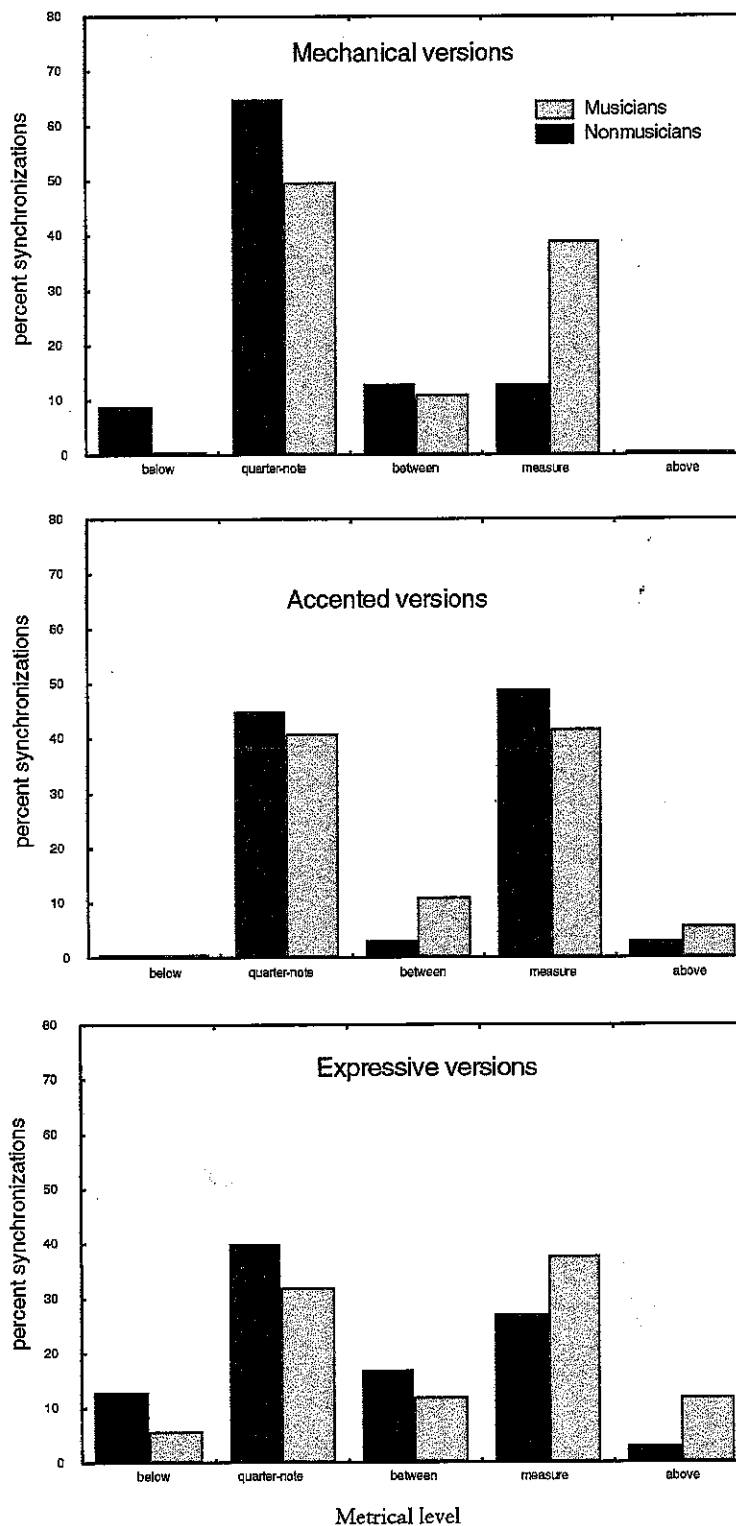


Fig. 4. Percentage of spontaneous synchronizations at each metrical level for the three versions, for the nonmusicians (dark bars) and musicians (light bars).

two groups. The pattern of results mirrors almost perfectly that observed for the mean synchronization IOI. Musicians tended to tap at higher hierarchical levels for the expressive versions, followed by the accented versions, and finally the mechanical versions. However, these effects were not significant with the χ^2 test. Nonmusicians tended to tap at higher hierarchical levels for the accented and expressive versions, compared with the mechanical versions, χ^2 (2, $N = 95$) = 6.19; $p < .05$ for the quarter-note level and below category, χ^2 (2, $N = 95$) = 11.5; $p < .01$ for the measure level and above category). The trend for musicians to tap at higher hierarchical levels than nonmusicians was only true for the mechanical and expressive versions: Musicians tapped at higher levels than nonmusicians for the mechanical (musicians, 0.78; nonmusicians, 0.17) and expressive versions (musicians, 1.32; nonmusicians, 0.57), but at similar levels for the accented versions (musicians, 1.12; nonmusicians, 1.16).

Theoretical Meter

In foregoing analyses, synchronizations were considered correct whatever their period or phase (as long as they satisfied the $\pm 10\%$ criterion relative to note onsets). For example, synchronizations at the dotted-quarter note level in 2/4 measures were included (theoretically incorrect period), as were synchronizations at the measure level falling on the second beat of 2/4 measures (theoretically incorrect phase; see Figure 5). Here, the period was considered theoretically correct if it corresponded to one level of the piece's hierarchical metric structure. The phase was considered theoretically correct if taps were in phase with the corresponding metrical level of the piece's hierarchical metric structure. Thus, if the period was correct, the phase could either be correct or incorrect. Because the periods of the eighth- and quarter-note levels are always theoretically correct, we concen-

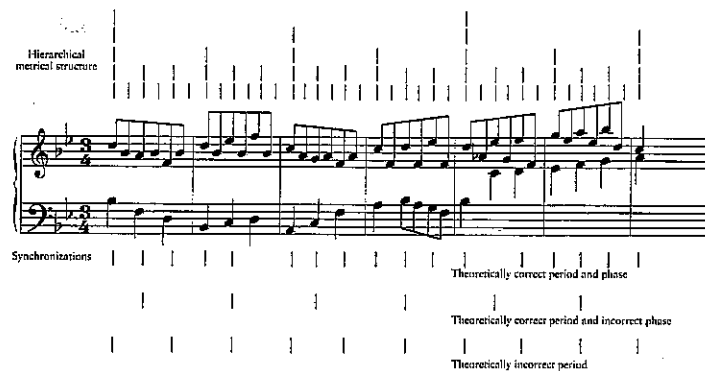


Fig. 5. Examples of synchronizations with theoretically correct/incorrect period and phase.

trated on synchronizations *above the quarter-note level*, where some ambiguity about the theoretically correct period and phase may arise.

In general, synchronizations above the quarter-note level agreed with the pieces' hierarchical metric structure, having both the correct period and correct phase (period = 82%, period and phase = 76%). It was more frequent for musicians (period = 87%, period and phase = 82%) than nonmusicians (period = 75%, period and phase = 68%). Surprisingly, synchronizations agreed with the theoretical metric structure more frequently with the mechanical version (period = 81%, period and phase = 81%) and the accented version (period = 90%, period and phase = 79%) than with the expressive versions (period = 74%, period and phase = 69%). However, these differences were not significant with χ^2 tests, probably because of the low number of observations involved.

In sum, the rate at which participants spontaneously synchronized varied systematically as a function of musical expertise and type of performance microstructure. First, musicians tapped spontaneously more slowly than nonmusicians did, which corresponded to a tendency to choose a higher metrical level, and musicians were more likely to respect the theoretical metrical structure of the excerpt. Thus, musical expertise led to a modification in the way listeners apprehended the music: in particular, they focused spontaneously on slower hierarchical levels, using a slower referent level. Second, synchronization rate varied with version. It was slowest with the expressive versions, intermediate with the accented versions, and fastest with the mechanical versions. In each case, slower tapping corresponded to higher metrical levels. Of note was the fact that the slower tapping of musicians compared with nonmusicians was valid only for the mechanical and expressive versions: Both groups tapped at the same rates with the accented versions, and only the musicians slowed further in the presence of performance microstructure.

FOCAL ATTENDING (FASTER AND SLOWER SYNCHRONIZATIONS)

After synchronizing with music at their spontaneous synchronization rate, participants were asked to synchronize faster and faster still, and slower and slower still. The same synchronization criteria were adopted here as for the spontaneous synchronizations. These nonspontaneous synchronizations shed light on the ability to focus on other hierarchical levels and thus provide an indication of focal attending.

Synchronizing at One Faster and One Slower Rate

A first measure of focal attending indicating the ease with which participants were able to move away from their referent level is provided by participants' ability to synchronize with at least one hierarchical level above

(slower = future-oriented attending) and below (faster = analytic attending) their spontaneous level. In terms of the general synchronization criterion defined earlier, focal attending was easier for musicians (95% of all excerpts) than for nonmusicians (85%), $\chi^2(1, N = 402) = 13.7$; $p < .001$. It was easier for the mechanical (92%) and accented versions (94%) and harder for the expressive versions (83%), $\chi^2(2, N = 402) = 10.4$; $p < .01$. Participants were more able to synchronize with at least one slower (94%) than with at least one faster (86%) level, $\chi^2(1, N = 402) = 7.11$; $p < .01$.

Number of Faster and Slower Synchronizations

A second measure of focal attending indicating the extent of the perceived hierarchical structure is provided by the number of levels at which each participant was able to synchronize above and below their spontaneous synchronization rate (see Figure 6). On average, participants were able to synchronize at 2.7 rates in addition to their spontaneous synchronization rate. An ANOVA was run on the number of additional synchronizations with the variables of group, version, and synchronization type (subdivision versus multiplication). Musicians were able to synchronize at more rates (3.2 rates) than nonmusicians (2.2 rates), $F(1, 34) = 21.23$; $p < .0001$. The number of synchronizations varied with version, $F(2, 68) = 22.05$; $p < .0001$: There were more synchronizations with mechanical and accented versions (3.0 rates in both cases) than with expressive ones (2.0 rates), $F(1, 34) = 35.7$; $p < .01$. Finally, participants were able to synchronize at more rates above their spontaneous synchronization rate than below (above = 1.4, below = 1.2), $F(1, 34) = 8.37$; $p < .01$.

Theoretical Meter

Synchronizations at hierarchical levels other than the spontaneous synchronization level frequently agreed with the pieces' hierarchical metric structure, both in terms of period and phase (period = 77%, period and phase = 62%). This was more frequent for the musicians (period = 82%, period and phase = 71%) than for the nonmusicians (period = 71%, period and phase = 51%), period: $\chi^2(1, N = 324) = 5.51$, $p < .02$; period and phase: $\chi^2(1, N = 324) = 13.2$; $p < .001$. Thus, the period was respected by about three quarters of the participants, whereas only half of the nonmusicians' synchronizations respected the phase: they adopted an appropriate rate, but they did not tap on the first beat of the measure. The period and phase were respected more in the expressive versions (period = 88%, period and phase = 65%) and accented versions (period = 79%, period and phase = 88%) than in the mechanical versions (period = 68%, period and phase = 50%), period: $\chi^2(2, N = 324) = 11.7$, $p < .01$; period and phase: $\chi^2(2, N = 324) = 16.0$; $p < .01$.

analytic attend-
synchronization critere-
s (95% of all
13.7; $p < .001$. It
94%) and harder
; $p < .01$. Partici-
power (94%) than
1; $p < .01$.

extent of the per-
of levels at which
y their spontane-
participants were able
ous synchroniza-
synchronization
synchronization
synchronization type (sub-
synchronize at more
11.23; $p < .0001$.
(68) = 22.05; $p <$
cal and accented
s (2.0 rates), $F(1,$
synchronize at more
below (above =

spontaneous syn-
rarchical metric
7%, period and
(period = 82%,
d = 71%, period
.02; period and
was respected by
only half of the
adopted an ap-
e measure. The
rsions (period =
period = 79%, pe-
(period = 68%,
 $p < .01$; period

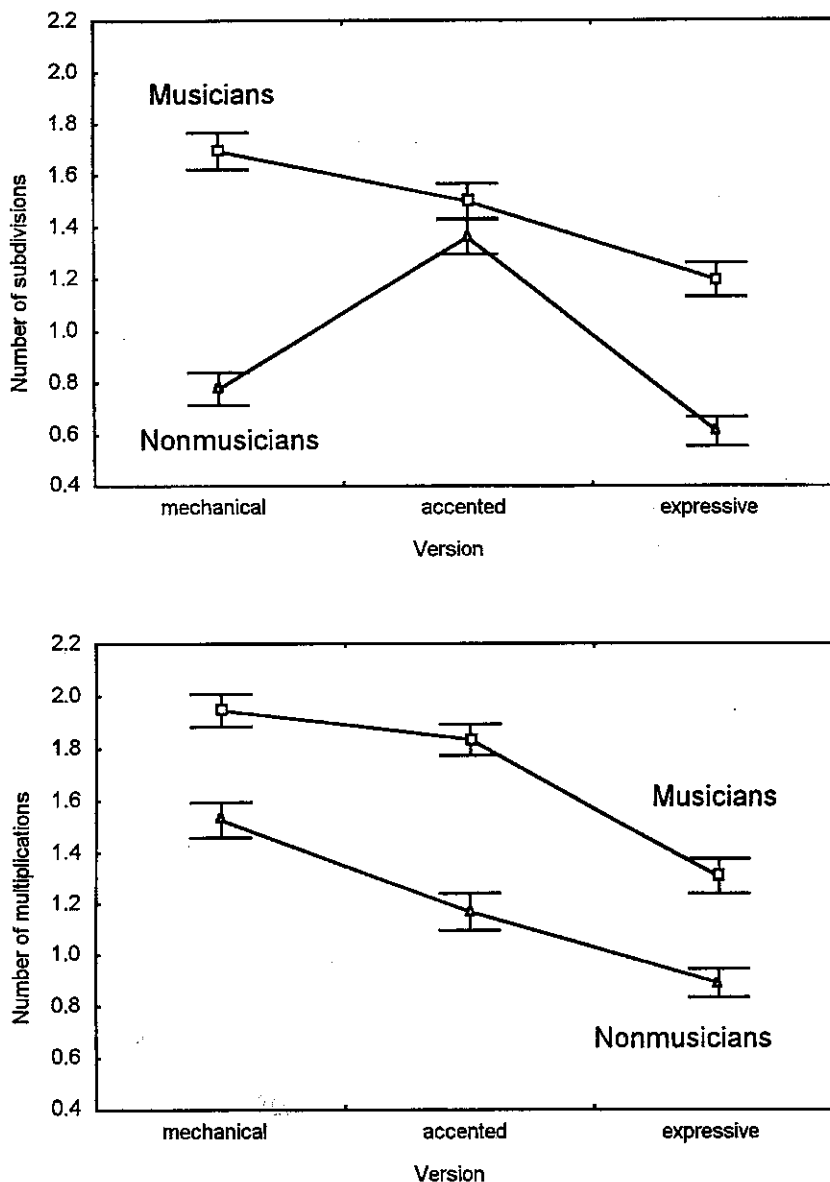


Fig. 6. Mean number of subdivisions and multiplications, for the three versions and two groups of participants.

In sum, enhanced focal attending abilities have been demonstrated in musicians compared with nonmusicians: Musicians were more able to synchronize with levels other than their referent level, and they accessed a wider range of rates. Musicians synchronized more frequently than did nonmusicians with the theoretically correct metrical period and phase. Dif-

ferences were also observed among the three versions: The performance microstructure of the expressive versions made it harder to synchronize with levels other than the referent level and restrained the range of synchronization levels. However, participants synchronized more frequently with the theoretically correct metric structure in the expressive versions than the mechanical versions. Multiplications were easier and more frequent than subdivisions.

Discussion

Tapping in time with music is a common, everyday behavior that has been the object of relatively few experimental investigations. Our experimental paradigm has demonstrated that tapping to music is relatively easy for nonmusicians and for musicians, confirming the appropriateness of this task for studying the way both nonmusicians and musicians perceive the temporal structure of music. We measured the different ways in which participants spontaneously synchronized with a variety of musical excerpts and how well they were able to synchronize with other hierarchical levels. We first discuss the results in relation to the psychological processes involved in accomplishing these tasks, focusing in particular on the role of musical expertise. We also investigated the influence of certain types of performance microstructure on these processes by comparing synchronization with different versions of the same pieces. The influence of this microstructure on each of the processes is then discussed.

PROCESSES INVOLVED IN SYNCHRONIZATION: ROLE OF MUSICAL TRAINING

An indication of *attunement*, that is, an individual's ability to entrain his/her internal rhythms to those in the environment, was provided by three measures of the ability to synchronize taps with the musical excerpts: general, initial, and optimal synchronization. As expected, attunement was better in musicians than in nonmusicians. Although we cannot reject entirely the possibility that these differences were due to musicians' enhanced motor skills, the level of motor skill involved was minimal, and tapping constitutes a commonly observed spontaneous behavior.

An indication of the *referent level* was obtained by measuring participants' spontaneous synchronization rates and the metrical levels with which they are aligned. Musicians tend to spontaneously synchronize at slower rates (corresponding to higher metrical levels) than nonmusicians do, providing support for the hypothesis that musicians use slower referent levels than nonmusicians use. This reflects musicians' ability to perceptually organize events over longer time spans than nonmusicians can, perhaps re-

sulting in the perception of longer units. This effect of musical expertise is much less obvious than the preceding one (that musicians synchronize better than nonmusicians do), but it is important and confirms previous results showing that musicians tap more slowly on average than nonmusicians do (Drake et al., in press).

An indication of *focal attending*, that is, the ability to focus attention on levels other than the referent level, was provided by asking participants to synchronize faster and slower than their spontaneous synchronization rate. This task proved to be quite easy: Most participants were able to produce at least one faster and one slower synchronization (86% and 94%, respectively). These abilities were enhanced in musicians compared with nonmusicians: Musicians were more able to produce one faster and one slower synchronization than were nonmusicians, and musicians were able to produce more faster and slower synchronizations than nonmusicians were able to produce. Musicians' synchronizations respected the metric structure more frequently than nonmusicians' synchronizations did so. Thus, musicians have easier access to levels other than their referent level, they have access to a wider range of hierarchical levels, and they better apprehend the theoretical metric structure. Musicians are more able to extract the metrical hierarchy from music, reflecting the existence of a more complete hierarchical representation.

These results confirm those obtained by Palmer and Krumhansl (1990) with goodness-of-fit judgments of temporal patterns in metrical contexts, which showed that mental representations for meter were enriched by musical training. Finer differentiation between hierarchical levels was achieved by musicians than by nonmusicians, and in particular, musicians were better able to perceive subdivisions of metrical levels.

Multiplications of the referent level reflect future-oriented attending, subdivisions reflect analytic attending. Our results confirm the prediction of Jones' Dynamic Attending Theory, which stipulates that future-oriented attending requires less effort than analytic attending requires: Multiplications were easier and more frequent than subdivisions.

ROLE OF PERFORMANCE MICROSTRUCTURE

The main aim of this study was to investigate the role of performance microstructure on the processes described in the preceding section. To our knowledge, this is the first demonstration that the presence of timing and intensity variations influences listeners' synchronization with music. Indeed, most of the measures examined in this study varied with version, sometimes interacting with musical expertise. These results are generally in agreement with the predictions derived from both the beat and meter induction models and the structural communication approach to performance.

Attunement was poorer with the expressive versions than with the mechanical and accented versions. This observed pattern is not consistent with predictions derived from the structural communication approach to music performance, which predicts that performance microstructure should facilitate the creation of an appropriate mental representation of metrical structure, either by direct or indirect communication. In contrast, the observed pattern is coherent with predictions of beat and meter induction models, namely that the more temporally regular a musical sequence, the easier it is to extract the underlying beat. This makes sense, as the underlying beat is temporally regular by definition. Thus, although meter might be communicated either directly or indirectly by performance microstructure, it is extracted more easily when the sequence is temporally regular, to the extent that synchronized tapping can be expected with beat extraction. Compared with the mechanical versions, the presence of metrical intensity accents in the accented versions did not significantly improve synchronization performance, suggesting that temporal regularity, rather than regularly occurring intensity accentuation, is the fundamental feature enabling participants to synchronize spontaneously.

The referent level was also influenced by the microstructure, but in different ways for musicians and nonmusicians. Musicians' referent levels were slower and higher in the metrical hierarchy for expressive versions than for accented and mechanical versions. Nonmusicians' referent levels were slower and higher in the hierarchy with accented and expressive versions compared with mechanical versions. Thus, participants in both groups used slower referent levels with expressive than with mechanical versions.

One could argue that tapping more slowly results from uncertainty about "when to tap" with temporally fluctuating music. Although this explanation cannot be completely excluded, several facts prevent us from adopting it. First, musicians used slower referent levels than do nonmusicians, and this cannot be explained by uncertainty about "when to tap." On the contrary, this difference suggests that the use of slower referent levels reflects the ability to organize events perceptually over longer time spans. Second, although musicians and nonmusicians synchronized in the same way with accented versions (slower referent levels for musicians than for nonmusicians were only observed with mechanical and expressive versions), performance microstructure present in the expressive versions further slowed down the tapping rate only for musicians. In contrast, nonmusicians slightly speeded up their tapping rate, although the effect was not significant. Thus, we believe that performance microstructure enables musicians to organize events perceptually over longer time spans and perhaps to perceive longer units. This result is coherent with a structural communication approach to performance.

Another finding of interest is that musicians and nonmusicians spontaneously synchronized in the same way with accented versions. Intensity

han with the me-
ot consistent with
pproach to music
ucture should fa-
ation of metrical
contrast, the ob-
meter induction
cal sequence, the
e, as the underly-
gh meter might be
e microstructure,
lly regular, to the
beat extraction.
metrical intensity
ove synchroniza-
rather than regu-
feature enabling

ucture, but in dif-
ferent levels were
versions than for
levels were slower
ve versions com-
both groups used
cal versions.

uncertainty about
ugh this explana-
t us from adopt-
o nonmusicians,
a to tap." On the
referent levels re-
nger time spans.
d in the same way
sicians than for
pressive versions),
ns further slowed
musicians slightly
significant. Thus,
to organize events
longer units. This
n to performance.
musicians sponta-
versions. Intensity

accents seem to represent a common perceptual cue. This finding was indeed suggested by previous work (Drake, 1993a): In an experiment investigating factors that influence the ease of reproduction of short musical rhythms, reproduction was found to be better for rhythms that had intensity accents on important metrical positions. This finding was true for adult musicians and nonmusicians, and also for 5-year-old and 7-year-old children. In the present experiment, other variations present in the expressive versions further slowed down the tapping rate only for musicians. This result is coherent with Sloboda's (1983) findings: Among the six performers Sloboda recorded, one of the least experienced performers confined his metrically related variations to the intensity domain, whereas the most experienced performer varied all three aspects of performance (timing, articulation, and intensity of events) in relation to their metrical position.

Focal attending to higher or lower metrical levels was harder with expressive versions than with mechanical and accented versions. Thus, as for the ability to synchronize spontaneously, focal attending was easiest with temporally regular sequences, in accordance with predictions of the beat and meter induction models. However, following the structural communication approach, the difficulty in focally attending with expressive variations could be attributed to the success of the expressive versions in communicating one particular metrical level to listeners, preventing them from accessing other levels. Thus, microvariations in the expressive versions may play a constraining role, focusing attention on one particular hierarchical level. Future studies will examine this issue further by testing whether or not the degree of tapping variability was lower in the expressive versions, as would be expected if performance microstructure reduces ambiguity about the metrical level.

Whereas the performance microstructure contained in the expressive versions appears to constrain synchronization behavior as regards the ease and range of synchronizations, it did improve synchronization behavior in another way: Synchronizations corresponded more frequently to the metric structure of the musical excerpts. Performance microstructure seems to have been as successful in communicating theoretical meter as systematic intensity accents.¹

References

- Bamberger, J. (1986). Cognitive issues in the development of musically gifted children. In R. Sternberg & J. Davidson (Eds.), *Conceptions of giftedness*. Cambridge: Cambridge University Press.

1. The authors thank Luca Stefan, who conducted the experiment, and David Ramet who performed the music. Thanks also to Caroline Palmer, Bruno Repp, Kate Stevens, and an anonymous reviewer for helpful comments on an earlier version of this article.

- Bigand, E., Lerdahl, F., & Pineau, M. (1994, July). Deux approches expérimentales des 4 composantes de la Théorie Générative de la Musique Tonale. In *Proceedings of the Third International Conference for Music Perception and Cognition* (pp. 259-260). Liège, Belgium: ICMPC.
- Clarke, E. F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 1-25). Oxford: Science Publications.
- Desain, P. (1992). A (de)composable theory of rhythm perception. *Music Perception*, 9, 439-454.
- Drake, C. (1993a). Reproduction of musical rhythms by children, adult musicians and adult nonmusicians. *Perception and Psychophysics*, 53, 25-33.
- Drake, C. (1993b). Influence of age and musical experience on timing and intensity variations in reproductions of short musical rhythms. *Psychologica Belgica*, 33, 217-228.
- Drake, C., Jones, M. R., & Baruch, C. (in press). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*.
- Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception*, 10, 343-378.
- Duke, R. A. (1989). Musicians' perception of beat in monotonic stimuli. *Journal of Research in Music Education*, 37, 61-71.
- Duke, R. A. (1994). When tempo changes rhythm: The effect of tempo on nonmusicians' perception of rhythm. *Journal of Research in Music Education*, 42, 27-35.
- Duke, R. A., Geringer, J. M., & Madsen, C. K. (1991). Performance of perceived beat in relation to age and music training. *Journal of Research in Music Education*, 39, 35-45.
- Fraisse, P. (1963). *Psychology of time*. New York: Harper.
- Gabrielsson, A., Bengtsson, I., & Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandinavian Journal of Psychology*, 24, 193-213.
- Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 3-18.
- Jones, M. R. (1976). Time, our last dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323-355.
- Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics*, 41, 631-634.
- Jones, M. R. (1990). Learning and development of expectancies: An interactionist approach. *Psychomusicology*, 2, 193-228.
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96, 459-491.
- Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: a correlational analysis. *Acta Psychologica*, 60, 173-191.
- Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception*, 11, 115-128.
- Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 331-346.
- Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 728-741.
- Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11, 409-464.
- Penel, A., & Drake, C. (1998). Sources of timing variations in music performance: A psychological segmentation model. *Psychological Research*, 61, 12-32.
- Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2, 411-440.
- Repp, B. H. (1999). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior*, 31, 145-164.
- Rosenthal, D. (1992). Emulation of human rhythm perception. *Computer Music Journal*, 16, 64-76.

- Shaffer, L. H., Clarke, E. F., & Todd, N. P. (1985). Metre and rhythm in piano playing. *Cognition*, 20, 61-77.
- Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, 35A, 377-396.
- Sloboda, J. A. (1985). Expressive skill in two pianists: Metrical communication in real and simulated performances. *Canadian Journal of Psychology*, 39, 273-293.
- Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception*, 6, 555-569.
- Vos, P. G., van Dijk, A., & Schomaker, L. (1994). Melodic cues for metre. *Perception*, 23, 965-976.