Rate Limits of On-Beat and Off-Beat Tapping With Simple Auditory Rhythms: 2. The Roles of Different Kinds of Accent

Bruno H. Repp
Haskins Laboratories

The relative difficulty of on-beat and off-beat finger tapping with simple auditory rhythms was assessed in four experiments with musically trained participants. The rhythms consisted of cyclically repeated TT0 or TTT0 patterns, where T denotes the presence and 0 denotes the absence of a tone. The tasks were to tap in synchrony with one of the T ("on-beat") positions or with the 0 ("off-beat") position. Experiments 1–3 used an adaptive procedure that determined the fastest tempo at which each task could be accomplished. Experiment 1 demonstrated that it is easier to tap on tones that carry a rhythmic grouping accent (T_2 in TT0, T_1 and T_3 in TTT0) than on other tones or in the 0 position. Off-beat tapping was more difficult in TT0 than in TTT0 sequences. Experiment 2 showed that a dynamic (+ pitch) accent on one of the tones facilitates synchronization with that tone and impedes synchronization with adjacent tones. Off-beat tapping was less affected by accent location. Experiment 3 required participants to "hear" different T positions as metrically accented (i.e., to construe them as the downbeat) while carrying out the various tapping tasks. Most participants found it difficult to maintain a cognitive downbeat at fast tempi when it did not coincide with their taps. However, when such a downbeat could be maintained, it did not seem to increase the difficulty of tapping (with one exception). This suggests a unidirectional dependence of metrical structure on action. In Experiment 4, the same tasks were presented at more moderate tempi, and the dependent measure was the variability of asynchronies. Metrical downbeat location still did not have any significant effect. Thus, synchronization difficulty seems to be affected only by a rhythm's physical structure, not by the cognitive interpretation that is given to that structure.

Received April 5, 2004, accepted August 24, 2004

The present study addresses two main questions: First, what is the fastest tempo at which musically trained individuals can coordinate periodic movements (finger taps) with a rhythm in either on-beat or off-beat mode? Second, to what extent do different kinds of accent (grouping, dynamic, metrical) facilitate or impede synchronization? The first question has to do with rate limits of sensorimotor and cognitive processes that may impose constraints on coordination in musical ensemble performance. The second question concerns aspects of musical structure (rhythm) and of cognitive interpretation (meter) that may have a bearing on the ease of rhythmic coordination.

Rate Limits of Synchronization

When playing in an ensemble, musicians must coordinate their actions with great temporal precision. This can present challenges in several ways. One form of challenge arises when the music is rhythmically complex, for obvious reasons. A second kind of challenge occurs when the tempo is very slow, because long interval durations are associated with large timing variability in both perception (Friberg & Sundberg, 1995) and action (Peters, 1989), and because events separated by intervals longer than about 2 seconds are difficult to anticipate (Mates, Radil, Müller, & Pöppel, 1994). A third type of challenge is posed by very fast tempi, at which even simple rhythms can become difficult to follow. The present study is concerned with this third situation, which may shed light on some temporal processing limits of the brain that are relevant to music perception and performance.

Musical ensemble playing is a form of sensorimotor synchronization in which both auditory and visual information are important. Here, however, only the auditory component is being investigated. In the laboratory, sensorimotor synchronization is commonly studied in the form of finger tapping with a computer-generated auditory sequence. Tapping in synchrony with isochronous sequences composed of identical clicks or tones (i.e., metronomes) has been studied extensively, and it is well known that on-beat (synchronized, in-phase) tapping, where the taps coincide

Music Perception volume 23, issue 2, pp. 165-187, issn 0730-7829, electronic issn 1533-8312

approximately with tone onsets, is easier than off-beat (syncopated, anti-phase) tapping, where the taps fall near the midpoints of the tone inter-onset intervals (IOIs). This difference is most evident when the sequence tempo is fast. People (at least those with music training) generally have no difficulty with on-beat synchronization even at very fast tempi, as long as they can move their finger or wrist quickly enough. That is, the rate limit of 1:1 on-beat tapping (i.e., one tap per tone) is close to the biomechanical limit and occurs typically at IOIs of 150-200 ms (Peters, 1989; Repp, 2005). This limit is in agreement with studies that have determined how fast people can tap without a pacing sequence (Keele & Hawkins, 1982; Keele, Pokorny, Corcos, & Ivry, 1985; Peters, 1980, 1985; Todor & Kyprie, 1980; Truman & Hammond, 1990).

However, when participants are asked to synchronize taps with selected tones in an isochronous sequence (e.g., 1:4 on-beat tapping: Repp, 2003) or to make just a single on-beat tap (Bartlett & Bartlett, 1959), thereby avoiding the biomechanical rate limit, an even faster rate limit of a perceptual or sensorimotor nature is revealed: Synchronization is difficult to maintain when the IOIs between sequence tones are shorter than about 120 ms. Typically, phase drift is observed (i.e., participants no longer tap at the correct tempo), and it also becomes difficult to perceive whether or not one's taps are in synchrony. It is currently unclear whether this rate limit reflects a purely perceptual limit (i.e., a difficulty of perceiving individual tones as distinct events) or a sensorimotor limit (i.e., a difficulty of processing sensory feedback about phase errors or period mismatches).

Both the biomechanical and the perceptual/sensorimotor rate limits for on-beat tapping are far below the IOI durations at which off-beat tapping either breaks down or switches involuntarily to on-beat tapping. Off-beat tapping thus seems to be subject to a more severe limit of a more cognitive nature. Specifically, the rate limit for off-beat tapping may be set by the ability to mentally subdivide temporal intervals, because the taps must be coordinated with the estimated midpoints of the IOIs between tones. This limit may be related to the fastest possible rate of the main beat (tactus) in music, whereas the on-beat tapping limit may reflect the fastest possible pulse rate (i.e., metrical subdivision; see London, 2002, 2004).

Information on the rate limit for off-beat tapping is scarce in the literature. Various finger tapping studies have included off-beat tapping tasks but have not used challenging sequence tempi (e.g., Chen, Ding, & Kelso, 2001; Repp, 2001, 2002; Semjen, Schulze, & Vorberg, 1992; Vos & Helsper, 1992). Researchers who take a dynamic systems approach to motor control (e.g., Engström, Kelso, & Holroyd, 1996; Kelso, DelColle, & Schöner, 1990) are typically interested in the transition from antiphase to in-phase coordination but not in the rate limit as such. They usually instruct the participants not to resist the transition. Also, the tasks usually require free flexions and extensions of limbs rather than taps on a surface, so that the results are difficult to compare with those of tapping studies. Moreover, the participants often have little music training. The tapping task, which often provides auditory feedback, is clearly more relevant to music performance than is free limb movement, which is closer to dance.

Fraisse and Ehrlich (1955), in what seems to be the earliest study of off-beat tapping, tested an unselected group of students and found that many of them already had difficulties when the IOIs became shorter than 1 s. Some participants, however, were still able to tap in antiphase when the IOIs were as short as 275 ms. Success at this very fast rate is surprising in view of the fact that the late Jeff Pressing, an experienced professional pianist and percussionist, considered 375 ms close to his own limit for off-beat tapping (Pressing, 1998). One likely reason for this discrepancy is that Pressing used sequences that were more than 10 times as long as those of Fraisse and Ehrlich. Thus, his personal limit applies to stable prolonged off-beat tapping, whereas some of the participants in the French study may have just been lucky to get through a short series of 25 off-beat taps. Moreover, Fraisse and Ehrlich considered off-beat tapping successful as long as taps and tones alternated, which is a rather lenient criterion.

Volman and Geuze (2000) tested three age groups of children (7, 9, and 11 years) with sequences whose frequency increased continuously from 1 to 3 Hz, in order to determine the critical frequency at which off-beat tapping turned into on-beat tapping or phase drift. The average critical frequency increased with age from 1.17 to 1.61 Hz, and the most successful participants in each age group were able to maintain off-beat tapping up to 1.61, 1.94, and 2.34 Hz, respectively. This corresponds to IOIs of 621, 515, and 427 ms, respectively. Thus, performance improved with age and probably had not yet reached adult levels.

In a recent experiment that led to the present study, Repp (2005) tested a small group of musically trained individuals with sequences of increasing tempo and observed that off-beat tapping became difficult for most participants around IOIs of 350 ms. Experiment 1 of the present study used an adaptive staircase procedure to determine the rate limit of off-beat tapping more

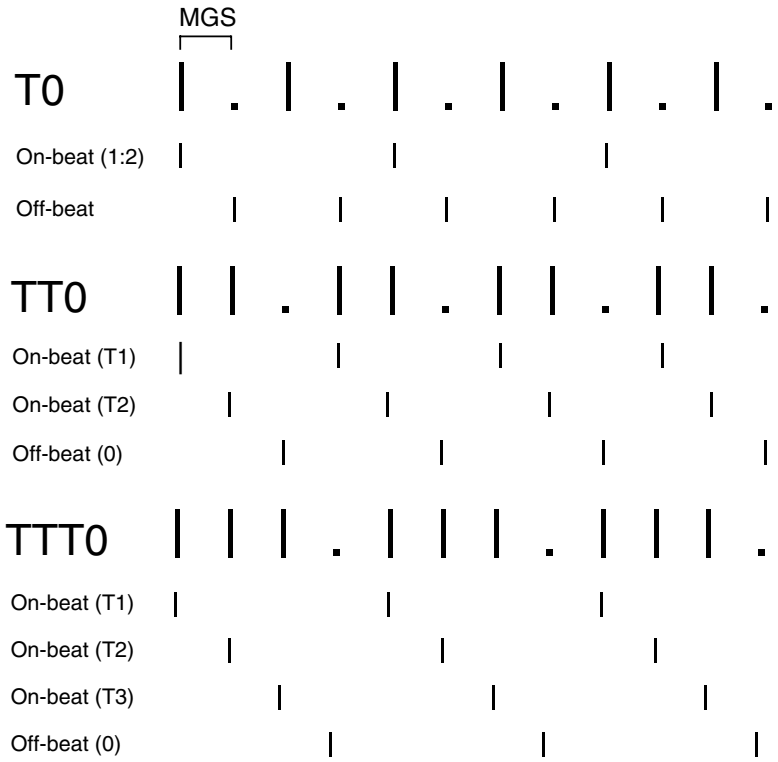

FIG 1. Schematic diagram of the three sequences, showing a few cycles of each, and of the nine tapping tasks. A large vertical bar stands for a tone onset (T), a dot for a silent beat (0), and a small vertical bar for a tap. MGS = metrical grid spacing.

precisely for a comparable group of participants. That limit presumably is relevant to musicians' ability to play extended syncopated passages, although in realistic musical contexts there are often additional auditory cues, such as dynamic accents or slower periodicities created by melodic pitch patterns, that may facilitate such a task (Keller & Repp, 2004).

The main purpose of Experiment 1, however, was to investigate whether the rate limits revealed in tapping with isochronous sequences also apply to coordination with somewhat more complex rhythms. Thus, as in Repp (2005), the standard on-beat and off-beat tapping tasks were extended to rhythmic patterns consisting of groups of two or three tones separated by a longer interval, but—going beyond the previous study—an adaptive algorithm was used to estimate rate limits.

The three sequence types used in Experiment 1 are illustrated schematically in Figure 1. Each consisted of repeated cycles containing tones present (T) and absent (0), with the timing of these events and nonevents being governed by an underlying isochronous metrical grid. The three sequence types were T0, TT0, and TTT0. Thus, in this context an isochronous sequence (T0) is considered as an alternation of tones and silences, which makes sense when the task is off-beat tapping, because the taps must coincide with the silent IOI midpoints. (For the sake of consistency, an isochronous sequence was still regarded as being composed of T0 cycles when the task was on-beat tapping, although the IOI midpoint played no role in that case.) Generalizing from T0 to TT0 and TTT0 sequences, on-beat and off-beat tapping were defined as follows: *On-beat tapping* requires tapping with one of the tones (T target, one tap per cycle), whereas *off-beat tapping* requires tapping at the silent midpoint of the long between-group IOI (0 target).[1] These tasks are also illustrated in Figure 1.

[1]The terms seem appropriate, given that the sequences were presented at rather fast tempi. At such tempi, the main beat (tactus, downbeat) of the strongly metrical TT0 and TTT0 patterns is located at the level of the cycle (Parncutt, 1994). In tapping with a T target, the downbeat is likely to be construed so as to coincide with the target (hence on-beat tapping), whereas in tapping with the 0 target the downbeat is most likely construed to be on a tone, not at the IOI midpoint (hence off-beat tapping). It is understood that exceptions may occur. Metrical interpretation is discussed more thoroughly in connection with Experiments 3 and 4, and in the General Discussion.

The individual tones in each cycle of TT0 and TTT0 sequences are identified by subscripts. The within-group IOI is determined by the underlying metrical grid and therefore is referred to as *metrical grid spacing* (*MGS*), which serves as the measure of sequence tempo. The between-group IOI is twice as long. In a T0 sequence, the MGS is equal to IOI/2. The mean inter-tap intervals (ITIs) increase from T0 to TT0 to TTT0 sequences because of the increasingly longer cycle durations (2*MGS, 3*MGS, and 4*MGS, respectively).

Duration of the ITIs could be a factor that affects the relative difficulty of tapping with these sequence types, as it is well known that variability increases with ITI duration (e.g., Peters, 1989). However, this applies mainly to intervals that are not subdivided, and the occurrence of additional tones within the ITIs may well attenuate or eliminate this factor (cf. Repp, 2003). Two other factors were expected to be much more important than ITI duration in affecting the relative difficulty of the various tapping tasks. One, of course, was whether tapping occurs in synchrony with a tone (on-beat tapping) or with the 0 position (off-beat tapping): On-beat tapping was expected to be generally easier than off-beat tapping because tones that are in phase with the taps offer stronger physical support (i.e., more accurate perceptual information on which error correction can be based) than silence (in which case the tones surrounding the silent tapping target, which are out of phase with the taps, serve as temporal references). The second important factor expected to affect the relative difficulty of the different on-beat tapping tasks in TT0 and TTT0 sequences was rhythmic grouping accent. This leads us to the second theme of the present study, namely the role that different kinds of accent may play in synchronization with nonisochronous rhythms.

Three Kinds of Accent

An accent is whatever makes one tone more salient than another. Accents can arise from the physical structure of rhythms (Lerdahl & Jackendoff, 1983, called those "phenomenal" accents), or they can be imposed in a top-down fashion by a listener, in which case the accent is subjective and does not have a direct physical correlate. Two kinds of phenomenal accent were considered in this study: grouping accent and dynamic accent. (A pitch accent accompanied the dynamic accent but was not investigated independently.) The third kind of accent studied was of the top-down kind: metrical accent.

It is known since the research of Povel and Okkerman (1981) that listeners perceive accents in temporal groups of identical tones. In groups of two tones (TT0), T_2 tends to be perceived as accented unless the between-group IOI is rather short (much less than twice the within-group IOI). In groups of three tones (TTT0), T_1 and T_3 are perceived as accented (Povel & Essens, 1985). The accented tones are perceptually more prominent and, if they are regularly spaced (as they were in the present sequences), induce the feeling of a metrical beat. This leads to the following predictions: On-beat tapping should be easier with T_2 than with T_1 in TT0 sequences, and it should be easier with T_1 and T_3 than with T_2 in TTT0 sequences. "Easier" means here a lower rate limit, that is, a shorter MGS duration at which on-beat tapping breaks down (according to some criterion). Preliminary support for the hypothesis concerning TTT0 sequences was obtained by Repp (2005), but his results concerning TT0 sequences were equivocal. The predictions were tested again in Experiment 1, using a different procedure.

Experiments 2-4 investigated the roles of two other kinds of accent, in addition to the grouping accents which are inherent in the temporal configuration of each rhythmic pattern. Experiments 2 and 4 explored the effect of a dynamic accent on one of the tones of TT0 and TTT0 patterns. (The accent was implemented as increased intensity and pitch in Experiment 2, but as an accompanying low tone in Experiment 4.) The predictions were straightforward: A dynamic accent should facilitate tapping with the accented tone and impede tapping with other tones. It was unclear whether it should have any effect on off-beat tapping.

Experiments 3 and 4 addressed the more complex issue of metrical accent whose discussion will be postponed until the introduction to those experiments. At this point, it will suffice to say that the subjective location of the downbeat was not controlled in Experiments 1 and 2, nor were participants questioned about it. They were free to consider any tone in a cycle (or, for that matter, the 0 position) as the downbeat.

Experiment 1

The purpose of Experiment 1 was to provide estimates, for each participant, of the fastest tempo (expressed as MGS duration) at which each tapping task could be executed. To that end, an adaptive staircase method was used that zeroed in on the rate limit of each participant in each task. Similar procedures have long been in use in psychophysics (see Leek, 2001), but they are novel

in the context of sensorimotor synchronization tasks.[2] The algorithm was simple, in order not to fatigue participants with overly long runs of trials, and two estimates were obtained for each task to get some indication of the reliability of the method and of any effects of practice.[3]

Method

PARTICIPANTS

Ten paid volunteers (7 women, 3 men), aged 18-30, participated in addition to the author who was 57 at the time. Two of the volunteers and the author had participated in the experiment described in Repp (2005) and were practiced in finger tapping tasks. The other participants were novices in tapping experiments but had substantial music training. The group included six professional-level classical musicians (one percussionist, two violinists, two violists, and a clarinetist, all graduate students at the Yale School of Music), four amateur pianists with 10 or more years of formal training, and one participant with only 4 years of basic piano instruction.[4] All participants were right-handed.

MATERIALS AND EQUIPMENT

The sequences consisted of high-pitched digital piano tones (E_7, MIDI pitch 100, 2637 Hz) that were produced on a Roland RD-250s digital piano via an Opcode Studio Plus Two musical instrument digital interface (MIDI) translator under control of a program written in MAX 3.0 which ran on a Macintosh Quadra 660AV computer.[5] The tones had sharp onsets and a nominal duration of 20 ms. (There was some residual decay and ringing following the nominal offset.) All tones were produced at the same nominal intensity (MIDI key velocity).

Participants sat in front of a computer monitor on which the current trial number was displayed, listened to the sequences over Sennheiser HD540 II earphones at a comfortable loudness level, and tapped on a Roland SPD-6 percussion pad which they held on their lap. They were allowed to use their preferred way of tapping. The majority rested the wrist and other fingers of the right hand on the surface of the pad and tapped by moving the index finger only; some, however, most notably the percussionist, tapped "from above" with the middle finger by moving the wrist and elbow of the unsupported arm. The impact of the finger on the rubber pad provided some direct auditory feedback (a thud), in proportion to the tapping force.

PROCEDURE

There were nine tasks, which were presented in a fixed order. Participants first tapped with every other tone in the T0 sequence (1:2 on-beat tapping). This task served both as warm-up and as a baseline for comparison with the next task, which was (1:1) off-beat tapping with the T0 sequence. Subsequently, participants tapped in turn with T_1, T_2, and 0 in TT0 sequences, and finally with T_1, T_2, T_3, and 0 in TTT0 sequences.

Each task comprised a variable number of trials called a run. A run typically lasted between 5 and 10 minutes. At the beginning of a run, the initial MGS duration was set. It was generally 200 ms for the on-beat tapping tasks, except for 1:2 tapping with T0 sequences, where the starting MGS was 150 ms. For off-beat tapping, the starting MGS was set to 200 ms but was changed to 250 ms if a participant had difficulty with 200 ms. (In general, whenever a participant had difficulties at the beginning of a run, the run was started over with the MGS increased by 50 ms.) Participants initiated a trial by pressing the space bar of the computer keyboard. After the sequence had started playing, they were free to start tapping in synchrony with the designated synchronization target whenever they felt ready.[6] The computer

[2]The method, with a different algorithm, was first used by Keller and Repp (2005).

[3]Although these estimates are more precise than the qualitative observations of Repp (2005), they are obviously not as accurate as those typically reported in psychophysical studies using adaptive threshold procedures, where trials typically are much shorter and less fatiguing than synchronization trials.

[4]The experiment served as a screening test for individuals who had volunteered to become regular participants in tapping experiments throughout the academic year. Several individuals, none of whom had much music training, performed so poorly already in the initial tasks with the T0 sequence that they were considered to have failed the test and were excused at that point.

[5]With this set-up, the tempo of the output was about 2.4% faster than specified in the MIDI instructions (as shown by earlier acoustic measurements), and participants' taps were registered at a correspondingly slower rate. All millisecond values are reported here as they appeared in the MAX environment.

[6]One factor that was not varied in the present study was the starting point of the sequences; they always started with T_1. It might be argued that this increased the relative salience of T_1 in TT0 and TTT0 sequences by biasing the metrical interpretation in favor of the downbeat falling on T_1. However, there is evidence from classic studies of temporal pattern perception that the starting point of a cyclically repeated auditory sequence has little effect on its perceptual organization (Royer & Garner, 1966), and in classical music it is quite common for a piece to start with one or two upbeats, perhaps more so than to start with a downbeat that is followed by one or two afterbeats. Therefore, the starting point of the sequences was not believed to be an important factor to control.

kept track of the number of taps made and of their asynchronies relative to the temporal target.

For a trial to be considered successful, each of 40 successive taps had to be within ±MGS/2 of the target. In other words, each tap had to be closer to its target (tone onset or IOI midpoint) than to the preceding or following tone onset or IOI midpoint. After 40 successful taps, the sequence stopped, a positive message appeared on the computer screen, and the MGS for the next trial was reduced by Δ ms, where Δ = 10 ms initially. After any unsuccessful tap, the sequence stopped immediately, a negative message appeared on the computer screen, the MGS for the next trial was increased by Δ ms, and Δ was reduced by 2 ms. After five unsuccessful trials, Δ reached zero, and the run was terminated with a chime. The MGS duration of the last successful trial was taken down as the threshold estimate. No other tapping data (asynchronies or ITIs) were recorded. (See Repp, 2005, for a more detailed analysis of synchronization behavior in the same tasks.)

Runs for all nine tasks were completed in a single session which lasted between 1 and 1.75 hours. An identical but typically shorter session took place about one week later. (Only the author ran himself on two successive days.) The initial MGS duration in each run was the same in Session 2 as in Session 1.

Results and Discussion

The results are shown as a box plot in Figure 2. Repeated-measures ANOVAs with the variables of task and session were conducted on the results for each sequence type. The Greenhouse-Geisser correction was applied where appropriate. (The ε values reported below pertain to that correction.) In none of these analyses was there a significant effect of session or a Task × Session interaction. Although some individual participants showed signs of improvement across sessions, overall there was no reliable evidence of a practice effect.

The results for the T0 sequence show that 1:2 on-beat tapping was much easier than off-beat tapping, $F(1,10) = 190$, $p < .0001$, with the mean threshold for on-beat tapping being at MGS = 91 ms (IOI = 182 ms) but that for off-beat tapping being at MGS = 175 ms (IOI = 350 ms). In other words, at the on-beat tapping threshold the sequence was about twice as fast as at the off-beat tapping threshold. However, because the functional MGS in on-beat tapping was really the IOI rather than IOI/2, the thresholds for on-beat and off-beat tapping were actually quite similar in terms of the functional MGS. Thus, the rate limit of mental subdivision in off-beat tapping was about the same as the limit of perceptual beat tracking in on-beat tapping.

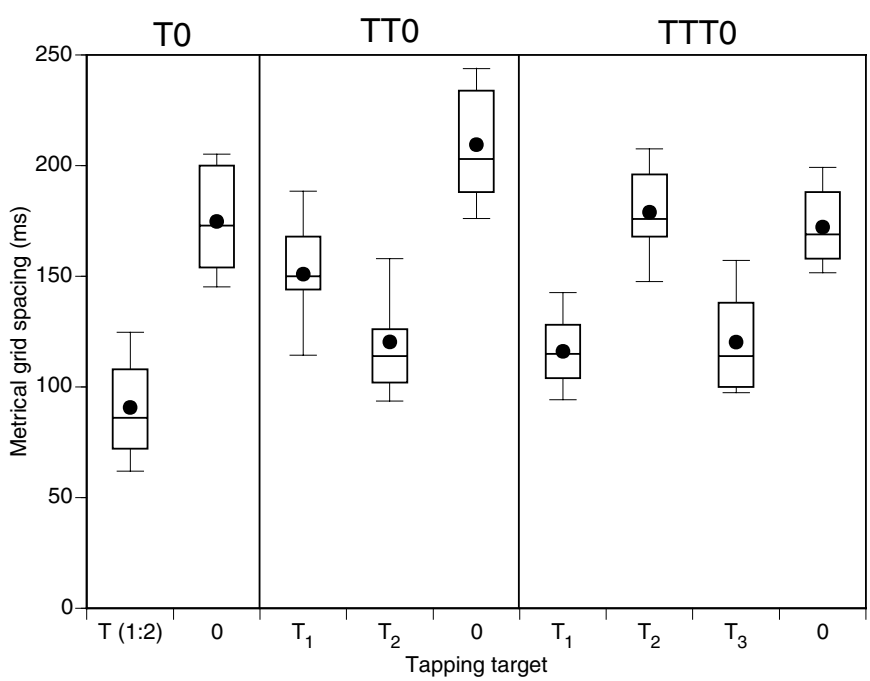

FIG 2. Results of Experiment 1. Each circle represents a mean threshold. The horizontal sides of a box indicate the quartiles of the distribution of 22 threshold estimates (11 participants × 2 sessions), and the whiskers span the 10th and 90th percentiles of the distribution.

The mean threshold value of IOI = 350 ms for off-beat tapping is in accord with the informal observations made in Repp (2005) and also with Pressing's (1998) stated personal limit of 375 ms for prolonged off-beat tapping. There were considerable individual differences, however, even within the present group of musically trained individuals, with individual mean thresholds ranging from IOI = 270 ms (the author) to 436 ms (a violinist).

The mean threshold for 1:2 on-beat tapping, IOI = 182 ms, is substantially higher than the previous estimate of IOI = 123 ms for 1:4 on-beat tapping (Repp, 2003). This difference is due to a stricter criterion for successful synchronization in the present experiment: Taps had to fall within ±IOI/4 of the target tones, whereas in the earlier study the criterion was ±IOI/2. Individual mean thresholds ranged from IOI = 122 ms (the percussionist) to 260 ms (the same violinist as above).[7]

The thresholds for the three tasks with TT0 sequences showed large differences, $F(2,20) = 72$, $p < .0001$, $\varepsilon = .99$. A separate comparison of the two on-beat tapping tasks showed tapping with T_1 to be more difficult than tapping with T_2, $F(1,10) = 17.7$, $p < .003$, as predicted on the basis of Povel and Okkerman's (1982) findings concerning perceived grouping accent. Off-beat tapping in turn was more difficult than tapping with T_2, $F(1,10) = 61$, $p < .0001$.

The thresholds for the four tasks with TTT0 sequences were also reliably different, $F(3,30) = 61$, $p < .0001$, $\varepsilon = .78$. As predicted from the observations of Povel and Essens (1985) concerning grouping accent in groups with more than two elements, tapping with T_1 or T_3 was easier than tapping with T_2 or in the 0 position. The two tasks within each of these pairs were about equally difficult.

It is noteworthy that, despite shorter ITI durations, tapping with T_2 in TT0 sequences was no easier than tapping with T_1 or T_3 in TTT0 sequences. Also, among the three off-beat tapping conditions, tapping in the 0 position of TT0 sequences was the most difficult, $F(2,20) = 17.3$, $p < .001$, $\varepsilon = 1$, even though the ITIs were shorter than those for off-beat tapping with TTT0 sequences. This difference was most likely due to the ternary subdivision of the beat of TT0 sequences. Off-beat tapping with T0 and TTT0 sequences, both of which have binary subdivisions, was equally difficult, even though the ITIs differed by a factor of 2. Therefore, ITI duration as such did not seem to play a role.

In summary, Experiment 1 confirmed the predictions regarding the influence of grouping accent on synchronization difficulty: In both TT0 and TTT0 sequences it was easier to tap with an accented than with an unaccented tone. The experiment also confirmed that off-beat tapping is more difficult than on-beat tapping, although tapping with T_2 in TTT0 sequences was no less difficult than tapping in the 0 position, almost as if T_2 were absent. The critical MGS duration for off-beat tapping in T0 and TTT0 sequences seemed to be about 170 ms (= 340 ms in terms of between-group IOI), but in TT0 sequences the mean off-beat synchronization threshold was above 200 ms.

Experiment 2

Experiment 2 employed only the TT0 and TTT0 sequences and introduced dynamic accents that either reinforced or competed with the rhythmic grouping accents. The dynamic accent consisted of a simultaneous increase in intensity and pitch of one of the tones in each cycle. There were two possible accent locations in TT0 sequences, and three in TTT0 sequences. For each of these accent locations, all on-beat and off-beat tapping tasks were carried out. This resulted in $2 \times 3 = 6$ tasks for TT0 sequences and $3 \times 4 = 12$ tasks for TTT0 sequences. Dynamic accents were expected to facilitate on-beat synchronization when the tapping target carried the accent, but to impede synchronization when another tone carried the accent. At the same time, the effects of grouping accent on on-beat tapping were expected to persist. As to off-beat tapping, it was not clear whether the location of the dynamic accent should be expected to make any difference.

Method

PARTICIPANTS

There were 8 participants (5 women and 3 men, including the author) in this experiment, all of whom had participated in Experiment 1. Between the two experiments, however, they had performed a number of other synchronization tasks. The previous participants who were no longer available included the percussionist, one violinist (the one who performed so poorly), and one violist.

MATERIALS AND EQUIPMENT

Only TT0 and TTT0 sequences were used. One tone was given a dynamic accent, implemented as a higher

[7]That violinist seemed to be poorly motivated and soon stopped being a participant in synchronization experiments. Interestingly, the percussionist excelled in on-beat tapping tasks but performed merely at an average level in off-beat tapping tasks. This supports the hypothesis that these tasks make different cognitive demands.

intensity and pitch relative to the other tone(s). The MIDI velocities of accented and unaccented tones were 60 and 30, respectively, a difference of about 10 dB. Their pitches were E_7 (MIDI pitch 100, 2637 Hz) and $C\#_7$ (MIDI pitch 97, 2217 Hz) respectively, a difference of 3 semitones. The nominal duration of each tone was 50 ms.

PROCEDURE

The experiment comprised four sessions, each lasting between 1 and 1.5 hours and typically spaced one week apart. In Session 1, participants first carried out the six tasks with TT0 sequences, which were blocked by dynamic accent location, and then the three off-beat tapping tasks with TTT0 sequences. In Session 2, the nine on-beat tapping tasks with TTT0 sequences were performed, also blocked by accent location. The order of blocks was approximately counterbalanced across participants, as was the order of tasks within blocks. Sessions 3 and 4 were replications of Sessions 1 and 2, respectively, but with the order of tasks reversed. Otherwise, the procedure was the same as in Experiment 1.

Results and Discussion

The results of Experiment 2 are shown in Figure 3. The mean results of Experiment 1 are also drawn in the figure as dotted horizontal lines, although comparisons must be made with caution because the participants were fewer in number and more practiced than in Experiment 1.

A 2 × 2 × 2 repeated-measures ANOVA on the TT0 on-beat tapping data with the variables of tapping target, (dynamic) accent location, and session revealed only one significant effect, the Tapping Target × Accent Location interaction, $F(1,7) = 29.6$, $p < .001$. Separate tests on the T_1 and T_2 target conditions showed a significant effect of accent location for each, $F(1,7) = 16.1$, $p < .005$, and 37.1, $p < .001$, respectively: Tapping with T_1 was easier when T_1 was accented, whereas tapping with T_2 was easier when T_2 was accented, as predicted. However, there was no longer a main effect of tapping target (grouping accent) in this experiment: Tapping with T_1 was about as easy as tapping with T_2.

A 2 × 2 ANOVA on the TT0 off-beat tapping data revealed a significant improvement across sessions,

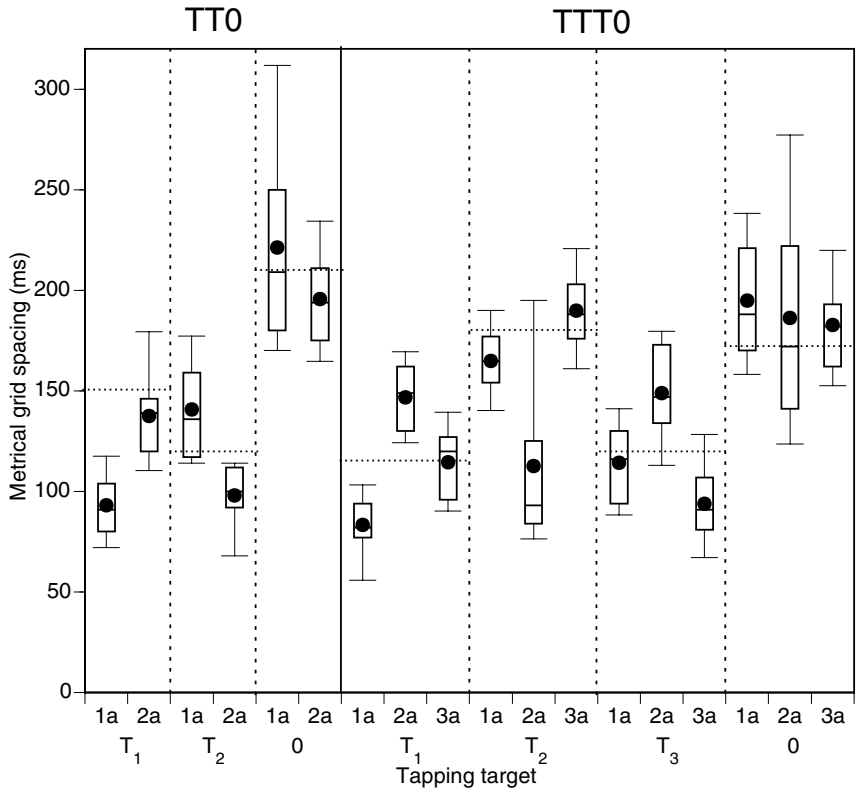

FIG 3. Results of Experiment 2, showing the distributions of 16 threshold estimates (8 participants × 2 sessions). 1a, 2a, 3a indicate the location of the dynamic accent. Dotted horizontal lines are the mean results of Experiment 1.

$F(1,7) = 19.6$, $p < .003$, as well as a small effect of accent location, $F(1,7) = 5.9$, $p < .05$: It was somewhat easier to tap in the 0 position when the dynamic accent was on T_2 than when it was on T_1.

A $3 \times 2 \times 2$ ANOVA to the TTT0 on-beat tapping data again showed the interaction between tapping target and accent location to be significant, $F(2,14) = 39.0$, $p < .001$, $\varepsilon = .52$. In addition, there was a significant main effect of tapping target, $F(2,14) = 37.7$, $p < .001$, $\varepsilon = .61$: Overall, tapping with T_2 was still more difficult than with either T_1 or T_3, which is the effect of grouping accent. Separate tests on each T target condition revealed a significant effect of accent location for each, $F(2,14) = 39.9$, $p < .001$, $\varepsilon = .85$, 23.9, $p < .001$, $\varepsilon = .71$, and 18.3, $p < .001$, $\varepsilon = .88$, respectively. For each target, tapping was easiest when the target tone was accented. In fact, when T_2 was accented, it was easier to tap with T_2 than with either T_1 or T_3, which suggests that the dynamic accent outweighed the grouping accent. However, there was large variability for tapping with an accented T_2, which indicates that for some participants the lack of a grouping accent still caused difficulties. It may also be noted that tapping on either T_1 or T_3 was less impeded when the dynamic accent was located on T_3 or T_1, respectively, than when it was located on the adjacent tone, T_2. In the former tasks, participants tapped in anti-phase with the dynamic accent, whereas in the latter tasks they tapped in $\pm$90-degree phase with it. Tapping with T_2 was somewhat easier when the accent was on T_3 than when it was on T_1.

A 3×2 ANOVA on the TTT0 off-beat tapping data revealed some improvement across sessions, $F(1,7) = 8.6$, $p < .03$, but no effect of accent location. However, there were large individual differences when T_2 was accented. One factor that may have contributed to this variability (as well as that observed when T_2 itself was the target) is that auditory streaming tended to occur at fast rates when T_2 was accented, and different participants may have been differentially susceptible to streaming. These sequences were not unlike the ones used by Van Noorden (1975) in his classic studies of stream segregation, considering that T_2 was not only louder but also higher in pitch than the other tones. Informal observations also suggested that some participants (such as the author) found it easier to tap in that condition because the off-beat taps were in anti-phase with the accent, whereas others (perhaps the ones who were more resistant to auditory streaming) did not seem to benefit from this relationship.

In summary, the results fully confirmed the predictions that on-beat tapping would be facilitated by dynamic accents when they coincide with the tapping target, but would be impeded when the dynamic accent is located on another tone, especially an adjacent one. Informal comparisons with the results of Experiment 1 (see Figure 3) suggest that both facilitation and inhibition occurred. The overall effect of grouping accent was still present in TTT0 sequences, but it disappeared in TT0 sequences, for unknown reasons. (Note that Repp, 2005, also did not find a consistent effect of grouping accent in TT0 sequences.) Dynamic accents did not have any strong effects on off-beat tapping. As in Experiment 1, there was no evidence for improvement with practice in the on-beat tapping tasks, but off-beat tapping did improve across sessions in Experiment 2.

Experiment 3

Experiment 3 investigated the possible effect of metrical interpretation on synchronization difficulty. Metrical accent (i.e., the feeling of a beat) is a purely subjective phenomenon that, however, depends strongly on rhythmic grouping and dynamic (as well as other phenomenal and structural) accents (Lerdahl & Jackendoff, 1983). Povel and Essens (1985) have shown that regularly spaced grouping accents induce perception of a beat that coincides with these grouping accents. Because of cyclic repetition, such regular spacing was present in the TT0 and TTT0 sequences of Experiments 1 and 2, which thus certainly were perceived as metrical structures. Given the range of tempi at which the sequences were presented, the main beat (tactus, downbeat) was almost certainly located at the level of the cycle (cf. Parncutt, 1994; Van Noorden & Moelants, 1999). As far as pure listening is concerned, TT0 sequences thus were most likely felt as having the downbeat on T_2, whereas TTT0 sequences were felt as having the downbeat on either T_1 or T_3, or possibly on both when the tempo was relatively slow.

In Experiment 2, the dynamic accent reinforced the preferred downbeat location when it coincided with a grouping accent, but when it fell elsewhere, it competed with the grouping accent, offering the possibility of hearing the downbeat on the tone carrying the dynamic accent. This may have resulted in greater variability of felt downbeat location from task to task and from participant to participant.

In both experiments, however, there was a third factor that may have been an important determinant of downbeat location, namely the tapping target. The author's experience as a participant was that, in order to carry out a given on-beat tapping task effectively, he needed to feel the downbeat on the tone that constituted the tapping target. When this was easy, the tapping task

was easy; when it was difficult, the tapping task was more difficult. Thus, for one participant at least, it seemed that the difficulty of an on-beat tapping task was closely related to the difficulty of intentionally assigning the downbeat to the target tone and keeping it there. In off-beat tapping, however, the author never felt compelled to think of his taps as the downbeats. In Experiment 1, he always considered the tone preceding the 0 position as the downbeat, and in Experiment 2, the dynamically accented tone.

Other participants may of course have used different strategies, but the author's self-observations raise an important question: Were the different levels of difficulty of the tasks in Experiments 1 and 2 a function of grouping accents and dynamic accents as such, as has been supposed so far, or were they rather a function of the relative difficulty of hearing the downbeat on the tapping target? These two possible determinants of task difficulty were confounded in Experiments 1 and 2 because participants were free to locate the downbeat wherever they liked. Experiment 3 attempted to dissociate metrical accent from other accents and from the tapping target by instructing participants to "hear" the downbeat in different locations, that is, to impose it intentionally. The author verified on himself that this can be done in principle, although it is not always easy. This divided on-beat tapping tasks into genuine on-beat (i.e., downbeat) and quasi-off-beat (i.e., upbeat or afterbeat) tapping tasks, but both kinds of task will still be referred to as on-beat (i.e., T-target) tapping, in accord with the definition given earlier. Participants were not asked to hear the downbeat in the 0 position during off-beat tapping, so off-beat tapping remained genuinely off-beat.

Experiment 3 thus investigated whether subjective metrical accents have an independent effect on synchronization difficulty when they are intentionally dissociated from grouping accents. It would have been convenient to investigate at the same time whether the effect of grouping accent persists when metrical accent is dissociated from it. However, because it is quite difficult to intentionally place and maintain a metrical downbeat on a tone that does not carry a grouping accent, it was decided to facilitate the task of downbeat placement by boosting the relative intensity of these tones somewhat. Thereby, the effect of grouping accent was diluted and could be assessed only to the extent that it was strong enough to overcome the intensity manipulation. If metrical accent were found to have an independent effect on the synchronization threshold, it might remain unclear whether there is an independent effect of grouping accent. If metrical accent were found to have no effect, however, then it could be inferred that the effects of grouping accent in Experiments 1 and 2 were indeed caused by grouping accents and not by coincident metrical accents.

Method

PARTICIPANTS

There were 9 participants, only two of whom (a clarinetist and the author) had participated in Experiments 1 and 2. The 7 newcomers were relative novices in tapping tasks.[8] They included three professional-level musicians (two violists and one violinist), one advanced amateur pianist, two amateur drummers, and one undergraduate who had had 5 years of piano instruction. In all, there were 5 women and 4 men, and ages ranged from 18 to 32 (except for the author who was 58 at the time). Two participants were left-handed, but only one of them tapped with the left hand.

MATERIALS AND EQUIPMENT

The TT0 and TTT0 sequences were used again, but an adjustment was made to the relative intensity of the tones that did not carry a grouping accent (T_1 in TT0, T_2 in TTT0), in order to facilitate hearing the downbeat on these tones. Their MIDI velocity was raised to 40, whereas that of the other tones was 30, a difference of about 3 dB. This difference was not heard as a distinctive accent and is not considered a dynamic accent in the present context. The pitch of all tones was $C\#_7$ (MIDI pitch 97, 2217 Hz), and their nominal duration was 50 ms. A real dynamic accent was present initially on one of the tones (see Procedure), implemented as a pitch of E_7 and a MIDI velocity of 60, as in Experiment 2. Except for a few initial sessions in which the old computer set-up was still used, this experiment was conducted using a new iMac G4 computer on which MAX 4.0.9 was installed. This eliminated the minor timing problem mentioned in Footnote 5.

PROCEDURE

The experiment comprised four sessions with the same design and procedure as Experiment 2. Each trial started with the sequence containing a dynamic accent (as in Experiment 2). However, as soon as the participant started tapping, the accent disappeared. Participants were instructed to consider the accented tone as the

[8]They were a new group of volunteers who had passed a screening test consisting of the tasks of Experiment 1, with the results of Experiment 1 (Figure 2) serving as the standard that had to be met. A number of additional volunteers failed this test, including some who had considerable music training. Three additional participants passed the screening test and started Experiment 3 but decided not to continue; their data were not included.

metrical downbeat and to keep "hearing" the downbeat on that tone throughout each trial. The dynamic accent in the initial cycle(s) was intended to help induce the downbeat as well as to remind participants of its required location, which was also conveyed by instructions prior to each run of trials. After each run, participants reported how easy or difficult it was to maintain the downbeat in the required location throughout the run, using a 5-point rating scale on which "1" stood for "easy" and "5" for "impossible." Because difficulties were most apparent at the fastest tempi at the end of a run, the higher ratings pertain mainly to these fast tempi. Participants were not allowed to tap their other hand or their foot on the intended downbeat; if such movements were observed, they were strongly discouraged.

Results and Discussion

The results are shown in Figure 4A, with the results of Experiment 1 again drawn in as horizontal lines for

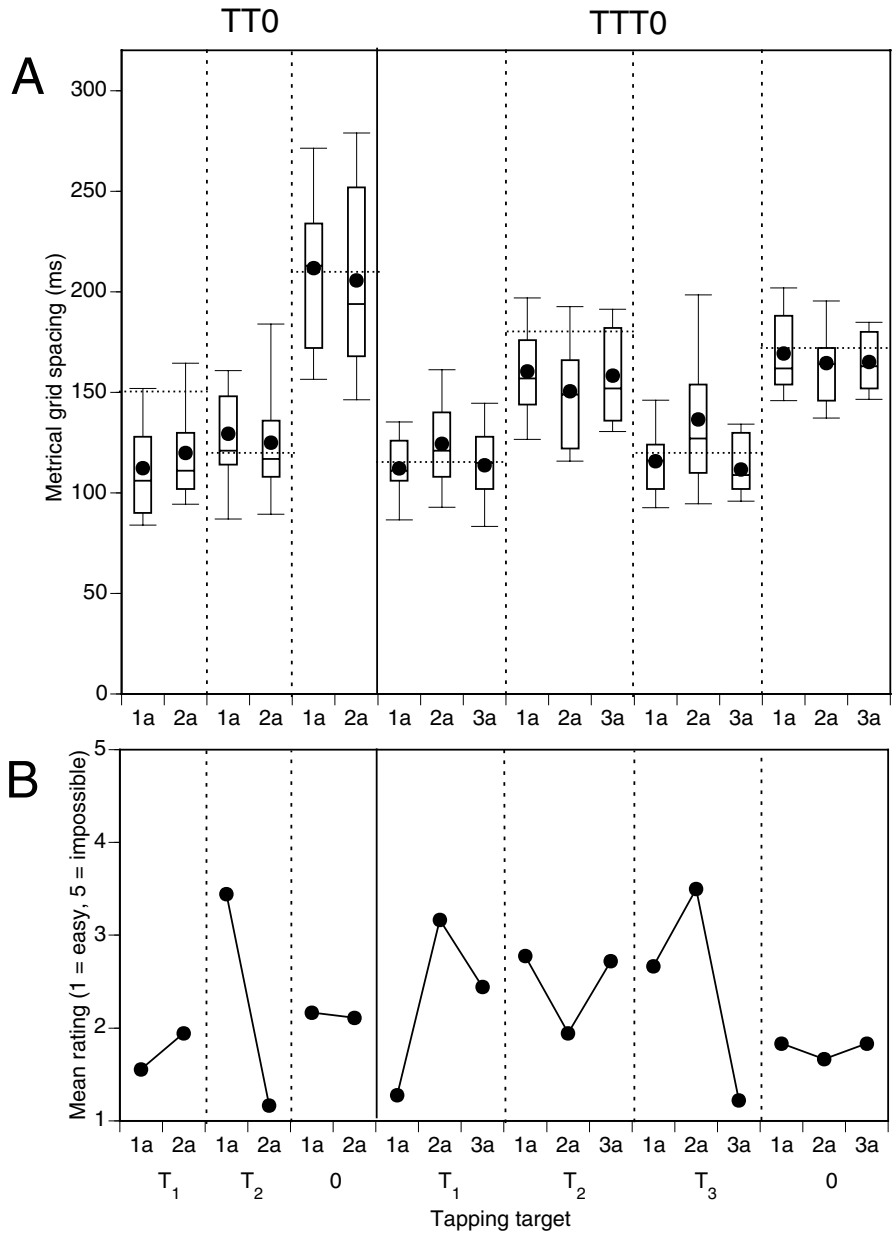

FIG 4. (A) Results of Experiment 3, showing the distributions of 18 threshold estimates (9 participants × 2 sessions). 1a, 2a, 3a indicate the location on which the metrical accent was supposed to be heard, according to instructions. Dotted horizontal lines are the mean results of Experiment 1. (B) Mean ratings of the difficulty of maintaining the downbeat in the required location.

informal comparison. On the abscissa, metrical accent location is indicated in the same way as dynamic accent location was indicated in Figure 3. If metrical accent had an independent effect on the synchronization threshold, the data should look similar to those in Figure 3. However, it is evident that metrical downbeat location had hardly any effect on tapping difficulty.

Separate ANOVAs on the TT0 on-beat and off-beat tapping data did not reveal any significant effects. In particular, the interaction between tapping target and downbeat location was far from significance. There was also no main effect of tapping target: Tapping with T_1 was no more difficult than tapping with T_2 and seemed facilitated relative to Experiment 1; this may be attributed to the increased relative intensity of T_1 in the current experiment. (However, it will be recalled that there was no effect of grouping accent in the TT0 sequences of Experiment 2 and in the experiment of Repp, 2005.)

The ANOVA on the TTT0 on-beat tapping data did reveal two significant effects. One was the main effect of tapping target, $F(2,16) = 13.1$, $p < .001$, $\varepsilon = .81$, because tapping with T_2 was still more difficult than tapping with either T_1 or T_3. The increased relative intensity of T_2 did seem to result in some facilitation relative to Experiment 1, but it was not enough to overcome the disadvantage of the target being in group-medial position. Thus, grouping accent still had some effect in these sequences. The other significant effect was the main effect of downbeat location, $F(2,16) = 4.8$, $p < .03$, $\varepsilon = .90$. It seems that, overall, tapping was somewhat more difficult when the downbeat was on T_2, perhaps because of the cognitive effort required to maintain it in that location. The crucial interaction between tapping target and downbeat location, however, did not reach significance, $F(4,32) = 2.2$, $p < .16$, $\varepsilon = .40$. The ANOVA on the TTT0 off-beat tapping data did not show any effects to be significant.

Although these results do not reveal any reliable effects of downbeat location on on-beat tapping, some small trends in the predicted direction can be seen in the means, suggesting that perhaps some individual participant(s) did show an effect. Before looking at some individual data, however, it may be asked whether the participants were able to hold on to the downbeat in the required locations. A failure to do so might have been responsible for the absence of any effect of downbeat location on tapping difficulty.

The participants' mean ratings of the difficulty of holding on to the downbeat are shown in Figure 4B. Error bars are not shown because there were very large individual differences in most conditions. The mean ratings indicate that, *on average*, maintenance of the downbeat in different conditions ranged from easy ("1") to moderately difficult ("3"). A mean rating of 3, however, reflects ratings ranging from 1 to 5. Interestingly, the mean ratings show almost exactly the pattern that had been predicted for the synchronization thresholds.

A 2 × 2 ANOVA on the ratings for TT0 on-beat tapping showed the interaction between tapping target and downbeat location to be significant, $F(1,8) = 25.6$, $p < .001$, whereas the two main effects fell just short of significance. Clearly, it was easy to hold on to the downbeat when it coincided with the tapping target, and more difficult when it did not. This was particularly true when the target was T_2, but much less so when the target was T_1. By contrast, downbeat location had no effect on the ratings for TT0 off-beat tapping, where it was equally easy to maintain the downbeat on either T_1 or T_2. This demonstrates that the participants were able to vary downbeat location as long as they were not tapping with a T target.

The ANOVA on the ratings for the TTT0 on-beat tapping conditions also showed the interaction between tapping target and downbeat location to be significant, $F(4,32) = 6.8$, $p < .003$, $\varepsilon = .67$. Again, it was easier to maintain the downbeat when it coincided with the taps, although even in that case it was harder to maintain it on T_2 than on T_1 or T_3, despite the boosted relative intensity of T_2. Downbeat location had no effect on the ratings for TTT0 off-beat tapping, where it was equally easy to maintain the downbeat on any tone.

When participants gave a high rating, indicating that it was difficult or impossible to maintain the downbeat in the required location, they were asked where they had heard the downbeat instead. In most cases of on-beat tapping, they said they heard the downbeat on the tone they were tapping with. In some cases, they said they heard it on another tone that carried a grouping accent. In off-beat tapping, there were a few instances where the downbeat was actually heard as coinciding with the tap.

As already mentioned, individual differences were large. Several participants, including highly trained musicians, found it almost impossible to maintain a mental downbeat that did not coincide with their taps in on-beat tapping.[9] Others experienced various degrees of difficulty, depending on the condition. Two

[9]The most extreme case was presented by one of the participants who did not complete the experiment (see Footnote 8). Although she had had extensive training as a violinist, she invariably heard the downbeat as coinciding with her taps, even in off-beat tapping. The experiment was discontinued for that reason.

participants deserve special attention, however. One amateur drummer (D.S.), although he had had only two years of percussion training, not only outperformed all other participants (including those in Experiments 1 and 2) in terms of synchronization thresholds, but also reported no difficulty in holding on to the downbeat in any condition. (Only three times he gave a rating of "2" because he felt some cognitive effort was needed to maintain the downbeat.) His results are shown in Figure 5A. They show generally high consistency across sessions and no effects of downbeat location, except when tapping with T_3 in the TTT0 pattern: There, the task was easiest for him when the downbeat coincided with the taps, and much more difficult when he thought of the downbeat as being on T_2. Otherwise, however, D.S. found it just as easy to tap on a tone when it constituted an upbeat or afterbeat than when it represented the downbeat.

The other participant of interest is the author (B.R.), because his results, shown in Figure 5B, were unique. They, too, show good consistency across sessions, but unlike D.S.'s data they reveal effects of downbeat

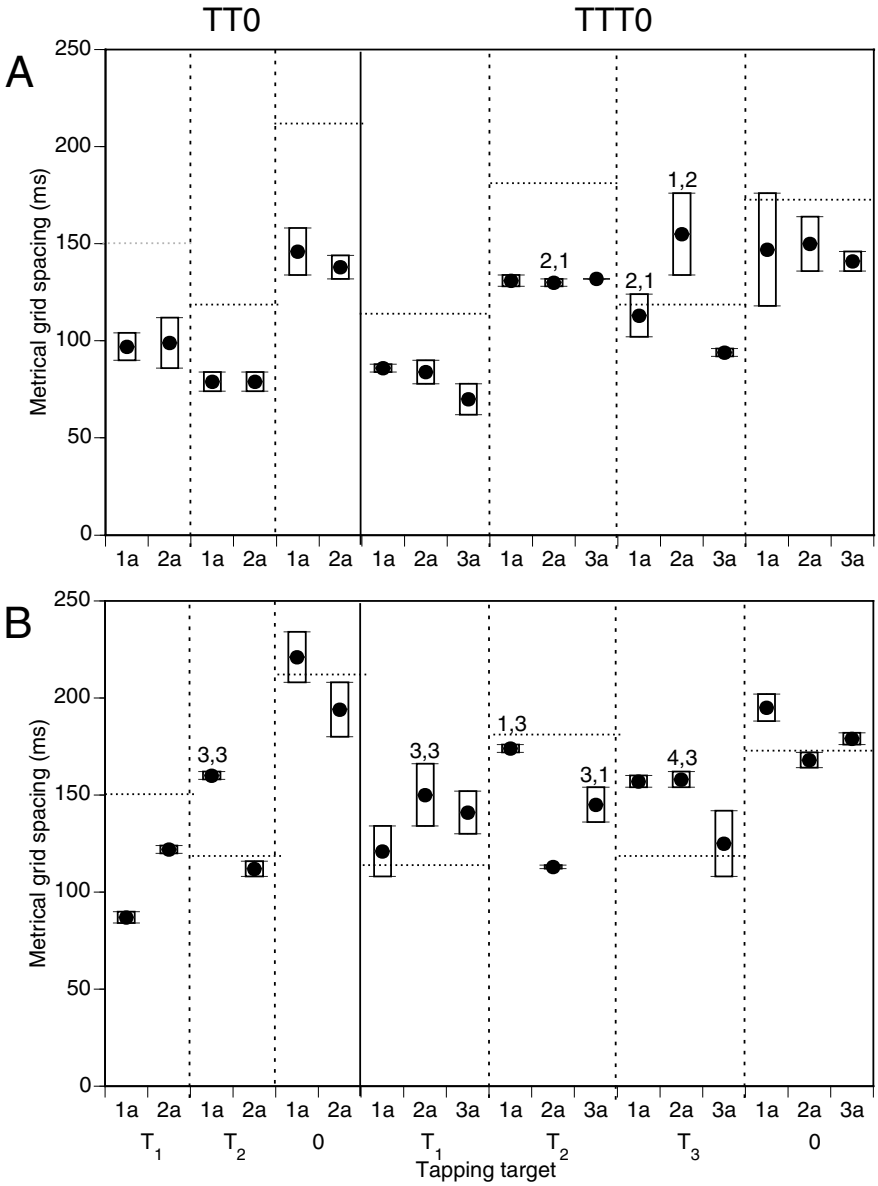

FIG 5. Results of two individual participants in Experiment 3: (A) D.S. (B) B.R. The horizontal sides of each box indicate the thresholds in the two sessions. Dotted horizontal lines represent the mean group results of Experiment 1. Digits indicate difficulty ratings in the two sessions for all conditions that did not receive ratings of "1" in both sessions.

location in all conditions. In both TT0 and TTT0 sequences, B.R. found on-beat synchronization substantially easier when the downbeat coincided with the taps than when it did not, which confirms his introspections during Experiments 1 and 2. He also found off-beat tapping with TT0 sequences easier with the downbeat on T_2 than on T_1, and off-beat tapping with TTT0 sequences easiest with the downbeat on T_2, because then the taps were in anti-phase with the downbeat. Clearly, the small trends in the expected direction in the mean data (Figure 4A) were due mainly to inclusion of B.R.'s results. As to the difficulty of maintaining the downbeat, B.R. experienced moderate difficulty (ratings of "3" and once of "4") in several conditions, but in most conditions he was able to maintain the downbeat with ease. The results of B.R. demonstrate that subjectively varied downbeat location *can* affect the synchronization threshold.

Difficulty in maintaining the downbeat can at best be a partial explanation for the overall negative findings, for some participants (like D.S.) were able to maintain the downbeat, yet they did not show any consistent effects. Also, B.R. still showed effects of downbeat location in those conditions where he had moderate difficulty in holding on to the downbeat, and the same might have been expected for others. One participant commented that, as a musician, he was motivated to do as well as possible in the tapping tasks, and so, in order to achieve a low synchronization threshold, he let the downbeat go where it wanted to go when the tempo got fast. Thus, some participants may not have exerted enough effort in holding on to the downbeat. However, this does not explain the absence of effects of downbeat location in cases such as D.S. where the downbeat seemed to be maintained successfully and without much effort.

There is also no reason to believe that participants had misunderstood the instructions, or that they did not tell the truth about their subjective metrical experience. Admittedly, it becomes increasingly difficult to tell where the downbeat is located as the tempo gets faster, and at the same time there is an increasing tendency of the downbeat to be attracted by the tapping target. It is possible that participants such as D.S. reported their good intentions or honest beliefs rather than their actual metrical interpretation at very fast tempi. However, if there had been a real effect of downbeat location, such very fast tempi should not have been reached when the downbeat did not coincide with the taps.

One participant, J.W., an amateur drummer playing in a rock band, pointed out that he did not necessarily think of the tone with the initial dynamic accent as the downbeat; rather he thought of it as a syncopation, especially because it also had a higher pitch initially (which reminded him of rim shots). In hindsight, it would perhaps have been better to mark the downbeat with a low pitch (as was done by Keller & Repp, 2005; also, Repp & Saltzman, 2001). However, J.W. had difficulty maintaining the downbeat, and his results resembled those of other, classically trained participants who had difficulty maintaining the downbeat in that none of them showed consistent effects of downbeat location.

It could be argued that only two participants, D.S. and B.R., met the requirements of this experiment in that they had sufficient control over their metrical interpretation to manipulate it independently of their action patterns and of the grouping accents inherent in the rhythms. This leaves open the possibility that one of them (D.S.) had such exceptional rhythmic skills that downbeat location truly did not make any difference for him, whereas the other one (B.R.) had lesser skills and therefore was susceptible to effects of downbeat location.

In summary, Experiment 3 suggests that metrical accent has no consistent effect on the synchronization threshold. This implies that the differences in synchronization thresholds observed in Experiments 1 and 2 were caused directly by grouping accents and dynamic accents, not by metrical accents that happened to coincide with these other accents. Indeed, an effect of grouping accent was still observed in the TTT0 sequences of Experiment 3, despite the increased relative intensity of T_2.

Experiment 4

One methodological problem of Experiment 3 was that it forced participants to perform at tempi at which it was difficult to maintain and distinguish between different metrical downbeat locations. It seemed prudent, therefore, to repeat the experiment using less extreme tempi, and taking the variability of asynchronies rather than the synchronization threshold as the dependent variable. This reinvestigation of the effect of metrical accent was the main purpose of Experiment 4. In addition, the experiment reinvestigated the effect of a dynamic accent on synchronization variability, albeit without dissociating dynamic accent from metrical accent. As in Experiment 3, a dynamic accent served to mark the tone that was to be considered as the metrical downbeat, but the accent now remained present during the whole first half of a trial. To prevent any metrical misconstrual of the dynamic accent (as by J.W. in Experiment 3), the accent was realized as a low-pitched tone accompanying the rhythmic pattern.

Method

PARTICIPANTS

The 8 participants were the same as in Experiment 3, except for the amateur pianist who was no longer available. About 5 months had elapsed, during which all participants had performed in various other synchronization experiments.

MATERIALS AND EQUIPMENT

These were the same as in Experiment 3, except for the following differences. Instead of the adaptive algorithm for estimation of the synchronization threshold, a fixed MGS duration was used in each of four successive blocks of trials: 240, 220, 200, and 180 ms. Thus, the tempo increased across the four blocks. Each block contained 6 conditions (2 metrical accent locations $\times$ 3 tapping targets) for TT0 sequences and 12 conditions (3 metrical accent locations $\times$ 4 tapping targets) for TTT0 sequences. The conditions occurred in different orders in each block, but conditions with the same accent location always followed each other. Each trial consisted of 80 cycles of the TT0 or TTT0 pattern. During the first 40 cycles, a dynamic accent was present on one of the tones, realized here not as an intensity and pitch increase of the tone itself, but as an accompanying low-pitched tone (E_3, MIDI pitch 52, 165 Hz, nominal duration MGS/10 ms, MIDI velocity 60). During the last 40 cycles, the low tone was absent.

PROCEDURE

Participants came for two 1-hour sessions, typically one week apart. In the first session, they did the four TT0 blocks and the first TTT0 block. In the second session, they did the remaining three TTT0 blocks. The computer monitor displayed the condition of the next trial (e.g., "tap1acc2," meaning "tap on the first tone while hearing the downbeat on the second tone"), and the participant started the trial by depressing the space bar of the computer keyboard. Participants were instructed to consider the low-pitched tone as marking the metrical downbeat and to maintain that metrical structure in their mind after the low-pitched tone disappeared. If the downbeat shifted to another location against their will, they were to stop tapping immediately and repeat the trial. (In view of these instructions, no difficulty ratings were collected.) Participants started tapping in the third cycle of each trial and thus made 78 or 79 taps per trial.

ANALYSIS

Only one participant, a violinist, experienced serious difficulties with holding on to the downbeat during the second halves of the trials, and her difficulty was restricted to hearing the downbeat on T_1 in TT0 sequences, even when T_1 was the tapping target. Because she felt completely unable to hear the downbeat on T_1 once the dynamic accent disappeared, she was encouraged to tap while hearing the downbeat on T_2; her data were not excluded. Three other participants had difficulty with the TT0 off-beat tapping task because some or all MGS durations fell below their synchronization threshold for that task. Their data from that condition were excluded because they showed phase drift and interruptions. Mean asynchronies and standard deviations were calculated separately for each half of each trial, including taps 3-38 (i.e., all but the first two) in the first half and taps 43-78 (i.e., all but the first four and any tap occurring after the end of the sequence) in the second half. Exceptionally large standard deviations and/or asynchronies (i.e., clear outliers) were excluded from calculation of means across participants and were replaced with values taken from adjacent cells in the design in the ANOVAs. The number of such substitutions was 1 in the TT0 on-beat tapping data (actually, a condition skipped by mistake), 7 in the TTT0 on-beat tapping data, 3 in the TTT0 off-beat tapping data, and 0 in the TT0 off-beat tapping data of the five participants who were able to manage off-beat tapping at all tempi.

Results and Discussion

STANDARD DEVIATIONS

The results for the standard deviations of the asynchronies are displayed in Figure 6 in a format similar to that of previous figures. Instead of a box plot, however, mean standard deviations and their standard errors are shown. Both statistics were first computed separately for each MGS value and then averaged across MGS values for graphic display. Separate repeated-measures ANOVAs were conducted on TT0 on-beat, TT0 off-beat, TTT0 on-beat, and TTT0 off-beat tapping data, with the variables of MGS duration (4 levels), accent condition (i.e., first vs. second halves of trials, 2 levels), accent position (2 or 3 levels), and tapping target (2 or 3 levels in on-beat tapping).

For TT0 on-beat tapping, the ANOVA revealed a significant decrease in variability as MGS duration decreased, $F(3,21) = 5.0$, $p < .03$, $\varepsilon = .57$. In addition, tapping on T_2 was significantly less variable than tapping on T_1, $F(1,7) = 19.4$, $p < .003$. Thus, the effect of grouping accent reappeared in TT0 sequences after having vanished in Experiments 2 and 3, even

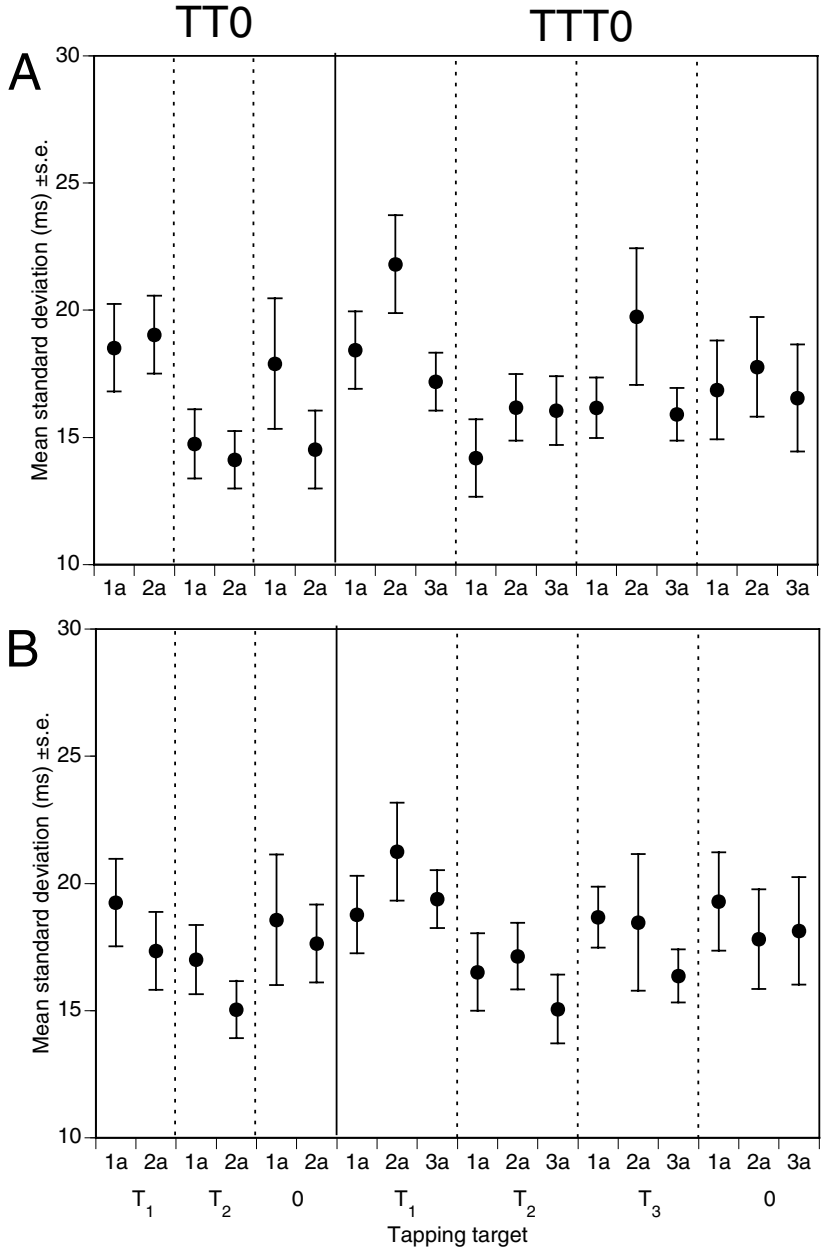

FIG 6. Mean standard deviations of asynchronies and their standard errors in the various conditions of Experiment 4, (A) when dynamic accents were present and (B) when only metrical downbeats were imagined in the same location. The data for TT0 off-beat tapping (0 target) are based on only 5 participants.

though T_1 was about 3 dB more intense than T_2, as in Experiment 3. However, the effect of greatest interest, the Accent Position × Tapping Target interaction, was nowhere near significance, $F(1,7) = 0.1$, and no other effects reached significance. Thus, accent position did not affect tapping variability, regardless of whether the metrical downbeats were reinforced by dynamic accents or whether they were merely present in participants' minds.

The analysis of the TT0 off-beat tapping data did not reveal any significant effects. Although comparisons with the on-beat tapping data must be made with caution because the data of three participants were omitted, the results indicate that off-beat tapping, when it could be managed, was no less variable than on-beat tapping. This contrasts with the higher synchronization thresholds for TT0 off-beat than for TT0 on-beat tapping in previous experiments.

The TTT0 on-beat tapping data showed two reliable main effects: of accent position, $F(2,14) = 14.7$, $p < .002$, $\varepsilon = .74$, and of tapping target, $F(2,14) = 8.9$, $p < .007$, $\varepsilon = .78$. Overall, tapping was more variable when the accent was on T_2 than when it was on T_1 or T_3. Surprisingly, however, tapping was most variable when the target was T_1 and least variable when the target was T_2. In other words, the effect of grouping accent observed in previous experiments, which always showed a disadvantage for T_2, was reversed. The crucial Accent Position × Tapping Target interaction was far from significance, $F(4,28) = 0.7$. Thus, there was again no evidence that either dynamic accentuation or metrical structure had any effect on synchronization variability.

Finally, the analysis of TTT0 off-beat tapping showed no significant effects. Figure 6 suggests that, as in TT0 sequences, off-beat tapping was about as variable as on-beat tapping. It should also be noted that the author, the only participant to show an effect of metrical accent location in Experiment 3, did not show any consistent effects of accent location in this experiment.

MEAN ASYNCHRONIES

The mean asynchronies are shown in Figure 7. They were of interest as a second variable that might be affected by metrical downbeat location, although they are not a measure of task difficulty. They should also not be interpreted as a measure of synchronization accuracy; rather, they probably represent the mean point of subjective synchrony in each task.

The analysis of TT0 on-beat tapping revealed four significant effects: of accent position, $F(1,7) = 9.0$, $p < .02$, of tapping target, $F(1,7) = 6.3$, $p < .05$, of Accent Position × Tapping Target, $F(1,7) = 8.1$, $p < .03$, and of MGS × Accent Position × Tapping Target, $F(3,21) = 4.4$, $p < .03$, $\varepsilon = .84$. It can be seen that participants tended to tap ahead of T_1 but right on T_2. Moreover, the anticipation tendency for T_1 was larger when the accent was on T_2 than when it was on T_1, whereas tapping on T_2 was not affected by accent position at all. This interaction was largest at the slowest tempo and diminished as the tempo increased. Note that the interaction did not depend on accent condition; if anything, it was slightly larger when metrical accents were maintained without dynamic accents.

The analysis of TT0 off-beat tapping did not reveal any significant effects, even though Figure 7 shows what seems like a huge effect of accent position. Embarrassingly, this difference was caused by a single participant (the author) who tapped far ahead of the 0 position (i.e., close to T_2) when the dynamic and/or metrical accent was on T_1.

The asynchronies of TTT0 on-beat tapping showed a main effect of MGS duration, $F(3,21) = 6.0$, $p < .03$, $\varepsilon = .45$, because asynchronies became smaller (less negative) as the tempo increased. The main effect of tapping target and the Accent Position × Tapping Target interaction both fell short of significance, but the triple interaction with accent condition was significant, $F(4,28) = 3.6$, $p < .04$, $\varepsilon = .66$. As can be seen in Figure 7, there was an anticipation tendency (negative mean asynchrony) when the tapping target was T_1 (as there was in TT0 on-beat tapping), whereas anticipation was negligible with T_2 and T_3 targets. Also with T_1, but not with T_2 and T_3 targets, accent position seemed to make a difference, such that the anticipation tendency was largest when the accent was on T_2. Moreover, these differences increased when the dynamic accent was absent.

The asynchronies of TTT0 off-beat tapping likewise became less negative as the tempo increased, $F(3,21) = 3.8$, $p < .05$, $\varepsilon = .70$. In addition, there were significant main effects of accent condition, $F(1,7) = 13.0$, $p < .01$, and of accent position, $F(2,14) = 5.0$, $p < .03$, $\varepsilon = .87$, as well as an interaction between these two variables, $F(2,14) = 6.8$, $p < .01$, $\varepsilon = .98$. Accent position had an effect only when the dynamic accent was present (Figure 7A), and the later the accent occurred in the rhythmic group, the later the taps occurred.

In summary, the results of this experiment confirm the negative results regarding metrical accent obtained in Experiment 3. Participants generally had no difficulty maintaining the downbeat, and downbeat location actually did have some effects on mean asynchronies, which is interesting and suggests that instructions were followed. Synchronization variability, however, seemed to be quite independent of metrical accent location. In other respects, however, the results of Experiment 4 contrast with those of previous experiments: Dynamic accent had no effect, in contrast to Experiment 2; the effect of grouping accent in TTT0 sequences was reversed, whereas the previously elusive effect of grouping accent in TT0 sequences resurfaced; and off-beat tapping was not much more variable than on-beat tapping. Synchronization performance at supra-threshold tempi thus seems to be subject to different constraints than performance near the synchronization threshold. In particular, it seems that tapping between two tones (i.e., with 0 in TT0 sequences, and with either T_2 or 0 in TTT0 sequences) results in equal or lower variability than tapping with a tone that precedes or follows a long between-group IOI, as long as the tempo is not too fast. This may be a kind of

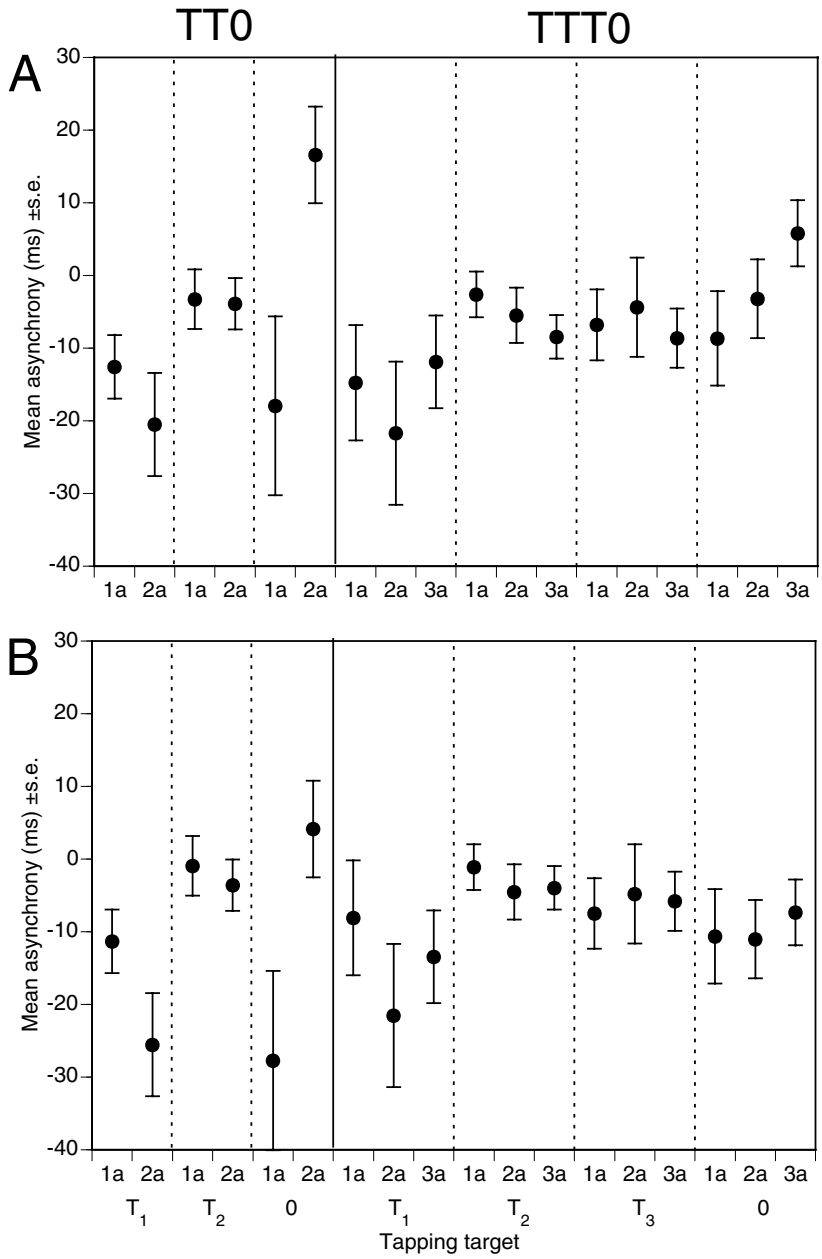

FIG 7. Mean asynchronies and their standard errors in the various conditions of Experiment 4, (A) when dynamic accents were present and (B) when only metrical downbeats were imagined in the same location. The data for TTO off-beat tapping (O target) are based on only 5 participants.

"subdivision benefit" obtainable when the MGS duration is longer than about 200 ms (Repp, 2003). MGS durations in Experiment 4 ranged from 180 to 240 ms and thus tended to be above this limit.

General Discussion

This study dealt with two main questions: (a) What is the fastest tempo at which various on-beat and off-beat synchronization tasks can be performed? (b) What effect do three types of accent have on synchronization difficulty?

The Synchronization Threshold

Experiment 1 used an adaptive procedure to estimate the synchronization threshold and started out by showing that the mean threshold for off-beat tapping with an

isochronous sequence was about twice that for 1:2 on-beat tapping. That is, the sequence had to be about twice as slow, but the rate of the alternating tones and taps was about the same as the rate of the tones in on-beat tapping. This suggests a connection between the two tasks, as hypothesized by London (2002, 2004). London presumed that a metrical beat must in principle allow for subdivision, so that the shortest possible beat duration is about twice the duration of the shortest possible subdivision duration. Off-beat tapping requires subdivision of beats, whereas 1:2 on-beat tapping involves superimposing beats (i.e., taps) on a rapid series of subdivisions. London suggested that the shortest possible durations might be 200 ms for beats and 100 ms for subdivisions, and Repp's (2003) results seemed to confirm this. The present results suggest a longer minimal duration of about 350 ms for beats that are not explicitly subdivided, which implies a limit of about 175 ms for mental subdivisions.

Following on the heels of an earlier qualitative study (Repp, 2005), Experiment 1 also extended the standard on-beat and off-beat tapping tasks to TT0 and TTT0 sequences. The synchronization threshold (within-group IOI) for on-beat tapping with tones in a rhythmic group that carried a grouping accent was about 120 ms on average, which does match Repp's (2003) findings for isochronous sequences and is close to London's (2002) hypothetical limit of metrical subdivision. However, the long between-group IOI in these rhythms may have facilitated synchronization, and it is not clear whether the results are directly comparable. The threshold for off-beat tapping with TTT0 sequences was similar to that for T0 sequences, but that for TT0 sequences was higher. This may be due to the ternary structure of these sequences. A disadvantage for ternary meter has been found occasionally in other studies. For example, Drake (1993) showed that rhythms with a ternary metrical structure are less accurately reproduced by children and adult nonmusicians than rhythms with a binary metrical structure. (Adult musicians showed a ceiling effect in that study.)

The synchronization threshold for on-beat tapping has been hypothesized to reflect an auditory temporal integration interval within which successive events form a group and can no longer be treated as individual events (Repp, 2003). Estimates of the duration of such an interval have come from various sources, for example from the electrophysiological studies of Yabe et al. (1997, 1998, 2001). In fact, the present findings for isochronous sequences are in better agreement with the findings of Yabe et al., which suggest an integration window of 150–200 ms duration, than are the results of Repp (2003). Other possible explanations of the synchronization threshold are that below the threshold there is not sufficient processing time for perceptual feedback about synchronization errors (phase correction), or that the inherent variability of finger taps begins to exceed the tolerance region for correctly placed taps, which shrinks proportionally with the IOI. Experiments need to be designed to distinguish between these different explanations.

Effects of Three Types of Accent on the Synchronization Threshold

GROUPING ACCENT

Experiment 1, and to a lesser degree Experiments 2 and 3, demonstrated that rhythmic grouping accent, as described by Povel and Okkerman (1981) and Povel and Essens (1985), has a strong effect on the synchronization threshold for on-beat tapping. This was true particularly in TTT0 sequences, where it was much easier to tap with the group-initial or group-final tone (T_1 or T_3) than with the group-medial tone (T_2). This difference held up even when the intensity of T_2 was about 3 dB higher than that of the other tones (Experiment 3). Surprisingly, however, Experiment 4 showed synchronization variability at moderate tempi to be lowest when T_2 was the target. Thus, the benefits of grouping accent emerged only at fast tempi, when synchronization performance became unstable. The likely reason for the T_2 advantage in Experiment 4 is that T_2 is flanked by other tones, which can serve as additional references for the temporal placement of taps. It is well known that variability is inversely related to interval duration, and only T_2 is both preceded and followed by a short IOI.

The effect of grouping accent on the synchronization threshold was less robust in TT0 than in TTT0 sequences. In Experiment 1, the predicted advantage for tapping with T_2 was found, but in Experiment 2 it was absent. Less surprisingly, it was also absent in Experiment 3, where T_1 was given an intensity boost. Repp (2005) also found no consistent advantage for T_2 over T_1. Thus, the difference in grouping accent between T_1 and T_2 in TT0 sequences seems to be less pronounced or consistent than that between T_2 and both T_1 and T_3 in TTT0 sequences. It may be more like the relationship between T_1 and T_3, because in each case the two tones delimit the group. In other words, T_1 in TT0 may actually carry a grouping accent by virtue of its group-initial position, but that accent may be somewhat weaker than the group-final accent on T_2. Unexpectedly, a reliable difference in favor of T_2

reemerged in Experiment 4 in terms of synchronization variability, even though T_2 was about 3 dB softer than T_1. The reason for this difference may lie in the relationship between IOI duration and variability: T_1 is preceded by an IOI that is twice as long as the one preceding T_2, and it is well known that variability increases with interval length. This variability is manifested mainly in the event that terminates an interval. The higher variability for T_1 than for T_2 and T_3 targets in TTT0 sequences can be explained in the same way.

Tapping with T_2 in TTT0 sequences was generally as difficult as off-beat tapping with either TTT0 or T0 sequences, which suggests that the physical presence of T_2 made little difference. Perhaps, if T_2 had been omitted, the same results would have been obtained. This would amount to 1:2 off-beat tapping with a T0 sequence, a task that was included together with 1:1 off-beat tapping in a recent study by Keller and Repp (2004). If anything, 1:2 off-beat tapping was easier (less variable) than 1:1 off-beat tapping in that study, which confirms indirectly that the physical presence of T_2 in TTT0 sequences did not facilitate tapping in that location.

DYNAMIC ACCENT

Experiment 2 investigated the impact of a dynamic accent on one of the tones on on-beat and off-beat synchronization thresholds. The dynamic accent was combined with a pitch accent (Drake & Palmer, 1993; Tekman, 1998; Thomassen, 1982) whose separate contribution was not examined, however. Although it is likely that pitch played a role, the intensity difference was probably the more important factor. The results of Experiment 2 showed quite clearly that a dynamic accent facilitates tapping on the accented tone but impedes tapping on other tones. Off-beat tapping was much less affected by the location of the dynamic accent. In Experiment 4, however, neither on-beat nor off-beat tapping was affected by dynamic accents (realized as an accompanying low tone). Like the effect of grouping accent, the effect of dynamic accentuation seems to emerge only at fast tempi that approach the synchronization threshold, where auditory streaming may also occur. Alternatively, it is possible (but unlikely) that accompanying low tones are ineffective as dynamic accents.

METRICAL ACCENT

Experiment 3 addressed the question of whether the effects of grouping accent and of dynamic accent were perhaps really due to metrical structure, because both types of accent are known to attract the metrica downbeat. The easiest tapping conditions in Experiments 1 and 2 were precisely those in which the perceived downbeat presumably coincided with the taps, or (if the downbeat always coincided with the taps) in which it was easiest to construe the downbeat as being in the same location as the taps. Experiment 3 attempted to dissociate metrical accent location from grouping accent, to see whether metrical accent has an independent effect on the synchronization threshold. The overall results suggest a negative answer, which means that the effects observed in Experiments 1 and 2 were indeed caused by grouping accents and dynamic accents, respectively, and not by correlated metrical accents. More generally, the results suggest that the relative difficulty of on-beat tapping near the synchronization threshold depends on physical signal properties (temporal structure and relative intensity, perhaps also pitch), not on the cognitive construction of a metrical framework.

However, the relative difficulty of maintaining the downbeat in a particular location depended not only on grouping accent (and presumably on dynamic accent as well, although this was not tested) but even more strongly on the motor action itself. Wherever the taps fell, according to instructions, there the downbeat was easiest to place and maintain, and there it tended to move even when the participant's intention was to place and maintain it in a different position. Some participants succumbed completely to that tendency; others had to exert considerable mental effort to counteract it. Thus there seems to be a unidirectional dependency of cognitive metrical organization on action: The location of the downbeat is strongly influenced by the taps, but the taps do not seem to be affected by the location of the downbeat (except in the case of one participant, the author).

These findings were confirmed in Experiment 4, where synchronization variability was found to be likewise insensitive to metrical accent location, at slower tempi that made few participants experience difficulties in placing metrical accents in their mind. Surprisingly, however, dynamic accents also had little effect on synchronization variability at these moderate tempi. Some effects of accents on mean asynchronies were noted, however.

The overall negative results concerning the effect of metrical accent on synchronization are in agreement with similarly weak or nonexistent effects of downbeat location in studies of rhythm production. In those studies, participants had to tap a rhythm, either in synchrony with a rhythm template or in free continuation. Repp and Saltzman (2001: Experiment 2) induced different downbeat locations in otherwise identical

rhythms by means of an accompanying low tone in the rhythm template with which participants synchronized their taps; during continuation tapping, the tone was absent and the metrical downbeat was maintained mentally. Repp, London, and Keller (2005) instructed participants to think of different downbeat locations in both synchronization and continuation tapping of uneven rhythms. One of the continuation conditions was paced by an isochronous sequence of downbeats, which should have helped maintain the downbeat in the intended location. Nevertheless, the timing of the taps was generally not affected by downbeat location in either study, although there were effects of metrical accents on tapping force and on asynchronies in synchronization.

How Is a Metrical Accent Generated?

This discussion concludes with some speculations on what internal generation of a metrical accent (downbeat) may involve. Beat induction through rhythmic sequences has been investigated in many studies (e.g., Drake, Penel, & Bigand, 2000; Povel & Essens, 1985; Snyder & Krumhansl, 2001; Toiviainen & Snyder, 2003), usually by asking participants to convey the perceived beat through overt movement such as finger tapping or hand clapping. In the present study, however, participants were asked to manipulate their metrical interpretation mentally, without conveying it through overt movement. The taps that were made sometimes coincided with the internal beat, and sometimes they did not. What went on in participants' heads?

One possibility is that intentional creation of a downbeat involves a periodically heightened focus of attention (cf. Large & Jones, 1999). This would be consistent with the subjective impression that a rhythm really *sounds* different when it is conceived in different metrical frameworks: The heightened attention may result in greater perceptual salience (metrical "accentedness") of the tones that are in focus.

A second possibility is that an auditory image of certain tones as being accented (i.e., being louder and/or having a different pitch) is created. This seems especially plausible in the present study because the downbeats were primed by dynamic (+ pitch) accents. Participants could have simply continued to imagine that the dynamic accent persisted after it disappeared. Is an imagined dynamic accent the same as a metrical accent? In general, the answer must be no, because it is possible to imagine accented tones that occur on weak metrical beats, as they often do in music. In the present study, however, an imagined dynamic accent may well have accompanied the downbeat, as no attempt was made to dissociate these two mental phenomena.[10]

A third possibility is that creation of a downbeat involves an internal simulation of appropriate actions such as foot tapping or hand clapping. In the brain, this may involve an increased activation of areas involved in action planning, as is generally found during covert action, motor imagery, or action simulation (Jeannerod, 1999, 2001). Simulated action might also be accompanied by imagery of auditory consequences of action such as thuds or claps.

Each of these possibilities may contain a grain of truth. However, the author's introspections suggest a fourth possibility, which combines aspects of the previous three. The hypothesis is that downbeat location is manipulated most effectively by generating an *internal rhythm* having a dynamic or pitch accent in the desired position and synchronizing that internal rhythm with the external rhythm. The important difference from the three preceding accounts is that the entire rhythm is generated internally, not just the downbeat. The internal rhythm is a combination of covert action and auditory image. Imagined speech can be effective. For example, the different downbeat locations in a TTT0 sequence could be created by imagining trisyllabic words or utterances with different accent patterns in synchrony with the rhythm (e.g., "petticoat," "banana," "à la carte"). The prosodic structure of the internal speech utterance thus is foisted on the percept of the external tone sequence. In fact, it is not necessary to imagine a specific utterance; it is sufficient to imagine its accent pattern as a speech melody, with increased pitch and/or loudness on the accented element. To the extent that this involves covert action, it is more likely vocal than gestural.

The resulting perceptual reorganization of the external rhythm is comparable to changing the perspective of a Necker cube: The same rhythm "sounds" entirely different when organized within a different metrical framework, even though it remains physically the same (and a listener may remain aware of the physical identity). Unlike a Necker cube, however, a rhythm can

[10]Yet another mental phenomenon that may have accompanied metrical framework generation is cognitive regrouping of the rhythm. For example, the TT0 rhythm might be conceived as cycles of TT0 when the downbeat is on T_1, but as cycles of T0T when the downbeat is on T_2. In other words, the rhythm would be organized into "measures," with a downbeat on the first element of each measure. If such cognitive regrouping occurred, it too was possibly confounded with downbeat location. It is unclear, however, whether cognitive regrouping is really possible, or whether it is just an illusion based on familiarity with musical notation.

often be organized in more than two ways, its different organizations are not equally easy to achieve and maintain (as we have seen in Experiment 3), and they can be controlled more effectively by pure mental effort than a Necker cube's perspectives (Toppino, 2003), at least by musically trained individuals. Rather than alternating more or less randomly between equally salient alternatives, as a Necker cube does, a rhythm usually favors one particular metrical organization, and other organizations must be achieved and maintained against a pull toward the favored organization. While it is well known that rhythmic grouping accent and dynamic accent are major determinants of the favored metrical organization, the present study shows that an even more important determinant is an overt motor action that singles out one of the tones of a rhythm (i.e., tapping). Clearly, it is easier to generate an internal rhythm whose accented element is in synchrony with the overt movement than one that conflicts with the movement. Some real mental effort, presumably reflecting a high level of attention, is felt in the latter case.

Obviously, introspection and speculation can only go so far, and it is hoped that neuroscientific investigations will soon shed more light on these intriguing issues.

Acknowledgments

This research was supported by NIH grants MH-51230, DC-03663, and HD-01994. A short report has appeared in conference proceedings (Repp, 2004).

Address correspondence to: Bruno H. Repp, Haskins Laboratories, 300 George Street, New Haven, CT 06511-6624. E-MAIL *repp@haskins.yale.edu*

References

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

Chen, Y., Ding, M., & Kelso, J. A. S. (2001). Origins of timing errors in human sensorimotor coordination. *Journal of Motor Behavior, 33*, 3-8.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception & Psychophysics, 53*, 25-33.

Drake, C. (1997). Motor and perceptually preferred synchronisation by children and adults: Binary and ternary ratios. *Polish Quarterly of Developmental Psychology, 3*, 43-61.

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

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception, 18*, 1-23.

Engström, D. A., Kelso, J. A. S., & Holroyd, T. (1996). Reaction–anticipation transitions in human perception–action patterns. *Human Movement Science*, *15*, 809-832.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America, 98*, 2524-2531.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Jeannerod, M. (1999). To act or not to act: Perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology, 52A*, 1-29.

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage, 14*, S103-S109.

Keele, S. W., & Hawkins, H. L. (1982). Explorations of individual differences relevant to high level skill. *Journal of Motor Behavior, 14*, 3-23.

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.

Keller, P. E., & Repp, B. H. (2004). When two limbs are weaker than one: Sensorimotor syncopation with alternating hands *Quarterly Journal of 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.

Kelso, J. A. S., DelColle, J. D., & Schöner, G. (1990). Action-perception as a pattern formation process. In M. Jeannerod (Ed.), *Attention and performance XIII* (pp. 139-169). Hillsdale, NJ: Erlbaum.

Leek, M. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics, 63*, 1279-1292.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

London, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses. *Music Perception, 19*, 529-550.

London, J. (2004). *Hearing in time*. New York: Oxford University Press.

Mates, J., Radil, T., Müller, U., & Pöppel, E. (1994). Temporal integration in sensorimotor synchronization. *Journal of Cognitive Neuroscience, 6*, 332-340.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception, 11*, 409-464.

Peters, M. (1980). Why the preferred hand taps more quickly than the non-preferred hand: Three experiments on handedness. *Canadian Journal of Psychology, 34*, 62-71.

Peters, M. (1985). Constraints in the performance of bimanual tasks and their expression in unskilled and skilled subjects. *The Quarterly Journal of Experimental Psychology, 37A*, 171-196.

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

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

Povel, D.-J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics, 30*, 565-572.

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

Repp, B. H. (2001). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 600-621.

Repp, B. H. (2002). Automaticity and voluntary control of phase correction following event onset shifts in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 410-430.

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits/costs of interval subdivision. *Journal of Motor Behavior, 35*, 355-370.

Repp, B. H. (2004). Effects of three kinds of accent on synchronization with simple rhythms. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the 8th Conference on Music Perception and Cognition* (CD-ROM). Causal Productions (http://www.causal-productions.com/).

Repp, B. H. (2005). Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 1. Qualitative observations. *Music Perception*, *22*, 479-496.

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

Repp, B. H., & Saltzman, E. L. (2001). *Influences of grouping and metrical structure on the kinematics of rhythmic finger tapping*. Unpublished manuscript. Downloadable from http://www.haskins.yale.edu/haskins/STAFF/repp.html.

Royer, F. L., & Garner, W. R. (1966). Response uncertainty and perceptual difficulty of auditory temporal patterns. *Perception & Psychophysics, 1*, 41-47.

Semjen, A., Schulze, H.-H., & Vorberg, D. (1992). Temporal control in the coordination between repetitive tapping and periodic external stimuli. In C. Auxiette, C. Drake, & C. Gérard (eds.), *Proceedings of the Fourth Rhythm Workshop: Rhythm Perception and Production* (pp. 73-78). Bourges, France: Imprimérie Municipale.

Snyder, J. S., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception, 18*, 455-489.

Tekman, H. G. (1998). Effects of melodic accents on perception of intensity. *Music Perception, 15*, 391-401.

Thomassen, J. M. (1982). Melodic accents: Experiments and a tentative model. *Journal of the Acoustical Society of America, 71*, 1569-1604.

Todor, J. I., & Kyprie, P. M. (1980). Hand differences in the rate and variability of rapid tapping. *Journal of Motor Behavior, 12*, 57-62.

Toiviainen, P., & Snyder, J. S. (2003). Tapping to Bach: Resonance-based modeling of pulse. *Music Perception, 21*, 43-80.

Toppino, T. C. (2003). Reversible-figure perception: Mechanisms of intentional control. *Perception & Psychophysics, 65*, 1285-1295.

Truman, G., & Hammond, G. R. (1990). Temporal regularity of tapping by the left and right hands in timed and untimed finger tapping. *Journal of Motor Behavior, 22*, 521-535.

Van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation, Eindhoven University of Technology. http://alexandria.tue.nl/extra1/PRF2A/7707058.pdf.

Van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pitch. *Journal of New Music Research, 28*, 43-66.

Volman, M. J. M., & Geuze, R. H. (2000). Temporal stability of rhythmic tapping "on" and "off the beat": A developmental study. *Psychological Research, 63*, 62-69.

Vos, P. G., & Helsper, E. L. (1992). Tracking simple rhythms: In-phase versus anti-phase performance. In F. Macar, V. Pouthas, & W. J. Friedman (eds.), *Time, action, and cognition: Towards bridging the gap* (pp. 287-299). Dordrecht: Kluwer.

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport, 8*, 1971-1974.

Yabe, H., Tervaniemi, M., Sinkkonen, J., Huotilainen, M., Ilmoniemi, R. J., & Näätänen, R. (1998). Temporal window of integration of auditory information in the human brain. *Psychophysiology, 35*, 615-619.

Yabe, H., Winkler, I., Czigler, I., Koyama, S., Kakigi, R., Sutoh, T., Hiruma, T., & Kaneko, S. (2001). Organizing sound sequences in the human brain: the interplay of auditory streaming and temporal integration. *Brain Research, 897*, 222-227.