



Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction

Masahiro Kawasaki^{1,2,3}, Yohei Yamada¹, Yosuke Ushiku¹, Eri Miyauchi¹ & Yoko Yamaguchi⁴

¹Rhythm-based Brain Information Processing Unit, RIKEN BSI-TOYOTA Collaboration Center, Japan, ²Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Japan, ³Department of Intelligent Interaction Technology, Graduate School of Systems and Information Engineering, University of Tsukuba, Japan, ⁴Neuroinformatics Japan Center, RIKEN Brain Science Institute, Japan.

Behavioral rhythms synchronize between humans for communication; however, the relationship of brain rhythm synchronization during speech rhythm synchronization between individuals remains unclear. Here, we conducted alternating speech tasks in which two subjects alternately pronounced letters of the alphabet during hyperscanning electroencephalography. Twenty pairs of subjects performed the task before and after each subject individually performed the task with a machine that pronounced letters at almost constant intervals. Speech rhythms were more likely to become synchronized in human–human tasks than human–machine tasks. Moreover, theta/alpha (6–12 Hz) amplitudes synchronized in the same temporal and lateral-parietal regions in each pair. Behavioral and inter-brain synchronizations were enhanced after human–machine tasks. These results indicate that inter-brain synchronizations are tightly linked to speech synchronizations between subjects. Furthermore, theta/alpha inter-brain synchronizations were also found in subjects while they observed human–machine tasks, which suggests that the inter-brain synchronization might reflect empathy for others’ speech rhythms.

Individual human behavioral rhythms in nature are independent but can be spontaneously synchronized and entrained to become a shared rhythm through interactions with others (i.e., social interactions) by both verbal and nonverbal communication^{1,2}. We experience daily synchronizations with others, such as with hand clapping or foot tapping and incidental coordination of speech frequencies in conversations^{3–6}. This unconscious, shared rhythm brings individuals close to each other, generates empathy, and coordinates performance^{7,8}. Indeed, mother–infant rhythmic coupling through the imitation of movement supports social and cognitive development, such as language acquisition and learning⁹.

Coordinated rhythms have been observed in the brain not only as local- and distant-regional synchronizations within an individual brain^{10,11} but also as inter-brain synchronizations between individuals¹². Human brain rhythms in specific electroencephalography (EEG) frequency bands have increased not only while performing synchronized behaviors but also while observing such behaviors^{13,14}. These brain activities are related to the understanding of others’ intentions (i.e., social cognition), in this case, the understanding of others’ behavioral rhythms¹⁵. Moreover, recent EEG studies that scanned multiple brains at the same time have shown phase synchronizations between individuals during the imitation of hand movements^{16–19} or with participation in cooperative games and actions^{20–22}. Synchronized changes between individual brains have also been reported in functional magnetic resonance imaging (fMRI; 23–26) and near-infrared spectroscopy (fNIRS; 27, 28) studies.

Although there is considerable evidence of behavioral synchronization and brain synchronization between individuals during social interactions through nonverbal communication, these relationships are not as clear during verbal communication. One EEG hyperscanning study revealed correlated brain activities between the speakers and listeners during verbal communication³¹, but this study did not address the behavioral rhythm (i.e., the speech rhythm) itself. Furthermore, it is poorly understood whether such synchronizations include both simultaneous common movement and turn-taking, the latter of which is typically non-simultaneous and requires unconscious alternation of behavior, such as between speakers during a conversation²⁹. Turn-taking requires interpersonal synchronization of speech rhythms, including timing, duration, interval, and speed of speech, along with the content and context of the conversation itself³⁰.

SUBJECT AREAS:

COOPERATION

EMPATHY

AGENCY

COGNITIVE NEUROSCIENCE

Received

27 September 2012

Accepted

3 April 2013

Published

22 April 2013

Correspondence and requests for materials should be addressed to M.K. (kawasaki@brain.riken.jp)



In the current study, we attempted to address whether inter-brain synchronizations on an EEG appear when speech rhythms are synchronized between two subjects in verbal communication. We conducted alternating speech tasks in which two subjects alternately and sequentially pronounced the alphabet during hyperscanning EEG recordings (human–human, Fig. 1A). The behavioral and brain synchronization between the subjects were evaluated in terms of the correlation of speech rhythms (duration and interval of pronunciation) and brain rhythms (EEG oscillatory amplitudes in specific frequency bands), respectively.

In addition, each subject participated in alternating speech tasks with a machine (human–machine, Fig. 1A), and these results were compared with the results of the human–human tasks to address the following questions: (1) How does inter-brain synchronization change after the two subjects' behavioral rhythms have been coordinated to common rhythms (i.e., the machine's rhythms), and (2) how does inter-brain synchronization between a subject performing and a subject observing the task with a machine (social cognition) differ from the inter-brain synchronization between subjects in the human condition (social interaction)?

Results

Speech rhythms: voice sounds. Each subject performed two types of alternating speech tasks: human–human tasks and human–machine tasks (Fig. 1A). They completed 14 sessions consisting of two pre-machine human–human sessions, 10 successive human–machine sessions [five voices (electronic, male, female, partner's, and subject's) at two paces (fixed and random)], and two post-machine human–human sessions (Fig. 1B). We recorded the sounds of human–human alternating speech tasks. We dissociated the durations of voices and intervals between the voices of the pair of subjects using short-time Fourier transformations (Fig. 2A).

During the human–human tasks, the durations and intervals of speech were correlated between the pair of subjects and not significantly different among the subjects (Fig. 3C, D). Notably, compared with the pre-machine tasks using one-factor repeated-measures ANOVA (pre-machine vs. post-machine), the pair-averaged correlations for the post-machine tasks were significantly higher (Fig. 3D;

duration, $F_{1,104} = 4.50, P = 0.036$ interval, $F_{1,104} = 6.19, P = 0.014$) and the differences were significantly lower (Fig. 3E; duration, $F_{1,104} = 5.49, P = 0.021$; interval, $F_{1,104} = 9.30, P = 0.003$). Pre-machine, the durations and intervals were not correlated between the two subjects in six pairs. Moreover, the durations were significantly different between the two subjects in four pairs, and the intervals were significantly different between the two subjects in three pairs. Post-machine, 17 of 18 pairs showed significant correlations in the duration and interval of the speech (example shown in Fig. 3B, C).

For human–machine tasks, the correlations between subject pairs were lower, and the differences between subject pairs were higher than those for the human–human tasks [Fig. 3C, D; correlation of duration, $F_{1,338} = 58.36, P = 0.001$; correlation of interval, $F_{1,338} = 127.01, P = 0.001$; difference of duration, $F_{1,338} = 8.09, P = 0.004$; difference of interval, $F_{1,338} = 9.04, P = 0.002$; one-factor repeated-measures ANOVA (human vs. machine conditions)].

Brain rhythms: EEG oscillations. To characterize brain oscillatory activity, we conducted wavelet analysis on the collected EEG data^{28,29}. The subject-averaged amplitudes of each frequency (ranging from 4 to 28 Hz) for performing the alternating speech tasks in both human–human and human–machine conditions were significantly higher than those for subjects who were observing the human–machine tasks [Fig. 4A; $P < 0.01$; one-factor repeated-measures ANOVA (human vs. machine conditions)]. Frequency ranges (4–28 Hz) were divided into two parts: theta/alpha bands (6–12 Hz), which showed higher amplitudes for human than machine conditions, and beta bands (20–28 Hz), which showed the opposite (machine > human).

The theta/alpha band amplitudes increased in the frontal regions during both human–human and human–machine tasks and extended to the central and parietal regions during human–human tasks (Fig. 4B). The theta/alpha amplitudes during human–human tasks were significantly higher than those during the human–machine tasks in the central and parietal regions (electrodes C3, T7, CP1, and CP2). For the human–human tasks, the post-machine activities were significantly higher than pre-machine activities in the Fp2, C3, and P4 electrodes.

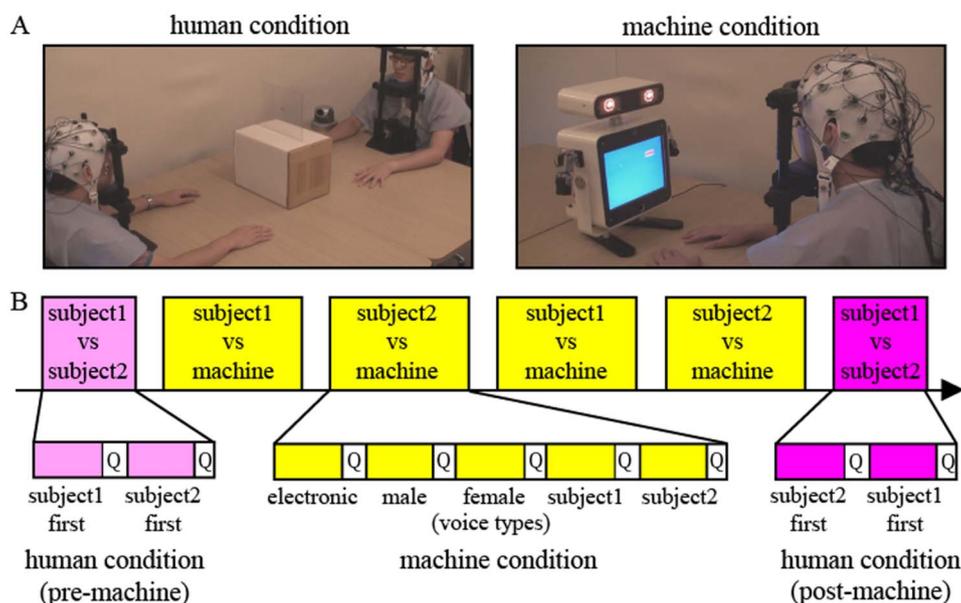


Figure 1 | Experimental setup of alternating speech tasks. (A) Alternating speech tasks between two human subjects (human–human) and one subject and a machine (human–machine). (B) Schematic illustration of the experiments. Each subject completed 14 sessions comprising two pre-machine human–human sessions, 10 successive human–machine sessions [five voices (electronic, male, female, partner's and subject's) at two paces (fixed and random)], and two post-machine human–human sessions. In each session, the subjects participated in alternating speech for 70 seconds. After each session, each subject rated the subjective qualities of the alternating speech (Q).

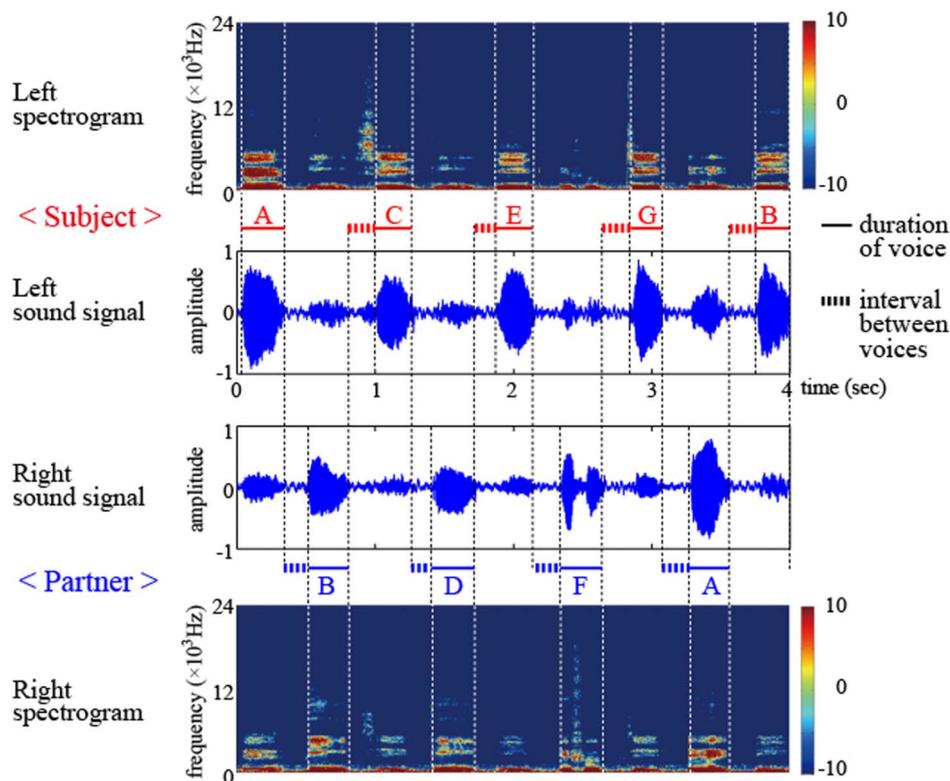


Figure 2 | Voice data and analyses from alternating speech tasks. Representative examples of voice data captured by a stereo recorder and spectrogram from the FFT analyses for the data. The durations of the subject’s and partner’s voices and the intervals between the voices (dashed lines) were dissociated and analyzed.

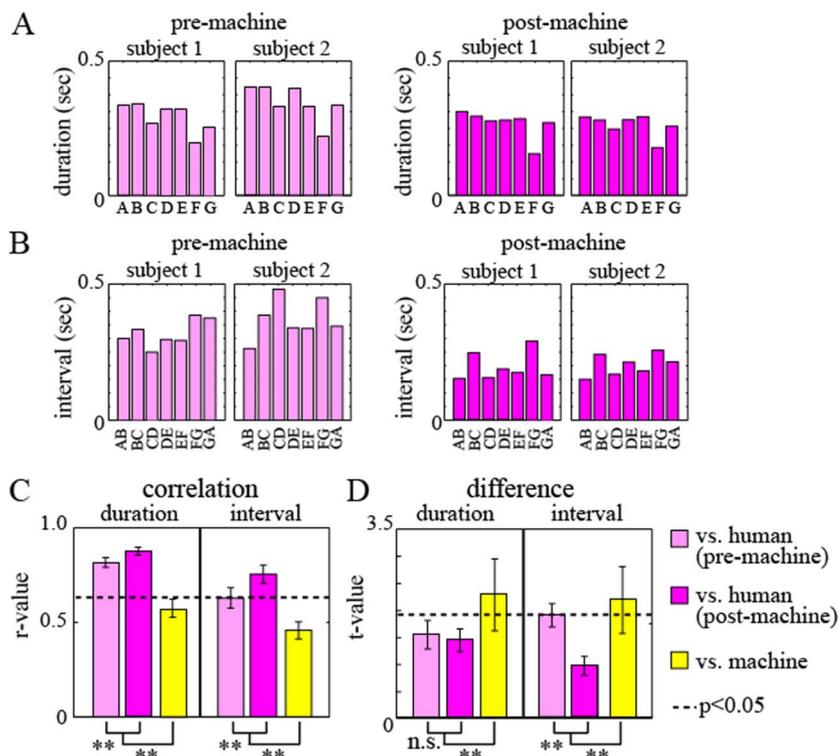


Figure 3 | (A) Averaged data and representative durations for each letter and (B) the intervals between each letter from one human–human pair before and after participating in the human–machine tasks. (C) Subject-averaged correlations and (D) differences in the durations and intervals between the subjects under each condition.

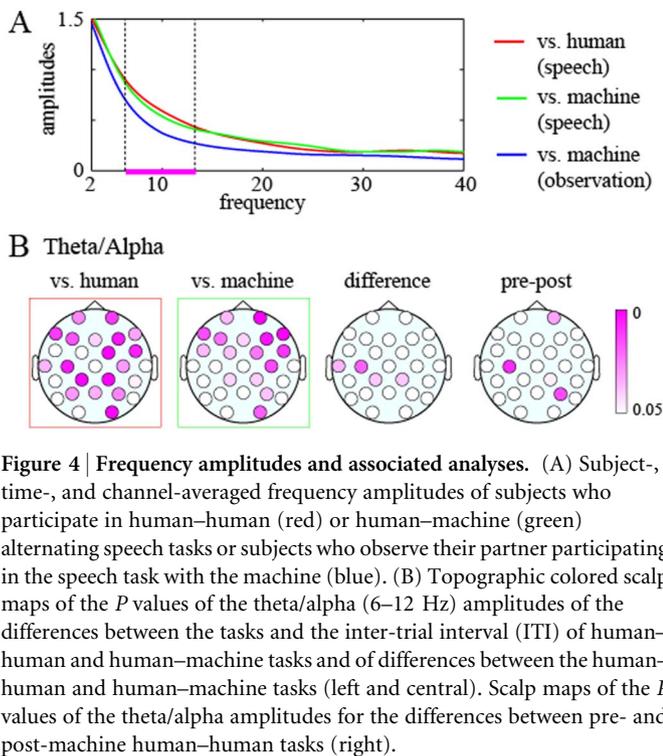


Figure 4 | Frequency amplitudes and associated analyses. (A) Subject-, time-, and channel-averaged frequency amplitudes of subjects who participate in human–human (red) or human–machine (green) alternating speech tasks or subjects who observe their partner participating in the speech task with the machine (blue). (B) Topographic colored scalp maps of the P values of the theta/alpha (6–12 Hz) amplitudes of the differences between the tasks and the inter-trial interval (ITI) of human–human and human–machine tasks and of differences between the pre- and post-machine human–human tasks (left and central). Scalp maps of the P values of the theta/alpha amplitudes for the differences between pre- and post-machine human–human tasks (right).

Similar to the theta/alpha amplitudes, the beta band amplitudes showed enhancement in the frontal regions but extended to the temporal and occipital regions under both human and machine conditions. However, there was no significant difference between the conditions for any electrode. Moreover, the beta activities showed

no significant difference between the post-machine and pre-machine human–human tasks.

To investigate inter-brain synchronization during alternating speech, we conducted cross-correlation analyses of the theta/alpha amplitudes from each pair of subjects' EEG data because these amplitudes were significantly different during the human–human and human–machine tasks. An example of the time course of a theta/alpha amplitude and the cross-correlation coefficients for one subject pair during a human–human alternating speech task is shown in Fig. 5A (left). In pair-averaged correlations of the theta/alpha amplitudes, high-peaked correlations were distributed in the temporal regions (electrodes F7, FC5, T7, T8, CP1, and CP2) for both pre- and post-machine using cross-correlation analyses that included both the pre- and post-machine conditions (Fig. 5B). For most of the electrodes, these values were significantly higher for the post-machine conditions [Fig. 5D; $P < 0.05$; one-factor repeated-measures ANOVA (pre-machine vs. post-machine) with post-hoc analyses (Bonferroni correction)].

The theta/alpha cross-correlation coefficients for the temporal and parietal electrodes were significantly correlated with the individual behavioral synchronizations between the subjects; in other words, there were high correlations between the speech durations and intervals between the two subjects (Fig. 5E; electrode measuring the peak statistic value, T7; correlation of duration, $r = 0.49$, $P = 0.046$; correlation of interval, $r = 0.44$, $P = 0.071$).

Finally, we investigated theta/alpha inter-brain synchronization between subjects who participated in alternating speech tasks with a machine and the subjects and observed the human–machine tasks (an example is shown in Fig. 5A, right). High-peaked correlations partially overlapped in the temporal regions, the same area in which significant correlations were seen in subjects during human–human tasks (compare Fig. 5B and Fig. 5C; electrodes F7, T7, CP1, and CP2). However, the theta/alpha cross-correlation coefficients for the

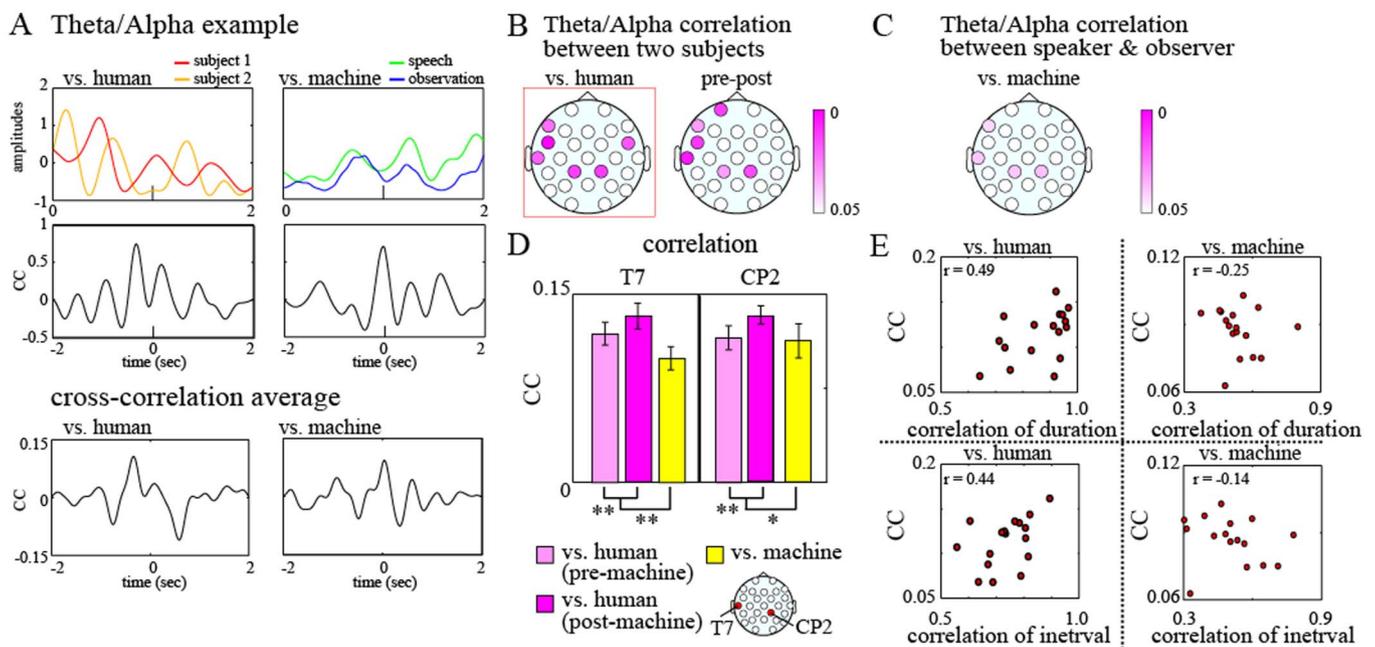


Figure 5 | (A) Examples of theta/alpha amplitudes on the temporal electrodes (T7) of two subjects and their time-course cross-correlation coefficients during the human–human and human–machine tasks (top and middle). Examples of averaged cross-correlation coefficients for human–human and human–machine tasks (bottom). (B) Scalp maps of the P values, which show significant correlations between the two subjects' theta/alpha amplitudes during the human–human tasks (left) and differences between the correlation coefficients in the pre- and post-machine human–human tasks (right). (C) Scalp maps of the P values, which show significant correlations between the two subjects' theta/alpha amplitudes during the human–machine tasks. (D) Subject-averaged correlation coefficients on the temporal (T7) and lateral parietal (CP2) electrodes during the human–human and human–machine tasks. (E) Scatter plots between the inter-brain correlation coefficients on the temporal (T7) electrodes and the correlation coefficients of the duration and interval of the speech rhythms between the subjects.



temporal regions were not significantly correlated with the individual behavioral synchronizations between subjects (Fig. 5E; electrode measuring the peak statistic value, T7; correlation of duration, $r = -0.25$, $P = 0.320$; correlation of interval, $r = -0.14$, $P = 0.591$).

Subjective ratings. After each alternating speech session, the subjects were asked to rate the following factors using a 5-point scale: “comfort,” “synchrony,” “speed,” “initiative,” and “humanity” (only for the human–machine tasks). For all of the factors except for humanity, ratings for the human partners were significantly higher than those for the machine partner [comfort, $F_{1, 502} = 82.54$, $P = 0.001$; synchrony, $F_{1, 502} = 80.75$, $P = 0.001$; speed, $F_{1, 502} = 7.59$, $P = 0.006$; initiative, $F_{1, 502} = 92.30$, $P = 0.001$; one-factor repeated-measures ANOVA (human vs. machine conditions)]. These results suggested that the subjects felt more comfortable, had better and faster synchronization, and had a higher initiative when performing alternating speech with a human partner than with the machine partner.

For the human–human tasks, post-machine ratings were significantly higher than pre-machine ratings for comfort ($F_{1, 142} = 30.71$, $P = 0.001$), synchrony ($F_{1, 142} = 4.45$, $P = 0.036$), and speed ($F_{1, 142} = 17.65$, $P = 0.001$), but the post-machine initiative was rated slightly lower (in favor of the partner’s initiative rather than the subject’s own initiative; $F_{1, 142} = 2.47$, $P = 0.118$).

Effects of the machine’s voice. Next, we examined the effects of machine voice type on the behavioral and EEG rhythms in human subjects participating in human–machine alternating speech tasks. In the analyses of speech rhythms (i.e., the correlations and differences of the durations and intervals), we conducted two-factor repeated-measures ANOVA for voice types and paces. For correlation of voices, the ANOVA showed a main effect for voice type (correlation of duration, $F_{4, 222} = 3.77$, $P = 0.005$; correlation of interval, $F_{4, 222} = 13.00$, $P = 0.001$) but no effect for pace (correlation of duration, $F_{1, 222} = 0.25$, $P = 0.62$; correlation of interval, $F_{1, 222} = 0.29$, $P = 0.59$). In contrast, for the difference in the voices, the ANOVA showed no effect of voice (difference of duration, $F_{4, 222} = 2.23$, $P = 0.066$; difference of interval, $F_{4, 222} = 2.35$, $P = 0.055$) or pace (difference of duration, $F_{1, 222} = 0.01$, $P = 0.920$; difference of interval, $F_{1, 222} = 0.04$, $P = 0.841$). Post-hoc analyses (Bonferroni correction) showed significant differences between voice types (Fig. 6A; $P < 0.05$; correlation of duration, male and female < electronic; difference of duration, self < male; correlation of interval, male < other voices; difference of interval, female and self < electronic and male).

The EEG analyses revealed theta/alpha activities that showed a main effect of voice but no effect of pace in some of the electrodes using a two-factor repeated-measures ANOVA. Post-hoc analyses revealed that the activities were higher with female and the partner’s voices than for the other voices ($P < 0.05$; Bonferroni correction). Moreover, only temporal electrodes registered significant correlations between the speakers and observers in response to the machine’s voices according to a two-factor repeated-measures ANOVA (Fig. 6B; T7 electrode vs. CP2 electrode). On the temporal electrodes, inter-brain correlations were higher for the female and partner’s voices than for the electronic and male voices ($P < 0.05$).

Along with the behavioral and EEG results, the subjective ratings showed significant differences based on voice type: for comfort, electronic < all other voices; for synchrony, electronic < self and female; for initiative, electronic and male < self; and for humanity, electronic < self and partner ($P < 0.05$; Bonferroni correction).

There were no differences in any of the subjective emotions, speech, or EEG rhythms between the condition with a fixed pace of machine speech and that with a random pace. This finding might have occurred because of the small range used for the random pace (400–600 msec) and its similarity to the fixed pace (500 msec). Indeed, most of the subjects did not notice the difference between

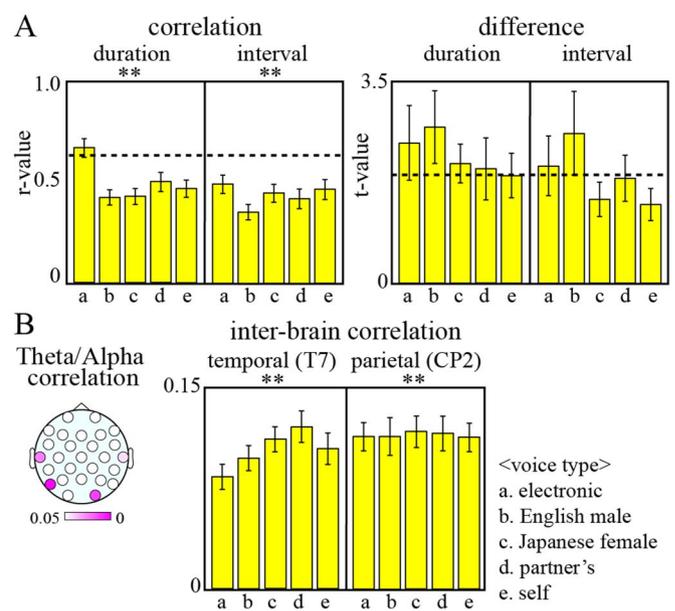


Figure 6 | Analyses of the speech rhythms among the voice types.

(A) Subject-averaged correlations and differences in the durations of the voices and intervals between the voices during the human–machine tasks; five voices were used [electronic (a), male (b), female (c), partner’s (d) and subject’s (e) voices]. (B) Scalp maps of the P values, which show significant main effects of the voices for correlations between the two subjects’ theta/alpha amplitudes under machine conditions (left). Subject-averaged correlation coefficients on the temporal (T7) and lateral parietal (CP2) electrodes for each machine voice.

the paces used. Although this study revealed no differences between the fixed and random paces of the machine’s speech, the possible effect of adaptation in human and machine communications will need to be clarified in future studies.

Discussion

The current study is the first to describe inter-brain synchronization along with synchronization of speech rhythms between subjects during an alternating speech task. Individual speech rhythms (i.e., duration of speech and intervals between two voices) were more likely to synchronize (high correlations and small differences), which is similar to the spontaneous synchronizations that were observed between humans in previous nonverbal studies^{2–4,13–16}. This phenomenon is specific to human–human communication because the same behavioral synchronizations were not found during the human–machine alternating speech tasks with the machine pronouncing letters at a fixed interval. The behavioral synchronization was also correlated with the theta/alpha oscillatory amplitudes, which were enhanced and synchronized between two subjects in the same bilateral temporal and lateral parietal regions during the human–human tasks. This result suggests that inter-brain synchronization tightly links to speech synchronization between subjects. Indeed, the inter-brain synchronization was enhanced as the speech rhythm coordination developed, with previous human–human and human–machine interactions strengthening subsequent human–human behavioral and brain coordination (comparing pre- and post-machine conditions).

Inter-brain synchronizations are of great importance to our cooperative communication. In previous studies¹⁹, inter-brain synchronizations have been reported in nonverbal communication, such as the imitation of hand movements^{16–18} or during participation in cooperative games and actions^{20–22}, and in verbal communication between listeners and speakers³¹. Such imitations, games, and speeches included the context of the subjects’ behaviors. The present



study indicated that inter-brain synchronizations are also found during simple communication, without the context of the subjects' behavior.

Numerous EEG studies have proposed that theta/alpha activities are involved in high cognitive functions, such as working memory^{11,32–34}. For example, theta and alpha oscillations in the temporal and parietal brain regions are modulated during auditory and visual working memory tasks, respectively¹¹. In this study, the theta/alpha modulation in the temporal regions during auditory communication is similarly related to the auditory working memory. Therefore, we propose that social interactions that lead to successful turn-taking require a working memory of others' speech rhythms. Similarly, previous studies have shown that alpha inter-brain synchronizations are also used in tasks that require working memory for others' behavior, such as imitation tasks¹⁶ and cooperative tapping tasks¹⁴.

In this study, theta/alpha inter-brain synchronization was also observed in the temporal and lateral parietal regions between speakers and observers (i.e., between subjects who participate in and who observe the human–machine tasks). These regions, which are in the vicinity of the temporal parietal junction (TPJ), have been proposed to be related to social cognition, such as understanding others' intentions, emotions, and behavior, including the theory of mind and the mentalizing process^{33–38}. The TPJ is associated not only with the observation of human behaviors but also with the observation of machine (i.e., robotic) behavior^{39–41}. Moreover, the theta/alpha amplitude modulation has been reported to be related to the coordination of self-behaviors for the observation of others' behaviors in tapping tasks^{13,14}. Thus, there are spatial and frequency overlaps between the neural mechanisms for social interaction and social cognition¹⁵. It should be noted, however, that the neural mechanisms for social cognition are enhanced during social interaction because the lateral parietal theta/alpha amplitudes during human–human tasks were higher and their synchronizations were greater than those observed during human–machine tasks.

There is a possibility that inter-brain synchronizations are related to jaw movements causing muscle artifacts in EEG data, especially within the temporal regions⁴². However, the theta/alpha enhancements in the temporal regions were not observed in human–machine tasks, although the tasks required subjects to pronounce letters in the same manner as in the human–human tasks. Moreover, in the human–human tasks, when one subject pronounced letters, the other subject did not pronounce letters. In other words, the two subjects did not move their jaws simultaneously. Finally, the human–machine tasks did not ask one subject (i.e., the observer) to pronounce letters when another subject (i.e., the speaker) participated in the task. Therefore, we concluded that the inter-brain connectivity was not due to the muscle movement artifacts.

In addition, our findings reveal the importance of the familiarity of the machine's voice (e.g., a Japanese-native voice for Japanese subjects) in human–machine speech coordination. Use of the subjects' voices, partners' voices, or a Japanese-native female voice showed higher correlations and smaller differences in the speech rhythms and higher theta/alpha synchronizations in the EEG than did other types of voices. Consistent with both behavioral and brain synchronizations, the use of familiar voices and human-like machine voices made the subjects feel comfortable and synchronized. Typically, the subjects' speech rhythms were more likely to be synchronized when the voices used were familiar, with positive subjective impressions, which underscores important differences in language and learning^{43,44}.

Interestingly, the effects of the machine's voice on the inter-brain synchronizations between the speakers and observers demonstrated functional dissociations within the temporal-parietal regions. The inter-brain synchronization of the temporal regions is susceptible to the machine's voice; greater synchronization was observed with the use of familiar voices. This finding might be due to the sensitivity

of the primary auditory areas in the temporal brain regions to auditory perceptual properties^{45,46}. However, the lateral parietal regions usually showed inter-brain synchronization, regardless of the machine's voice.

The existing relationships between the subjects might also affect the degree of empathy and their coordination of behavioral and brain rhythms, although there were no differences between the genders, ages, and relationships of the subject pairs who were acquaintances or strangers at the time of the study (e.g., more than half of the desynchronized pairs were already acquaintances). Future studies should address how behavioral and brain synchronization are influenced by personal relationships (e.g., mother–infant, husband–wife, boss–subordinate, teacher–student, etc.).

Methods

Subjects. Twenty pairs (N = 40) of healthy right-handed Japanese volunteers (nine female–female, seven male–male, and four female–male pairs; mean age = 21.57 ± 0.84 years, range 18–40 years) with normal or corrected-to-normal visual acuity, normal hearing acuity, and normal motor performance participated in the experiment. Six pairs met for the first time at the experiment, and 14 pairs were acquaintances. All of the subjects provided written informed consent prior to participation in the study. The study protocol was approved by the Ethical Committee of RIKEN (in accordance with the Declaration of Helsinki). One female–female pair was excluded from the analyses because of a lack of subjective ratings. Each subject completed a pretest, during which their pronunciation of the alphabet from “A” to “G” was recorded to program the machine's voice and define the spectrogram of their pronunciation.

Alternating speech tasks. Two types of tasks were performed: human–human tasks and human–machine tasks.

The human–human alternating speech task was performed by two human subjects, who were asked to pronounce the alphabet from “A” to “G” sequentially and alternately; for example, one subject said “A,” and then the other subject said “B,” and so on. Speech rhythms were not instructed. Each alternating speech session was conducted for 70 seconds. The interval between sessions was at least 10 seconds. The average frequency of loops (i.e., the number of times from “A” to “G”) was no fewer than 20 per 70-second session.

In the human–machine alternating speech tasks, one human subject participated in alternating speech with an anthropomorphic robot (PoCoBot; Business Design Company, Japan). The machine used five voices, electronic, male, female, the human subject's partner's voice, and the human subject's own voice, each of which was set at two different paces between pronunciations (e.g., between “A” and “C”): fixed intervals (1000 ms) and random intervals (800 to 1200 ms).

Each subject completed 14 sessions (Fig. 1B). First, the two human subjects performed the human–human task twice, with each subject starting the alternating speech. Next, each subject participated in 10 human–machine tasks (five types of voices × two voice paces). Lastly, the two human subjects repeated the task twice, again with each subject starting the alternating speech.

After each session, each subject evaluated the subjective qualities of “comfort,” “synchrony,” “speed,” and “initiative” for the human–human and human–machine alternating speech tasks and “humanity” for the human–machine tasks using a 5-point scale.

Task procedures. Each subject completed 14 sessions and underwent EEG recording, voice recording, and video capture during each session. The tasks were conducted in an electronic- and sound-shielded room.

The series of sessions were conducted as follows (Fig. 1B). First, the two human subjects performed the task twice, with each subject starting the alternating speech (human–human). The subjects were seated at a distance of 150 cm and faced each other eye-to-eye by placing their chins in chin rests. During the sessions, the subjects were asked to keep their chins in the chin-rests, lay their hands on the desk, keep their bodies still, and fix their eyes on a cross that was located on a clear partition placed between them (Fig. 1A, left).

Next, each human subject of the pair performed the speech task with an anthropomorphic robot (PoCoBot; Business Design Company, Japan) (human–machine). The machine remained stationary at a distance of 50 cm from the subject during the sessions, although it was able to make motions with its head and hands between tasks (Fig. 1A, right). The machine used five voices, electronic, male, female, the human subject's partner's voice, and the human subject's own voice, each of which was set at two different paces: fixed intervals and random intervals. The electronic voice included embedded machine sounds. The male voice was an English native speaker. The female voice was a Japanese native speaker. The subject's partner's voice and the subject's own voice were recorded during the pretest and filtered to remove background noise. The fixed intervals were set at 1,000 ms between pronunciations. The random intervals were shuffled from 800 to 1,200 ms between pronunciations, with a uniform distribution. Each subject alternately completed 10 human–machine sessions (five voices × two paces) in two batches. One batch included five voices and



alternated between the fixed and random paces. The orders of the conditions were counter-balanced across the subjects.

Lastly, the two human subjects repeated the task twice, again with each subject starting the alternating speech (Fig. 1B).

In our EEG and behavioral measurements, the subject who started the alternating speech was asked to press one key at the onset of each session to trigger a signal to the EEG systems. Then, 70 seconds later, another trigger signal was sent to the EEG systems that indicated the offset of each session. The experimenters then opened the door of the shielded room, and the subjects stopped the task and rested. We analyzed the EEG and speech data in the temporal windows between the onset trigger signal and offset trigger signal.

Subjective ratings. After each session, each subject used a 5-point scale to report their subjective ratings of “comfort” (5 was “very comfortable,” and 1 was “not comfortable”); “synchrony” (5 was “strong feelings,” and 1 was “no feelings” of synchronization with partner); “speed” (5 was “fast,” and 1 was “slow”); “initiative for alternating speech” (5 was “subject had higher initiative,” and 1 was “partner had higher initiative”), and “humanity” of the machine’s behavior (5 was “human-like,” and 1 was “machine-like”).

Voice sound recording and analyses. Voice sounds were recorded from a linear stereo PCM recorder using a PCM-D50 (SONY, Japan; sampling 96.00 kHz, 16 bit), which was placed centrally between the two human subjects or one human subject and one machine. Its right and left microphones were directed at each subject. We filtered out the background noise from the right and left recorded voice sounds (from 20,000 to 44,000 Hz). We first calculated the spectrogram of the sound data that were recorded during the pretest with short-time Fourier transforms and then identified the frequency ranges of each letter spoken by each subject. Referencing the frequency ranges of each letter, we applied short-time Fourier transforms for the right and left recorded data during the alternating speech tasks and dissociated the voices and intervals of the two subjects or one subject and one machine (Fig. 2).

The durations of pronunciation of each letter and the intervals between each letter were averaged for each session. For both durations and intervals for each session of each subject pair, we calculated the correlations and differences between the two subjects or one subject and one machine by using Pearson’s correlation analyses and ANOVA with post-hoc analysis (Bonferroni correction), respectively.

EEG recordings and analyses. The individual brain activities of the subjects in each pair were independently recorded by two EEG systems with 27 active scalp electrodes embedded in an electro cap (actiCAP) and BrainAmp ExG MR equipment (Brain Products, Germany). The sampling rate was 1,000 Hz. Reference electrodes were placed on the right and left ear lobes. Electrooculography (EOG) was performed with electrodes that were placed above and below the left eye to monitor eye blinks and vertical eye movements and electrodes that were placed 1 cm from the right and left eyes to monitor horizontal eye movements. To reduce or eliminate artifacts, we conducted infomax independent components analysis (ICA) on the EEG data. The ICA components with the most significant correlations with the vertical and horizontal EOG were rejected, and the ICA-corrected data were recalculated by regression of the remaining components⁴⁷.

To identify the time-frequency amplitudes during the alternating speech tasks, we applied wavelet transforms assuming that Morlet’s wavelets have a Gaussian shape in the time domain ($SD \sigma_t$) and the frequency domain ($SD \sigma_f$) around its center frequency (f)⁴⁸. The TF amplitude $E(t, f)$ for each time point of each session was calculated as the squared norm of the results of the convolution of the original EEG signals $s(t)$ with the complex Morlet’s wavelet function $w(t, f)$:

$$w(t, f) = (\sigma_t \sqrt{\pi})^{-1/2} \exp(-t^2/2\sigma_t^2) \exp(i2\pi ft)$$

$$E(t, f) = |w(t, f) \otimes s(t)|^2$$

where $\sigma_f = 1/(2\pi\sigma_t)$. The wavelet that we used was characterized by a constant ratio ($f\sigma_f = 7$), with f ranging from 0.5 to 40 Hz in 0.5-Hz steps. The amplitude was calculated by subtracting the baseline data measure in the ITI for each frequency band. The oscillatory amplitude of each frequency was averaged across times from the onset to the offset of each session for each condition.

Inter-brain synchronizations were evaluated by cross-correlation analyses of convolutions between the two subjects’ oscillatory amplitudes at each frequency during a 2-second window. The window was moved from the onset to the offset of each session with a 1-second overlap. The calculated cross-correlation coefficients were averaged among all of the windows for each session, and then, the maximum values were defined.

- Giles, H., Coupland, N. & Coupland, J. Accommodation theory: communication, context, and consequence. In Giles, H., Coupland, J., Coupland, N. (eds.) *Contexts of Accommodation: Developments in Applied Sociolinguistics*. (New York: Cambridge University Press, 1991).
- Kelso, J. A. S. *Dynamic Patterns: The Self-Organization of Brain and Behavior*. (Cambridge: MIT Press, 1995).
- Néda, Z. *et al.* The sound of many hands clapping - tumultuous applause can transform itself into waves of synchronized clapping. *Nature* **403**, 849–850 (2000).

- de Rugy, A., Sables, R., Oullier, O. & Temprado, J. J. A neuro-mechanical model for interpersonal coordination. *Biol. Cybern.* **94**, 427–443 (2006).
- Richardson, M. J., Marsh, K. L. & Schmidt, R. C. Effects of visual and verbal interaction on unintentional interpersonal coordination. *J. Exp. Psychol. Human* **31**, 62–79 (2005).
- Richardson, D., Dale, R. & Shockley, K. Synchrony and swing in conversation: coordination, temporal dynamics and communication. In: Wachsmuth I, Lenzen M, Knoblich G (eds.) *Embodied Communication*. (Oxford University Press, 2008).
- Feldstein, S. & Welkowitz, J. Conversational congruence: Correlates and concerns. In Siegman, A., Feldstein, S. (eds.) *Nonverbal Behavior and Communication*. (Hillsdale, NJ: Lawrence Erlbaum Associates, 1978).
- Street, Jr., R. L. Speech convergence and speech evaluation in fact-finding interviews. *Human Communication Research* **11**, 139–169 (1984).
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L. & Jasnow, M. D. Rhythms of dialogue in infancy: coordinated timing in development. *Monogr. Soc. Res. Child Dev* **66**, 1–132 (2001).
- Varela, F., Lachaux, J. P., Rodriguez, E. & Martinerie, J. The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* **2**, 229–239 (2001).
- Kawasaki, M., Kitajo, K. & Yamaguchi, Y. Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory. *Eur. J. Neurosci.* **31**, 1683–1689 (2010).
- Hasson, U. *et al.* Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trend. Cogn. Sci.* **16**, 114–121 (2012).
- Tognoli, E., Lagarde, J., de Guzman, G. C. & Kelso, J. A. The phi complex as a neuromarker of human social coordination. *Proc. Natl. Acad. Sci. USA* **104**, 8190–8195 (2007).
- Naeem, M., Prasad, G., Watson, D. R. & Kelso, J. A. Electrophysiological signatures of intentional social coordination in the 10–12 Hz range. *Neuroimage* **59**, 1795–803 (2012).
- De Jaegher, H., Di Paolo, E. & Gallagher, S. Can social interaction constitute social cognition? *Trend. Cogn. Sci.* **14**, 441–447 (2010).
- Dumas, G. *et al.* Inter-brain synchronization during social interaction. *PLoS ONE* **5**, e12166 (2010).
- Lindenberger, U., Li, S. C., Gruber, W. & Muller, V. Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neuroscience* **10**, 22 (2009).
- Dumas, G., Chavez, M., Nadel, J. & Martinerie, J. Anatomical connectivity influences both intra- and inter-brain synchronizations. *PLoS ONE* **7**, e36414 (2012).
- Konvalinka, I. & Roepstorff, A. The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Front. Hum. Neurosci.* **6**, 215 (2012).
- Babiloni, F. *et al.* Cortical activity and connectivity of human brain during the prisoner’s dilemma: an EEG hyperscanning study. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **4953–6** (2007).
- Astolfi, L. *et al.* Estimation of the cortical activity from simultaneous multi-subject recordings during the prisoner’s dilemma. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **1937–9** (2009).
- De Vico Fallani, F. *et al.* Defecting or not defecting: how to “read” human behavior during cooperative games by EEG measurements. *PLoS ONE* **5**, e14187 (2010).
- Hasson, U. *et al.* Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634 (2004).
- Montague, P. R. *et al.* Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage* **16**, 1159–1164 (2002).
- King-Casas, B. *et al.* Getting to know you: reputation and trust in a two-person economic exchange. *Science* **308**, 78–83 (2005).
- Anders, S. *et al.* Flow of affective information between communicating brains. *Neuroimage* **54**, 439–446 (2010).
- Funane, T. *et al.* Synchronous activity of two people’s prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *J. Biomed. Opt.* **16**, 077011 (2011).
- Cui, X., Bryant, D. M. & Reiss, A. L. NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* **59**, 2430–2437 (2012).
- Pickering, M. J. & Garrod, S. Towards a mechanistic psychology of dialogue. *Behav. Brain Sci.* **27**, 169–226 (2004).
- Wilson, M. & Wilson, T. P. An oscillator model of the timing of turn-taking. *Psychon. B. Rev.* **12**, 957–968 (2006).
- Kuhlen, A. K., Allefeld, C. & Haynes, J.-D. Content-specific coordination of listeners’ to speakers’ EEG during communication. *Front. Hum. Neurosci.* **6**, 266 (2012).
- Jensen, O., Gelfand, J., Kounios, K. & Lisman, J. E. Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb. Cortex* **12**, 877–882 (2002).
- Klimesch, W., Freunberger, R., Sauseng, P. & Gruber, W. A short review of slow phase synchronization and memory: evidence for control processes in different memory systems? *Brain Res.* **1235**, 31–44 (2008).
- Sauseng, P. *et al.* Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.* **19**, 1846–1852 (2009).



35. Adolphs, R. Social cognition and the human brain. *Trends Cogn. Sci.* **3**, 469–479 (1999).
36. Allison, T. *et al.* Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* **4**, 267–278 (2000).
37. Gallagher, H. L. & Frith, C. D. Functional imaging of “theory of mind.” *Trends Cogn. Sci.* **7**, 77–83 (2003).
38. Saxe, R. & Kanwisher, N. People thinking about thinking people The role of the temporo-parietal junction in “theory of mind” *Neuroimage* **19**, 1835–1842 (2003).
39. Press, C., Bird, G., Flach, R. & Heyes, C. Robotic movement elicits automatic imitation. *Cogn. Brain Res.* **25**, 632–640 (2005).
40. Gazzola, V., Rizzolatti, G., Wicker, B. & Keysers, C. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage* **35**, 1674–1684 (2007).
41. Oberman, L. M., McCleery, J. P., Ramachandran, V. S. & Pineda, J. A. EEG evidence for mirror neuron activity during the observation of human and robot actions: toward an analysis of the human qualities of interactive robots. *Neurocomputing* **70**, 2194–2203 (2007).
42. Goncharova, D. I., McFarland, D. J., Vaughan, T. M. & Wolpaw, J. R. EMG contamination of EEG: spectral and topographical characteristics. *Clin. Neurophysiol.* **114**, 1580–1593 (2003).
43. Nazzi, T., Jusczyk, P. W. & Johnson, E. K. Language discrimination by English-learning 5-month-olds: effects of rhythm and familiarity. *J. Mem. Lang.* **43**, 1–19 (2000).
44. Lee, C. S. & Todd, N. P. M. Towards an auditory account of speech rhythm: application of a model of the auditory ‘primal sketch’ to two multi-language corpora. *Cognition* **93**, 225–254 (2004).
45. Zatorre, R. J. & Belin, P. Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* **11**, 946–953 (2001).
46. Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music. Auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* **8**, 547–558 (2007).
47. Kawasaki, M. & Yamaguchi, Y. Frontal theta and beta synchronizations for monetary reward increase visual working memory capacity. *Soc Cogn Affect Neurosci* in press.
48. Tallon-Baudry, C., Bertrand, O., Delpuech, C. & Pernier, J. Oscillatory gamma band activity (30–70 Hz) induced by a visual search task in human. *J. Neurosci.* **17**, 722–734 (1997).

Acknowledgments

This research was supported by a Grant-in-Aid for Scientific Research on Innovative Areas (22118510 & 21120005). The authors would like to thank Naoko Tsukakoshi and Yuuki Koizumi for their support in the data acquisition and analysis and Hideyuki Takahashi for his support with the robotic design.

Author contributions

M.K., E.M. and Y. Y. designed the research; M.K., Y. Y. and E.M. performed the research; M.K., Y. Y. and Y.U. analyzed the data; and M.K. and Y. Y. wrote the paper.

Additional information

Competing financial interests: The authors declare no competing financial interests.

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

How to cite this article: Kawasaki, M., Yamada, Y., Ushiku, Y., Miyauchi, E. & Yamaguchi, Y. Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction. *Sci. Rep.* **3**, 1692; DOI:10.1038/srep01692 (2013).