Exp Brain Res (2005) 163: 226–238
DOI 10.1007/s00221-004-2159-8

RESEARCH ARTICLE

Aniruddh D. Patel · John R. Iversen · Yanqing Chen
Bruno H. Repp

The influence of metricality and modality on synchronization with a beat

Received: 27 February 2004 / Accepted: 18 October 2004 / Published online: 15 January 2005
© Springer-Verlag 2005

Abstract The great majority of the world's music is metrical, i.e., has periodic structure at multiple time scales. Does the metrical structure of a non-isochronous rhythm improve synchronization with a beat compared to synchronization with an isochronous sequence at the beat period? Beat synchronization is usually associated with auditory stimuli, but are people able to extract a beat from rhythmic visual sequences with metrical structure? We addressed these questions by presenting listeners with rhythmic patterns which were either isochronous or non-isochronous in either the auditory or visual modality, and by asking them to tap to the beat, which was prescribed to occur at 800-ms intervals. For auditory patterns, we found that a strongly metrical structure did not improve overall accuracy of synchronization compared with isochronous patterns of the same beat period, though it did influence the higher-level patterning of taps. Synchronization was impaired in weakly metrical patterns in which some beats were silent. For the visual patterns, we found that participants were generally unable to synchronize to metrical non-isochronous rhythms, or to rapid isochronous rhythms. This suggests that beat perception and synchronization have a special affinity with the auditory system.

Keywords Auditory perception · Motor skills · Music · Rhythm · Visual perception

A. D. Patel (✉) · J. R. Iversen · Y. Chen
The Neurosciences Institute,
10640 John Jay Hopkins Drive,
San Diego, CA 92121, USA
E-mail: apatel@nsi.edu
Tel.: +1-858-626-2085
Fax: +1-858-626-2099

B. H. Repp
Haskins Laboratories, 270 Crown St.,
New Haven, CT 06511, USA

Introduction

Most if not all human cultures have some form of music which contains or induces a *beat*—a perceived pulse that marks equally spaced points in time (Large and Palmer 2002; Nettl 2000). It is common for humans to synchronize their body movements with the beat, either as part of dance or simply as an individual response to the music. Beat perception and synchronization (BPS) typically occur quickly and require little conscious effort, which might lead one to believe that these are simple phenomena based on primitive cognitive and neural mechanisms. Several lines of evidence suggest otherwise. First, *Homo sapiens* is the only species known to spontaneously synchronize body movements with an auditory rhythmic pulse. Several insect and frog species synchronize their sound production with conspecifics during rhythmic chorusing, but do not show some key features of BPS. For example, human BPS does not necessarily involve sound production and is characterized by flexibility in terms of the coupling of sensory and motor systems: humans can move in phase or anti-phase with a beat, and can synchronize using different body parts.[1]

Second, human listeners can often synchronize at rates which are integer multiples or fractions of the basic beat. This indicates that the mind has access to several distinct levels of periodicity, one of which can be selected at any given time as the beat (Drake et al. 2000b; Parncutt 1994). Third, a beat can be perceived even when musical sequences have a good deal of systematic or random timing variability due to expressive factors

---

[1] Importantly, the human uniqueness of BPS is not due to a lack of accurate timekeeping mechanisms in other animals. For example, studies of rabbits (e.g., Moore et al. 1998) show that other mammals are capable of interval timing. Given this fact, it is curious that despite many decades of research in psychology and neuroscience in which animals have been trained to do complex motor tasks, there is to our knowledge not a single published case in which a non-human animal species has been successfully trained to tap or peck in synchrony with an auditory metronome.

(Large and Palmer 2002; Madison and Merker 2002). Fourth, a perceived beat can tolerate a certain amount of counterevidence such as accented events at non-beat locations and beat locations without event onsets (syncopation). These considerations suggest that beat perception reflects a rich set of psychological processes.

BPS is of scientific interest because its underlying cognitive and neural mechanisms touch on several key issues in psychology. These include mental timekeeping, the relationship between perception and action, and the coordination of different brain systems (auditory and motor). Fortunately, BPS has begun to attract interest from a variety of perspectives, including behavioral research (Drake 1997; Drake et al. 2000a; Snyder and Krumhansl 2001), computational modeling (Desain 1992; Desain and Honing 1999; Large 2000; Large and Kolen 1994; Large and Palmer 2002; Todd et al. 1999, 2002, Toiviainen and Snyder 2003), and neuroscience (Brochard et al. 2003; Snyder and Large, 2004, in press; cf. Thaut 2003). Thus the study of BPS has the potential to mature into an area where alternative computational models compete to explain a common body of behavioral and neural data. To achieve this level of sophistication, however, a foundation of basic behavioral and neural research on BPS is needed. In particular, behavioral and brain responses to beat-inducing sequences of varying degrees of complexity are needed to help constrain quantitative models of BPS (cf. Large 2000; Large et al. 2002). Although a long tradition of research exists on synchronization with simple metronomes (e.g., Dunlap 1910; Fraisse et al. 1958; Woodrow 1932), there is relatively little work on BPS in sequences whose rhythmic complexity has been systematically manipulated. This research is needed because real musical rhythms are more complex than a simple metronome. In particular, real musical rhythms often exhibit *meter*.

Meter and synchronization with a beat

Meter is a ubiquitous aspect of musical rhythm, and refers to multiple levels of periodicity in rhythmic structure, i.e., periodicity at multiple time scales (Cooper and Meyer 1960; Lerdahl and Jackendoff 1983). For example, a waltz has a meter in which the basic beat is accompanied by a higher level periodicity every three beats, in contrast to a march, which has a higher-level periodicity every two beats. Multiple levels of periodicity can be created in a number of ways. In a waltz, for example, the higher-level periodicity may be marked by intensity, pitch, or duration accents on every third beat. This physical accenting is not a necessary prerequisite for meter, however; metrical structure can also be created by temporal patterning alone in the absence of any pitch or intensity variation (e.g., Povel and Essens 1985, and our stimuli). In general, the number of levels of periodicity that can be perceived and synchronized with is limited to about three or four, e.g., a level above and a level below the most salient beat or "tactus" (Drake et al. 2000a; London 2002), which tends to be in the 400–900 ms (Parncutt 1994) or 300–700 ms range (van Noorden and Moelants 1999).

There are theories of rhythm perception which are explicitly concerned with regular temporal patterns at different time scales. For example, the dynamic attending theory of Jones (Jones 1976; Jones and Boltz 1989) posits that there are internal attentional rhythms which can become entrained to temporal regularities in the environment. An important aspect of this theory is the notion that different internal rhythms entrain to different levels of temporal structure. This idea has been implemented with computational models that employ oscillators to model internal rhythms (Large and Jones 1999; Large and Palmer 2002; McAuley 1995). The interaction of different internal rhythms is modeled via a bank of coupled noisy nonlinear oscillators tuned to different periods, which are excited by acoustic input. A single oscillator excited by a simple isochronous sequence produces output with a certain degree of variability. A metrical sequence, however, excites several oscillators with periods which are multiples of the basic beat period. Due to the coupling, these different oscillators are drawn into a stable relationship with each other. This results in lowered variability for the oscillator at the beat period compared with a situation in which only a single oscillator is engaged. Thus, such models predict improved synchronization to a beat when listening to metrical (vs. simple isochronous) sequences (E. Large, personal communication).[2] However, any improvement due to meter is likely to depend on the number and regularity of events that reinforce temporal structure at different time scales. Thus, an oscillator may be more strongly entrained when all cycles are marked by event onsets than when some event onsets are missing, e.g., due to syncopation (cf. Large and Jones 1999).

We investigate how BPS varies as a function of metrical structure by examining the synchronization accuracy to strongly metrical, weakly metrical, and isochronous sequences, all of which have the same beat period (800 ms).[3] (Definitions of strongly and weakly metrical are given in the Methods section.) This allows us to directly address whether a strong meter improves synchronization with a beat. In addition to the question of variability, we also ask if metrical structure leaves an imprint on the temporal pattern of synchronization. That is, we ask if taps to metrical sequences are different in terms of their higher-order patterning than taps synchronized to isochronous sequences of the same beat period. There are a few studies which examine higher-level patterns of tapping in response to rhythmic manipulations. Vorberg and Hambuch (1978) found

---

[2] We are grateful to Edward Large for confirming these predictions with the coupled-oscillator model of Large and Jones (1999) and with a newer version of this model, using our stimuli as input.

[3] Throughout this paper, the term "synchronization accuracy" refers to the variability of tapping as measured by the variability of tap-to-tone asynchronies or of inter-tap-intervals (ITIs), and not to absolute asynchronies or mean ITIs.

higher-level periodicities in continuation tapping after a beat was induced by a preceding isochronous synchronization sequence. Keller and Repp (2004) discovered that regular accents in an isochronous sequence modulated the tap-tone asynchronies in an anti-phase synchronization task. In contrast to these studies, we looked for effects of a much slower higher-level periodicity (circa 3 s, see below) on tapping patterns. Any such effects can help inform computational models of BPS, which must ultimately account for the effect of meter on both the variability and the temporal patterning of synchronized movement.

Meter and modality

Patterns of synchronization as a function of stimulus *modality* are also important to constrain theories of the neural mechanisms of BPS. To date, studies of BPS to complex rhythms have focused exclusively on sound: we do not know if beat perception can be supported by complex rhythmic patterns in other modalities such as vision. It is well known that synchronization with a visual metronome is worse than with an acoustic metronome (Chen et al. 2002; Kolers and Brewster 1985; Repp and Penel 2002; Semjen and Ivry 2001). It is also known that non-isochronous temporal sequences are better discriminated and reproduced when presented in the auditory than the visual modality (Handel and Buffardi 1969; Glenberg and Jona 1991). It is not known, however, if a beat can be extracted from metrical visual sequences. The question is worth asking because there is evidence that perceivers are sensitive to multiple levels of event structure in simple visual sequences (light flashes) (Holleran and Jones, unpublished data, 2001), suggesting that they may be able to extract lower-level and higher-level periodicities to infer a meter. Furthermore, one may note that it is common for musicians to follow the beat of a conductor, suggesting that BPS can occur in the visual modality.

We address this issue by presenting the same rhythmic patterns either as auditory sequences (tones) or visual sequences (light flashes). We chose a tempo for these sequences (one beat every 800 ms) at which people can synchronize comfortably with an auditory or visual metronome. The metrical sequences in our study included shorter intervals (minimum 200 ms) as well as longer periodicities which defined higher level temporal units (3.2 s, see Methods). Recent results on synchronization with visual sequences (Repp 2003) suggest that visual sequences containing short intervals might be difficult to synchronize with. This naturally raised the question of whether we should employ a slower tempo for our visual sequences. Two factors led us to keep the tempo the same in the two conditions. First, the study of Repp (2003) only used isochronous sequences, leaving open the question of whether metrical structure might overcome the difficulties posed by short visual intervals. Second, if we employed a slower tempo for the visual sequence (e.g., by making the shortest interval 400 ms), the higher level temporal units would have been lengthened to a duration which would likely fall outside the temporal envelope for meter perception (London 2002). Thus we decided to use the same tempo for our auditory and visual stimuli, which had the added advantage of direct comparability of results across domains.

Past work on BPS to complex sequences

Several studies in the literature have examined auditory BPS with complex stimulus sequences. For example, Handel and coworkers (Handel and Lawson 1983; Handel and Oshinsky 1981) used synchronized tapping to assess the perceived beat in polyrhythmic sequences. Van Noorden and Moelants (1999) had participants tap with the perceived beat of various musical pieces broadcast on the radio, in order to determine the most preferred beat frequency, which they found to be slightly faster than 2/s. Drake et al. (2000a, b) examined participants' ability to tap with the beat of musical pieces, and also at lower and higher metrical levels (subdivisions or multiplications of the beat), as a function of the presence versus absence of intensity accents or expressive timing variation. Although expressive timing variation reduced synchronization accuracy, it seemed to aid the recovery of metrical structure. Large and Palmer (2002) arrived at similar conclusions on the basis of computational modeling. Snyder and Krumhansl (2001) examined the accuracy of tapping with the beat of piano ragtime pieces as a function of the presence/absence of pitch variation and of the left-hand part. Whereas pitch had little effect, elimination of the left-hand part (which regularly marked the beat) made synchronization more variable (but see Toiviainen and Snyder 2003, for evidence that pitch structure influences BPS). Large (2000) reproduced these findings with his computational model of beat finding. In a recent study by Large et al. (2002), error correction in response to small perturbations at different metrical levels was examined in a task that required tapping in synchrony with different metrical levels of complex rhythms. Experiments by Repp (1999a, b, 2002) on synchronized tapping with piano music were concerned primarily with participants' ability to track expressive timing variations. There seems to be no precedent in the literature for our research, which addresses the influences of metrical structure and modality on the accuracy of synchronization with the beat of complex rhythms.

Methods

Task

Throughout the experiment (conducted at Haskins Laboratories), the participants had a single task: to synchronize their taps with the (prescribed) beat of rhythmic sequences. The beat always had a tempo of close to 1.25 Hz (one beat every 800 ms)[4], which was indicated at the beginning of each trial by an induction sequence of nine isochronous tones or light flashes, depending on the condition. The induction sequence led immediately into the rhythmic sequence, which lasted for 48 s (60 beats). The participants were instructed to synchronize with the induction sequence (one tap per event), starting with the second event, and then to continue tapping at the same tempo with the rhythmic sequence, using the structure of the sequence to aid synchronization of the taps with the beat. If the structure did not help and synchronization with the beat appeared difficult, participants were simply to try to maintain a constant rate of tapping. The extent to which they were able to maintain the tempo and appropriate phase of their tapping served as a measure of the success of beat perception and synchronization.

Materials

Seven types of rhythmic sequence were used, of which three were isochronous, three were metrical, and one was isochronous with gaps. They are described below along with the rationale for each. Schematic diagrams of the different sequence types are shown in Fig. 1.

1. Isochronous 800 (I-800)

This served as a baseline sequence in which one event occurred every 800 ms.

2. Isochronous 400 (I-400)

In this condition one event occurred every 400 ms, so that participants tapped on every other event. This condition was included in order to determine if any advantage of BPS in strongly metrical (vs. I-800) sequences might simply be due to explicit subdivision of the beat interval rather than to multiple periodicities both above and below the prescribed beat period. On the basis of previous results on interval subdivision (Repp 2003), we expected beat synchronization in I-400 sequences to be more accurate than in I-800 sequences.

3. Isochronous 200 (I-200)

In this condition one event occurred every 200 ms, so that participants tapped on every fourth event. Like the I-400 sequence, this sequence controlled for possible benefits of subdivision of the beat vs. advantages due to the coordination of lower and higher-level periodicities. However, we did not expect synchronization in I-200 sequences to be more accurate than in I-400 sequences because Repp (2003) found that subdivision aided synchronization only when the subdivisions were longer than 200 ms.

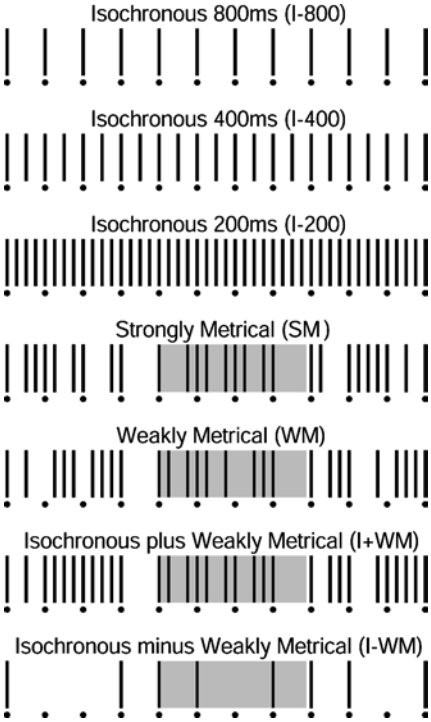

Fig. 1 Examples of different sequences. Only a portion of each sequence is shown. *Vertical bars* represent event onsets, and *dots* represent beat locations. In the non-isochronous sequences, the shading corresponds to one Povel and Essens (P&E) pattern of four beats. Examples of auditory and visual versions of the SM and WM sequences can be found at: http://www.nsi.edu/uers/patel/sound_examples/BT1

4. Strongly metrical (SM)

To construct SM sequences we used temporal patterns from Povel and Essens (1985). These authors investigated the reproduction of different rhythmic patterns. All patterns consisted of permutations of the sequence of integers [1 1 1 1 1 2 2 3 4], where each integer represents a multiple of a basic inter-onset-interval (IOI) of 200 ms; the longest interval was always in the final position. For example, [1 1 1 1 3 1 2 2 4] can be transcribed in the following way, where x represents an event (e.g., a tone onset or flash at the beginning of a 200-ms interval) and a dot represents a 200-ms silent interval: x x x x x . . x x . x . x . . .

Povel and Essens (P&E) presented each permutation in a repeating fashion (the same pattern chained together with itself) and found that some permutations were learned more quickly and reproduced more accu-

---

[4] The actual inter-beat interval was 781 ms because the software used for sequence generation (MAX 3.0 running on a Macintosh Quadra 660AV computer) produced output that was 2.4% faster than specified. Throughout the paper, we report the specified or recorded temporal values. Actual values can be obtained, if necessary, by multiplying by .976.

rately than others. They attributed this to the ease with which the different patterns induced an internal clock based on the temporal regularity of subjective accents,[5] and provided a mathematical model ranking the clock-inducing strength of each pattern. We used P&E's ranking of 35 permutations of the basic pattern to create our SM sequences. Specifically, we used their top 15 clock-inducing patterns, which are shown in Table 1 ("SM" patterns).

For visual convenience, each pattern's representation in terms of x's and dots is accompanied by an indication of the beat by a vertical bar just before the beat location (note that all P&E patterns are four beats long and always have a long silent interval after the fourth beat). As can be seen from Table 1, all 15 SM patterns have an event on each beat. Each SM sequence in our study had P&E patterns 1–15 chained together in random order, so that each pattern occurred just once in each SM sequence, creating non-isochronous non-repeating metrical sequences. This ensured that BPS was not based on a strategy of memorizing particular patterns and tapping on particular events, but instead relied on true extraction of a beat (cf. Large et al. 2002). Figure 2 shows a theoretical metrical analysis of a portion of one such sequence, suggesting the different levels of periodicity that are accessible to a listener (i.e., synchronization is possible at any of these levels).

As suggested by the figure, SM sequences are metrical because they have periodicity at multiple time scales. In particular, there is a salient higher-level periodicity every four beats, created by the long final gap in each P&E pattern. This lends a feeling of a "downbeat" to beat 1 of each P&E cycle, i.e., a long-duration periodicity that rides atop the basic periodicity of a beat every 800 ms. Furthermore, within the 60 beat periods of each SM sequence, 23 (38%) have an event which subdivides the beat into two 400-ms intervals.

It should be noted in passing that in their original study, Povel and Essens (1985) based their model of beat induction strength on the temporal patterning of subjective accents produced by their sequences, based on rules for assigning subjective accents to groups of tones (Povel and Okkerman 1981). We did not follow this scheme in assigning metrical strength to our sequences, but used a simplified scheme based on classifying each P&E pattern as strongly metrical or weakly metrical depending on whether it had an event at every beat position or not. Since each SM sequence contained the same 15 P&E patterns, the sequences were balanced for overall metrical strength according to Povel & Essens' ranking of the 15 patterns.

Table 1 Strongly metrical and weakly metrical patterns from Povel & Essens (1985)

<table>
<tr><td>SM1</td><td>| x x x x | x . . x | x . x . | x . . .</td></tr>
<tr><td>SM2</td><td>| x x x . | x . x x | x . . x | x . . .</td></tr>
<tr><td>SM3</td><td>| x . x x | x . x x | x . . x | x . . .</td></tr>
<tr><td>SM4</td><td>| x . x . | x x x x | x . . x | x . . .</td></tr>
<tr><td>SM5</td><td>| x . . x | x . x . | x x x x | x . . .</td></tr>
<tr><td>SM6</td><td>| x x x . | x x x . | x x . . | x . . .</td></tr>
<tr><td>SM7</td><td>| x . x x | x x . x | x . . x | x . . .</td></tr>
<tr><td>SM8</td><td>| x x . . | x x x x | x . x . | x . . .</td></tr>
<tr><td>SM9</td><td>| x x . . | x . x x | x . x x | x . . .</td></tr>
<tr><td>SM10</td><td>| x . x x | x . x x | x x . . | x . . .</td></tr>
<tr><td>SM11</td><td>| x x x . | x x . . | x x . x | x . . .</td></tr>
<tr><td>SM12</td><td>| x x . x | x x x . | x . . x | x . . .</td></tr>
<tr><td>SM13</td><td>| x x . x | x . x x | x x . . | x . . .</td></tr>
<tr><td>SM14</td><td>| x x . . | x x . x | x . x x | x . . .</td></tr>
<tr><td>SM15</td><td>| x . . x | x x . x | x x . x | x . . .</td></tr>
<tr><td>WM1</td><td>| x x x x | x . x x | . x . . | x . . .</td></tr>
<tr><td>WM2</td><td>| x x x x | . x . . | x x x . | x . . .</td></tr>
<tr><td>WM3</td><td>| x x x . | . x x . | x x x . | x . . .</td></tr>
<tr><td>WM4</td><td>| x . x x | x . . x | . x x x | x . . .</td></tr>
<tr><td>WM5</td><td>| x . x . | . x x x | x . x x | x . . .</td></tr>
<tr><td>WM6</td><td>| x x x x | . x . x | . . x x | x . . .</td></tr>
<tr><td>WM7</td><td>| x x . x | x x . x | . . x x | x . . .</td></tr>
<tr><td>WM8</td><td>| x x . x | . . x x | x . x x | x . . .</td></tr>
<tr><td>WM9</td><td>| x . x x | x x . x | . . x x | x . . .</td></tr>
<tr><td>WM10</td><td>| x . . x | x x x x | . x x . | x . . .</td></tr>
<tr><td>WM11</td><td>| x x x x | . x x x | . . x . | x . . .</td></tr>
<tr><td>WM12</td><td>| x x x x | . . x x | . x x . | x . . .</td></tr>
<tr><td>WM13</td><td>| x x . x | x x x . | . x x . | x . . .</td></tr>
<tr><td>WM14</td><td>| x x . x | . . x x | x x x . | x . . .</td></tr>
<tr><td>WM15</td><td>| x . x . | . x x x | . x x x | x . . .</td></tr>
</table>

SM strongly metrical, *WM* weakly metrical, *x* event onset, . silent position, | indicates that the following event or silent position is associated with a beat

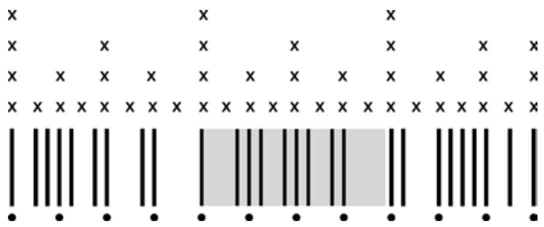

Fig. 2 A metrical analysis of a portion of an SM sequence. The beat prescribed to the listener (800-ms period) is shown by solid dots below the vertical lines, and corresponds to the third row of x's from the top. X's below this level show a subdivision of the beat at 400 ms, and x's at higher levels show higher level periodicities at 1,600 and 3,200 ms. The 3,200-ms period is the length of one P&E cycle and is marked by the presence of a long gap after the fourth beat of each cycle. As in Fig. 1, the shading indicates one P&E pattern of four beats. This hierarchical metrical analysis is our own, and not that of Povel and Essens (1985)

5. *Weakly metrical (WM)*

These sequences consisted of random concatenations of the 15 P&E patterns that were least likely to induce a beat in the study of Povel and Essens (1985) (labeled WM patterns in Table 1). As can be seen in Fig. 1, each WM pattern had one or two beats that were not marked by event onsets (always the second and/or third beat). As a result, within each WM sequence 19 of the 60 beats (32%) did not correspond to physical events. WM sequences were included in our experiment to verify that metricality affects synchronization accuracy. We expected that participants would be more accurate

---

[5] Accents perceived due to grouping rather than due to physical differences between events such as intensity differences.

synchronizing with SM sequences than with WM sequences. It should be noted that WM sequences *can* be perceived as metrical structures, especially when an induction sequence is provided, but the beat should be more difficult to maintain than in SM sequences because there is less structural support for it.

6. *Isochronous plus weakly metrical (I+ WM)*

This sequence type was included to address the question of whether the (expected) difficulty of synchronizing with WM sequences was due entirely to the occurrence of unmarked beats. I + WM sequences were constructed by "filling in" these unmarked beats, in other words by adding the I-800 sequence to each WM sequence. This operation was expected to make the sequence strongly metrical, and synchronization performance was expected to be similar to that with SM sequences.

7. *Isochronous minus weakly metrical (I− WM)*

Finally, for each WM sequence we generated an isochronous sequence with gaps by omitting from the I-800 sequence all those events that were missing in the same beat position in the WM sequence. Of course, the resulting sequences were not really isochronous any more. To the extent that missing beat events impair synchronization, we expected that synchronization with I–WM sequences would be less accurate than with I-800 sequences.

Equipment and procedure

The seven sequence types were arranged in blocks, in different random orders. Each block had one instance of each sequence type. Ten blocks were presented in each modality (though due to technical problems the number of blocks actually completed by each subject ranged from 8 to 10). The first block was treated as practice and not analyzed. The auditory blocks were administered first, and the visual blocks in a second session 1–12 days later.

All sequences were produced under control of customized software (MAX 3.0) running on a Macintosh Quadra 660AV computer. Auditory sequences were produced on a Roland RD-250s digital piano via an Opcode Studio Plus Two musical instrument digital interface (MIDI) translator and were heard over Sennheiser HD540 II earphones at a comfortable intensity. The sequences consisted of identical high-pitched (C8, 4,186 Hz) synthetic piano tones ("pings") which had sharp attacks reflecting mainly key impact noise and decayed within about 100 ms. (No "note offset" was specified in the MIDI instructions.) Visual sequences consisted of a flashing light, a circular green LED 3 mm in diameter, which was the "out" indicator of an Opcode II MIDI translator box that responded to MIDI messages (being connected to a second, inactive digital piano). Because the flashes were difficult to see in bright light, the room lights were extinguished and the computer screen dimmed and covered with a cardboard flap, so that only weak illumination emanated from a gap at the bottom of the flap. The box with the flashing light was placed on the rear edge of the computer keyboard. Under these conditions, the flashes were clearly visible and distinct from each other even at 200-ms IOIs.[6] The same MIDI instructions as for auditory sequences (consisting exclusively of "note on" MIDI messages) were used to activate the light.

Participants tapped with the index finger of their preferred hand on a white key of a Fatar Studio 37 MIDI controller (a quiet three-octave piano keyboard), which they held on their lap. The key depressions were recorded by the MAX program that also controlled presentation of the sequences. The response key on the MIDI keyboard moved about 10 mm, and the key depression was recorded about halfway during the downward movement. This may have added up to −20 ms to the measured asynchronies, depending on the force with which participants struck the response key (the greater the force, the smaller the added asynchrony). Since each subject tapped in a fairly stereotyped way, we believe this measurement error largely contributed to inter-subject differences in absolute asynchronies, and did not strongly influence measures of the variability of asynchronies or ITIs (which were the main focus of our study). In the few places where we report mean asynchrony data in this paper, this caveat should be kept in mind.

For most participants, there was no auditory feedback from the keyboard, but a few struck the key forcefully enough to make the impact noise audible.

Participants

The nine participants (five women, four men) included six paid participants, a research assistant, a postdoctoral researcher, and one of us (BHR). The paid participants and BHR had participated in many earlier synchronization experiments, the other two in a few. Their ages ranged from 20 to 31 years, except for BHR who was 56 years old. All but one were right-handed. A wide range of musical training was represented, from no training at all to professional level. The work was approved by the local ethics committee, and all participants gave informed consent.

Data analysis

The raw data produced by participants were a sequence of tap times, measured from the onset of each sequence,

---

[6] We did not have the equipment necessary to determine the precise luminance and duration of the flashes. However, we determined informally that a flicker sensation persisted up to rates of about 30 Hz, i.e., a 33-ms IOI.

from which asynchronies between taps and theoretical beats as well as inter-tap-intervals (ITIs) were calculated. Extreme ITI values due to missing and double taps (which occurred infrequently, $< 1\%$), were excluded by an automatic algorithm. On the basis of examination of the data, valid (non-extreme) ITIs were defined in the range of .55 to 1.75 times the target inter-beat-interval (IBI) of 800 ms (i.e., 440–1,400 ms). ITIs longer than this were considered missing taps and were replaced by the average ITI for that trial. The majority of double taps were due to key bounce in the keyboard. The second tap of a double tap was deleted and the data analyzed as if it had not occurred.

On some trials a participant was unable to synchronize with the beat of a sequence, and instead tapped at a different tempo, which resulted in phase drift. We identified drift trials with a strict criterion, i.e., a trial in which one or more taps had an absolute asynchrony of greater than 400 ms (half the beat period). Typically, the asynchronies continued to increase after such a large asynchrony. Such trials were excluded from further analysis of asynchronies, but they were included in the analysis of ITIs. In the auditory condition, drift occurred on only 12 trials out of 268. Drift during the visual condition was common, however, and is discussed below.

Our primary dependent measures of interest were the variability of asynchronies and ITIs for each trial, as well as measures of temporal patterns in the asynchrony and ITI time series. Wherever asynchrony variability data were available, they were emphasized in the analysis because they are a measure of synchronization accuracy, whereas the ITI variability is an indicator of tempo consistency. Mean asynchronies and ITIs were of subsidiary interest; the latter primarily served to indicate whether participants were able to maintain the prescribed tempo or whether they drifted off to another tempo.

In the sections that follow, results for the auditory and visual conditions are presented and analyzed separately. This decision is based on the fact that we observed very different behavior in the two conditions. Specifically, taps to visual sequences often became decoupled from the stimulus, with participants drifting off to tap at a different tempo. Within each condition, data are analyzed using a repeated measures analysis of variance with sequence type and trial number as nested within-participant factors. Since the effect of trial number is not of interest in this study, analysis focuses on the effect of sequence type. Post-hoc comparisons were conducted with Fisher's protected least significant difference test.

Results: auditory sequences

Mean ITIs and asynchronies

Overall, participants synchronized quite well with the beat of the auditory patterns. Analysis of ITIs showed that the average ITI (across participants and trials) was very close to 800 ms for each sequence type: no average deviated by more than 0.5 ms from the target value of 800 ms.

Analysis of mean asynchronies revealed differences between the seven sequence types (Fig. 3), $F_{(6,48)} = 6.83$, $p < 0001$. Asynchronies were negative, as commonly observed in synchronization tasks. They were most negative in I-800 and I–WM sequences, and least negative in I-200 and I + WM sequences, with subdivision in general (with the exception of WM) yielding significantly less negative mean asynchrony than I-800 sequences (I-400: $p < 0.011$; I-200: $p = 0.0001$; SM: $p = 0.002$; I + WM: $p < 0.0001$). This reflects a decrease in asynchrony with an increase in mean event rate, which is consistent with findings of Wohlschläger and Koch (2000) and of Repp (2003). Mean asynchrony for WM sequences was intermediate between SM an I-800, but was not significantly different from either.

Variability of asynchronies

There were marked differences in tapping variability between sequence types, $F_{(6,48)} = 3.54$, $p < .01$ (Fig. 4). Post-hoc tests confirmed the following equalities or differences:

1) Variability in SM sequences was not significantly lower than in I-800 sequences ($p = .54$), which answers one of our main questions.
2) It was, however, significantly lower than in WM sequences ($p = .01$), which confirms the validity of the metrical strength manipulation in complex sequences.
3) Variability was lower in I-400 than in I-800 sequences, though this difference fell short of significance ($p = .07$). In contrast, the difference in variability between I-200 and I-800 sequence did not even approach significance ($p = .43$). This is very similar to the "subdivision benefit" at 400 ms but not at 200 ms found by Repp (2003).

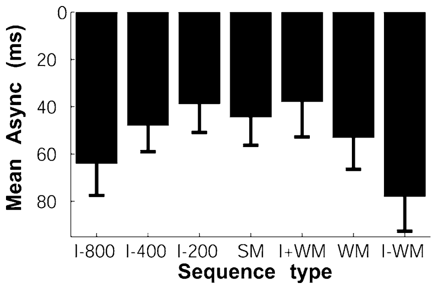

Fig. 3 Mean and s.e. of tap-to-tone asynchronies in the auditory condition. Abbreviations on the x-axis indicate sequence type: I-800 (isochronous 800-ms period); I-400 (isochronous 400-ms period), I-200 (isochronous 200-ms period), SM (strongly metrical, I + WM (isochronous plus weakly metrical), WM (weakly metrical), I-WM (isochronous minus weakly metrical)

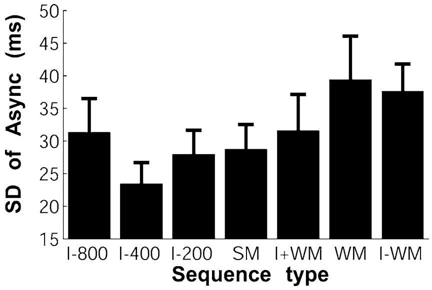

Fig. 4 Mean and s.e. of variability (standard deviation) of asynchronies in the auditory conditions

4) Variability in I + WM sequences was not significantly different than in SM sequences, and was lower than that in WM sequences, though this difference did not reach significance ($p = .07$).

5) Variability in I–WM sequences was 20% greater than in I-800 sequences, but due to variability among participants this difference was not significant ($p = .14$).

Temporal patterning of taps

Although a strong meter in SM sequences did not reduce the overall variability of tapping compared with an isochronous sequence of the same beat period, it did influence the temporal pattern of tapping. This was manifested in a smaller average asynchrony to taps on beat 1 of each P&E cycle vs. on beats 2, 3, and 4. We quantified this measure as "beat 1 relative asynchrony", which measures how close the tap on beat 1 of each cycle is to the actual beat, relative to how close the other taps in the cycle are to their corresponding beats (i.e., beats 2, 3, and 4). Mathematically, it is defined as $\left|\overline{beat1}\right| - \left|\overline{beat234}\right|$, where $\left|\overline{beat1}\right|$ indicates the absolute value of the mean of asynchronies for beat 1, and $\left|\overline{beat234}\right|$ is the absolute value of the mean of asynchronies to beats 2, 3, and 4 within a given trial. If the first tap of each four-beat cycle is closer to the beat than the other taps, the beat 1 relative asynchrony will be negative. Figure 5a shows an asynchrony time-series illustrating the calculation of this measure. In this trial, asynchronies to beat 1 (indicated by *dark markers*) are closer to zero than asynchronies to the other beats of each 4-beat cycle.

Figure 5b shows the mean value of beat 1 relative asynchrony in the different sequence types. There was a large, and highly significant, effect of sequence type, $F_{(6,48)} = 5.8$, $p = 0.0001$. Beat 1 relative asynchrony was not significantly different from zero for any of the isochronous sequences (where beat 1 was defined arbitrarily as the beat corresponding to beat 1 in SM sequences). For SM sequences, however, the mean beat 1 relative asynchrony was −5.6 ms, which was significantly different from zero, $t_{(8)} = -4.1$, $p = 0.0001$, indicating that on average, taps to beat 1 were more accurate (closer to zero asynchrony by about 5 ms). This tendency was even more pronounced in I + WM sequences, which suggests that these sequences were effectively SM sequences. The beat 1 relative asynchrony for WM sequences was not significantly different from that for SM sequences ($p = 0.13$) although the effect was smaller and the difference from zero was just barely significant ($p = 0.05$).

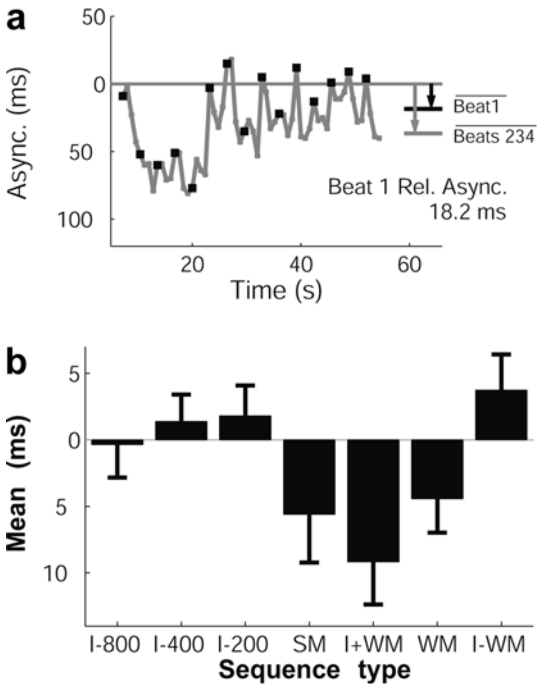

Fig. 5a An example of an asynchrony time series from the strongly metrical condition. The *horizontal line* at 0 asynchrony represents perfect synchrony of tap and tone. *Dark squares* indicate taps associated with the first beat of each P&E cycle. "Beat 1 relative asynchrony" is the absolute value of the mean of asynchronies associated with beat 1, minus the absolute value of the mean of the asynchronies associated with beats 2, 3, and 4. **b** Mean and s.e. of beat 1 relative asynchrony in the different conditions. Note the large negative values for SM and I + WM, indicating that taps were closer to the beat on beat 1 than on beats 2,3, and 4 in these conditions

Results: visual sequences

Mean ITIs and asynchronies

By a number of different measures, BPS in rhythmic visual sequences was considerably more difficult than in auditory sequences. A tendency to drift away from the beat was very common, especially in sequences containing intervals shorter than 800 ms. Because of considerable variation across participants, Fig. 6 shows the mean ITI of each sequence type for each participant.

As can be seen, only in I-800 and I-WM sequences did the ITI consistently match the prescribed 800-ms

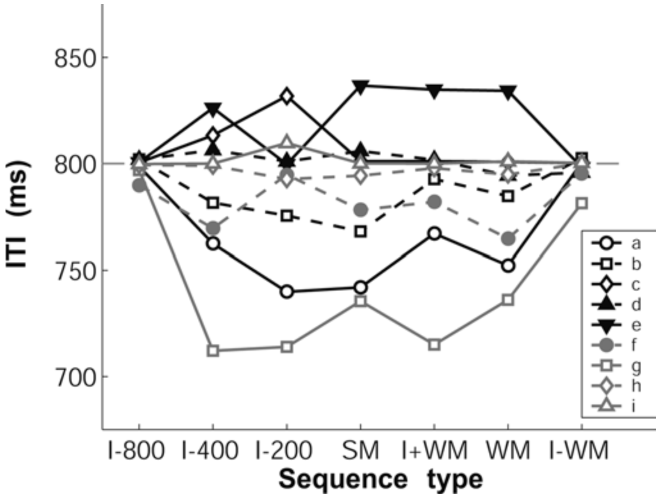

Fig. 6 Mean of ITIs in the visual conditions for individual participants (a–i)

beat period. In the other sequence types, which all contained intervals of 400 and 200 ms, individuals diverged markedly from the beat period (ranging from approximately 710 to 840 ms), and each participant tended to adopt a consistent tempo across sequence types. Mean ITIs (ms) across participants were: I-800: 798.8, I-400: 785.6, I-200: 784.3, SM: 784.6, I + WM: 788.0, WM: 784.7, I-WM: 797.1. These means conceal a good deal of individual variation, however, as seen in Fig. 6.

A related measure of synchronization difficulty are the percentages of drift trials in the seven sequence types, which were: I-800 (11%), I-400 (75%), I-200 (96%), SM (58%), I + WM (55%), WM (64%), and I-WM (15%). The sequences with large numbers of drift trials were precisely those in which participants had ITIs which were substantially different from the beat period. Together with the ITI results, this suggests that tapping in sequences with short intervals had become decoupled from the stimulus, with each participant tapping at their own preferred period, perhaps constrained by memory of the target beat period presented in the induction sequence. The difficulty in synchronizing with sequences containing short IOIs is consistent with Repp's (2003) finding that synchronization with isochronous sequences of light flashes tends to break down when IOIs get shorter than about 460 ms.

The disruptive effect of short intervals was not insurmountable, however: two participants were able to synchronize with SM sequences in every trial, and with I + WM and WM sequences, too. One of these participants was author BHR; the other one was his research assistant. It is likely that they used a strategy whereby they utilized the regular long interval at the end of each four-beat pattern as a guide and interpolated their taps within this four-beat interval (see "Temporal patterning of taps", below). One of the two was also able to synchronize in all but one trial with I-400 sequences.

For non-drift trials, the average asynchrony in I-800 sequences was −75 ms, larger (in absolute value) than for the auditory condition, which is also in agreement with Repp's (2003) previous findings. For I-WM sequences, the mean asynchrony was very similar, −69 ms.

Variability of ITIs

Because of the large percentage of drift trials, analysis of tapping variability focused on ITIs rather than asynchronies. The ITI standard deviations are shown in Fig. 7, plotted together with data from the auditory condition for comparison (note that the auditory data do not contain the few drift trials).

It can be seen that variability was larger with visual than with auditory sequences, $F_{(1,8)} = 5.23$, $p = .05$. The Modality × Sequence Type interaction did not reach significance. In a separate ANOVA on the visual condition, the main effect of sequence type was not significant $F_{(6,48)} = 1.84$, $p = .11$, though there was a weak tendency for tapping to be less variable in isochronous sequences than in sequences with variable intervals (see Fig. 7). Unlike the auditory results, there was no benefit of subdivision in isochronous sequences (i.e., tapping to I-400 was not less variable than to I-800), and metricality (SM and I + WM vs. WM) played no role at all. Furthermore, omission of beat events in I–WM sequences barely changed the variability of ITIs (a nonsignificant increase in variability of 5%, $p = .44$).

Temporal patterning of taps

For the two participants who were able to synchronize to the SM sequences, it is of interest to know if the regular long interval at the end of each cycle influenced the temporal patterning of taps in a way similar to that seen with auditory patterns (cf. auditory results section, above). Thus we examined the beat 1 relative asynchrony in SM sequences for these two participants, to see if taps to beat 1 were closer to the beat than taps on other beats. This analysis revealed no evidence for beat 1 taps to be closer to the beat than taps on other beats.

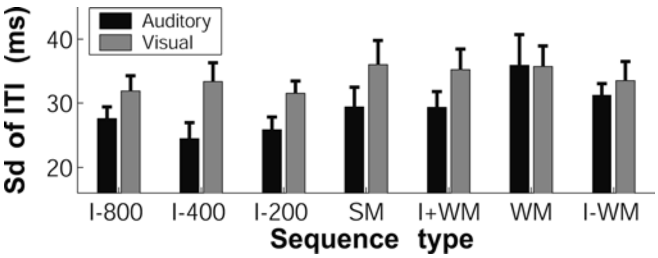

Fig. 7 Mean and s.e. of ITI variability (standard deviation) in the visual and auditory conditions

Discussion

Beat perception and synchronization (BPS), which appear to be unique to humans, involve strong coupling between different brain systems and between perception and action. Two salient features of BPS are its relationship to meter and to the auditory system, yet the influence of meter and modality on BPS have received little direct study. In terms of meter, it is notable that music seldom presents a listener with a single level of periodicity. Instead, periodicities are nested within each other to form metrical structures, and it is these structures that normally elicit BPS. In this paper we sought to answer two questions: (1) how does synchronization with a metrical structure differ from synchronization with a single periodicity?; and (2) in terms of modality, we know BPS occurs with complex acoustic rhythms, but can a beat also be extracted from equally complex visual rhythms? We addressed these questions by measuring synchronization to rhythmic patterns which were either isochronous or metrical and were either auditory or visual. We first discuss the results for auditory sequences, and then turn to the visual sequences.

Auditory sequences

For the auditory sequences, we found that tapping to strongly metrical patterns was no less variable than tapping to a metronome of the same beat period of 800 ms, suggesting that the presence of metrical structure does not necessarily improve synchronization. It is unlikely that this lack of improvement was due to a floor effect, since simple isochronous subdivision of the 800 ms metronome led to a reduction in variability. While this reduction just escaped significance in this study ($p = 0.07$), it is very similar to the subdivision benefit found by Repp (2003) using similar sequences.

Thus, if BPS involves the entrainment of coupled oscillators with different periods (e.g., Large and Jones 1999), it appears that these oscillators gain no benefit (in terms of reduced variability) from the coordination of periodicities above and below the level of the beat. This finding poses a challenge to oscillator models which predict improved synchronization with metrical vs. isochronous sequences (cf. Introduction, “Meter and synchronization with a beat”), and raise the question of whether such models will need to be modified to accommodate these results.[7]

Why would synchronization be improved by simple subdivision of an isochronous sequence but not by metrical structures of the kind used here? Assuming an attentional oscillator based framework, one reason for this may be that an isochronous sequence provides consistent physical support for its main periodicity (as well as for higher-level periodicities such as the 800-ms beat period in I-400 and I-200 sequences), whereas a complex rhythm provides only intermittent (and irregular) physical support for periodicities below the main beat (SM and I + WM sequences) or even for the main beat itself (WM sequences). To achieve maximum accuracy, participants in a synchronization task may rely on the oscillator with the strongest physical support rather than on coupling between oscillators of different periods. In I-400 sequences, this is the 400-ms level, and because perceptual variability decreases as interval duration decreases down to about 300 ms (Friberg and Sundberg 1995; Hibi 1983), the variability of the taps decreases as well. That is, participants automatically track a 400-ms beat and tap with every other beat. In SM and I + WM sequences, the most consistently supported level is the 800-ms beat, and so participants tap with it, as they do in I-800 sequences. In WM and I–WM sequences, even the main beat is weakened by the intermittency of beat events. In I-200 sequences, participants probably track the 200-ms periodicity, which is relatively more variable than the 400-ms periodicity, and tap with every fourth cycle; hence there is an increase in variability relative to I-400 sequences.

Although strong metricality did not influence tapping variability, it did influence the temporal patterning of taps. For strongly metrical sequences there was a significant tendency for the asynchrony on beat 1 of each P&E pattern to be smaller than on beats 2, 3, and 4. That is, taps were physically closer to the beat on beat 1 of each P&E pattern than on the other beats. Since our metrical sequences consisted of P&E patterns chained together, the interval formed by the tap to beat 4 in one cycle and the tap to beat 1 in the next cycle defined the final ITI of each four-beat cycle (cf. Fig. 1). Thus another way to state our finding about the smaller asynchrony on beat 1 is that there was a tendency to lengthen the final ITI of each four-beat cycle. Final lengthening of groups is a phenomenon well known from speech and music (e.g., Penel and Drake 1998; Repp 1998; Wightman et al. 1992), although the groups in our study extended over a rather long time span compared with those in earlier studies. Nevertheless, these periodic patterns suggest that the metrical sequences were perceptually segmented into four-beat groups, i.e., that listeners were sensitive to the higher-level periodicities in these temporal structures. It is likely that the most salient cue to this higher-level periodicity was the long gap at the end of each P&E pattern, as gaps are known to act as a powerful segmentation cue in auditory sequences (Garner and Gottwald 1968).

Visual sequences

By presenting the same temporal patterns used in the auditory modality as temporal sequences of light flashes,

---

[7] It should be noted that Povel and Essens (1985), who provided the patterns we used in our study, were interested in a model of rhythm perception based on different coding strategies for metrical and non-metrical sequences, and were not working in an entrainment / attentional-oscillator based framework. Nevertheless, their sequences are useful for exploring predictions of this framework.

we examined whether a beat could be extracted from complex visual rhythms at these rates. The answer to this question was clearly negative. Participants often had difficulty in synchronizing with visual rhythms, and succeeded best when there were long temporal intervals between flashes (I-800 and I–WM sequences). In sequences with shorter intervals, individuals tended to tap at a different tempo and did not maintain a consistent phase relationship with the theoretical beat. This suggests that for these participants the visual metrical sequences were simply meaningless patterns of light. While it is already known that synchronization with isochronous sequences is poorer with visual flashes than with auditory tones, our data show that metrical structure does not improve synchronization in the visual case.

It is notable that two of our participants *were* able to synchronize with visual metrical sequences, possibly by using their knowledge of the four-beat periodicity in the auditory rhythms to extract the stable first and last beat of each P&E pattern, which delimited the recurrent empty 800-ms interval. These observations suggest that a beat can be imposed on sequences of flashing lights with effort, but only by using special strategies (e.g., consciously focusing on a known or discovered underlying pattern). It seems that our rhythmic visual sequences do not easily induce a beat on their own, let alone a multi-level metrical structure.[8]

Why were our visual sequences so difficult to synchronize with? Since the individual flashes were easily discriminable at the shortest intervals used in our temporal patterns (200 ms), the results cannot simply be due to the nervous system failing to receive distinct impulses associated with each stimulus. Instead, the problem appears to be with the coordination of motor activity with the incoming visual information (Fraisse 1948). Recently, Repp (2003) has observed that synchronization with isochronous visual patterns of the kind used here breaks down at a sequence rate corresponding to IOIs of about 460 ms, almost four times slower than the corresponding limit for isochronous auditory patterns (see also Bartlett and Bartlett 1959). The 200-ms and even the 400-ms intervals in our rhythmic visual sequences were below this ‘‘synchronization threshold’’ in the visual modality, and we found that the presence of metrical structure did not help compensate for this. An obvious idea would be to slow the visual patterns down. However, if patterns were slowed down so that the shortest interval is above the synchronization threshold, they probably would be too slow to make good metrical patterns. The fastest periodicity would function as the beat in that case, and at best there could be additional slower periodicities above that level. A genuine metrical structure, however, always allows for at least one level of beat subdivision (London 2002). Thus it remains to be seen if the auditory system is unique in its ability to represent temporal structure on multiple time scales in a way that can be flexibly used for synchronization. It may be that we did not use the optimal visual stimuli for linking temporal properties of visual patterns to motor actions, and the answer to this question awaits further research.

Possible neurobiological reasons for the superiority of the auditory system in BPS

While more experiments are needed to determine if audition has a genuine advantage over vision in BPS, one may still ask if there are neurobiological reasons for expecting that synchronization will never be as good to visual stimuli as to auditory ones. There are at least two types of reasons that might explain why synchronization with auditory stimuli is superior. The first type concerns the precision with which temporal information is encoded in neural activity in the two domains. It is known that the ascending auditory pathway is specialized for dealing with fine temporal information (e.g., Carney 1999), and that primary auditory and visual cortex have important differences in the physiology and connectivity of neurons (Read et al. 2002). Thus it may be that the temporal information reaching the cortex is more accurate for auditory than for visual stimuli, and/or that the cortical cells which process the incoming information are better able to encode temporal structure in the auditory system. One way to test this idea might be to present ‘‘smeared’’ versions of auditory patterns, e.g., tones with gradual onsets or some temporal jitter between them, to see if this impairs synchronization to auditory patterns. McAnally (2002) has shown that participants can tap quite consistently with continuous tones that are frequency-modulated according to a cosine function at rates of 2–4 Hz, although variability was greater than for click sequences. It remains to be seen whether variability would be as large as for visual sequences. Synchronization with temporally jittered auditory sequences has been studied repeatedly (e.g., Hary and Moore 1987; Repp 2002; Repp and Penel 2004; Schulze 1992). There the ITIs tend to mimic the sequence inter-onset intervals at a lag of 1, as a consequence of automatic phase correction. This makes comparison with perfectly isochronous visual sequences difficult.

The second type of reason for superior synchronization to auditory stimuli concerns the efficacy of sensorimotor coupling, that is, the degree to which temporal information in either modality is able to drive actions. There is some evidence for weaker sensorimotor coupling of vision vs. audition in temporal tasks. For example, Repp and Penel (2002, 2004) observed that phase correction of tapping after a temporal perturbation to a sequence was less effective with visual than with auditory sequences. Progress on this issue may require

---

[8] It is likely that the recurring long interval also aided synchronization with auditory sequences, especially WM sequences. The difference between auditory SM and WM sequences might have been even more striking if that long interval had not been present. We are currently investigating this issue.

neurophysiological studies which can assess the degree of dynamic coupling between auditory-motor vs. visual-motor systems during synchronization.

If, after trying many different kinds of visual stimuli, it obtains that synchronization is always superior with auditory stimuli, this fact may reflect a basic functional difference between the auditory and visual systems. The auditory system constantly deals with multi-timescale structures. Melody perception, for example, involves tracking the sizes of pitch intervals as well as the more slowly unfolding patterns of melodic contour. Speech also has structure at different time scales, with phoneme identity relying on brief spectro-temporal cues (typically <100 ms) while prosodic patterns occupy hundreds of milliseconds to seconds. Understanding these acoustic sequences involves extracting information at different time scales and integrating it in real time. The visual system, in contrast, constantly deals with patterns that contain distinct information at different *spatial* scales, as when the details of an image rely on higher spatial frequencies than the general outline of the figures involved. This division of labor in terms of spatial scale is reflected in the cortical visual pathway, where neurons in higher areas respond to larger spatial scales (i.e., they have larger receptive fields, Maunsell and Newsome 1984; Oram and Perrett 1994). It may be that the visual system simply cannot neurally support processing structures at multiple time scales in a way that can guide action. Thus further cross-modal research is needed to help address the extent to which BPS relies on mechanisms that are unique to the auditory system.

Acknowledgements We thank Mari Jones for insightful comments. This research was supported by Neurosciences Research Foundation as part of its program on music and the brain at The Neurosciences Institute, where ADP is the Esther J. Burnham fellow and JRI is the Karp Foundation Fellow, by the H.A. and Mary K. Chapman Charitable Trust, and by National Institutes of Health Grant MH-51230 to BHR.

References

Bartlett NR, Bartlett SC (1959) Synchronization of a motor response with an anticipated sensory event. Psychol Rev 66:203–218

Brochard R, Abecasis D, Potter D, Ragot R, Drake C (2003) The "ticktock" of our internal clock: direct brain evidence of subjective accents in isochronous sequences. Psychol Sci 14:362–366

Carney LH (1999) Temporal response properties of neurons in the auditory pathway. Curr Opin Neurobiol 9:442–446

Chen Y, Repp BH, Patel AD (2002) Spectral decomposition of variability in synchronization and continuation tapping: comparisons between auditory and visual pacing and feedback conditions. Hum Mov Sci 21:515–532

Cooper G, Meyer LB (1960) The rhythmic structure of music. University of Chicago Press, Chicago

Desain P (1992) A(de)composable theory of rhythm perception. Music Perception 9:439–454

Desain P, Honing H (1999) Computational models of beat induction: the rule-based approach. Journal of New Music Research 28:29–42

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, Jones MR, Baruch C (2000a) The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. Cognition 77:251–288

Drake C, Penel A, Bigand E (2000b) Tapping in time with mechanically and expressively performed music. Music Perception 18:1–24

Dunlap K (1910) Reactions to rhythmic stimuli, with attempt to synchronize. Psychol Rev 17:399–416

Fraisse P (1948) Rythmes auditifs et rythmes visuels [Auditory and visual rhythms]. L'Année Psychologique 49:21–41

Fraisse P, Oléron G, Paillard J (1958) Sur les repères sensoriels qui permettent de controler les mouvements d'accompagnement de stimuli périodiques. [On the sensory reference points that make possible the control of movements accompanying periodic stimuli.] L'Année Psychologique 58:322–338

Friberg A, Sundberg J (1995) Time discrimination in a monotonic, isochronous sequence. J Acoust Soc Am 98:2524–2531

Garner WR, Gottwald RL (1968) The perception and learning of temporal patterns. Q J Exp Psychol 20:97–109

Glenberg A, Jona M (1991) Temporal coding in rhythm tasks revealed by modality effects. Mem Cognit 19:514–522

Handel S, Buffardi L (1969) Using several modalities to perceive one temporal pattern. Q J Exp Psychol 21:256–266

Handel S, Lawson GR (1983) The contextual nature of rhythmic interpretation. Percept Psychophys 34:103–120

Handel S, Oshinsky JS (1981) The meter of syncopated auditory polyrhythms. Percept Psychophys 30:1–9

Hary D, Moore GP (1987) Synchronizing human movement with an external clock source. Biol Cybern 56:305–311

Hibi S (1983) Rhythm perception in repetitive sound sequence. J Acoust Soc Jap (E) 4:83–95

Jones MR (1976) Time, our lost dimension: toward a new theory of perception, attention, and memory. Psychol Rev 83:323–355

Jones MR, Boltz M (1989) Dynamic attending and responses to time. Psychol Rev 96:459–491

Keller PE, Repp BH (2004) Staying offbeat: sensorimotor syncopation with structured and unstructured auditory sequences. Psychol Res (in press)

Kolers PA, Brewster JM (1985) Rhythms and responses. J Exp Psychol Hum Percept Perform 11:150–167

Large EW (2000) On synchronizing movements to music. Hum Mov Sci 19: 527–566

Large EW, Jones MR (1999) The dynamics of attending: How we track time-varying events. Psychol Rev 106:119–159

Large EW, Kolen JF (1994) Resonance and the perception of musical meter. Connection Science 6:177–208

Large EW, Palmer C (2002) Perceiving temporal regularity in music. Cognitive Science 26:1–37

Large EW, Fink P, Kelso JAS (2002) Tracking simple and complex sequences. Psychol Res 66:3–17

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

London J (2002) Cognitive constraints on metric systems: some observations and hypotheses. Music Perception 19:529–550

Madison G, Merker B (2002) On the limits of anisochrony in pulse attribution. Psychol Res 66:201–207

Maunsell JHR, Newsome WT (1984) Visual processing in monkey extrastriate cortex. Ann Rev Neurosci 10:363–401

McAnally KI (2002) Timing of finger tapping to frequency modulated acoustic stimuli. Acta Psychologica 109:331–338

McAuley JD (1995) Perception of time as phase: Toward an adaptive-oscillator model of rhythmic pattern processing. Unpublished doctoral dissertation, Indiana University, Bloomington

Moore JW, Choi J-S, Brunzell DH (1998) Predictive timing under temporal uncertainty: the time derivative model of the conditioned response. In: Rosenbaum DA, Collyer CE (eds) Timing of behavior: neural, psychological, and computational perspectives. MIT Press, Cambridge, MA, pp 3–34

Nettl B (2000) An ethnomusicologist contemplates universals in musical sound and musical culture. In: Wallin NL, Merker B, Brown S (eds) The origins of music. MIT Press, Cambridge, MA, pp 463–472

Oram MW, Perrett DI (1994) Modeling visual recognition from neurobiological constraints. Neural Networks 7:945–972

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

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

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. Percept Psychophys 30:565–572

Read HL, Winer JA, Schreiner CE (2002) Functional architecture of auditory cortex. Curr Opin Neurobiol 12:433–440

Repp BH (1998) Variations on a theme by Chopin: relations between perception and production of deviations from isochrony in music. J Exp Psychol Hum Percept Perform 24:791–811

Repp BH (1999a) Control of expressive and metronomic timing in pianists. J Mot Behav 31:145–164

Repp BH (1999b) Relationships between performance timing, perception of timing perturbations, and perceptual-motor synchronization in two Chopin preludes. Australian J Psychol 51:188–203

Repp BH (2002) The embodiment of musical structure: effects of musical context on sensorimotor synchronization with complex timing patterns. In: Prinz W, Hommel B (eds) Common mechanisms in perception and action: Attention and Performance XIX. Oxford University Press, Oxford, UK pp 245–265

Repp BH (2003) Rate limits in sensorimotor synchronization with auditory and visual sequences: the synchronization threshold and the benefits and costs of interval subdivision. J Mot Behav 35:355–370

Repp BH, Penel A (2002) Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. J Exp Psychol Hum Percept Perform 28:1085–1099

Repp BH, Penel A (2004) Rhythmic movement is attracted more strongly to auditory than to visual rhythms. Psychol Res 68:252–270

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

Semjen A, Ivry RB (2001) The coupled oscillator model of between-hand coordination in alternate-hand tapping: a reappraisal. J Exp Psychol Hum Percept Perform 27:251–265

Snyder JS, Krumhansl CL (2001) Tapping to ragtime: cues to pulse finding. Music Perception 18:455–490

Snyder JS, Large EW (2004) Gamma band activity reflects the metric structure of rhythmic tone sequences. Cogn Brain Res (in press)

Thaut MH (2003) Neural basis of rhythmic timing networks in the human brain. Annals of the NY Academy of Sciences 999:364–373

Todd NP McA, O'Boyle DJ, Lee CS (1999) A sensory-motor theory of rhythm, time perception and beat induction. J New Music Res 28:5–28

Todd NP McA, Lee CS, O'Boyle DJ (2002) A sensorimotor theory of temporal tracking and beat induction. Psychol Res 66:26–39

Toiviainen P, Snyder JS (2003) Tapping to Bach: resonance-based modeling of pulse. Music Perception 21:43–80

van Noorden L, Moelants D (1999) Resonance in the perception of musical pulse. J New Music Res 28:43–66

Vorberg D, Hambuch R (1978) On the temporal control of rhythmic performance. In: Requin J (ed) Attention and Performance VII. Erlbaum, Hillsdale, NJ, pp 535–555

Wightman CW, Shattuck-Hufnagel S, Ostendorf M, Price, PJ (1992) Segmental durations in the vicinity of prosodic boundaries. Journal of the Acoustical Society of America 91:1707–1717

Wohlschläger A, Koch R (2000) Synchronization error: an error in time perception. In: Desain P, Windsor L (eds) Rhythm perception and production. Swets & Zeitlinger, Lisse, The Netherlands, pp 115–128

Woodrow H (1932) The effect of rate of sequence upon the accuracy of synchronization. J Exp Psychol 15:357–379