

# Synchronizing to auditory and tactile metronomes: a test of the auditory-motor enhancement hypothesis

Paolo Ammirante<sup>1</sup> · Aniruddh D. Patel<sup>2</sup> · Frank A. Russo<sup>1</sup>

Published online: 31 May 2016  
© Psychonomic Society, Inc. 2016

**Abstract** Humans show a striking advantage for synchronizing movements with discretely timed auditory metronomes (e.g., clicking sounds) over temporally matched visual metronomes (e.g., flashing lights), suggesting enhanced auditory-motor coupling for rhythmic processing. Does the auditory advantage persist for other modalities (not just vision)? Here, nonmusicians finger tapped to the beat of auditory, tactile, and bimodal metronomes. Stimulus magnitude and rhythmic complexity were also manipulated. In conditions involving a large area of stimulation and simple rhythmic sequences, tactile synchronization closely matched auditory. Although this finding shows a limitation to the hypothesis of enhanced auditory-motor coupling for rhythmic processing, other findings clearly support it. First, there was a robust advantage with auditory information for synchronization with complex rhythm sequences; moreover, in complex sequences a measure of error correction was found only when auditory information was present. Second, higher order grouping was evident only when auditory information was present.

**Keywords** Sensorimotor synchronization · Vibrotactile perception

Several studies have shown that, given the task of synchronizing movements with a discretely timed metronome (e.g., tapping a finger), humans have a striking advantage with auditory

metronomes (e.g., clicking sounds) over visual ones (e.g., flashing lights; see, e.g., Bartlett & Bartlett, 1959; Chen, Repp, & Patel, 2002; Dunlap, 1910; Jäncke, Loose, Lutz, Specht, & Shah, 2000; Patel, Iversen, Chen, & Repp, 2005; Repp & Penel, 2002, 2004).<sup>1</sup> Interestingly, this same auditory versus visual advantage for tapping with discretely timed metronomes is not seen in monkeys, suggesting that it points to something specific about human brain organization (e.g., enhanced auditory-motor coupling for rhythmic processing; Patel, 2014; Zarco, Merchant, Prado, & Mendez, 2009). One suggestion currently being investigated is that the capacity for precise and flexible auditory-motor rhythmic synchronization may be specific to species that engage in vocal learning (Patel, Iversen, Bregman, & Schulz, 2009; Schachner, Brady, Pepperberg, & Hauser, 2009).

Does the auditory advantage persist with other modalities (not just vision) for synchronizing to discretely timed rhythms? The answer to this question can help constrain theories about the degree of auditory-motor specialization in human brains. The tactile modality is a logical testing ground for addressing this question. First, we speak of “feeling the beat” in music, and we often experience pulse-like tactile rhythmic patterns (e.g., when listening to speakers with deep bass at a concert). Second, unlike visual cues, tactile information acquired from touching a sound source consists of the same physical (vibratory) energy used by the auditory system.

✉ Paolo Ammirante  
paolo.ammirante@ryerson.ca

<sup>1</sup> Department of Psychology, Ryerson University, 350 Victoria Street, Toronto, ON M5B 2K3, Canada

<sup>2</sup> Department of Psychology, Tufts University, Medford, MA, USA

<sup>1</sup> “Discretely timed” refers to the fact that metronome events are brief with no stimulus information between events. Recent findings have shown that synchronization with a continuous visual metronome (e.g., an animation of a bouncing ball with a clear point of impact) can be as good as synchronization with a discretely timed auditory metronome (Gan, Huang, Zhou, Qian, & Wu, 2015; Iversen, Patel, Nicodemus, & Emmorey, 2015).

Furthermore, mechanoreceptors in the skin and the ear are structurally similar and exhibit similar response characteristics, such as loudness summation for tones closely spaced in frequency (Marks, 1979; Verrillo & Gescheider, 1975). Thus, vibration may evoke low-level processes that are similar between modalities. In fact, it has been shown that vibrotactile input can stimulate the auditory cortex (Caetano & Jousmaki, 2006) and that tactile and auditory stimuli can be confused when presented simultaneously or in alternation (Bekesy, 1959; Gescheider & Niblette, 1967).

We are aware of six studies involving finger tapping that have compared auditory and tactile synchronization (see Giordano & Wanderley, 2015, for a musical context). Although the results across studies are inconsistent, all report at least a modest (and sometimes significant) auditory advantage. All applied tactile stimulation to the fingertip of the non-dominant hand (Müller et al., 2008, also applied stimulation to the toe). Most used the standard deviation (*SD*) of the asynchronies between taps and metronome events (averaged across trials) as the dependent measure of synchronization. Among these studies, Müller et al. (2008) used simple metronome sequences (tapping to an isochronous beat) and found no statistical difference between auditory and tactile (fingertip) but did find a significant advantage for auditory over tactile (toe). Another two studies also appear to show an auditory advantage for simple sequences but do not report statistical comparisons (Elliot, Wing, & Welchman, 2010, Fig. 2b; Elliot, Wing, & Welchman, 2011, Fig. 2a [0 ms jitter condition in both studies]). However, Aronow, Austin, Iversen, and Patel (2015) found significant auditory advantages for tapping to the beat of simple metronome sequences (with an isochronous beat) and complex sequences (modeled on the “strongly metrical” condition of Patel et al., 2005, which contained half and quarter divisions of the beat).

Using standard deviations of the durations of intertap intervals (ITI) as the dependent measure, Kolers and Brewster (1985) also found a significant auditory advantage for simple sequences. Using this same measure, Brochard, Touzalin, Despres, and Dufour (2008) found no statistical modality difference for either simple or complex sequences. However, it should be noted that the standard deviation of ITI may be a better index of tempo maintenance than of synchronization (Patel et al., 2005). Indeed, it is possible to obtain low standard deviation of ITI by maintaining a steady beat that bears no relation to the metronome signal. For example, in an auditory/visual study, Repp (2003, Experiment 1) found little difference between modalities on standard deviation of ITI despite substantial difficulties with visual synchronization.

Most previous studies (except Elliot et al. 2010, 2011) delivered auditory stimulation to both ears via headphones, in contrast to tactile stimulation that was unilateral and spatially constrained. This raises the question of whether the auditory stimulus was more immersive and thus perceptually salient. It

also seems possible that headphones incidentally delivered vibrotactile stimulation. In the current study, headphones were not used for auditory stimulation. We also considered the possibility that a larger contact area might improve tactile synchronization, perhaps to the extent that it would exceed auditory performance. For example, a spatial summation effect has been shown in vibrotactile perception, such that detection thresholds are lowered with presentation to a larger area of skin (Verrillo, 1992). To address this issue, tactile stimuli were presented bilaterally to the lower back, and the size of the contact area was manipulated.

In addition, a bimodal auditory-tactile condition was included to investigate multisensory integration. The modality appropriateness hypothesis argues that perception gives precedence to the sense most appropriate to a task. For example, audition tends to dominate temporal tasks while vision tends to dominate spatial tasks (e.g., Lukas, Philipp, & Koch, 2014). The maximum likelihood estimation (MLE) model (Ernst & Banks, 2002) quantifies this account, predicting that independent multisensory cues are optimally integrated by weighting them according to their reliability (Ernst & Banks, 2002). Thus, the MLE model predicts reduced variance in multimodal tasks relative to unimodal.

Elliot et al. (2010, 0 ms jitter condition) found an auditory-tactile bimodal advantage for the standard deviation of asynchronies consistent with the MLE model’s prediction. In that study, auditory and tactile cues were spatially separated (auditory = ears; tactile = fingertip stimulation) and physically distinct (auditory = vibratory; tactile = nonvibratory tap). The current bimodal condition involved an ecological situation, such as might occur when listening to loud music, in which cues were neither spatially separated nor physically distinct. We considered the possibility that our arrangement might violate the assumption of the MLE model of independence of the timing cues. For this reason, we derived predictions for optimal integration using the classical MLE model (Ernst & Banks, 2002) and a model that allows for correlated cues (Oruç, Maloney, & Landy, 2003).

Finally, we examined timing dependencies between ITIs. Certain patterns are commonly observed in the autocorrelation structure of synchronization trials that are thought to reflect specific underlying timing processes. For example, in the context of synchronization (as opposed to self-paced rhythmic tapping), negative Lag 1 autocorrelation (i.e., the tendency for a shorter ITI to follow a longer one and vice versa) is thought to reflect an error correction process (Iversen et al., 2015). Autocorrelations at longer lags are typically zero but may arise from slow drift in ITI (positive at lags greater than 1) or from the imposition of a metrical framework. For example, explicit duple meter tapping has been shown to elicit a negative correlation at Lag 3 (Vorberg & Hombuch, 1978).

Interestingly, in Iversen et al.’s (2015) study (see Note 1), although the standard deviation of asynchronies did not

significantly differ between tapping to an auditory metronome and a (silent) moving image of a bouncing ball, a significantly negative Lag 3 autocorrelation only emerged in the former. This finding suggests that, despite equivalent performance across modalities, only auditory synchronization made use of higher order grouping and/or error correction, even in the absence of instructions to do so. We aimed to determine whether such higher order timing patterns were also present in the tactile domain. The absence of these patterns in tactile performance would further support a unique auditory-motor connection in rhythmic processing.

## Method

### Participants

Twenty-one undergraduate students from Ryerson University (16 females; mean age = 21.52 years,  $SD = 5.27$ ) participated for course credit. Number of years of private music lessons ranged from 0 to 10 years ( $M = 1.33$ ,  $SD = 2.37$ ); none of the participants were professional musicians. Three participants (two female) did not complete the study because of equipment failure.<sup>2</sup>

### Apparatus

All stimuli were presented through voice coils embedded in a form-fitting chair (Emoti-Chair; Karam, Russo, & Fels, 2009). The voice coils were one inch in diameter and arranged in two columns of eight so that they made contact with the left and right sides of the lumbar region of the back (see Fig. 1). When driven by an auditory signal, the voice coils produce felt vibration and suprathreshold sound. Participants' finger taps were recorded from a Roland HPD 10 MIDI drum pad.

### Stimuli

Metronome tones were 40-ms sine waves (10 ms rise and decay) presented at 196 Hz. Tones were organized into simple and complex rhythm trials, as illustrated in Fig. 2. Trials of both type began with an induction phase consisting of eight isochronous metronome events with an interbeat interval (IBI) of 800 ms. Simple trials contained an additional 60 events.

Complex trials contained 15 rhythmic patterns used in previous studies: these were the “strongly metrical” patterns in Patel et al. (2005; cf. Brochard et al., 2008, Fig. 2, for a visual representation), based on Povel and Essens (1985). Each pattern contained four beat-level events (again presented with an IBI of 800 ms) with additional events at quarter, half, and three-quarter subdivisions of the beat (i.e., interonset intervals of 200, 400, and 600 ms, respectively). For all complex patterns, the first three beats were subdivided and the final beat was not. Complex trials were constructed by concatenating the 15 patterns in random order (several different randomizations were created). Thus, simple and complex trials each contained 68 beat-level events (eight induction plus 60 postinduction). There was no variability in magnitude between events in either the auditory or tactile modality.

### Procedure

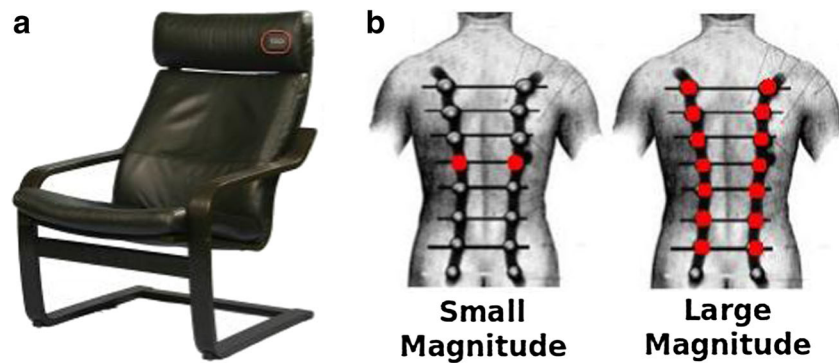
Simple and complex trials were presented under auditory, tactile, and auditory-tactile conditions. In the tactile and auditory-tactile conditions, participants sat in the Emoti-Chair. In the tactile condition, sound emitted by the voice coils was masked by loud white noise presented over headphones at a fixed level (~85 dBA). Potential bone conduction of sound was masked using white noise presented by vibrotactile stimulation using Tactaid VBW32 transducers placed on the left and right mastoids (Russo, Ammirante, & Fels, 2012). In the auditory condition, participants sat facing the empty Emoti-Chair in another chair of similar construction at a distance of ~100 cm from seat to seat. The number of active voice coils was also manipulated to yield two stimulus magnitudes (small = 2 voice coils; large = 14 coils). For both the auditory and auditory-tactile conditions, this amounted to differences in sound level of ~3 dB.

Participants were instructed to tap along with beat-level metronome events but not their subdivisions. After a brief practice period, participants completed 24 trials: 3 modalities (auditory, tactile, auditory-tactile) × 2 stimulus magnitudes (small, large) × 2 rhythmic complexities (simple, complex) × 2 repetitions. Trials were blocked by modality. Trials within a block and the order of block presentation were counterbalanced. For each trial, a custom Csound script (<http://www.csounds.com/>) presented stimuli and recorded the timing of responses.

## Results

Timing of the postinduction phase responses were analyzed. Rare instances where the intertap interval (ITI) exceeded 1,300 ms (0.36 % of all responses) were likely due to unregistered taps and were deleted from the time series.

<sup>2</sup> Although a larger sample size would have been preferable, we note that small sample sizes are not unusual in tapping studies, which can be data intensive. Our final sample size ( $n = 18$ ) was larger than all but 1 of the 13 tapping studies in our reference list that were published since 2000 (mean  $N = 12.6$ ). Two studies in our list investigated modality effects in tapping and report effect sizes (Gan et al., 2015; Iversen et al., 2015). If we take the average of the effect sizes that they report for a main effect of modality on tapping accuracy, .05 significance level, and .8 power, then a sample size of 20 is recommended.



**Fig. 1** a. Emoti-Chair. b. Approximate contact areas on the back of the two rows of voice coils embedded in the chair. Active voice coils for small magnitude (left panel) and large magnitude (right panel) conditions are shown in red. (Color figure online)

### Standard deviation of asynchronies

To assess timing precision, standard deviation (*SD*) of the asynchronies between finger taps and beat-level metronome tones were calculated for each trial. Each participant's best trial for each condition (i.e., the trial with lowest *SD*) was identified. Figure 3 shows data from these trials averaged across participants.

Our expectation that tactile performance should improve with a larger contact area size was supported. Standard deviation of asynchronies collapsed across simple and complex conditions was significantly lower for large magnitude stimuli ( $M = 94$  ms) than small ( $M = 111$  ms),  $t(17) = 2.22$ ,  $p = .04$ . However, despite a concomitant increase in intensity, there was no significant performance gain in either the auditory or auditory-tactile large magnitude conditions ( $p > .51$ ). A repeated-measures ANOVA (3 modality  $\times$  2 stimulus magnitude  $\times$  2 rhythmic complexities) confirmed that this Modality  $\times$  Stimulus Magnitude interaction was significant,  $F(2, 34) = 4.2$ ,  $p = .02$ .

We partially replicated previous findings of no auditory over tactile advantage (Brochard et al., 2008; Muller et al., 2008). In the simple rhythm condition (see the left panel of Fig. 3), tactile performance ( $M = 70$  ms) did not significantly differ from either auditory ( $M = 64$  ms) or auditory-tactile ( $M = 63$  ms,  $p > .31$ ). However, in the complex condition (right panel of Fig. 3), tactile performance ( $M = 134$  ms) was considerably poorer than both auditory ( $M = 86$  ms,  $p = .02$ ) and auditory-tactile ( $M = 86$  ms,  $p = .02$ ). The ANOVA confirmed this: significant main effects of Modality,  $F(2, 34) = 7.08$ ,  $p = .003$ , and complexity  $F(1, 17) = 15.99$ ,  $p = .001$ , were qualified by a significant Modality  $\times$  Complexity interaction,  $F(2, 34) = 3.76$ ,  $p = .03$ . Interestingly, no such interaction was reported in Aronow et al. (2015), who found an auditory advantage for both simple and complex stimuli using the same dependent measure, based on stimulating one fingertip in the tactile condition. The current findings, when considered alongside those of Aronow et al. (2015), suggest that simple but not complex tactile synchronization improves with the

salience of tactile stimulation. There were no other significant main effects or interactions.

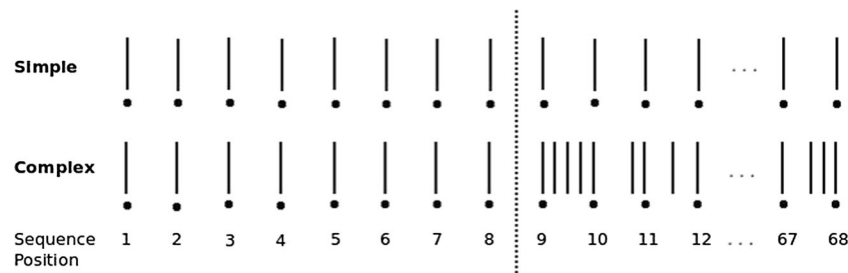
### Multisensory integration

Although there were no significant differences in standard deviation of asynchronies between auditory and bimodal conditions, this need not rule out multisensory effects, which may have been small and/or masked by between-subjects variability. To test whether multisensory cues were optimally combined according to MLE, we used unimodal performance to predict bimodal performance.

Following Elliot et al. (2010), we assumed that unimodal performance included (1) modality-dependent sources of variance related to the sensory registration of events (auditory =  $\sigma_{MDA}^2$ ; tactile =  $\sigma_{MDT}^2$ ) and sensorimotor coupling and (2) modality-independent sources of variance ( $\sigma_{MI}^2$ ). The latter may include noise originating from an internal timekeeper and noise associated with the motor implementation of taps (Vorberg & Wing, 1996). An empirical estimate of  $\sigma_{MI}^2 = 1,024$  ms<sup>2</sup> was obtained from a previous study (Zelaznik & Rosenbaum, 2010).<sup>3</sup> Modality-dependent variance estimates were obtained by subtracting  $\sigma_{MI}^2$  from total variance (averaged across participants) in the unimodal auditory condition ( $\sigma_{MDA}^2$ ) and from the unimodal tactile condition ( $\sigma_{MDT}^2$ ), respectively. An estimate of integration according to MLE, in which cues are optimally weighted according to their reliability, was obtained as follows (Elliot et al., 2010; Ernst & Banks, 2002):

$$\sigma_{MLE}^2 = \sigma_{MI}^2 + \frac{\sigma_{MDA}^2 \sigma_{MDT}^2}{\sigma_{MDA}^2 + \sigma_{MDT}^2} \quad (1)$$

<sup>3</sup> Zelaznik and Rosenbaum (2010, Fig. 2, tapping without auditory feedback condition) also tested undergraduate students. For self-paced tapping trials following the removal of an 800-ms IOI metronome, they report a coefficient of variation (*SD*/mean) of 4%. With an ITI of ~800 ms (2010, Fig. 1), this is equivalent to  $\sigma^2$  of 1,024 ms<sup>2</sup> or *SD* of 32 ms.



**Fig. 2** Schematics of simple and complex rhythmic sequences. The solid vertical lines indicate event onsets and the dots indicate beat locations. The dotted line marks the end of the induction phase at beat 8. Following

this, the complex condition contained subdivisions of the beat. The complex sequence shown is one of several different sequences used

Variance from (1) was converted to standard deviation by taking the square root. MLE estimates were compared with obtained standard deviation values from the bimodal conditions. We report findings collapsed across stimulus magnitude, which yielded qualitatively similar results to separate analyses on this factor. As shown in Fig. 4, obtained standard deviation was higher than MLE estimates for both simple and complex rhythms, suggesting suboptimal integration. This contrasts with Elliot et al.'s (2010, 0 ms jitter condition) finding of a good fit of the MLE model to auditory-tactile performance.

We considered the possibility that optimal integration was taking place but that the MLE model's assumption of independence of timing cues was violated. Where processing noise associated with the two cues is highly correlated, there is less gained from the combination of cues, and variance should be higher than what is predicted by the MLE model (Wing, Dumas, & Welchmann, 2010). In the current study, auditory and tactile cues were correlated insofar as they were neither spatially separated nor physically distinct, and thus may have activated a common source of variance.

To test this possibility, the same estimates of  $\sigma_{MDA}^2$ ,  $\sigma_{MDT}^2$ , and  $\sigma_{MI}^2$  were applied to the following model, which allows for correlated cues (Oruç et al., 2003):

$$\sigma_{MLE}^2 = \sigma_{MI}^2 + \frac{\frac{1}{\sigma_{MDA}^2} + \frac{1}{\sigma_{MDT}^2} - 2\rho\sqrt{\frac{1}{\sigma_{MDA}^2}\frac{1}{\sigma_{MDT}^2}}}{1-\rho^2} \quad (2)$$

where  $\rho$  is the correlational coefficient between  $-1$  and  $1$  that minimizes the squared difference between obtained and estimated variance. As shown in Fig. 4, this alternative model closely approximated the obtained data. Moreover, consistent with our prediction that bimodal performance involved the optimal integration of correlated cues,  $\rho$  was .73 and .61 for simple and complex rhythms, respectively.

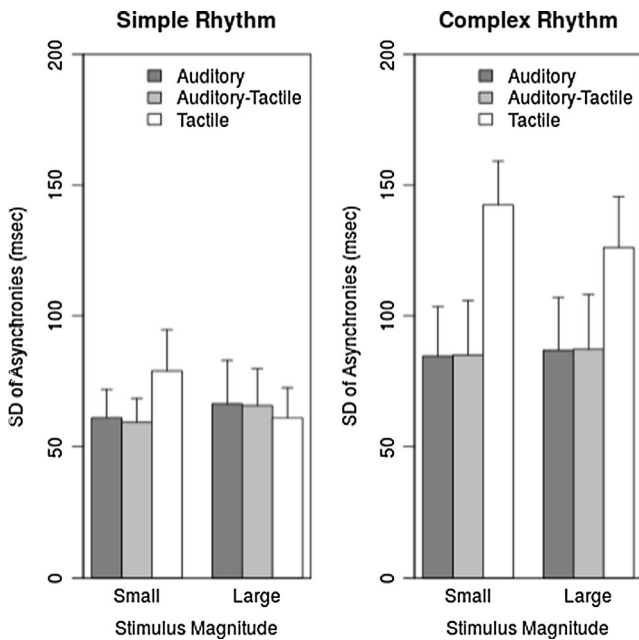
### Autocorrelations at Lags 1–4

To assess timing dependencies between intertap intervals (ITI), we examined the autocorrelation structure of each trial's

time series (see Iversen et al., 2015, Section 2.4 for a discussion of this type of analysis of tapping data in relation to models of timekeeping). For each participant and each condition, autocorrelation coefficients were calculated at Lags 1 to 4 for the best trial as identified by the standard deviation of asynchronies. Figure 5 shows these coefficients averaged across participants. We report two separate ANOVAs: Lag 1 autocorrelations and higher order autocorrelations (Lags 2–4). This is a convention used in other studies (see, e.g., Iversen et al., 2015) and is theoretically motivated—the negative autocorrelations found at Lag 1 are thought to reflect a different timing process than at higher order lags.

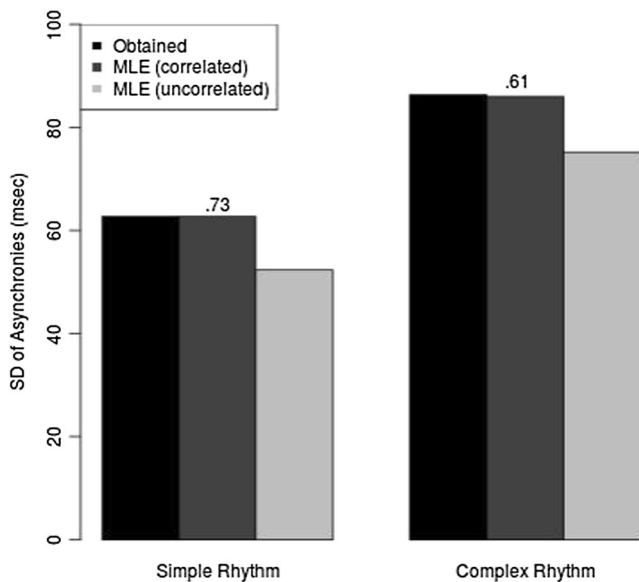
A common observation in synchronized tapping with an auditory metronome is a negative autocorrelation at Lag 1, thought to reflect an error correction process (Repp, 2005). As shown in Fig. 4, mean Lag 1 autocorrelation coefficients were significantly negative for all conditions except tactile-complex. A repeated-measures ANOVA on Lag 1 autocorrelation coefficients confirmed this interaction. Similar to standard deviation of asynchronies, a significant main effect of complexity,  $F(1, 17) = 15.99, p = .001$ , and a marginal main effect of modality,  $F(2, 34) = 3.26, p = .051$ , were qualified by a significant Complexity  $\times$  Modality interaction,  $F(2, 34) = 3.76, p = .03$ .

At higher order lags (i.e., lags greater than 1), coefficients at or near zero are often observed. This was the case across modalities in the simple rhythm condition but not in the complex one. First, the significantly positive coefficients across modalities at Lag 4 (bottom two panels in Fig. 5) were likely cued by the four-beat grouping structure of the rhythmic patterns (recall that the fourth beat in each pattern was never subdivided while the other three beats always were; cf. Vorberg & Hambuch, 1978). Thus, it appears that participants were sensitive to exogenous rhythmic cues to metrical structure in auditory and tactile domains. A repeated-measures ANOVA on the coefficients at higher order Lags 2–4 supported this with a Complexity  $\times$  Lag interaction,  $F(2, 34) = 8.47, p = .001$ . Whereas no difference was found between lags in the simple rhythm condition ( $p = .52$ ), in the complex condition, coefficients were significantly more positive at Lag 4 ( $M = .11$ ) than Lag 2 ( $M = .01, p <$



**Fig. 3** Standard deviation of asynchronies for simple (*left panel*) and complex (*right panel*) rhythmic sequences

.001) and Lag 3 ( $M = -.02, p < .001$ ), but did not differ between Lags 2 and 3 ( $p = .18$ ). On the other hand, in the complex-small magnitude condition (bottom left panel in Fig. 5), tactile coefficients were positive at *all* lags, suggesting slow drift (i.e., tendency to speed up or slow down over the course of a trial). Thus, it is possible that, with less



**Fig. 4** A comparison of obtained SD of asynchronies in the bimodal condition with estimates based on the classical MLE model—MLE (uncorrelated), Ernst and Banks (2002)—which assumes independence of timing cues, and an alternative model that allows for correlated cues—MLE (correlated), Oruç et al. (2003). Correlational coefficients for best fits of the latter model are indicated above the bars

salient tactile stimulation, participants were insensitive to rhythmic structure at both beat and metrical levels.

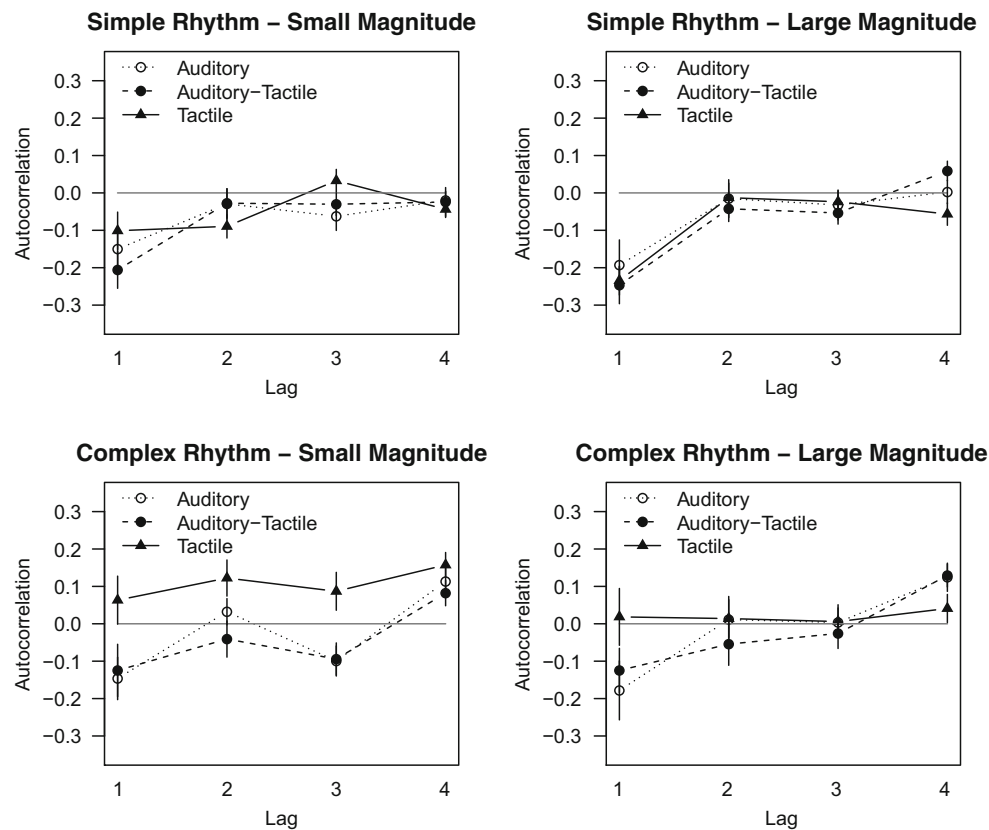
Finally, consistent with Iversen et al.’s (2015) previous finding, we found evidence for an auditory-specific negative dependency at Lag 3 in a Modality  $\times$  Lag interaction,  $F(4, 68) = 3.03, p = .02$ . At Lag 3, auditory ( $M = -.05$ ) and auditory-tactile ( $M = -.05$ ) coefficients were each significantly more negative than tactile ( $M = .03, p < .03$ ) but did not differ from each other. There were no modality differences at Lags 2 or 4 ( $p > .3$ ). Following Vorberg and Hambuch (1978), Iversen et al. (2015) proposed that negative Lag 3 dependency reflects the imposition of duple meter. The fact that Lag 3 dependency was only apparent with the availability of auditory information suggests that higher order beat perception may be more readily available in the auditory domain. Moreover, as shown in Fig. 5, Lag 3 coefficients only significantly different from zero with complex rhythm small magnitude sequences (auditory  $M = -.1, p = .02$ ; auditory-tactile  $M = -.09, p = .04$ ). This suggests that participants used a grouping strategy under challenging conditions (i.e., when auditory beat information was less salient).

**Discussion**

The current data suggest that if tactile information is made salient enough, the accuracy of synchronization to a tactile metronome can equal synchronization to an auditory metronome, further challenging the idea of an auditory advantage over all other modalities for synchronization to discretely timed rhythmic stimuli. With a tactile metronome presented over a large area of the back (instead of just to the fingertip or toe, as in previous studies), the accuracy of tactile synchronization with a moderately paced metronome equaled auditory performance, at least for simple rhythmic sequences.<sup>4</sup> This raises the possibility that if even more of the body’s surface were stimulated, synchronization to a tactile metronome could be significantly better than to an auditory metronome.

While this aspect of our finding shows some limitations to the hypothesis of enhanced auditory-motor coupling for rhythmic processing (Patel et al., 2005), other aspects clearly support the hypothesis. First, we observed a robust advantage with auditory information for synchronization with complex rhythm sequences, replicating the finding from Aronow et al. (2015). Further analysis showed the absence of error correction based on the previous tap period in complex tactile sequences (as delineated by its ITI and indexed by negative Lag

<sup>4</sup> Future studies might also manipulate metronome rate. For example, Repp (2003) found the fastest rate limit for synchronization was much lower for auditory beeps than flashing lights, suggesting different limitations in temporal processing. It remains unclear whether such differences exist between auditory and tactile modalities.



**Fig. 5** Autocorrelation of time series ITI for simple (*top row*) and complex (*bottom row*) and small (*left column*) and large (*right column*) magnitude metronome sequences

1 autocorrelation). Second, consistent with previous findings (Iversen et al., 2015), we found evidence of higher order error correction (negative Lag 3 autocorrelation) only when auditory information was present, and in a new finding, particularly when it was less salient (complex rhythm, small magnitude condition). The current study contributes to a growing literature demonstrating that the auditory advantage in sensorimotor synchronization, first observed in comparisons between (auditory) metronome ticks and (visual) flashing lights (e.g., Bartlett & Bartlett, 1959; Chen et al., 2002; Dunlap, 1910; Jäncke et al., 2000; Patel et al., 2005; Repp & Penel, 2002, 2004), may be more nuanced than previously thought. For example, recent studies have shown that visual performance for simple rhythms improves to auditory levels when visual information is continuous rather than discrete (Gan et al., 2015; Iversen et al., 2015). Here, we show that, when made perceptually salient, tactile synchronization to simple rhythms can equal auditory synchronization. We also show that in a real-world multimodal setting in which auditory and tactile cues are neither spatially separated nor physically distinct, that the integration of these cues may involve a common source of sensory variance. One potential source may be auditory belt areas, which are shown to be activated by both auditory and vibrotactile rhythmic pulsed stimuli (see Occelli, Spence, & Zampini, 2011, for a review).

Our findings suggest that what appears to remain privileged to the auditory modality is a flexibility with respect to the hierarchical processing of rhythmic stimuli. Individuals retain the ability to track a sounded beat when divisions of the beat are present in the stimulus (complex rhythm condition) and to impose a higher order grouping even when it is not present in the stimulus (simple rhythm condition). The neural basis for this flexibility may lie in the connectivity between auditory cortex and motor regions, the latter of which are suggested to be richly coupled to auditory regions in beat perception even in the absence of overt movement (see Patel & Iversen, 2014, for a review). We speculate that this auditory-motor connectivity, including core areas of the auditory cortex, may be less extensive in the auditory belt areas activated in tactile perception.

This line of research, which contrasts synchronization to the beat of temporal patterns presented in different modalities, may have important implications for understanding the evolution of neural mechanisms involved in musical beat processing. In particular, our findings suggest the need to understand how mechanisms supporting synchronization to the beat in temporally complex sequences differ from those involved in simple 1:1 synchronization with a metronome. It has been suggested that the ability to flexibly and predictively synchronize with a beat has a relationship with the neural circuitry for vocal learning or vocal motor control (Fitch, 2015; Patel, 2006; Patel

2014). Perhaps this relationship only applies to beat processing in complex sequences that require hierarchical temporal processing (cf. Hove, Fairhurst, Kotz, & Keller, 2013).

Relevant to this issue, future studies might investigate whether beat tracking with complex tactile sequences (as assessed by synchronization) is experience dependent. For example, Iversen et al.'s (2015) recent auditory/visual study showed that deaf participants synchronized better than hearing individuals to a flashing light metronome. In fact, many members of the deaf community describe being engaged with strongly rhythmic music through felt vibration (Darrow, 1993; Good, Reed, & Russo, 2014). Thus, it would be of interest to test if tactile synchronization with complex sequences is better in deaf individuals than the general population. Similar effects might be expected in other individuals with regular exposure to felt musical rhythms, including instrumental musicians (Goebel & Palmer, 2008) and frequent attendees of loud concerts and/or dance parties.

**Acknowledgements** We thank John Iversen for helpful comments on an earlier draft and Domenica Fanelli for assistance with data collection.

## References

- Aronow, B., Austin, J., Iversen, J., & Patel, A. D. (2015, August). *Synchronization to auditory and tactile rhythms*. Poster presented at the biennial meeting of the Society for Music Perception and Cognition, Nashville, TN.
- Bartlett, N. R., & Bartlett, S. C. (1959). Synchronization of a motor response with an anticipated sensory event. *Psychological Review*, *66*, 203–218.
- von Békésy, G. (1959). Similarities between hearing and skin senses. *Psychological Review*, *66*, 1–22.
- Brochard, R., Touzalin, P., Despres, O., & Dufour, A. (2008). Evidence of beat perception via purely tactile stimulation. *Brain Research*, *1223*, 59–64.
- Caetano, G., & Jousmaki, V. (2006). Evidence of vibrotactile input to human auditory cortex. *NeuroImage*, *29*, 15–28.
- Chen, Y., Repp, B. H., & Patel, A. D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science*, *21*, 515–532.
- Darrow, A. A. (1993). The role of music in deaf culture: Implications for music educators. *Journal of Research in Music Education*, *41*, 93–110.
- Dunlap, K. (1910). Reactions to rhythmic stimuli, with attempt to synchronize. *Psychological Review*, *17*, 399–416.
- Elliott, M. T., Wing, A. M., & Welchman, A. E. (2010). Multisensory cues improve sensorimotor synchronisation. *European Journal of Neuroscience*, *31*, 1828–1835.
- Elliott, M. T., Wing, A. M., & Welchman, A. E. (2011). The effect of ageing on multisensory integration for the control of movement timing. *Experimental Brain Research*, *213*, 291–298.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.
- Fitch, W. T. (2015). The biology and evolution of musical rhythm: An update. In I. Toivonen, P. Csúri, & E. van der Zee (Eds.), *Structures in the mind: Essays on language, music, and cognition in honor of Ray Jackendoff* (pp. 293–324). Cambridge, MA: MIT Press.
- Gan, L., Huang, Y., Zhou, L., Qian, C., & Wu, X. (2015). Synchronization to a bouncing ball with a realistic motion trajectory. *Scientific Reports*, *5*, 11974.
- Gescheider, G. A., & Niblette, R. K. (1967). Cross-modality masking for touch and hearing. *Journal of Experimental Psychology*, *74*, 313–320.
- Giordano, M., & Wanderley, M. M. (2015). *Follow the tactile metronome: Vibrotactile stimulation for tempo synchronization in music performance*. Proceedings of the SMC Conference, Maynooth, Ireland.
- Goebel, W., & Palmer, C. (2008). Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research*, *186*(3), 471–479.
- Good, A., Reed, M. J., & Russo, F. A. (2014). Compensatory plasticity in the deaf brain: Effects on perception of music. *Brain Sciences*, *4*, 560–574.
- Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. E. (2013). Synchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness. *NeuroImage*, *67*, 313–321.
- Iversen, J. R., Patel, A. D., Nicodemus, B., & Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition*, *134*, 232–244.
- Jäncke, L., Loose, R., Lutz, K., Specht, K., & Shah, N. J. (2000). Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cognitive Brain Research*, *10*, 51–66.
- Karam, M., Russo, F. A., & Fels, D. I. (2009). Designing the model human cochlea: An ambient crossmodal audio-tactile display. *IEEE Trans Haptics*, *2*, 160–169.
- Kolers, P. A., & Brewster, J. M. (1985). Rhythms and responses. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 150–167.
- Lukas, S., Philipp, A. M., & Koch, I. (2014). Crossmodal attention switching: Auditory dominance in temporal discrimination tasks. *Acta Psychologica*, *153*, 139–146.
- Marks, L. E. (1979). Summation of vibrotactile intensity: An analog to auditory critical bands? *Sensory Processes*, *3*, 188–203.
- Müller, K., Aschersleben, G., Schmitz, F., Schnitzler, A., Freund, H.-J., & Prinz, W. (2008). Inter- versus intramodal integration in sensorimotor synchronization: A combined behavioral and magnetoencephalographic study. *Experimental Brain Research*, *185*(2), 309–318.
- Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in temporal perception. *Psychonomic Bulletin and Review*, *18*, 429–454.
- Oruç, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Research*, *43*, 2451–2468.
- Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, *24*(1), 99–104.
- Patel, A. D. (2014). The evolutionary biology of musical rhythm: Was Darwin wrong? *PLoS Biology*, *12*(3), e1001821.
- Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The action simulation for auditory prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*, 57.
- Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, *19*, 827–830.
- Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricity and modality on synchronization with a beat. *Experimental Brain Research*, *163*, 226–238.
- Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*(4), 411–440.
- Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, *35*(4), 355–170.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969–992.



- Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1085–1099.
- Repp, B. H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, *68*, 252–270.
- Russo, F. A., Ammirante, P., & Fels, D. (2012). Vibrotactile discrimination of musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 822–826.
- Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, *19*, 831–836.
- Verrillo, R. (1992). Vibration sensation in humans. *Music Perception*, *9*, 281–302.
- Verrillo, R., & Gescheider, G. A. (1975). Enhancement and summation in the perception of two successive vibrotactile stimuli. *Perception & Psychophysics*, *18*(2), 128–136.
- Vorberg, D., & Hambuch, R. (1978). On the temporal control of rhythmic performance. In J. Requin (Ed.), *Attention and performance VII* (pp. 535–555). Hillsdale, NJ: Erlbaum.
- Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action* (Motor skills, Vol. 2, pp. 181–262). London, England: Academic Press.
- Wing, A. M., Dumas, M., & Welchman, A. E. (2010). Combining multisensory temporal information for movement synchronisation. *Experimental Brain Research*, *200*(3–4), 277–282.
- Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, *102*, 3191–3202.
- Zelaznik, H. N., & Rosenbaum, D. A. (2010). Timing processes are correlated when they share a salient event. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1565–1575.