OXFORD

Article

Effects of metronomic sounds on a self-paced tapping task in budgerigars and humans

Yoshimasa SEKI* and Kenta TOMYTA

Department of Psychology, Aichi University, Machihatacho, Toyohashi, Aichi, Japan

*Address correspondence to Yoshimasa Seki. E-mail: yoshimasa.seki@gmail.com

Abstract

The origin of rhythmic synchronization or entrainment to a musical beat in animals has been widely discussed. Parrots are suitable animals to examine the relationship between the capability of vocal learning and spontaneous rhythmic synchronization. In this study, budgerigars *Melopsittacus undulatus* learned to tap (peck) 2 keys alternately at a self-paced rate. Then, the metronomic sounds were played in the background during test sessions while the birds were performing the key pecking task, although they were not required to synchronize tap timing with the metronome. We found modest but significant effects of the metronome rhythms on the tap timing in some subjects. We also tested humans *Homo sapiens* using almost the same method. In contrast to the birds, a number of human subjects synchronized tap timing to the onset of the metronome without verbal or documented instructions. However, we failed to find an effect of the metronome on self-paced tap timing in some human subjects, although they were capable of rhythmic synchronization. This is the first report describing the effects of metronomic sounds on self-paced tapping in nonhuman vocal learners. This study introduces a new method that can be used in future research comparing birds that differ in vocal learning capacities, social structure, age, sex, hormonal status, and so on as part of examinations of the evolutionary foundations of beat processing.

Key words: operant conditioning, parrots, rhythm, synchronization, vocal learning

Occasionally, humans listening to certain rhythms, such as marching music, will spontaneously synchronize their movements with those rhythms. Many animal species innately control the timing of specific behaviors with the behavior of other individuals; for example, in finches (during calling, e.g., Benichov et al. 2016), frogs (during calling), and even invertebrates such as fireflies (during lighting) and crabs (during waving) during courtship displays (Ravignani 2015). However, the time range and type of motor control are not likely to be extended beyond those innate behaviors. Therefore, a report (Patel et al. 2009) of a dancing parrot that can synchronize body movements with various musical beats, which supports the "vocal learning and rhythmic synchronization hypothesis" (Patel 2006), had a strong impact on researchers. Schachner et al. (2009) also supported this hypothesis. Subsequently, Hasegawa et al. (2011)

examined the capacity for rhythmic synchronization in budgerigars, another parrot species, under an operant conditioning experiment, which also supported the hypothesis.

However, there is a counterargument to the hypothesis. Cook et al. (2013) reported that a sea lion, a nonvocal learning species, could synchronize body movements with various musical beats. In connection with this question, Hattori et al. (2013) reported that a chimpanzee, which is also a nonvocal learning species, tapped synchronously with an isochronous sound sequence at 600-ms intervals; however, the chimpanzee did not demonstrate tempo-flexibility. Similarly, Large and Gray (2015) detected spontaneous entrainment and synchronization to drum strikes, within a certain range around a spontaneous motor tempo, in a bonobo (watching a human drummer striking). These primate studies suggested that social interaction

is a key factor in the origin of rhythmic synchronization. Thus, further comparative studies are necessary to advance discussion of spontaneous synchronization to purely auditory rhythms.

The capability of budgerigars for rhythmic synchronization was initially documented by a study (Hasegawa et al. 2011) in which audiovisual metronomes were presented to birds. The budgerigars were required to peck a key (or "make a hit") within a certain time window around the onset of each stimulus; a hit sequence rather than a single hit was required for a reward. Although the birds also could simply peck the key in response to each stimulus, the analysis of the results revealed that the birds did not do so. Instead, they anticipated the timing of the incoming stimulus and synchronized the tap timing with the onset of the metronome, even though it was not a necessary condition for the reward. This finding was interesting; however, it is unclear whether the birds need explicit synchronization training or whether synchronization can be spontaneous.

Therefore, in this study, we attempted to investigate the effects of an isochronous sound sequence presented to budgerigars as an "attractor" (as opposed to "distractor") for self-paced tapping under controlled experimental conditions. A further purpose of this study was to investigate spontaneous or nonvolitional synchronization by vocal learning animals to a rhythmic sound sequence. As the budgerigar is a species capable of vocal learning throughout its lifetime, it has been widely used in laboratory experiments investigating the capability for vocal learning. In addition, as this small species is relatively easier to train using operant conditioning apparatus than larger parrot species, they were suitable subjects for this study. We next compared the behavior of budgerigars with that of humans to elucidate the spontaneity of the synchronization.

Budgerigars were trained to tap 2 keys (left and right) alternately using an operant conditioning method with food reinforcements. One of the keys was illuminated to prompt the subject to tap. When the subject tapped the key, the illumination disappeared, and the other key illuminated immediately. The keys were illuminated until the bird tapped them. Eight taps were required to obtain a food reward. We did not present any other stimuli, and the bird was not specifically trained to adjust its response timing, so that the tap intervals were completely dependent on the bird; this ensured that we obtained genuine self-paced tapping. Then, an isochronous sound sequence, in which the inter-stimulus onset-interval (IOI) corresponded to the median of the inter-tap-intervals (ITIs) and the variations therein (the IOIs spanned −10%, +10%, and +20% of the original sound sequence), were presented in the background of the task. This custom-made metronome facilitated the stimuli to act as an attractor to a greater degree than ready-made (i.e., potentially hard to follow) stimuli. To compare the capability with that of humans, an almost identical experiment was performed in human subjects (Ss). The Ss were instructed via a paper document to tap 2 pads alternately at their own pace. Stimulus sound sequences were created as in the bird experiment.

Our method was similar to, but differed slightly from, Repp (2006), in which Ss were instructed to tap a single pad with a certain metronome rhythm. Then, they were asked to maintain the tap timing intervals without the metronome. Metronomes with various intervals were presented during the task to determine the effects on tap timing. Thus, the stimuli in their study acted as a distractor. However, in this study, we did not require anything from the budgerigars and humans other than alternate tapping, meaning that it was unnecessary for the subjects to follow any rules. Thus, the sound sequences could be an attractor. Of course, if the subjects were to maintain their self-pacing and the timing were out of sync with the stimuli, then the auditory stimuli would be a distractor. Another possibility is that the subjects may keep tapping and ignore the playback sounds altogether.

Our study also paralleled a study done in chimpanzees and humans (Hattori et al. 2015) using 2 keys on a keyboard instrument. The experiment presented certain ready-made variations of metronomic sound to determine the effects of the distractor on self-paced tapping. The original aim of this study was to observe whether the budgerigars and humans would spontaneously synchronize tap timing with the onset of a metronome (i.e., to determine whether the metronome would work as an attractor); however, if the metronome could affect the tap timing and if the birds did not follow the stimuli, the metronome could be understood as a "distractor," which would be similar to the study performed by Hattori et al. (2015).

In summary, our research question was whether the budgerigars, as a vocal learning species, would spontaneously synchronize to a metronomic rhythm played in the background. Thus, we hypothesized that (1) the timing of the birds' tapping would center around (or slightly faster than) the sound onset if budgerigars had a strong tendency to spontaneously entrain to a rhythmic sound sequence. Even if this were not the case, (2) a faster IOI would lead to a shorter ITI and, likewise, a slower IOI would lead to a longer ITI in these birds.

Materials and Methods

Bird experiment
Subjects
Five budgerigars (∼5 years old) were used. The birds were kept in an animal-rearing room, maintained at 25°C under a 12L:12 D photoperiod at Aichi University, Aichi Prefecture, Japan. During the experimental period, food volume was kept constant so that the body weight of the subjects was maintained at around 85% of the free-feeding condition. All experimental procedures and the housing conditions were approved by the Animal Experiments Committee of Aichi University (approval no. 201501), and all experiments were performed in accordance with the Fundamental Guidelines for Proper Conduct of Animal Experiments and Related Activities in Academic Research Institutions under the jurisdiction of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Apparatus
A wooden cage was placed in a sound-attenuated box. Two response keys (light-emitting diodes (LEDs)) were located on a side wall of the cage. The distance between the keys was 2.5 cm. The tap signals generated by a piezo sensor attached on each key were digitized and sent to a PC via an I/O device (DIO-1616RYX-USB; Contec, Osaka, Japan). A grain feeder, which was controlled with the I/O device, delivered millet seed into a cup as a food reward. A small loudspeaker was placed above the keys to play back the auditory stimuli (stereo Windows PCM files; See "Stimuli" section, below) which were played back from the left channel of the stereo output from another independent PC. A trigger, a short rectangle wave that was used to signal the onset of each metronomic sound to a multichannel recorder (sampling rates of 1000 Hz, 1401micro; CED, Cambridge, UK), was outputted via the right channel of the stereo output of the PC. The tap signals were also recorded independently with the recorder (sampling rate of 10 kHz) so that we could use it to calculate the difference between the timing of the taps and the sound stimuli from the records (Supplementary Figure S1).

Figure 1. A bird tapping a left key and a right key alternately during a training session without auditory stimuli. Also, see Supplementary data, Movies S1–S4.

Stimuli

Sound stimuli were metronomic sounds (3 kHz pure tone, 44.1 kHz, 50 ms in duration each) that were played back at 64 dB at the level of the bird's head.

Procedures

The birds learned to tap an illuminated key to obtain a reward during their initial training. Only one of the 2 keys was illuminated at a given time. Then, when a subject tapped the illuminated key, the illumination was terminated immediately, and the other key was illuminated. The system ignored taps on a nonilluminated key; thus, the birds tapped 2 keys alternately (Figure 1). The number of taps needed for a reward was increased from 1 to 8 in the initial training period (e.g., one tap were required for the first 50 rewards, then, 2 taps were required for the next 50 rewards, and so on); thus, we finally obtained 400 taps (8 taps × 50 trials) on the illuminated keys during each regular training session. The median tap interval (except between the first and second taps) was calculated for each session. The birds completed several sessions with metronomic sounds in the middle of the training period, which familiarized them with the playback of auditory stimuli during the task. Then, the test sessions began when the experimenter judged that the variation in the median tap intervals had been stabilized. For the test sessions, we used 5 types of custom-made sound stimuli for each subject; standard (hereafter, STANDARD); the Inter Simulus onset Interval (ISI) corresponded to the shortest median ITI obtained during the training sessions without metronome sounds for each bird in the final phase of the training period), 10% faster (FAST), 10% and 20% slower (SLOW − 1 and SLOW − 2) than STANDARD, and an irregular (RAND) metronome, in which the intervals appeared randomly at 100, 200, 300, or 400 ms. All metronomic sounds were prepared as 30-min audio files (.wav Windows PCM files, 44.1 kHz) with the MATLAB program (MathWorks, Natick, MA, USA). Therefore, in total, there were 7 test conditions, STANDARD, SLOW − 1, SLOW − 2, FAST, without sound, RAND, and second STANDARD (which helped us confirm the stability of the tap intervals), presented in this order for all subjects. The playback of the sound stimuli resumed when the LED illumination signaled the beginning of each trial, and the playback paused when the bird was rewarded. Each condition was repeated twice, so the total number of taps for each condition was 800 (8 taps × 50 trials × 2 sessions).

Human experiment

Subjects

We used 13 subjects (Ss; age 20–24 years) who had normal hearing. All experimental procedures were approved by the Local Ethics Committee of Aichi University (approval no. 201602).

Apparatus

To obtain rhythmic signals generated by the Ss, we used 2 pads on which tapping areas were indicated by circles (each 12.5 cm in diameter; the distance between the centers of the circles was 18.0 cm). Ss tapped the pads alternately with a wooden stick (length, 18.0 cm). The tap signals from each pad were sent to a personal computer (hereafter, PC_A) via an electric drum (CLIPHIT; KORG, Tokyo, Japan) as hi-hat sounds (the mean latency between hit timing and audio output was 5.2 ms; SD 2.0 ms), which were recorded at a 44.1 kHz sampling rate and analyzed with the software Avisoft-SASLab Pro (Avisoft Bioacoustics, Berlin, Germany). The sound outputs from the 2 electric drums were combined into a single channel and recorded via the left channel of a stereo input from PC_A. The metronomic stimuli were played back from another computer (PC_B), and a stereo output channel was connected to the right channel of the stereo input of PC_A. This setting enabled us to record both tapping signals and metronomic stimuli in parallel, and simultaneous with PC_A. Another channel of the stereo output from PC_B was connected to a loud speaker (AT-SP151; Audio-Technica, Tokyo, Japan) located in front of the subjects at a distance of 50 cm to play back the metronomic stimuli. The sound amplitude at the Ss position was 62 dB, and the background noise was 37 dB. We confirmed that the subjects could hear the sound while completing the task.

Stimuli

We obtained the median of the self-paced tapping intervals of each subject with no sound stimuli (See *Procedures* section for details). Then, we created sound stimuli (STANDARD, FAST, SLOW – 1, and SLOW − 2) from the intervals, as in the bird experiment. These stimuli were used to investigate relevance to the task, because their relevance was initially unknown.

Procedures

The Ss were instructed to tap the pads in alternation at their own pace, even when sounds were played back from the loud speaker. The Ss were given as much time as they wanted to practice the task. Then, they performed alternate tapping 40 times (8 taps × 5 trials with 5-s pause between trials) with no sound stimuli. We recorded the tap signals, and the ITIs were analyzed with the *pulse train analysis* function of SASLab. We immediately calculated the median of the ITIs and created isochronous sound stimuli that had the median ITIs, and the same variations therein, as in the bird experiment. Then, the Ss proceeded to the test trials. The order of the test conditions was the same as in the bird experiment. We obtained 40 taps

under each test condition, as in the preceding no-stimulus condition. Playback of the sound sequence started at the beginning of the task and was maintained until end of the task.

Ten of the 13 subjects completed an additional experiment in which they were instructed to synchronize to the rhythm of the STANDARD metronome (hereafter, SYNC condition) after performing all other conditions to confirm their synchronization performance.

Analyses

We examined synchronicity between tap timing and sound onset of the metronome using a Rayleigh test with a single specified direction with eighteen 20-degree bins each (with Bonferroni correction) to ascertain whether the metronome worked as an attractor. Then, we generated a simple bootstrap ($n = 1,000$) of the data to obtain confidence intervals (CIs) for the mean direction μ, the 95% CI, and the concentration parameter k, which were calculated using the maximum likelihood estimates from a von Mises distribution (Ruxton 2017). In addition, we used a Rayleigh test with unspecified direction to determine whether the metronome worked as a distractor. We also examined the effects of the metronomic sounds on the ITIs among conditions by pairwise comparison using the Wilcoxon rank-sum test (with Bonferroni correction). All statistical tests were performed with R software (R Development Core Team, Vienna, Austria) and the *circular* package for the circular statistics.

Results

Bird experiment

When the experimenter judged that the tap interval curve of each bird was almost flat during the training session (Sessions 52–64; Supplementary Figure S2), the test sessions began. The intervals became increasingly short, with repeated up and down fluctuations. Because the interval showed such fluctuations and we had no prior information to determine the best time point to begin the test metronome sessions, those test sessions began when the experimenter postulated the curve of the tap interval stabilized. However, after the test sessions, we calculated the correlations between the median ITIs and the session numbers during the test sessions to confirm that the curve of ITIs no longer showed the downward trends (i.e., shortening of ITIs); the results showed no such trend, except for the bird G results (Supplementary Figure S3). The mean IOI of the metronome sounds under the STANDARD condition was 225 ms (range: 175–295 ms; SD = 44.7 ms).

A significant bias ($P < 0.05$ with Bonferroni correction, test statistics are shown in Supplementary Table S1) of the tap timing toward the specified directions was seen for bird A (STANDARD, 230–310 degrees; FAST, 310–10 degrees) and bird M (FAST, 270–330 degrees) according to the Rayleigh test (Figure 2). The test alone was not enough to show the home direction for the distribution; however, the μ values and the CIs calculated via the bootstrap method likely support the results of the Rayleigh test (STANDARD of bird A, $\mu = -45.3$ ms, CI $= -59.0$ to -31.1 ms; FAST of bird A, $\mu = -10.2$ ms, CI $= -24.3$ to 3.2 ms; $\mu = -29.2$ ms, CI $= -43.6$ to -14.2 ms; Figure 2). Moreover, the bootstrap method showed larger k values under STANDARD (0.22) and FAST (0.22) conditions than under SLOW – 1 (0.08) and – 2 (0.09) for bird A, and larger k values were seen under FAST (0.22) compared with STANDARD (0.05), SLOW – 1 (0.07), and SLOW – 2 (0.04) for bird M, which supports the results of the Rayleigh test. No bias appeared under the

SLOW conditions in these birds. This is consistent with the results of the Rayleigh tests with an unspecified direction (significant only in STANDARD and FAST conditions with bird A [test statistic = 0.109, $P < 0.001$; test statistic = 0.106, $P < 0.001$, respectively] and the FAST condition with bird M [test statistic = 0.112, $P < 0.001$]). Supplementary Figure S4 presents additional analyses of the data. No other birds showed biased tap timing under any of the other conditions.

Bird P was among the birds to show no significant tap timing bias; however, bird P showed significantly shorter ITIs under the FAST condition than under the STANDARD and SLOW conditions (W = 297,500, $P < 0.001$; Figure 3, left), suggesting that the faster metronome caused the bird to tap the keys faster. This was supported by the distribution of the peck-timing data of this bird, which reflected a peak shift of ITIs (Supplementary Figure S5, upper panel). In addition, bird P showed significantly longer ITIs under the SLOW – 2 condition than under the STANDARD condition (W = 219, 770, $P < 0.001$; Figure 3, left), suggesting that the slower metronome caused the bird to tap the keys more slowly. In contrast, bird A showed significantly longer ITIs under the FAST condition than under the STANDARD and SLOW conditions (W = 189,960, $P < 0.001$; Figure 3, right). The peak of the distribution of the peck-timing data of this bird under the FAST condition was almost the same as that under the STANDARD condition (Supplementary Figure S5, lower panel). More details are shown in Supplementary Information (Supplementary Figure S6).

The difference in the coefficient of variation (C.V.) of the ITIs of the last training session without sound stimuli just before the metronome session did not differ significantly from that of the STANDARD condition (Wilcoxon signed-rank test, V = 1, $P = 0.125$; Supplementary Figure S7).

Human experiment

The mean IOI of the metronome under the STANDARD condition was 537 ms (range: 295–720 ms; SD = 125.7 ms). There seemed to be individual differences in the effects of the sound stimuli; however, the Rayleigh test revealed that the tap timing of 6 Ss (#03, #06, #08, #10, #11, and #12; hereafter, Group 1) was significantly and spontaneously biased toward the stimulus onset of the metronome under all conditions (Figure 4; test statistics are presented in Supplementary Table S2). Tap timing by 3 Ss (#02, #07, and #09) showed a bias only under the STANDARD condition. S. #04 and #13 showed a bias only under the FAST condition, #01 showed a bias under the STANDARD and SLOW – 1 conditions, and #05 showed a bias under the STANDARD, SLOW – 1, and FAST conditions. The Rayleigh test revealed that all subjects exhibited a significant tap timing bias under at least one condition. All 10 Ss who underwent the SYNC test showed a significant bias under that condition. More examples are presented in the Supplementary Information (Supplementary Figure S8).

ITIs under the FAST condition were shorter than those under the STANDARD condition in Ss #04 (W = 941, $P < 0.001$, with Bonferroni correction), #05 (W = 885, $P < 0.05$), and all 6 Ss of Group 1 (#03, W = 965.5, $P < 0.001$; #06, W = 1186, $P < 0.001$; #08, W = 1048.5, $P < 0.001$; #10, W = 1108, $P < 0.001$; #11, W = 1164, $P < 0.001$; #12, W = 1200, $P < 0.001$; Figure 5, left), as well as for bird P. However, no Ss exhibited longer ITIs under the FAST condition, although this was observed in bird A. Significantly longer ITIs were observed under the SLOW – 1 and – 2 conditions than those under the STANDARD condition in #01 (W = 230.5, $P < 0.001$; W = 295, $P < 0.01$), #09 (W = 214, $P < 0.001$; W = 325,
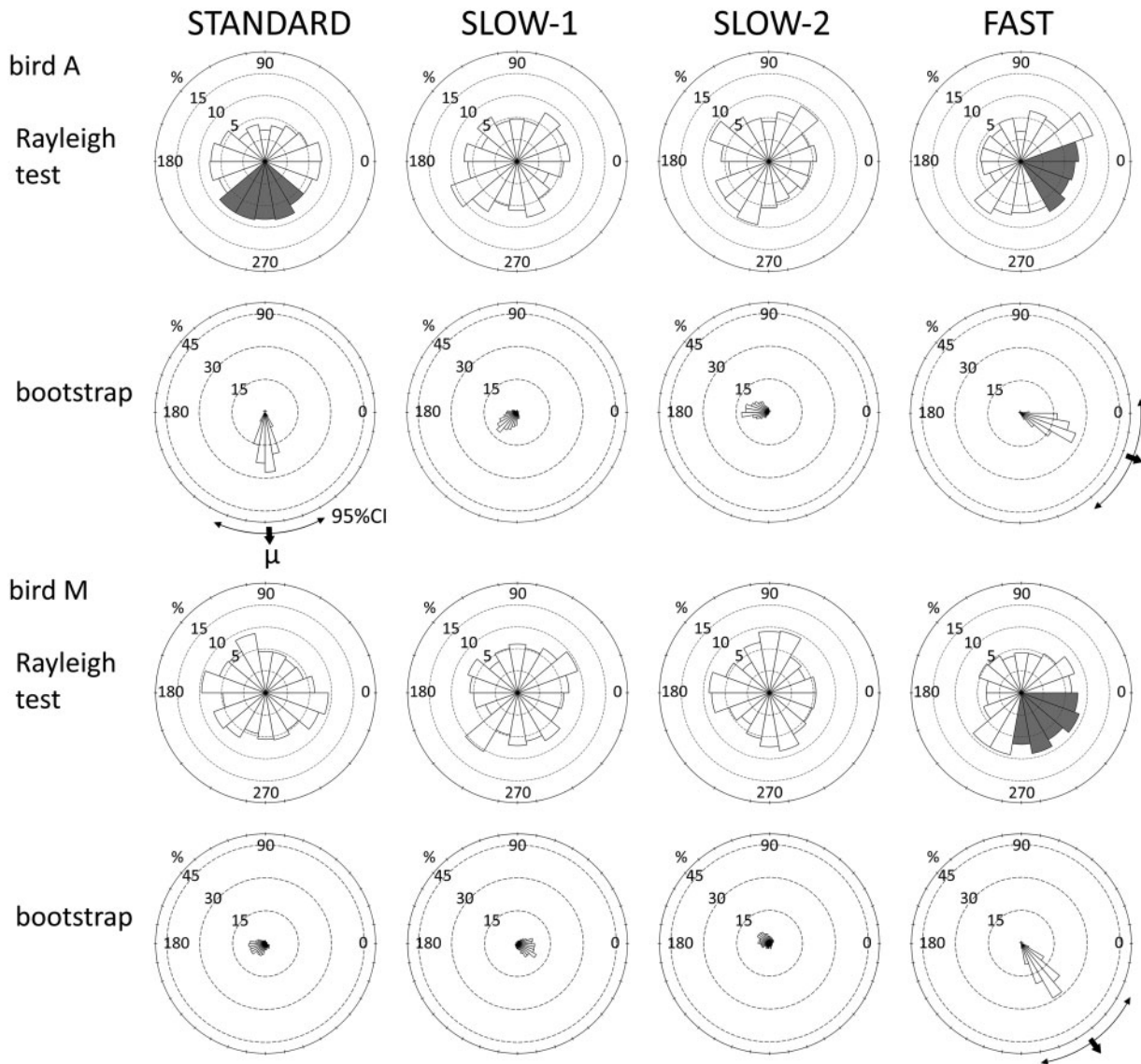
Figure 2. Distribution of tap timing in 2 birds: 0° indicates the onset of the metronome sound. Shaded areas indicate the time period showing significant bias according to the Rayleigh test ($P < 0.05$ with Bonferroni correction; test statistics are shown in Supplementary Table S1).
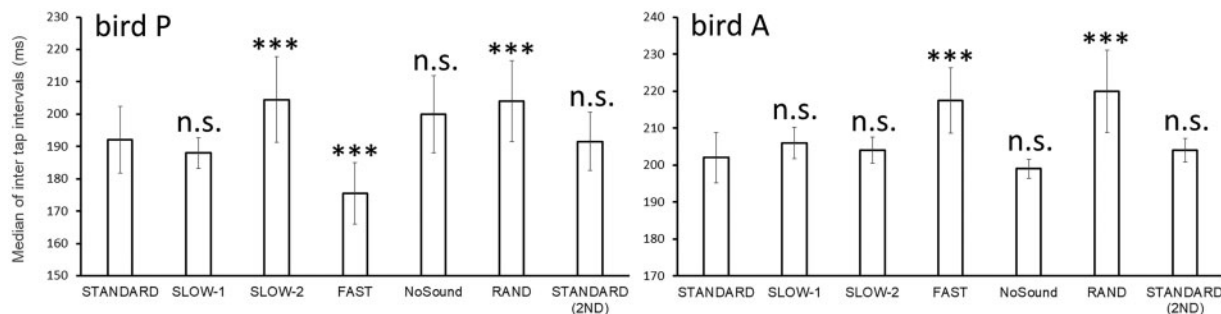

Figure 3. Median ITIs of birds P and A. Bird P showed shorter intervals under the FAST condition. Bird A had longer intervals under the FAST condition. Error bars indicate standard errors. ***$P < 0.001$ versus STANDARD.

$P < 0.01$) and 5 Ss of Group 1 (#03, $W = 75$, $P < 0.001$, $W = 1$, $P < 0.001$; #08, $W = 93$, $P < 0.001$, $W = 16$, $P < 0.001$; #10, $W = 58$, $P < 0.001$, $W = 0$, $P < 0.001$; #11, $W = 43$, $P < 0.001$, $W = 0$, $P < 0.001$; #12, $W = 28$, $P < 0.001$, $W = 0$, $P < 0.001$;

Figure 5, right). Significantly longer ITIs were observed in #13 ($W = 363$, $P < 0.05$) and the remaining subject in Group 1 (#06, $W = 0$, $P < 0.001$) under the SLOW $-$ 2 condition than those under the STANDARD condition.
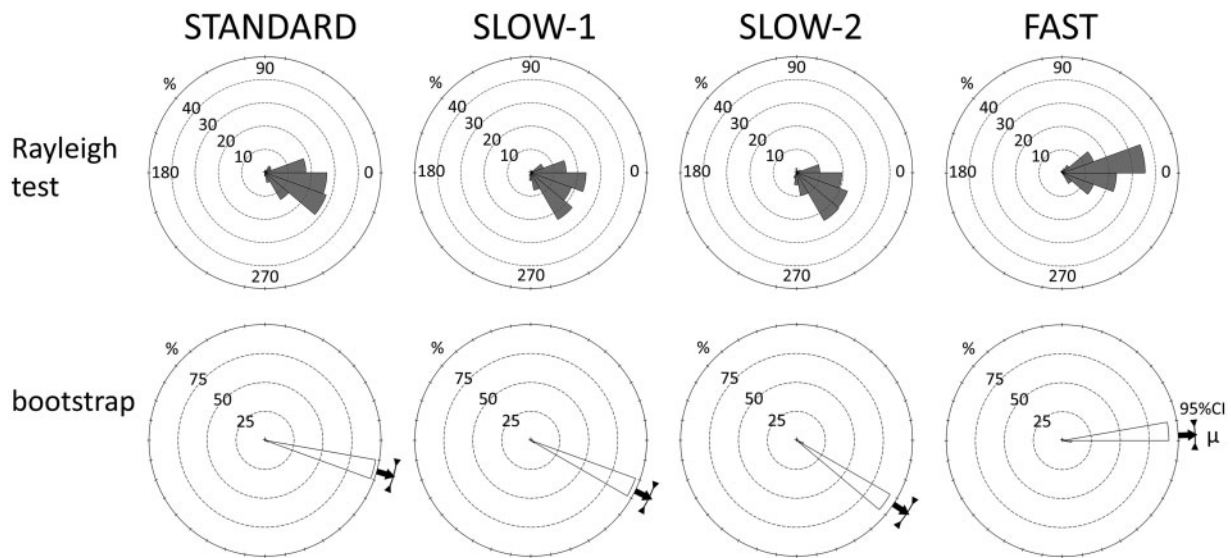
Figure 4. Distribution of tap timing: pooled Group 1 data of the human experiment. Here, 0° indicates the onset of the metronome. Shaded areas indicate the time period showing significant bias according to the Rayleigh test ($P < 0.05$ with Bonferroni correction; test statistics are presented in Supplementary Table S2).
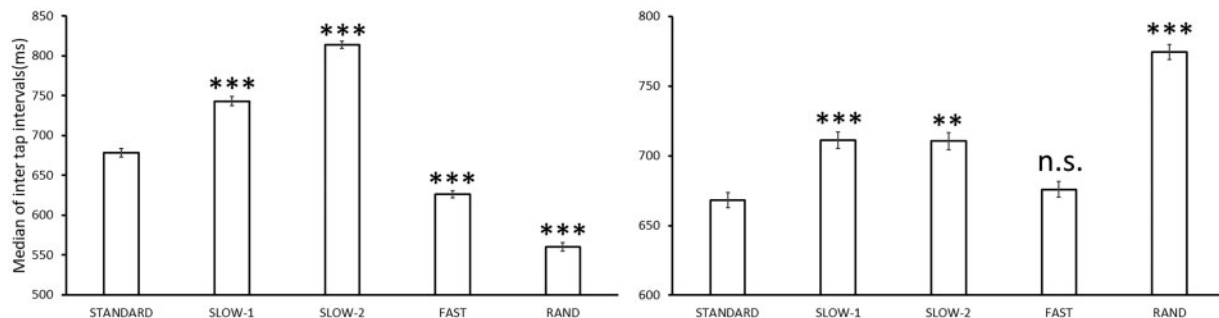

Figure 5. Example ITIs of the human subjects (left, #10; right, #01). The interval of #10 was longer under the SLOW conditions and shorter under the FAST condition, typical of the results observed in Group 1; this suggested that tap timing was affected by the metronome rhythm. The tap interval of #01 was significantly longer under the SLOW − 2 condition, suggesting that tap timing was affected by the metronomic sounds, although #01 showed no tap timing bias for stimulus onset according to the Rayleigh test, as was seen for bird P. ***$P < 0.001$, **$P < 0.01$ against STANDARD.

Similar to birds, humans did not exhibit more constant (less variable) tapping intervals with the metronome (Supplementary Figure S9). The difference in the C.V. of the ITIs between the STANDARD and no-sound condition sessions was not significant (Wilcoxon signed-rank test, $V = 24$, $P = 0.147$; Supplementary Figure S10), although the C.Vs. of humans were smaller than those of birds.

Discussion

The results suggested that rhythmic sounds modestly but significantly affected the tap timing in 2 of 5 budgerigars. Some of the results are consistent with our predictions for the birds: (1) the peck timing was influenced by the metronome, and the timing shifted toward the stimulus onset of the metronome in bird A under the FAST condition; and (2) the ITIs became shorter with the playback of the faster metronome and become longer with the playback of the slower metronome (but only for bird P). However, the results should be interpreted with caution. Especially in the latter case, the effects may be explained not only by sensorimotor coordination but also by several physical characteristics or features of the autonomic nervous system. For example, the heart beat and respiratory rates could have

been affected by the metronome, which would have accelerated the tap speed (Haas et al. 1986).

Significant results were found under the FAST condition, which might be related to the method itself; the behavior originated during a food reward operant task. We speculate that the birds wanted to obtain the reward as quickly as possible: this would bias the movement toward the faster side, such that a shorter IOI might lead the birds to truncate the movement, as seen in bird P (Figure 3, left). However, if the tap speed had already reached a speed limit under the STANDARD condition, the bird would not have been able to further accelerate the tap speed under the FAST condition (Supplementary Figure S5, lower panel). In such a case, the bird might be confused by the sound stimuli and thus decelerate the tap speed, as seen in bird A (Figure 3, right). The ITIs under the RAND condition were longer than those under the STANDARD condition in birds P and A. This could represent further evidence that the ITIs were affected by the rhythm of the sound sequences in those birds. In the circular analyses, significant results were mainly found under the FAST condition, as well as those found in the ITI analyses. Birds A and M did not demonstrate perfect synchronization to the metronome. Nevertheless, the tap timing bias appeared 10–30 ms before

the sound stimulus onset under the FAST condition supporting the above idea. However, the effects of the metronome sounds were not universally observed among the birds, which was unexpected because we had assumed that the rhythm of the metronome would lead the tap timing to be much more constant (but, interestingly, a similar tendency was also observed in humans).

The remaining concern is that bird M showed a tap timing bias, or was "synchronized" with the metronome, under the FAST condition (Figure 2); however, the ITI under the FAST condition did not differ from that under the STANDARD condition (Supplementary Figure S3), which is not easy to explain. Therefore, a question whether the "synchronization" reflects a chance phase alignment, or not, could arise. Because it was possible for the IOI of the metronome to be consistent with the preferred ITI of the birds, if bird M anticipated the timing of an incoming sound cue and aligned the timing of the first tap in a trial, the timing of the remaining 7 taps should be close to the timing of the sound sequence (however, of course, this is not very likely because the bird did not have any reason to do so). In this case, the ITI should be affected by the sound sequences as well, but the actual data do not show this result. The actual result can be explained by the fact that several, but not all, taps in a single trial were affected by the metronome sound. Of course, there may be other possible counterarguments; however, we believe that the data exhibited some effects of metronomic rhythm on tap timing.

The finding that the rhythmic sound stimuli affected tap timing in the budgerigars is consistent with the "vocal learning and rhythmic synchronization hypothesis" (Patel 2006), although we do not argue that the finding represents strong evidence supporting the hypothesis. The effects were limited and weaker than we expected based on the findings of parrot (cockatoo) studies; cockatoos exhibit a high capacity for spontaneous rhythmic behavior (Patel et al. (2009) in sulphur-crested cockatoo; Heinsohn et al. (2017) in palm cockatoo). Seki et al. (under review) found that the Bengalese finch, one of the vocal learning bird species, had poorer rhythmic synchronization performance than budgerigars during a task, suggesting that there are differences in rhythmic synchronization among vocal learning bird species even if the vocal learning and rhythmic synchronization hypothesis is true. Besides, budgerigars, unlike other parrot species, were not listed in the animals showing spontaneous rhythmic entrainment in YouTube videos (Schachner et al. 2009), suggesting budgerigars may perform worse on rhythmic synchronization tasks than other parrot species. It might be interesting to use cockatoos, another parrot species, in a similar study.

Comparison of the results between the 2 species

The results of budgerigars were not consistent, as the subjects showed large individual differences. Although we cannot, at present, explain these individual differences, even Snowball, the cockatoo, seemed to be an exceptional individual capable of dancing to a synchronized musical beat (Patel et al. 2009). Thus, this raised questions about why almost all humans showed nonvolitional synchronization in a similar task. In general, most adult humans seem to be capable of synchronizing movements to various rhythmic sound stimuli, mainly because they have learned to do so by cultural mechanisms (e.g., dancing or marching to music; cf. Kirschner and Ilari 2014), in addition to the innate tendencies (Kirschner and Tomasello 2009; Zentner and Eerola 2010). In our study, human subjects could synchronize tap timing when they were instructed to do so (seen in Supplementary Figure S8 as examples). However, some of the human subjects did not synchronize tap timing to the rhythm under several of the conditions. Thus, without explicit instructions, some Ss did not choose to synchronize their movements with rhythmical sounds during the task. Of course, some Ss may have resisted this intentionally. If so, it is interesting that they exhibited tap timing bias under at least one condition; moreover, this bias might have originated involuntarily. Another possibility is that some of the Ss may have guessed the purpose of the experiment and thus synchronized their tapping with the rhythm of the isochronous sound sequences to help the experimenter. This could have occurred in the cognitive process of the Ss on both explicit and implicit levels. These might have been ad hoc ideas, but they also might indicate that the experimenters, and even the Ss themselves, were fully cognizant of exactly what took place during the cognitive processing, which could also be partially true for the birds.

We can compare another point between the 2 species. We used the same sound sequence (100–400 ms) under the RAND condition as in the bird experiment to equalize the conditions, but the stimuli had shorter IOIs for average ITIs (537 ms) in humans. This may have shortened the ITIs under the RAND condition versus the STANDARD condition, as seen in subject #10 (Figure 5, left). In contrast, some Ss showed longer ITIs under the RAND condition, as seen in #01 (Figure 5, right), similar to the birds.

Questions for future studies

One of the most difficult problems in these types of experiments is that neither verbal nor paper-based instructions can be used in the context of nonhuman animals, unlike human psychological experiments. In this task, birds obtained a food reward using their preferred (i.e., the easiest) and/or most effective method. Therefore, even if they were capable of synchronizing or entraining to the rhythm of the isochronous sound sequences, it was not necessary to show this.

Another concern is that we used pure tones for the metronomic sounds to determine the generalized (or musical) rhythmic performance of the birds. Future studies could use the vocal sounds of birds as the metronomic stimuli, as these sounds may be more salient to birds.

In addition, because fewer trials were needed to obtain the required data, the humans performed fewer trials than the birds. The results might have differed if humans had performed the same number of trials as birds.

In this study, we proposed an experimental system to investigate the rhythmic behavior of animals that could be utilized in future studies. The system had 2 response keys and the ability to play back the sound stimuli in the background to obtain additional data on the origin of rhythmic synchronization. Such a system could be used to determine differences in the rhythmic behavior between vocal learning and nonvocal learning birds, between solitary and social/cooperative birds, and among birds at different vocal learning stages. Also, the system would enable us to examine the capability of small animals to maintain an internal beat (See the monkey saccade experiment performed by Takeya et al. (2017)).

Of course, humans have a distinctive capability for rhythmic synchronization, and our results might indicate that the capability for vocal learning might not be enough for rhythmic synchronization. Nevertheless, as previous studies have shown, there is a strong correlation between vocal production and rhythmic behavior in humans (Ejiri 1998; Dalla Bella et al. 2015). Thus, the origin of the capabilities of humans should be considered from the mutually compatible perspectives of vocal learning, social cooperation, and so on.

Supplementary material

Supplementary material can be found at https://academic.oup.com/cz.

Funding

Author contributions

Y.S. designed the study, carried out the bird experiment, analyzed the data, and wrote the manuscript. T.K. conducted the human experiment and analyzed the data.

References

Benichov JI, Benezra SE, Vallentin D, Globerson E, Long MA et al., 2016. The forebrain song system mediates predictive call timing in female and male zebra finches. *Curr Biol* **26**:309–318.

Cook P, Rouse A, Wilson M, Reichmuth C, 2013. A California sea lion *Zalophus californianus* can keep the beat: motor entrainment to rhythmic auditory stimuli in a non-vocal mimic. *J Comp Psychol* **127**:412–427.

Dalla Bella S, Berkowska M, Sowiński J, 2015. Moving to the beat and singing are linked in humans. *Front Human Neurosci* **9**:663.

Ejiri K, 1998. Relationship between rhythmic behavior and canonical babbling in infant vocal development. *Phonetica* **55**:226–237.

Haas F, Distenfeld S, Axen K, 1986. Effects of perceived musical rhythm on respiratory pattern. *J Appl Physiol* **61**:1185–1191.

Hasegawa A, Okanoya K, Hasegawa T, Seki Y, 2011. Rhythmic synchronization tapping to an audio-visual metronome in budgerigars. *Sci Rep* **1**:120.

Hattori Y, Tomonaga M, Matsuzawa T, 2013. Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Sci Rep* **3**:1566.

Hattori Y, Tomonaga M, Matsuzawa T, 2015. Distractor effect of auditory rhythms on self-paced tapping in chimpanzees and humans. *PLoS ONE* **10**: e0130682.

Heinsohn R, Zdenek CN, Cunningham RB, Endler JA, Langmore NE, 2017. Tool-ssisted rhythmic drumming in palm cockatoos shares key elements of human instrumental music. *Sci Adv* **3**:e1602399.

Kirschner S, Ilari B, 2014. Joint drumming in Brazilian and German preschool children: cultural differences in rhythmic entrainment, but no prosocial effects. *J Cross Cull Psychol* **45**:137–166.

Kirschner S, Tomasello M, 2009. Joint drumming: social context facilitates synchronization in preschool children. *J Exp Child Psychol* **102**: 299–314.

Large EW, Gray PM, 2015. Spontaneous tempo and rhythmic entrainment in a bonobo *Pan paniscus*. *J Comp Psychol* **129**:317–328.

Patel AD, 2006. Musical rhythm, linguistic rhythm, and human evolution. *Music Percept* **24**:99–104.

Patel AD, Iversen JR, Bregman MR, Schulz I, 2009. Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr Biol* **19**: 827–830.

Ravignani A, 2015. Evolving perceptual biases for antisynchrony: a form of temporal coordination beyond synchrony. *Front Neurosci* **9**:339.

Repp BH, 2006. Does an auditory distractor sequence affect self-paced tapping? *Acta Psychol* **121**:81–107.

Ruxton GD, 2017. Testing for departure from uniformity and estimating mean direction for circular data. *Biol Lett* **13**:20160756.

Schachner A, Brady TF, Pepperberg IM, Hauser MD, 2009. Spontaneous motor entrainment to music in multiple vocal mimicking species. *Curr Biol* **19**:831–836.

Takeya R, Kameda M, Patel AD, Tanaka M, 2017. Predictive and tempo-flexible synchronization to a visual metronome in monkeys. *Sci Rep* **7**:6127.

Zentner M, Eerola T, 2010. Rhythmic engagement with music in infancy. *Proc Nat Acad Sci USA* **107**:5768–5773.