



Rhythmic synchronization tapping to an audio–visual metronome in budgerigars

Ai Hasegawa^{1,2}, Kazuo Okanoya^{1,2,3}, Toshikazu Hasegawa¹ & Yoshimasa Seki^{2,3}

¹Department of Cognitive and Behavioral Science, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, 153-8902, Japan, ²Biolinguistics Laboratory, Brain Science Institute (BSI), Riken, Wako, Saitama, 351-0198, Japan, ³Japan Science and Technology Agency, ERATO, Okanoya Emotional Information Project, Wako, Saitama, 351-0198, Japan.

SUBJECT AREAS:
ANIMAL BEHAVIOUR
ZOOLOGY
THEORY
BEHAVIOUR

Received
15 February 2011

Accepted
3 October 2011

Published
17 October 2011

Correspondence and
requests for materials
should be addressed to
Y.S. (yseki@brain.
riken.jp)

In all ages and countries, music and dance have constituted a central part in human culture and communication. Recently, vocal-learning animals such as parrots and elephants have been found to share rhythmic ability with humans. Thus, we investigated the rhythmic synchronization of budgerigars, a vocal-mimicking parrot species, under controlled conditions and a systematically designed experimental paradigm as a first step in understanding the evolution of musical entrainment. We trained eight budgerigars to perform isochronous tapping tasks in which they pecked a key to the rhythm of audio–visual metronome-like stimuli. The budgerigars showed evidence of entrainment to external stimuli over a wide range of tempos. They seemed to be inherently inclined to tap at fast tempos, which have a similar time scale to the rhythm of budgerigars’ natural vocalizations. We suggest that vocal learning might have contributed to their performance, which resembled that of humans.

Dancing to the rhythm of music is such a natural and spontaneous human behaviour that we are typically unaware that the ability to align movement to external rhythmic stimuli is not ubiquitous in nature. For example, rhesus monkeys (*Macaca mulatta*) had great difficulty synchronizing their finger taps with both external visual and auditory stimuli in isochronous tapping tasks¹, while human children with no musical training could perform similar tasks quite easily². To date, there have been no other scientific attempts to train animals to move in synchrony with external rhythmic stimuli. However, spontaneous motor entrainment to music has been demonstrated in several parrot species and one elephant species^{3,4}, supporting the vocal learning and rhythmic synchronization hypothesis, which posits that vocal learning provides a neurobiological foundation for auditory–motor entrainment⁵. These reports suggest that there are additional spontaneous dancers in the natural world other than humans, but they consist of case studies of one individual per species and analyses of YouTube videos, whose recording conditions are unknown. To examine this topic, we systematically investigated the generation of rhythmic motor patterns synchronized with external stimuli in budgerigars (*Melopsittacus undulatus*; Fig. 1), a vocal-learning parrot species, using audio–visual metronome-like stimuli instead of complex music. This study is an important first step towards understanding the timing control mechanism in vocal learners.

We trained budgerigars (four males, four females) to perform isochronous tapping tasks. Isochronous tapping is a task originally designed to investigate cognitive timing control in humans, in which a subject taps his or her finger in synchronization with external rhythmic stimuli⁶. In a previous study with rhesus monkeys¹, three members of this species were trained to perform isochronous tapping by an operant conditioning method. The monkeys and human subjects were required to produce four taps to a visual or an auditory metronome followed by three continuation taps in the absence of stimulus presentation. After more than a year of extensive training, the monkeys finally learned the task, but their performance differed from that of humans in that they did not synchronize their movements to the stimulus onset as human subjects did, instead tapping around 300 ms after stimulus onsets. In addition, there was no auditory dominance in their tapping behaviour, while humans performed noticeably better under the auditory condition than under the visual condition. Using similar tasks, we attempted to determine whether the budgerigars’ performance resembles that of humans or if it parallels that of non-vocal-learning rhesus monkeys.

Through an operant conditioning procedure, the budgerigars learned to peck at an LED key during the 300-ms stimulus presentation or during an acceptable period before and after stimulus presentation, and such responses were counted as ‘hits’. The hit period before stimulus onset accounted for 20% of the inter-stimulus onset interval (IOI) (under the 450-ms- and random-IOI conditions, it was 50 ms and 180 ms, respectively), whereas the hit period after onset accounted for 20% of the IOI length plus the 300-ms stimulus presentation (Fig. 2; see Methods). We tested the birds under seven different IOIs: 450 ms, 600 ms, 900 ms, 1,200 ms, 1,500 ms,



Figure 1 | Our bird participants. Budgerigars are an easily maintained parrot species and are also called budgies or parakeets.

1,800 ms, and random numbers of milliseconds. The random-IOI condition, in which three possible IOIs (i.e., 900 ms, 1,200 ms, and 1,500 ms) appeared in random order, served to obtain estimated reaction time (ERT) to the stimuli. Under each condition, the birds were required to make 50 key-pecking sequences, each consisting of six successive hits, and we analysed the birds' tapping behaviour during these sequences. All birds learned to perform the tapping task (see Supplementary Videos S1 and S2 online for examples). Additionally, we recorded the warble songs of our participants and compared the intervals between warble-song elements with the IOIs at which they created precise intervals in the tapping tasks.

The present article describes the following two experimental datasets obtained in the tapping experiments: data on the time difference between tap onset and stimulus onset (asynchrony), which were used to analyse the key pecking timing using circular statistics, and data on the length of interresponse intervals (IRIs). We first offer evidence for entrainment in budgerigars by comparing the real birds' performance with that of various simulated birds. We also clarify the behavioural difference between tapping under fast-tempo (450 ms and 600 ms) and slow-tempo (900 ms, 1,200 ms, 1,500 ms,

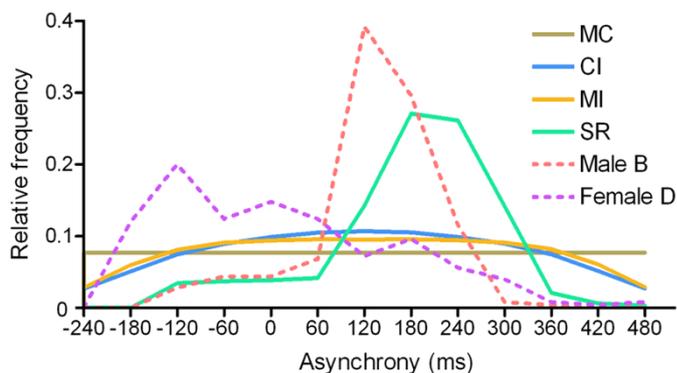


Figure 3 | Asynchrony distribution in simulated and real birds. The relative frequency distribution of asynchrony in the MC, CI, MI, and SR simulations and two real subjects (Male B and Female D) under the 1,200-ms-IOI condition is shown (bin width, 60 ms). The figure shows that the real birds directed their pecks at a certain point, and the phase relationship remained constant although they could have accomplished the tasks using the MC, CI, and MI strategies.

and 1,800 ms) conditions, discussing this difference from an evolutionary perspective by analysing the budgerigars' warble songs.

Results

We investigated whether budgerigars are capable of motor 'entrainment' to an audio-visual and, ultimately, to an auditory metronome. The essential element of entrainment, a phenomenon also referred to as beat perception and synchronization (BPS)⁵, is that a subject extracts periodicities from rhythmic stimuli and moves in a consistent phase relationship with them based on temporal anticipation. The temporal relationship between movement and stimuli should be steady when the rhythm is constant.

Using the tap-asynchrony data from the second to the sixth pecks in 50 six-hit sequences (data obtained from female D under the short IOI conditions were excluded; see Data analysis), we confirmed that the movements of real birds in all 46 tapping experiments (6 constant-IOI conditions \times 8 individuals – 2 excluded) maintained a consistent phase relationship with the stimuli (Rayleigh test with unspecified mean direction⁷, $P < 0.0001$), with a peak in asynchrony distribution (see Fig. 3 for examples of two individuals).

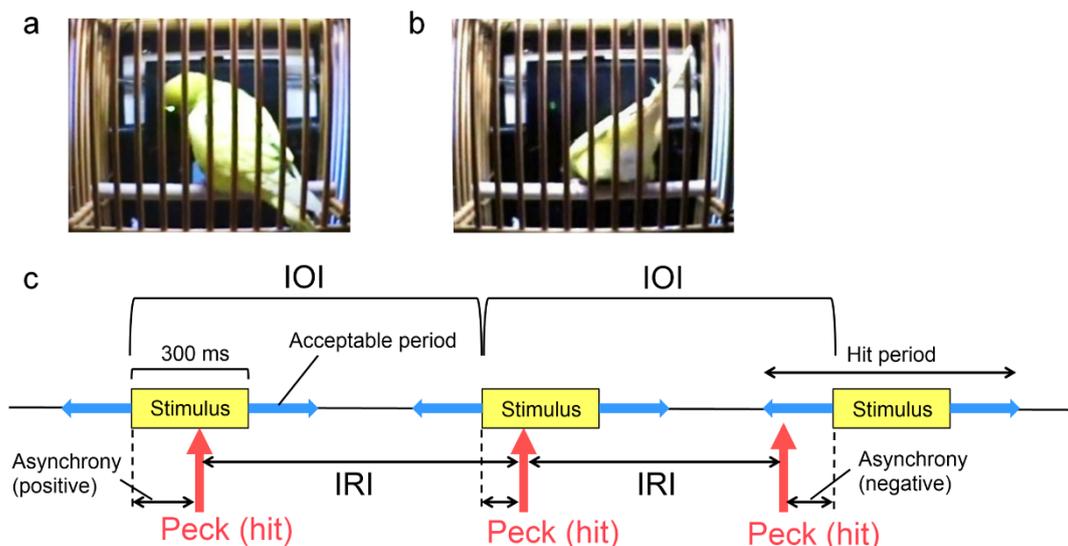


Figure 2 | Budgerigars' tapping experiments. Every time the birds produced a six-hit sequence (a), they obtained a reward (b). A frame-by-frame depiction of the tapping experiments is shown (c).



Phase-matched tapping with stimulus onset. A supplemental but nonetheless important feature of entrainment is phase matching (i.e., synchronization) with the beat (in this article, we regard phase matching with the stimulus onset as phase matching with the beat). It is not a necessary condition for entrainment, but it is typically seen in human tapping⁷, and is considered to be the outcome of temporal anticipation and thus likely to represent entrainment.

In our budgerigars, phase-matched tapping with stimulus onset was significant in one of seven subjects under the 450-ms-IOI condition, in two of seven under the 600-ms-IOI condition, in five of eight under the 900-ms-IOI condition, in seven of eight under the 1,200-ms-IOI condition, and all eight subjects showed significant synchronized tapping under the 1,500-ms- and 1,800-ms-IOI conditions (Rayleigh test with a specified mean direction of zero⁸, $P < 0.01$; Table 1, see Fig. 4 for examples under the 1,200-ms-IOI condition).

In 6 of 46 tapping experiments, the birds' movements were phase-matched with stimulus onset and not with the ERT (Supplementary Table S1, see Figure 4 for examples). The ERT was 150 ms and was used as the average response latency (150.4 ms) and the modal class value of the frequency distribution histogram under the random-IOI condition (see Methods section). In addition, in 32 of 46 experiments, the centre of the range in the circular distribution according to which a significant number of the pecks were directed ($P < 0.01$) lied before ERT (Supplementary Table S1, also see Figure 4). These results suggest that the birds' pecks were directed toward stimulus onset rather than toward ERT.

Negative mean asynchrony. In isochronous tapping tasks performed by humans, taps tend to precede the actual beat by a few tens of milliseconds⁷, which is known as negative mean asynchrony

(NMA). NMA suggests temporal anticipation even more strongly than does phase matching with stimulus onset. We calculated the mean asynchrony of the second through sixth pecks in each of the 50 six-hit sequences under each constant-IOI condition for each individual. We observed the NMA under all constant-IOI conditions except for that under the 450-ms-IOI condition (Table 1).

Possibilities other than entrainment. However, other possible explanations may be capable of explaining these results. Using computer simulations involving 10,000 simulated birds for the six constant-IOI conditions, we examined the following four possibilities that could have made the birds appear entrained.

First, the birds may have been pecking randomly. To test this possibility, we conducted a Monte Carlo (MC) simulation. Second, the birds may have continued to produce intervals of arbitrary length regardless of the actual IOI. This possibility was tested by the constant-interval (CI) simulation. Third, the birds may have stored the actual IOI length in their short-term memory and reproduced it in a serial fashion without attending to the stimuli. We tested this possibility using the memorised IOI (MI) simulation. The fourth possibility was that the birds may have been simply reacting to the stimuli. This possibility was tested by the simple reaction (SR) simulation in which the birds' response time was around the ERT of the real subjects (see Methods section for details of each simulation).

We analysed the data in the same way in each simulation as we did for the real subjects. We then compared the performance of simulated subjects with that of real birds in terms of three aspects: the rate at which phase-matching with stimulus onset occurred, the rate at which NMA appeared, and the number of failures in the completion of a six-hit sequence (i.e., one to five successive hit sequences) before producing 50 six-hit sequences.

Table 1 | Comparison between real and simulated subjects.

IOI (ms)		Phase matched ($P < 0.01$) w/ stimulus onset (w/ ERT)	NMA	Number of failures (range)	
450	Real birds	1/7 (7/7)	0/350	144.4 (3–506)	
	Simulated ones	MC	60/10,000 (170/10,000***)	8/500,000	3,908.9 (2,046–6,302) *
		CI	5/10,000 * (9,772/10,000)	299/500,000	267.0 (136–450) *
		MI	160/10,000 (6,705/10,000)	780/500,000	15.3 (3–38) **
		SR	0/10,000 ** (10,000/10,000)	0/500,000	36.2 (13–65) *
600	Real birds	2/7 (7/7)	8/350	75.6 (16–237)	
	Simulated ones	MC	95/10,000 ** (114/10,000***)	164/500,000 ***	3,965.8 (1,923–6,309) *
		CI	387/10,000 (9,816/10,000)	1,313/500,000 ***	260.8 (134–423) ***
		MI	1,042/10,000 (6,816/10,000)	2,958/500,000 **	14.7 (3–32) ***
		SR	0/10,000 *** (10,000/10,000)	5/500,000 ***	1.0 (0–7) ***
900	Real birds	5/8 (6/8)	79/400	101.4 (18–324)	
	Simulated ones	MC	87/10,000 *** (66/10,000***)	495/500,000 ***	16,133.5 (9,381–25,081) ***
		CI	1,974/10,000 * (9,829/10,000*)	2,434/500,000 ***	424.0 (224–711) ***
		MI	2,107/10,000 (8,057/10,000)	4,503/500,000 ***	32.6 (10–62) *
		SR	0/10,000 *** (10,000/10,000 ***)	5/500,000 ***	0 ***
1,200	Real birds	7/8 (5/8)	115/400	80.6 (164–645)	
	Simulated ones	MC	104/10,000 *** (81/10,000***)	961/500,000 ***	35,929.5 (19,060–55,657) ***
		CI	6,465/10,000 (9,862/10,000***)	3,946/500,000 ***	558.0 (310–917) ***
		MI	3,449/10,000 * (8,609/10,000)	6,381/500,000 ***	46.0 (18–89)
		SR	59/10,000 *** (10,000/10,000 ***)	7/500,000 ***	0 ***
1,500	Real birds	8/8 (7/8)	116/400	104.1 (253–833)	
	Simulated ones	MC	104/10,000 *** (110/10,000***)	1,551/500,000 ***	60,680.1 (32,568–102,881) ***
		CI	6,084/10,000 (9,879/10,000)	5,776/500,000 ***	666.5 (307–1,119) ***
		MI	5,039/10,000 * (9,006/10,000)	8,376/500,000 ***	56.0 (25–104)
		SR	9,962/10,000 (10,000/10,000 **)	2/500,000 ***	0 ***
1,800	Real birds	8/8 (8/8)	108/400	190.4 (32–411)	
	Simulated ones	MC	69/10,000 *** (91/10,000***)	2,218/500,000 ***	87,654.1 (50,098–139,610) ***
		CI	7,401/10,000 (9,877/10,000)	7,878/500,000 ***	756.8 (396–1,181) ***
		MI	6,051/10,000 (9,144/10,000)	10,837/500,000 ***	63.8 (26–114) ***
		SR	10,000/10,000 (10,000/10,000)	2/500,000 ***	0 ***

Compared with real birds using Fisher's exact test (for the ratio of phase-matched subjects and that of sequences with negative mean asynchronies) and Wilcoxon signed-rank test (for the number of failures) with the significance level adjusted by Bonferroni correction for four independent tests, *: $P < 0.0125$, **: $P < 0.0025$, ***: $P < 0.00025$.

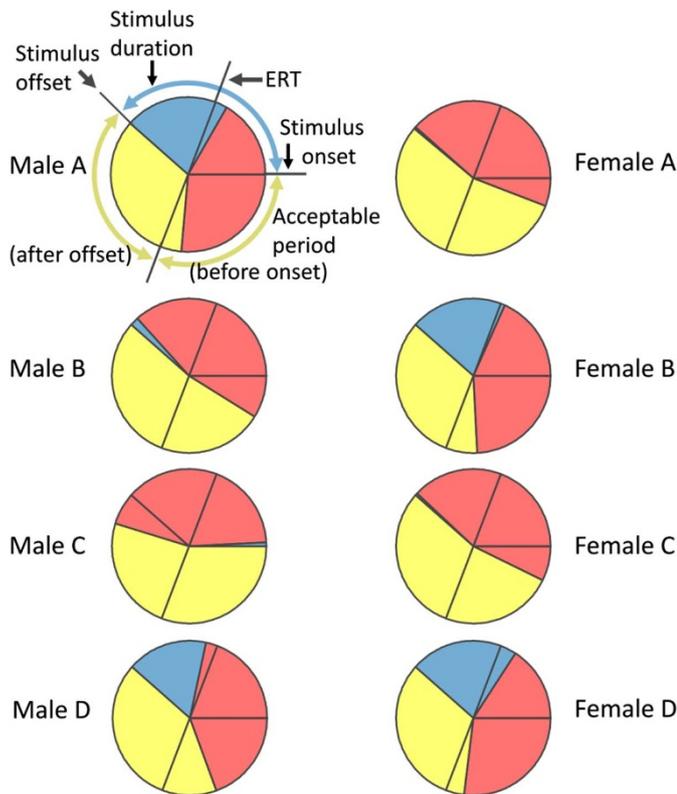


Figure 4 | Examples of the circular distribution of the pecks. Each circle shows the distribution of the peck responses along the time axis in the hit period of each subject under the 1200-ms-IOI condition. We conducted the Rayleigh test with a specified mean direction for each degree (from 0 to 360 degrees). The red zone indicates the range according to which a significant number of the pecks were directed ($P < 0.01$). The blue zone indicates the 300-ms stimulus presentation, with a line in the middle showing ERT. The remaining yellow area represents two acceptable periods before and after the stimulus onset. Pecks of Male A, Female B, and Female D clearly occurred near the stimulus onset rather than according to the ERT.

Monte Carlo simulation. The probability of observing periodicity in the MC-simulated birds was under 1% (Rayleigh test with unspecified mean direction at the $P < 0.01$ level). Additionally, the Gini coefficients⁹ of the asynchrony distribution of the simulated subjects, which are a measure of inequality in a distribution, were smaller than the real birds under all conditions (Kruskal–Wallis test, $P < 0.0001$; Dunn’s multiple comparison test, $P < 0.001$), suggesting that no consistent phase relationship with the stimuli existed (Fig. 3, Supplementary Table S2). Thus, the MC model lacks the crucial prerequisite condition for entrainment.

Additionally, the simulated subjects were much less likely to show phase matching with the stimulus onset and NMA (Fisher’s exact test, $P < 0.0025$ except for the 450-ms-IOI condition under which no NMA was observed in real subjects; Table 1), producing many more failures than the real birds (Wilcoxon signed rank test, $P < 0.0125$; Table 1). These results indicate that the budgerigars’ periodicity and phase-matched tapping did not occur by chance and that they may have been behaving according to a certain set of rules.

Constant interval simulation. In the CI simulation, the phase-matching rate did not differ from that of the real subjects under four conditions ($P > 0.0125$; Table 1). However, the Gini coefficients of the asynchrony distribution of the simulated birds were smaller than the real birds under all conditions (Kruskal–Wallis

test, $P < 0.0001$; Dunn’s multiple comparison test, $P < 0.001$), suggesting a fluctuating phase relationship with the stimuli over the course of the trials (Fig. 3, Table S2). Additionally, with the exception of the 450-ms-IOI condition, the simulated subjects produced NMA less frequently ($P < 0.00025$; Table 1) and the simulated birds failed more frequently ($P < 0.0125$). Thus, the CI model also failed to explain the real birds’ performance.

Memorised IOI simulation. In the MI simulation, the simulated birds’ pecks showed ‘accurate’ periodicity. Additionally, their phase-matching rate did not differ from that of real subjects under four conditions ($P > 0.0125$; Table 1). However, the smaller Gini coefficients of the asynchrony distribution of the simulated birds under all conditions (Kruskal–Wallis test, $P < 0.0001$; Dunn’s multiple comparison test, $P < 0.001$) indicated that the phase relationship with the stimuli did not remain constant over the trials (Fig. 3, Table S2).

Additionally, with the exception of the performance under the 450-ms-IOI condition, the rate at which NMA appeared was lower than that in trials with the real subjects ($P < 0.0025$). Furthermore, the simulated birds had fewer failures than did the real ones under four conditions ($P < 0.0125$). Thus, the MI model could not serve as the appropriate model for the real birds’ behaviour.

Simple reaction simulation. It is difficult to distinguish the SR strategy from true entrainment because the SR simulated birds maintained a consistent phase relationship with the stimuli over the course of the trials (Fig. 3).

However, the phase-matching rates of the real and simulated birds differed under the 450–1,200-ms-IOI conditions, although the pecks could phase-match with the stimulus onset under the 1,500-ms- and 1,800-ms-IOI conditions (Table 1). As the IOI elongated, the hit period became longer and both stimulus onset and ERT were located in a relatively narrow range. Therefore, the circular distribution logically concentrated on the period including both ERT and stimulus onset. Nevertheless, the responses of one real bird remained phase-matched with stimulus onset but not with ERT under the 1,500-ms-IOI condition, showing that the rate at which real and simulated birds phase-matched with ERT differed.

Moreover, NMA was extremely rare in the simulated birds, and the simulated subjects had far fewer failures than did the real birds ($P < 0.0125$).

Taken together, none of the four possible models provided a sufficient explanation for the real birds’ behaviour. The budgerigars’ pecks maintained a consistent phase relationship with the stimuli based on anticipation, and this pecking pattern was not the outcome of simple reactions. Consequently, the original idea, that the budgerigars anticipated the stimuli and entrained to the metronome, survived.

Tapping to auditory stimuli. In the main part of our study, we presented the budgerigars with auditory and visual cues simultaneously to facilitate the association between the LED key and the reward as per the operant conditioning protocol^{10–12}. However, it is important to test entrainment to auditory stimuli alone because this is the modality of interest in the vocal learning and rhythmic synchronization hypothesis⁵.

To this end, isochronous tapping to auditory stimuli alone was tested in three female budgerigars. Tapping was measured as they were alternately exposed to IOIs of 600 ms and 1,500 ms. The experiment was repeated twice (see Methods section).

The birds quickly learned to perform the task without special training. Their pecks were phase matched with the stimulus onset in five of the six tapping experiments (Rayleigh test, $P < 0.01$). This finding indicated that the auditory cue alone was sufficient for entrainment.



Behavioural difference between tempos. As shown, the budgerigars entrained to the metronome at a wide range of tempos and under different modalities. However, their behaviour differed under the fast- and slow-tempo conditions.

First, we calculated the mean tap asynchronies under each IOI condition (see Table S3 for the means and the error analysis). The mean asynchronies under the slow-tempo conditions were smaller than the reaction latencies under the random-IOI condition ($N = 35$ under the fast-tempo conditions, $N = 40$ under the slow-tempo and random-IOI conditions; Dunn's multiple comparison test, $P < 0.05$), but those under the fast-tempo conditions were not different from those under the random-IOI condition ($P > 0.05$; Fig. 5, Table S4).

Second, we calculated the SDs of the last four IRIs (i.e., for the last five pecks) in the six-hit tapping sequence under each IOI condition as an indicator of the birds' tapping variability. The SDs were much smaller and decreased slightly with long IOIs under the fast-tempo conditions ($N = 1,400$ for the fast-tempo conditions, $N = 1,600$ for the slow-tempo conditions; 450 ms: 87.0; 600 ms: 77.0; 900 ms: 132.8; 1,200 ms: 156.0; 1,500 ms: 188.6; 1,800 ms: 221.2). On the other hand, the SDs increased as the IOIs increased under the slow-tempo conditions (see Fig. 6 for the frequency distribution in IRIs).

Third, we examined the accuracy of the IRIs of the budgerigars' tapping in terms of whether it differed from the actual length of IOIs. For this analysis, we divided the four IRIs of the last five pecks into the first and latter two, and investigated the accuracy of the IRIs in each pair. The difference from the actual IOIs in the first and latter two IRIs under the fast-tempo conditions did not differ from zero (one-sample t -test, $t = 0.720$, $d.f. = 1,399$, $P = 0.472$ for the first two IRIs; $t = 1.233$, $d.f. = 1,399$, $P = 0.218$ for the latter two IRIs). On the other hand, the deviation from the actual IOI was larger than zero for the first two IRIs ($t = 6.464$, $d.f. = 3,199$, $P < 0.0001$), but did not differ from zero for the latter two IRIs under the slow-tempo conditions ($t = 1.467$, $d.f. = 3,199$, $P = 0.142$; Fig. 7). This indicated that IRIs maintained a constant precision under the fast-tempo conditions, whereas the budgerigars first created a rhythm faster than the actual IOI and then modified it to approach the actual IOI during the process of pecking under the slow-tempo conditions.

As a means of investigating these results from an evolutionary perspective, we analysed the budgerigars' warble songs, which consist of various elements with a complex temporal structure¹³. Both males and females incorporate learned sounds in their warble songs¹⁴. We recorded 111 warble songs produced by three males (range, 14–62).

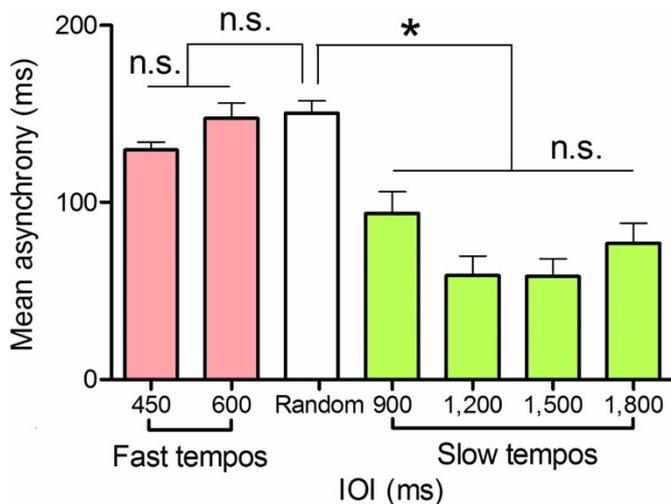


Figure 5 | Mean asynchronies under the seven IOI conditions. Mean asynchronies of the eight birds for the second to the sixth pecks in the reinforced six-hit sequences. Error bars represent one standard error of the mean. The results of the statistical tests are displayed (see also Table 1).

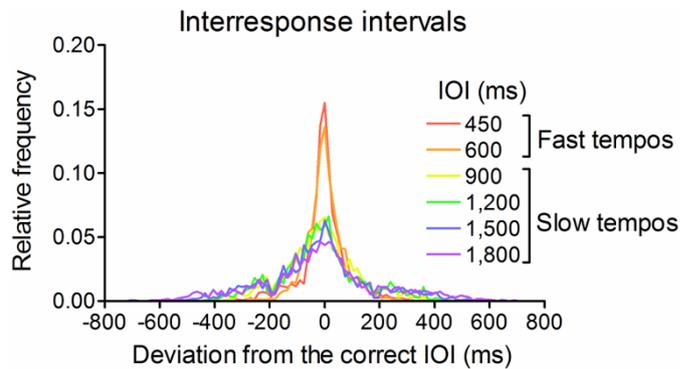


Figure 6 | Variability in interresponse intervals. Relative frequency distribution of deviations from the correct IOI for the eight birds' last four IRIs (i.e., IRIs for the last five pecks of each six-hit sequence) under the six constant-IOI conditions.

Relative frequency distributions of the interval lengths of song elements were calculated for each individual and averaged among the three males. We found that 83.9% of all the intervals between elements were under 600 ms, with a peak at around 100–200 ms ($N = 18,173$). Only 10.6% demonstrated durations longer than 900 ms. This tendency is consistent with a previous report¹⁴.

We observed spontaneous tapping behaviour at the previous tempo at the start of training for a new IOI. Just after the birds started tapping training with a new IOI, they kept tapping at almost the same tempo as that in the previous or penultimate IOI condition. We observed this behaviour eight times in the 40 between-trial transitions of constant IOIs (5 transitions \times 8 individuals), and it was observed in five of our subjects and between various IOI combinations (Table S5, Supplementary Videos S3 and S4). For example, male B continued tapping at the previous 600-ms IOI under the 1,800-ms-IOI condition (in which two successive hits were required for reinforcement). In this case, more IRIs were around 600 ms (450–750 ms) than in the long range (750–1,050 ms) during the first 20 reinforced sequences, but that this tendency disappeared in the last 30 reinforced sequences (χ^2 test, $\chi^2 = 6.352$, $d.f. = 1$, $P = 0.0117$). However, in the other cases, such behaviours disappeared very rapidly and the birds quickly began to tap at the new IOI. We compared the occurrence rate of such IOI-retaining behaviours between the 600-ms-IOI condition (4 cases/23 subsequent IOI combinations) and the other four IOI conditions (4 cases/92 subsequent

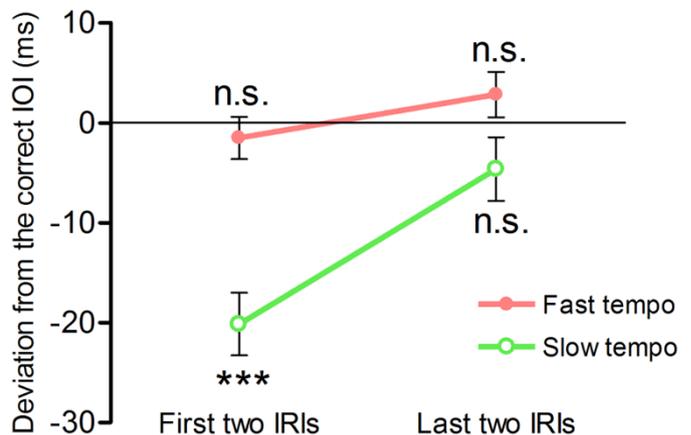


Figure 7 | Accuracy of interresponse intervals. Differences from the correct IOI in the first and last two IRIs of the last five pecks in the six-hit sequence under the fast- and slow-tempo conditions (mean \pm standard errors of mean). The results of the one-sample t -test show significant ($P < 0.05$) or non-significant differences from zero.



IOI combinations; the 450-ms-IOI condition was excluded because it was the last IOI condition for all subjects). We found that the 600-ms IOI tended to be retained with greater frequency than the other IOIs (Fisher's exact test, $P = 0.0497$).

Discussion

Overall, our results demonstrated that the budgerigars can entrain to external stimuli. Phase-matched tapping with the stimulus onset and the frequent observation of NMA by the budgerigars strongly suggested that they were truly entrained to the metronome, even though we did not require them to move rhythmically.

Through computer simulations, we discarded four possible models by which the birds might have appeared to be entrained; the MC simulation, in which randomly pecking birds were simulated, ruled out the possibility that the budgerigars' performance occurred by chance. The lack of periodicity in the MC subjects implied that if the subjects were behaving randomly, selecting six-hit samples would not be sufficient to get a false positive in terms of evidence of entrainment. The CI and MI models were more accurate than the MC model, but they did not meet the requirement for entrainment. The SR model was the most competitive but also failed to explain the real birds' behaviour. If the MI and SR models were operative, the budgerigars could have achieved the task more efficiently, but apparently they could not help but entrain to the rhythm. Thus, we conclude that the budgerigars truly entrained to the metronome.

Interestingly, we found several behavioural differences under the fast-tempo and slow-tempo conditions. The greater asynchrony under the fast-tempo conditions could be explained by synchronization with the perceptual centre of stimuli rather than the physical onset of the stimuli, as in humans¹⁵; however, this notion needs to be tested using shorter tones. Under the fast-tempo conditions, the perceptual centre of stimuli was likely to lag behind stimulus onset because the ratio of stimulus duration (300 ms) to IOI was relatively high. Additionally, under the fast-tempo conditions, the IRIs the budgerigars created was most accurate and least variable. The SDs of IRIs increased as a function of IOIs under the slow-tempo conditions as observed in humans and monkeys¹, but they decreased slightly under the fast-tempo conditions, which could not be explained by linear prolongation of the acceptable period (see Methods section).

The notion that budgerigars might be inherently inclined to tap at fast tempos is corroborated by our finding that the motor pattern at the 600-ms IOI tended to be retained most frequently, while spontaneous tapping at slower tempos disappeared quite rapidly. Thus, the IOI around 600 ms seems to be the preferred tempo of budgerigars.

We also found that the birds first created IRIs shorter than the actual IOIs and elongated these as the pecking proceeded under the slow-tempo conditions (this may be comparable to a phenomenon known as negative lag-one in humans¹⁶). The phenomenon to approach the preferred motor tempo in musical entrainment has also been reported in a cockatoo¹⁷. The fact that the birds eventually changed even the preferred tempo to the correct tempo supports entrainment because it indicates that budgerigars performed live monitoring of the stimuli and real-time error correction to match the tempo of the stimuli, which occurred within a few stimulus presentations.

The fact that both the preferred tapping tempo and interval length in warble songs shared a similar temporal range under 600 ms suggests that they are based on the common internal time scale. This assumption is also supported by the similar rhythm of other natural behaviours of budgerigars, such as the nudging (mean speed: about one nudge every 480 ms) and pumping (mean speed: about one pump every 250 ms) associated with courtship¹⁸.

Indeed, the budgerigars' behaviour differed from that of humans with respect to several features, such as the only occasional NMA, the

forced tapping using food reinforcement, and the short run of taps resulting from the experimental design. However, the budgerigars showed more behavioural similarities with humans than with rhesus monkeys¹. First, the budgerigars performed phase-matched tapping to the stimulus onset at a wide range of tempos. In contrast, rhesus monkeys always tapped after stimulus onset and never synchronized with it even when the duration of stimulus presentation was as short as 33 ms and the perceptual centre of the stimuli was located near the site of stimulus appearance¹. Second, budgerigars could entrain to auditory stimuli alone with little training. Studies investigating modalities related to synchronization in humans are in agreement about auditory dominance^{19–22}, but monkeys show a clear preference for visual stimuli and require considerable training to learn to tap to auditory stimuli¹. Although our findings did not prove auditory dominance in budgerigars, they are pertinent to the vocal learning and rhythmic synchronization hypothesis, which explores auditory–motor synchronization.

Our results support the hypothesis that vocal learning is the cause of the budgerigars' more human-like performance in isochronous tapping compared to rhesus monkeys. The phenomenon of interest for the aforementioned vocal learning and rhythmic synchronization hypothesis is motor entrainment to music⁵. The entrainment to an auditory metronome demonstrated in the budgerigars represents an important milestone to support this hypothesis. Our findings suggest that vocal-learning animals are not only better at musical entrainment, but also at simpler rhythm production and perception tasks in general when the rhythm is conveyed in the auditory domain. For example, vocal-learning starlings (*Sturnus vulgaris*) could discriminate tones varying in tempo²³ and in rhythmic and arrhythmic patterns²⁴, and vocal-learning dolphins (*Tursiops truncatus*) could discriminate different 4-s rhythms²⁵. On the other hand, non-vocal-learning pigeons (*Columba livia*) were only capable of tempo discrimination and could discriminate neither rhythm nor meter despite task simplification²⁶. Thus, the budgerigars' more human-like performance compared with monkeys in isochronous tapping might be explained by their vocal learning ability.

Because this study was designed as a controlled experiment, we excluded all social factors. However, given that a study on humans reported the social facilitation of rhythmic entrainment among children²⁷ and in view of the social nature of budgerigars²⁸, a more social experimental environment may improve the birds' performance.

This study is the first report on synchronization tapping in a vocal-mimicking species. Our study may provide a key to the unsolved riddle of why some individual parrots express musical entrainment behaviour and others of the same species do not. In our experiments, we showed that budgerigars generally have the ability to entrain to external stimuli when they are encouraged to do so. This entrainment is concerned not only with dancing, but also with music making and drumming that are synchronised with an external beat. We can infer from our finding that such rhythmic sensorimotor synchronization ability is likely present in all individual parrots, but latent in many of the conspecifics. In conclusion, our findings provide additional support for the hypothesis that rhythmic synchronization evolved as a by-product of selection for vocal learning.

Methods

Animals. Adult budgerigars (four males and four females, 1 year old) participated in the experiments and recordings of warble songs were used. The birds' access to food was restricted and they were maintained at an average of 84% of their free-feeding body weights for the operant procedure. All experimental procedures involving the animals complied with a protocol approved by the RIKEN Animal Experiments Committee and the RIKEN BSI guidelines.

Apparatus. We used a sound-attenuated chamber (inside dimensions: D 40.9 × W 58.9 × H 37.0 cm) for the tapping experiments and song recordings. We mounted an operant conditioning cage (15.5 × 30 × 22 cm high) in this chamber, which accommodated a custom-built response panel consisting of one piezoelectric sensor to detect pecking vibrations mounted behind a green light-emitting diode (LED) key. The tapping movements were registered with a device with accuracy in the



milliseconds. A 100-ms refractory period followed each response recording to avoid registering multiple time points for one response. A loudspeaker was placed above the roof of the cage to play back auditory stimuli: 3-kHz pure tones (a frequency close to budgerigars' natural contact call frequency²⁹) with a sound pressure level of 70 dB. The birds received one or two seed grains from a food tray under the key over a 2-s interval immediately following a correct response. The system was controlled by a personal computer and custom-made software created on a platform (Visual Studio 2005; Microsoft, USA).

Procedures. We trained the birds to peck the LED using operant conditioning techniques incorporating positive reinforcement. First, we trained the birds to peck the lit key. As soon as they pecked the key, the light disappeared and they obtained a reward. Then, we trained the birds to peck the blinking key during stimulus presentation. We shortened the stimulus duration gradually from 1,000 ms to 300 ms. The IOI length was either 3, 4, or 5 s and was randomly chosen at the beginning of each new sequence (i.e., after the bird obtained a reward). This phase ended when the ratio of the number of pecks during stimulus presentation (i.e., reinforcements) in each trial (which should be 50) to the number of total stimulus presentations exceeded 75% on three occasions when the stimulus duration was 300 ms.

The tapping training phase in which auditory and visual stimuli were simultaneously presented for 300 ms followed. We used the same 'acceptable periods' prior to stimulus onset and after stimulus offset. The first response for a stimulus during stimulus presentation and the acceptable period was counted as a 'hit'. Any following responses during the period and a response outside of this period were regarded as an 'error'. If no response occurred from the beginning of the acceptable period before stimulus onset to the beginning of the acceptable period before the next stimulus onset, a 'miss' was recorded. We used six constant-IOI conditions (450, 600, 900, 1,200, 1,500, and 1,800 ms) as well as a random-IOI condition. Under the random-IOI condition, three IOIs appeared in random order (900, 1,200, or 1,500 ms). The training order of the conditions was randomised among individual birds, but every bird received the random and the 450-ms-IOI conditions in the penultimate and final positions, respectively. The acceptable period was 20% of the IOI length under all conditions except the 450-ms- and random-IOI conditions in which it was 50 ms and 180 ms, respectively. Under each condition, the number of successive hits required for reinforcement was gradually increased from one to six. If the birds produced an error or a miss response just after making several hit responses, the accumulated hit count returned to zero. After 50 six-hit sequences under each IOI condition, the bird proceeded to the next IOI condition. These six-hit sequences were used for the analyses.

Three birds (Females A, B, and C) were tested using auditory stimuli alone. In order to accomplish this, the luminance of the LED was gradually reduced during two or three successive sessions. Then, the electric current to the LED was cut off. The other experimental conditions were the same as those used in the audio-visual experiment. Two IOI conditions (600 ms and 1500 ms; representing fast and slow tempo) were used, and the birds were alternately exposed to each IOI twice.

Recording of warble songs. We placed two plastic cages (15.5 × 30 × 22 cm high) facing towards a microphone in the sound-attenuated chamber. Birds were separately placed into the chamber (for almost 3 days per bird) in random pairs. We recorded songs using software Avisoft-RECORDER (Avisoft Bioacoustics, Germany) at a sampling frequency of 22.05 kHz. Songs were high-pass-filtered at 1 kHz to eliminate low-frequency noise produced by a ventilating fan. Each data point consisted of a continuous recording longer than 20 s. Warble elements were automatically separated and the interval between the onset of adjacent elements were measured using the 'automatic measurement' function of Avisoft SASLab Pro as the first-to-last points in the spectrogram exceeding a fixed threshold. This threshold was established by manual adjustment so that the segmentation pattern corresponded almost perfectly with that of human observers.

Simulations. We simulated four models: Monte Carlo (MC), constant interval (CI), memorised IOI (MI), and simple reaction (SR). Six simulated pecks were generated following the rule for each simulation (see below) in the stimulus-presentation sequence consisting of five intervals. Intervals between pecks were at least 100 ms, which corresponded with the 100-ms refractory period of experiments in real birds. Then, each peck was judged, in order, from the onset of the first stimulus presentation as a 'hit', 'error', or 'miss' in the same manner as in the real experiment. When an error or a miss was detected, the pecking series was regarded as a failure at that time point, and the next six key pecks were generated. The second to the sixth pecks were analysed as in the real experiment. We generated data for 10,000 simulated subjects under six constant-IOI conditions, and the simulations continued until each subject produced 50 six-hit sequences.

In the CI and MI simulations, we appended fluctuations to the periodic response intervals to create more natural models. To do this, we referred to the scalar timing theory³⁰ and to studies on the discrimination of short temporal intervals by humans and birds^{31,32}. Briefly, we randomly sampled a value as a response interval for each peck on a normal distribution using the Box-Muller method³³. The distribution was created based on the original interval length created according to the rule for each simulation (see below) and the SD, which was estimated from a fixed Weber fraction value (0.05) and the original interval³¹.

Monte Carlo simulation. Random key pecks were generated using the pseudo-random number generator in Visual Studio. These pecks were treated as described above.

Constant interval simulation. The first and the second peck were produced at random. When both responses fell within the hit ranges, the interval between those pecks was maintained at the original interval and used for the remainder of the series, with fluctuations as described above.

Memorised IOI simulation. We modelled each simulated bird memorising the correct IOI length under each IOI condition and producing pecks following the interval using their short-term memory with response timings irrelevant to the stimulus presentation. The intervals in a pecking sequence fluctuated as described above.

Simple reaction simulation. The simulated subjects reacted following the reaction times of the real subjects to the stimuli. The simulated birds produced pecks based on the appearance ratio of each modal class value in the ERT histogram in random order (see below).

Data analysis. We discarded all first-peck data from the reinforced sequence data analyses consisting of six hits in the timing analyses of the real and simulated subjects because we could not distinguish between first-peck data that were related to timing anticipation (e.g., the beginning of synchronization after several stimulus presentations or the first hit after making several error responses) and first-peck data that were independent of rhythm (e.g., direct responses to the first stimulus presentation after an interval between stimulus sequences).

Also excluded from the analyses were the tapping data obtained from female D under the 450-ms- and 600-ms-IOI conditions, including the data obtained during the 600-ms-IOI audio-visual and visual-only conditions. This bird pecked the key in a manner that clearly differed from that of the other seven bird subjects.

Statistical tests were conducted with two-tailed tests using R 2.13.0 (R foundation for Statistical Computing). For the Rayleigh test, we regarded each hit period as a cycle. The time each peck occurred was compared to the stimulus onset and assigned a relative phase in radian. To estimate ERT, we created a histogram according to Sturges' rule (number of bins = $\log_2(\text{number of data}) + 1$)³⁴ based on data obtained from the real subjects under the random IOI condition.

- Zarco, W., Merchant, H., Prado, L. & Mendez, J. C. Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *J. Neurophysiol.* **102**, 3191–3202 (2009).
- McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M. & Miller, N. S. The time of our lives: lifespan development of timing and event tracking. *J. Exp. Psychol. General* **135**, 348–367 (2006).
- Patel, A. D., Iversen, J. R., Bregman, M. R. & Schulz, I. Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* **19**, 827–830 (2009a).
- Schachner, A., Brady, T. F., Pepperberg, I. M. & Hauser, M. D. Spontaneous motor entrainment to music in multiple vocal mimicking species. *Curr. Biol.* **19**, 1–6 (2009).
- Patel, A. D. Musical rhythm, linguistic rhythm, and human evolution. *Music Percept.* **24**, 99–104 (2006).
- Dunlap, K. Reactions on rhythmic stimuli, with attempt to synchronize. *Psychol. Rev.* **17**, 399–416 (1910).
- Repp, B. H. Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev.* **12**, 969–992 (2005).
- Fisher, N. I. *Statistical Analysis of Circular Data* (Cambridge University Press, 1993).
- Gini, C. *Variabilità e mutabilità (Variability and Mutability)* (Bologna, 1912).
- Bilbrey, J. & Winokur, S. Controls for and constraints on autoshaping. *J. Exp. Anal. Behav.* **20**, 323–332 (1973).
- Cleland, G. G. & Davey, G. C. L. Autoshaping in the rat: the effects of localizable visual and auditory signals for food. *J. Exp. Anal. Behav.* **40**, 47–56 (1983).
- Winokur, S. & Boe, R. Effects of an auditory CS in automaintenance training with pigeons. Paper presented at the 16th Annual Meeting of the Psychonomic Society, Denver. (1975).
- Farabaugh, S.M. Brown, E.D. & Dooling, R.J.. Analysis of warble song of the budgerigar *Melopsittacus undulatus*. *Bioacoustics* **4**, 111–130. (1992).
- Eda-Fujiwara, H. & Okumura, H. The temporal pattern of vocalizations in the budgerigar (*Melopsittacus undulatus*). *J. Yamashina Inst. Ornithol.* **24**, 18–31 (1992).
- Vos, P. G., Mates, J. & Kruysbergen, N. W. The perceptual centre of a stimulus as the cue for synchronization to a metronome: evidence from asynchronies. *Quart. J. Exp. Psychol.* **48A** (4), 1024–1040 (1995).
- Wing, A. M. & Kristofferson, A. B. Response delays and the timing of discrete motor responses. *Percept. Psychophys.* **14**, 5–12 (1973).
- Patel, A. D., Iversen, J. R., Bregman, M. R., and Schulz, I. Avian and human movement to music: two further parallels. *Commun. Integr. Biol.* **2**, 1–4 (2009b).
- Brockway, B. Ethological studies of the budgerigar: reproductive behavior. *Behaviour* **23**, 294–324 (1964).



19. Chen, Y., Repp, B. H. & Patel, A. D. Spectral decomposition of variability in synchronization and continuation tapping: comparisons between auditory and visual pacing and feedback conditions. *Hum. Mov. Sci.* **21**, 515–532 (2002).
20. Repp, B. H. 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 (2003).
21. Repp, B. H. & Penel, A. Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *J. Exp. Psychol.* **28**, 1085–1099 (2002).
22. Repp, B. H. & Penel, A. Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychol. Res.* **68**, 252–270 (2004).
23. Hulse, S., Humpal, J. & Cynx, J. Processing of rhythmic sound structures by birds. *Ann. N Y Acad. Sci.* **423**, 407–419 (1984).
24. Hulse, S. & Kline, C. The perception of time relations in auditory tempo discrimination. *Learn. Behav.* **21**, 281–288 (1993).
25. Harley, H., Odell, K., Outnam, E. A., Goonen, C. & DeLong, C. M. Dolphins perceive rhythm: a belated ode to Stewart Hulse. Paper presented at the 9th International Conference on Comparative Cognition. (2002).
26. Hagmann, C. A. & Cook, R. G. Testing meter, rhythm, and tempo discrimination in pigeons. *Behav. Process.* **85**, 99–110 (2010).
27. Kirschner, S. & Tomasello, M. Joint drumming: social context facilitates synchronization in preschool children. *J. Exp. Child Psychol.* **102**, 299–314 (2009).
28. Brockway, B. Ethological studies of the budgerigar: non-reproductive behavior. *Behaviour* **22**, 193–222 (1964).
29. Dooling, R. J. Perception of vocal signals by the budgerigar (*Melopsittacus undulatus*). *Exp. Biol.* **45**, 195–218 (1986).
30. Allan, L. G. The influence of the scalar timing model on human timing research. *Behav. Processes* **44**, 101–117 (1998).
31. Thompson, J. G., Schiffman, H. R. & Bobko, D. J. The discrimination of brief temporal intervals. *Acta Psychologica* **40**, 489–493 (1976).
32. Weisman, R., Brownlie, L., Olthof, A., Njegovan, M., Sturdy, C. & Mewhort, D. Timing and classifying brief acoustic stimuli by songbirds and humans. *J. Exp. Psychol.: Anim. Behav. Process.* **25**, 139–152 (1999).
33. Box, G. E. P. & Muller, M. E. A note on the generation of random deviates. *Ann. Math. Statist.* **29**, 610–611 (1958).
34. Sturges, H. A. The choice of a class interval. *J. Am. Stat. Assoc.* **21**, 65–66 (1926).

Acknowledgements

We thank Hitomi Abe for help in maintaining the birds and conducting the experiments.

Author contributions

A.H. and Y.S. designed the experiments, analysed the data, and wrote the paper. Y.S. constructed the apparatus and developed the computer programs. A.H. conducted the experiments. K.O. and T.H. provided ideas and supervised the project.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors have no conflicts of interests to declare.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

How to cite this article: Hasegawa, A., Okanoya, K., Hasegawa, T. & Seki, Y. Rhythmic synchronization tapping to an audio-visual metronome in budgerigars. *Sci. Rep.* **1**, 120; DOI:10.1038/srep00120 (2011).