Brain activity supporting alternating speech for semantic words: simultaneous magnetoencephalographic recording

Risa Anada[a], Hayato Watanabe[b,c], Atsushi Shimojo[b,d], Hideaki Shiraishi[e], Koichi Yokosawa[b]

[a] Graduate school of Health Sciences, Hokkaido University, N-12, W-5, Kita-ku, Sapporo 060-0812, Hokkaido, Japan

[b] Faculty of Health Sciences, Hokkaido University, N-12, W-5, Kita-ku, Sapporo 060-0812, Hokkaido, Japan

[c] Department of Child Studies, Toyooka Junior College, Tobera 160, Toyooka 668-8580, Hyogo, Japan

[d] Department of Pediatrics, Graduate School of Medicine, Hokkaido University, N-15, W-7, Kita-ku, Sapporo 060-8638, Hokkaido, Japan

[e] Department of Pediatrics, Hokkaido University Hospital, N-14, W-5, Kita-ku, Sapporo 060-8648, Hokkaido, Japan

Corresponding author: Koichi Yokosawa

(Address) Faculty of Health Sciences, Hokkaido University,

N-12, W-5, Kita-ku, Sapporo 060-0812, Hokkaido, Japan

(Tel) +81-11-706-2828

(E-mail) yokosawa@med.hokudai.ac.jp

Running title: Brain activity supporting alternating speech

Abstract

Communication, especially conversation, is essential for human social life. Many previous studies have examined the neuroscientific underpinnings of conversation, i.e., language comprehension and speech production. However, conversation inherently involves two or more people, and unless two people actually interact with one another, the nature of the conversation cannot be truly revealed. Therefore, in this study, we used two magnetoencephalographs that were connected together, and simultaneously recorded brain activity while two people took turns speaking in a word association/alphabet completion task. We compared the amplitude modulation of the alpha- and beta-band rhythms within each of 62 brain regions under semantic (word association; less predictable) and non-semantic (alphabet completion; more predictable) conditions. We found that the amplitudes of the rhythms were significantly different between conditions in a wide range of brain regions. Additionally, significant differences were observed in nearly the same group of brain regions after versus before each utterance, indicating that a wide range of brain areas is involved in predicting a conversation partner's next utterance. This result supports the idea that mentalizing, e.g., predicting another person's speech, plays an important role in conversation, and suggests that the neural network implicated in mentalizing extends over a wide range of brain regions.

Key words

Brain imaging, Verbal communication, Conversation, Mentalizing, Theory of mind

1. Introduction

Communication is essential to human social life. In particular, verbal communication plays an important role in human communication. In verbal communication, speech is produced by the complex movements of the vocal organs, e.g., movement of the tongue or lips, which are controlled by the brain. Similarly, when we hear words spoken by another person's via our auditory system, the brain perceives and recognizes words and understands their content. Various invasive and non-invasive measurements of brain function have revealed the processes involved in speech and language perception from a neuroscientific perspective. However, communication is not a stand-alone process; it is an interaction between two or more people. Therefore, to clarify the neural basis of communication, a one-person (1PN) setup, in which one subject is directed to speak or to listen to stimuli, is not sufficient (Sassa et al. 2007). Instead, a two-person (2PN) setup, in which two or more subjects take turns speaking and listening, is necessary (Hari et al. 2015; Redcay and Schilbach 2019). Ahn and colleagues conducted a study with a 2PN set-up and reported larger suppression (or referred to as desynchronization) of the alpha-band rhythm in the left temporal and right parietal regions during a social interaction compared with that with no interaction. They also reported significant gamma-band enhancement (synchronization) in the left temporal and frontal regions (Ahn et al. 2018). Furthermore, Kawasaki and colleagues conducted a study using alternating human-to-human and human-to-machine alphabet utterances and found that brain rhythms were synchronized when listening to a human-to-human utterance (Kawasaki et al. 2013).

During verbal communication, speakers and listeners mutually predict each other's mental state and plan their next utterance based on this prediction (Hari and Kujala 2009; Friston and Frith 2015a, 2015b). However, to the best of our knowledge, little attention has been paid to the prediction of mental states, i.e., the mentalizing function, during communication (Bögels et al. 2015; Kuhlen et al. 2017; Schurz et al. 2021). To clarify the neural basis of communication, an experimental setup

should include not only two people but also a task in which the participants are asked to mutually predict the utterances of one another (Sebanz et al. 2006; Hari and Kujala 2009; Jiang et al. 2021).

The simultaneous recording of multiple brain functions is called hyperscanning (Montague et al. 2002). Although many previous studies have conducted hyperscanning using electroencephalography, near-infrared spectroscopy, and functional magnetic resonance imaging (Balconi and Vanutelli 2017; Pérez et al. 2017; Hirsch et al. 2018; Liu et al. 2018; Redcay and Schilbach 2019; Czeszumski et al. 2020; Ayrolles et al. 2021; Kelsen et al. 2022; Pérez and Davis 2023), few hyperscanning studies have used magnetoencephalography (MEG) (Baess et al. 2012; Hirata et al. 2014; Ahn et al. 2018) because few research institutions own multiple MEG devices. Considering the temporal scale of listening and speaking during verbal communication, MEG hyperscanning is the suitable method for examining the neural basis of communication. This is because it has both high temporal and spatial resolutions, making it especially useful for estimating the source and tracking time variation of brain activities. In this study, we used a MEG hyperscanning system at Hokkaido University, in which two MEG devices were directly connected by optical fibers, to simultaneously record the brain activities of two subjects who conversed in a face-to-face manner (Watanabe et al. 2022).

The purpose of this study was to clarify the neural basis of predictions regarding the speech of a conversation partner as an important component of communication. To this end, we compared brain activity for speech prediction that was easy (non-semantic, few variations) and difficult (semantic, many variations). In order to exclude brain activities and motion artifacts associated with speech motions, we adopted a turn-taking paradigm in this study. Although inter-brain analysis is also possible from the MEG data obtained, single-brain analysis was performed in this study. We compared brain activities between that during thinking about own speech and that during waiting for partner's speech.

2. Materials and Methods

2. 1. Participants

Ten pairs of healthy Japanese speakers (11 women, 9 men), aged 22.3 ± 1.0 (mean ± SD) participated in this experiment. The participants in each pair knew one another. The sex of the dyads (same-sex or opposite-sex) was not controlled. All participants had no history of neurological disorders, and were right-handed according to the Edinburgh Handedness Inventory. The experimental protocol involving human subjects described in this report was approved by the Ethics Committees of the Faculty of Medicine, Hokkaido University. Written informed consent was obtained from each participant prior to the experiment.

2. 2. Devices

We used an MEG hyperscanning system composed of a 101-channel MEG system (site A; custom type, Elekta Neuromag) and a 306-channel MEG system (site B; VectorView, Elekta Neuromag) installed at Hokkaido University. The systems were located 500 m away from one another. An audiovisual interface and direct connections with fiber optic cables enabled simultaneous MEG recording during virtual face-to-face communication between each pair of participants (Watanabe et al. 2022). The transmission of trigger and auditory signals had a very minimal delay (1.95–3.90 μs and 3 ms, respectively) and video signals were transmitted with a short latency (80.1 ± 5.94 ms, range: 65–94 ms). Given reports that communication feels ‘unnatural’ when audio signals (voice) are heard before video signals (Kurita et al. 1994), the auditory signals were delayed by 90 ms to synchronize them with the video signals. Additionally, a half-mirror apparatus allowed the participants to communicate using direct eye contact (see Appendix A). The system was thus optimized for natural face-to-face communication.

2. 3. Experimental protocol

There were two experimental conditions, Natural and Patterned. In the Natural condition, the participants took turns uttering semantic words. In the Patterned condition, they uttered non-semantic syllables according to a set of directions.

Figure 1 shows a schematic diagram of one trial in the Natural condition. Each trial was triggered by a prompt, which was a word that appeared on a red or blue banner for 0.75 s. The prompt overlapped the eyes of the partner on the video image. The first participant was expected to utter a word associated with the prompt 5 s after of the prompt onset. The second participant then uttered a word associated with the word spoken by the first participant. The two participants spoke in an alternating pattern with a 5-s interval, which was indicated by gray fixation cross for 0.3 s. They could say a noun, adjective, verb, or adverb, “apple” (prompt) – “crisp” – “potato-chips” – “fry”, for instance (the participants spoke in Japanese in the real experiment; see Appendix B). The color of the prompt banner indicated which participant should speak first. For instance, red indicated the participant at site A and blue indicated the participant at site B, although this was counterbalanced. The alternate-speaking component of the trial continued until the presentation of a green fixation cross. The number of alternate speaking was set randomly between 4 and 6. Although MEGs were recorded during the whole experiment, only those from the beginning of the trial to the 4th utterance were analyzed. In the Natural condition, it was difficult to predict the partner’s next utterance because each word had many possible associations.

In the Patterned condition, a segment of the Japanese alphabet was presented in the prompt. The participants were asked to utter 3 or 4 consecutive syllables in the Japanese alphabet (not shown in the figure). Because there are no variations in the alphabet, it was easy to predict the partner’s next utterance. One block consisted of 16 trials, and 4 blocks were conducted, i.e., there were 32 trials in the Natural condition and 32 trials in the Patterned condition.

2. 4. Stimuli

The prompt stimuli consisted of 32 semantic words in the Natural condition and 32 non-semantic syllables in the Patterned condition. The semantic words were selected based on Miyazaki et al. (Miyazaki et al. 2001), who translated the ANEW (Affective Norms for English Words (Bradley and Lang 1999) system to Japanese, and assigned an affective valence to the nouns. In this study, only neutral words were selected. Four native Japanese speakers who were not participants in the main MEG recording task, but were similar in terms of age, thought up words associated with each noun and listed them independently. For the experiment, we selected 36 nouns for which the associated words uttered by the four participants were not expected to overlap. Four nouns were used for a training session that took place before the task, and the remaining 32 nouns were used as prompts in the Natural condition of the experiment. The non-semantic syllables comprised 3 or 4 consecutive syllables from the Japanese alphabet; 32 syllables with no specific meanings were used as the prompts in the Patterned condition. All words and syllables for the training and real tasks are listed in Appendix B.

The visual stimuli were projected from a projector located outside the magnetic shield room onto a back-projection screen inside the magnetic shield room. The distance between the screen and the participant's eyes was approximately 1.1 m at Site A and approximately 1.5 m at Site B. The screen was large enough to project the entire face of each participant at the opposite site. The size of the participant's face on the screen was approximately 15 cm high (visual angle of 5.7°–7.8°) and 11 cm wide (4.2°–5.7 °). The banners used to present the prompts were 3 cm high (1.1°–1.6°) and 5–8.5 cm wide (1.9°–4.4°), depending on the number of letters. To ensure natural communication, a camera was located behind a half-mirror to allow direct eye contact (see Appendix A). A non-magnetic speaker was placed in front of each participant, and a microphone was placed in front

or beside each participant. The sound volume was adjusted to a comfortable level.

2. 5. Experimental procedure

The participant pairs completed a short training session until they could perform the task without any problems. After that, they moved to their respective sites. At each site, head position indicator (HPI) coils were attached to the participant's head. The spatial position of the attached coils and 3 landmarks, i.e., the nasion (root of the nose) and the preauricular areas (left and right ear) were digitized, along with the outer shape of the head. Electrocardiogram and electrooculogram electrodes were also attached to the limbs and areas around the eyes to detect artifacts. Following electrode placement, each participant was seated on a measurement chair in a magnetically shielded room and briefed on the precautions to be taken during the experiment. Each participant completed a short training session to confirm that they were able to perform the task without any difficulty, and then began the first session. The participants took a short break between sessions if needed. All uttered words and syllables were recorded via a voice recorder for later analysis. Microphones outputs were acquired, together with the MEG data, to determine the onset and offset times of each utterance. After completing all of the sessions, each participant was asked to complete a questionnaire about the difficulty of the task using a 5-point scale that ranged from easy (1) to difficult (5).

2. 6. Data acquisition and processing

The frequency passbands ranged from 0.1 to 200 Hz, and signals were sampled at 600 Hz for both MEG devices. All MEG data processing was performed using Brainstorm (Tadel et al. 2011), which is freely available for download online under the GNU general public license (http://neuroimage.usc.edu/brainstorm). The MEG data were filtered with a 1–40 Hz passband and cleaned via independent component analysis (ICA) and Signal Space Projection (SSP). ICA was

applied to eliminate magnetocardiograms and/or ambient noise with 0–2 components, while SSP was applied to eliminate artifacts associated with eye blinking. Digitized points showing the outer shape of the participant's head with three landmarks were coregistered to the common template brain ICBM152 [Fonov et al., 2011] provided by Brainstorm. Source brain activities were estimated via minimum-norm estimation (MNE) with 15,002 vertices on the cortex of the template brain. The MNE algorithm and the standard arrangement of the vertices were also provided by Brainstorm. The noise covariance matrix was calculated using the 300-s window of MEG data recorded without the participant. The source brain activity on each vertex was band-pass filtered for the alpha-band (8–12 Hz) and beta-band (15–29 Hz), and the time traces of their amplitudes, $x(t)$, were calculated via Hilbert transformation. The amplitudes were normalized to the mean amplitude of the baseline, μ, to be x_{std} according to Eq. (1). The baseline was set as the instant 1 s after the presentation of the green fixation cross (see Fig. 1).

$$x_{std} = \frac{x(t)-\mu}{\mu} \times 100\ (\%) \quad (1)$$

2. 7. Data analysis

2. 7. 1. Behavior

We compared the number of characters in the participant utterances between the conditions (Natural/Patterned) using a paired t-test. The values of the difficulty ratings given in the questionnaire were also compared. The significance threshold was set at 5%, and statistical analysis was performed using SPSS ver. 22.

2. 7. 2. Rhythms

The normalized amplitudes of the alpha- and beta-band rhythms at the 15,002 vertices were averaged to survey the time course of whole brain activity.

For our main analysis of the alpha-band and beta-band rhythms, we averaged the normalized amplitudes x_{std} of each band within two time windows, a Think state and a Wait state, at each vertex. By surveying the voice recordings of the utterances, we found that each utterance ended at 1.7 ± 0.4 s (mean ± SD) from the onset of the gray fixation cross. Considering that auditory evoked brain responses could last for 0.5 s after the offset of an utterance by the participant or their partner (Hari and Puce 2017), we considered the possibility that brain activities associated with the processing of the auditory stimulus itself could be contaminated for 2.2 ± 0.4 s after the onset of the gray fixation cross. To avoid contamination, we eliminated the data collected within the 3 seconds after the onset of gray fixation cross, and selected the time windows 2 seconds before the self utterance (Think state) and that 2 seconds before the partner's utterance (Wait state) for further analysis. Note that the Think state for one participant was the Wait state for the partner (see Fig. 1).

We averaged the normalized amplitudes of the alpha- and beta-band rhythms over the vertices within each of the 62 brain regions based on the Mindboggle atlas (Klein et al. 2017). Even though two different MEG devices were used in this work, the signal sources estimated via MNE have been confirmed to coincide at least within a brain region in the Mindboggle atlas (Watanabe, under preparation). The mean normalized amplitude of each brain region was compared between conditions (Natural/Patterned) and states (Think/Wait) via a 3-way repeated measures (RM) ANOVA (condition (2) × state (2) × brain region (62)). The significance threshold was set at 5% and multiple comparisons were corrected via the Bonferroni method.

Finally, we calculated the time courses of the normalized amplitudes of several characteristic brain regions in which the normalized amplitudes were significantly different between the conditions.

3. Results

3. 1. Behavior

We counted the numbers of characters in the utterances from the 1st to 4th exchange, i.e., the targets of the rhythm analysis for each block. The results of 56 ± 6.1 characters for the Natural condition and 55 ± 4.2 characters for the Patterned condition (average ± SD) showed no significant difference between conditions ($p = 0.48$) according to a paired t-test. Besides, the means scores on the questionnaire about task difficulty were 3.1 ± 0.9 for the Natural condition and 2.7 ± 0.9 for the Patterned condition (average ± SD); there was no significant difference between the conditions ($p = 0.13$) according to a paired t-test.

3. 2. Alpha-band rhythm

The 3-way RM ANOVA with condition (Natural/Patterned) × state (Think/Wait) × brain region (62) revealed significant interactions between condition and brain region ($F_{(5.323, 101.146)} = 7.956$, $p < 0.001$) and between state and brain region ($F_{(3.222, 61.221)} = 11.679$, $p < 0.001$). The degrees of freedom were corrected via the Greenhouse-Geisser method. A post-hoc t-test showed significant suppression of the alpha-band rhythm in the Natural condition in 42 brain regions (Table 1, Fig. 2 (a)). In addition, another post-hoc t-test showed significant suppression of the alpha-band rhythm during the Wait state in 41 brain regions (Table 2, Fig. 2 (b)). Several examples of the time courses of normalized amplitudes at specific brain regions are shown in Fig. 3. The normalized amplitudes of the alpha-band rhythms in each condition were averaged among all participants ($n = 20$). In these brain regions, significant differences were observed for both the condition factor and state factor.

(Insert Table 1 and Table 2 close to Fig. 2)

3. 3. Beta-band rhythm

The 3-way RM ANOVA for condition (Natural/Patterned) × state (Think/Wait) × brain region (62) revealed significant interactions between condition and state ($F_{(1,19)} = 4.497$, p = 0.047) and condition and brain region ($F_{(6.622,125.814)} = 4.798$, p < 0.001). The degrees of freedom were corrected via the Greenhouse-Geisser method. A post-hoc t-test was conducted for each factor. In terms of the relationship between condition and state, no significant differences were observed between conditions in either the Think or Wait state. However, the normalized amplitude of the beta-band rhythm was larger for the Wait state than for the Think state in both the Natural and Patterned conditions (p < 0.001 for both conditions). In terms of the relationship between condition and brain region, we found that the normalized amplitudes of the beta-band rhythm were lower in the Natural condition compared with those in the Patterned condition in 4 brain regions. Conversely, the normalized amplitudes of the beta-band rhythm were higher in the Natural condition compared with those in the Patterned condition in 6 brain regions. The results of the post-hoc t-tests for the condition and brain region are shown in Table 3 and Fig. 4. Time courses of the normalized amplitude of the beta-band rhythm for some noteworthy brain regions are shown in Fig. 5. The normalized amplitudes of the beta-band rhythms in each condition were averaged among all participants (n = 20).

(Insert Table 3 close to Fig. 4)

4. Discussion

We found no significant differences in the number of characters of the uttered words/syllables between the conditions (Natural/Patterned). We also found no differences in the task difficulty ratings of the participants. Therefore, the different brain activities between the conditions appear to have relied on some higher brain functions, e.g., semantic processes or prediction of a partner's

utterances.

4. 1. Alpha-band rhythm

Post-hoc analysis of the normalized amplitude of the alpha-band rhythm revealed significant differences between conditions (Natural/Patterned) in 42 brain regions. For all of these brain regions, the normalized amplitudes of the alpha-band rhythm were lower in the Natural condition than in the Patterned condition (Table 1 and Fig. 2 (a)). Similarly, the normalized amplitudes of the alpha-band rhythm were lower during the Wait state than during the Think state in 41 brain regions (Table 2 and Fig. 2 (b)). The 42 and 41 brain regions, respectively, almost overlapped. Considering prior studies, suppression, or, event related desynchronization (ERD), of the alpha-band rhythm can be thought to reflect activity in relevant brain regions, and a greater number of regions were activated in the Natural condition versus the Patterned condition. At the same time, the regions activated during the Wait state compared with the Think states were almost the same.

Focusing on the middle temporal gyrus (Left) and supramarginal gyrus (Left), which roughly correspond to Wernicke's area (Dronkers et al. 2004), we observed clear suppression after hearing (approximately −4.3 s; see Fig. 3 (a)-L, (b)-L) and after speaking (approximately 0.7 s), regardless of condition. When hearing the utterance of their partner, the alpha-band rhythm of the participants underwent a similar degree of transient suppression in both the Natural and Patterned conditions. Although this suppression recovered within 1–2 s in the Patterned condition, it lasted longer in the Natural condition.

Ahn and colleagues reported that the left temporal and right parietal regions were activated (alpha-band rhythms were suppressed) under an interactive condition (Ahn et al. 2018). These regions were included in the 42 brain regions (Table 1 and Fig. 2 (a)) in which alpha-band rhythms were suppressed in the Natural condition compared with those in the Patterned condition in the

present study. There were both meanings and wide variations in the spoken words uttered in the Natural condition, but no meanings and less variations in the syllables uttered in the Patterned condition. This result suggests that verbal interactions or word exchanges with wider variation recruits brain regions that are associated with mentalizing (Apps et al. 2016) related to predicting a partner's utterances. That is, these results are consistent with Ahn and colleagues' previous work (Ahn et al. 2018).

Additionally, alpha-band rhythms were suppressed in the Wait state compared with those in the Think state in 41 brain regions that were almost the same as the 42 brain regions above. Interestingly, a larger number of brain regions were activated in the Wait state that in the Think state. In the Think state, while the participant processed their partner's utterances, the brain regions involved with semantic processing were likely activated. When the subject took care to make it easier for the partner to think about the next utterance, we expected that the subject's brain activity in the Think state would be higher (or wider), but this was not observed. In the Wait state, in which the participant was predicting their partner's next utterance, the brain regions associated with mentalizing were widely recruited.

Taken together, the higher brain activity in 42 brain regions observed in the Natural condition compared with the Patterned condition may have been related to both semantic processes and mentalizing. Indeed, while linguistic processes may be inseparable from mentalizing, the higher brain activity in 41 brain regions observed in the Wait state compared with the Think state could be attributed to mentalizing. The overlap between the 42 and 41 brain regions in the present work suggests that higher brain activity in the Natural condition is also mostly attributable to mentalizing.

4. 2. Beta-band rhythm

As described in the results section, the post-hoc t-tests for the condition and brain region revealed

that four brain regions had higher normalized beta-band rhythm amplitudes for the Patterned condition than for the Natural condition, while six brain regions had higher normalized amplitudes for the Natural condition than for the Patterned condition (Table 3 and Fig. 4). The four brain regions with higher amplitudes in the Patterned condition encompassed the occipital visual-related regions (Mckiernan et al. 2003), i.e., the bilateral cuneus and pericalcarine, while the six brain regions with higher amplitudes for the Natural condition were mainly right lateral regions.

The normalized beta-band rhythm amplitude was larger during the Wait state than the Think state. Consistent with the normalized alpha-band rhythm amplitude, this result suggests that the whole brain was more activated during the Wait state.

Examples of the time course of the normalized amplitude of the visual-related brain regions (bilateral cuneus, Fig. 5 (a)) show that they were larger for the Patterned condition during the Think state, but not different during the Wait state. In contrast, those of the right lateral regions (Fig. 5 (b)-R, (c)-R) were larger for the Natural condition during the Wait state, but not different during the Think state. These results suggest that the beta-band rhythm amplitude is modulated by multiple factors. The beta-band rhythm in the visual-related regions is known to be modulated by attention and top-down control of sensory input (Mckiernan et al. 2003). Therefore, the larger amplitude in the Patterned condition during the Think state observed in the occipital visual-related regions might be explained by higher attention or top-down control required in the Patterned condition. Previous studies on the relationship between beta-band rhythms and mentalizing are limited. However, the larger amplitude observed in the right lateral regions for the Natural condition during the Wait state could be associated with mentalizing functions (Kennedy and Adolphs 2012; Ciaramidaro et al. 2014; Bilek et al. 2015; Stolk et al. 2015; Alkire et al. 2018). Previous work has suggested that the prefrontal cortices are involved in mentalizing (Amodio and Frith 2006; Frith and Frith 2007; Stolk et al. 2015; Kuhlen et al. 2017). However, our results did not show any difference between

conditions in the prefrontal cortices. Nevertheless, the brain regions involving in mentalizing have not commonly been described in previous works. The mentalizing function in the prefrontal cortices is less likely to be recruited in linguistic tasks.

5. Conclusion

Our results indicate that mentalizing is stronger while waiting for a conversation partner to speak (Wait state), rather than when thinking about speaking (Think state), during an alternative speaking task modeled on a conversation. This likely reflects predictions by each conversant about the other's next utterance. Alpha-band rhythm suppression could be related to mentalizing in large brain regions, including the bilateral parietal, occipital, and base regions, while beta-band rhythm was enhanced in the right lateral regions. Previous studies have adopted joint attention, moral judgments, false beliefs, and others as mentalizing tasks (Schurz et al. 2021). This study demonstrates that brain activities relating to mentalizing can also be observed in an alternative-speaking task that is similar to daily communication or conversations. This study reports the involvement of some brain regions that have not been previously related to metalizing, including the pericalcarine, lingual gyrus, and cuneus. It is likely that mentalizing is not represented in specific brain regions but that it relies on a network involving multiple brain regions. Future intra/inter brain correlation/causality analysis is needed to reveal these dynamics (Lotter et al. 2023). The brain regions identified in this study are candidates for the seeds in future network analyses.

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP20H04496. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We thank emeritus prof. Shinya Kuriki of Hokkaido University, Japan and Dr. Jared Boasen of HEC Montréal, Canada for their useful discussion and suggestions. We also thank Tsuyoshi Sonehara for his help with the experimental setup, especially the direct-eye-contact system (Appendix A). We also thank our colleagues and the graduate and undergraduate students for their help and discussion; especially Prof. Takuya Saito and Dr. Atsuhito Toyomaki at Hokkaido University Hospital, Prof. Kazuyori Yagyu at Health Sciences University of Hokkaido, and Dr. Yui Murakami at Japan Healthcare University, Sapporo, Japan.

Corresponding author: Koichi Yokosawa
(Address) Faculty of Health Sciences, Hokkaido University, N-12, W-5, Kita-ku, Sapporo 060-0812, Hokkaido, Japan
(Tel) +81-11-706-2828
(E-mail) yokosawa@med.hokudai.ac.jp

Conflict of Interests

The authors have no conflicts of interest to declare.

References

Ahn S, Cho H, Kwon M, Kim K, Kwon H, Kim BS, Chang WS, Chang JW, Jun SC. Interbrain phase synchronization during turn-taking verbal interaction—a hyperscanning study using simultaneous EEG/MEG. Hum Brain Mapp. 2018:39:171–188.

Alkire D, Levitas D, Warnell KR, Redcay E. Social interaction recruits mentalizing and reward systems in middle childhood. Hum Brain Mapp. 2018:39:3928–3942.

Amodio DM, Frith CD. Meeting of minds: The medial frontal cortex and social cognition. Nat Rev Neurosci. 2006:7: 268–277.

Apps MAJ, Rushworth MFS, Chang SWC. The anterior cingulate gyrus and social cognition: Tracking the motivation of others. Neuron. 2016:90: 692–707.

Ayrolles A, Brun F, Chen P, Djalovski A, Beauxis Y, Delorme R, Bourgeron T, Dikker S, Dumas G. HyPyP: A hyperscanning python pipeline for inter-brain connectivity analysis. Soc Cogn Affect Neurosci. 2021:16:72–83.

Baess P, Zhdanov A, Mandel A, Parkkonen L, Hirvenkari L, Mäkelä JP, Jousmäki V, Hari R. MEG dual scanning: A procedure to study real-time auditory interaction between two persons. Front Hum Neurosci. 2012:6:Article 83.

Balconi M, Vanutelli ME. Cooperation and competition with hyperscanning methods: Review and future application to emotion domain. Front Comput Neurosci. 2017:11:Article 86.

Bilek E, Ruf M, Schäfer A, Akdeniz C, Calhoun VD, Schmahl C, Demanuele C, Tost H, Kirsch P, Meyer-Lindenberg A. Information flow between interacting human brains: Identification, validation, and relationship to social expertise. Proc Natl Acad Sci USA. 2015:112:5207–5212.

Bögels S, Barr DJ, Garrod S, Kessler K. Conversational interaction in the scanner: Mentalizing during language processing as revealed by MEG. Cerebral Cortex. 2015:25:3219–3234.

Bradley MM, Lang PJ. Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. Technical report C-1, The center for research in Psychology, University of Florida (FL); 1999.

Ciaramidaro A, Becchio C, Colle L, Bara BG, Walter H. Do you mean me? Communicative intentions recruit the mirror and the mentalizing system. Soc Cogn Affect Neurosci. 2014:9:909–916.

Czeszumski A, Eustergerling S, Lang A, Menrath D, Gerstenberger M, Schuberth S, Schreiber F, Rendon ZZ, König P. Hyperscanning: A valid method to study neural inter-brain underpinnings of social interaction. Front Hum Neurosci. 2020:14:Article 39.

Dronkers NF, Wilkins DP, van Valin RD, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. Cognition. 2004:92:145–177.

Fonov VS, Evans VS, Botteron K, Almli VS, McKinstry RC, Collins DL, the Brain Development Cooperative Group. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage. 2011:54:390–405.

Friston KJ, Frith CD. Active inference, Communication and hermeneutics. Cortex. 2015b:68:129–143.

Frith CD, Frith U. Social cognition in humans. Current Biology. 2007:17: R724–R732.

Hari R, Henriksson L, Malinen S, Parkkonen L. Centrality of Social Interaction in Human Brain Function. Neuron. 2015:88:181–193.

Hari R, Kujala M. Brain Basis of Human Social Interaction: From Concepts to Brain Imaging. Physiol Rev. 2009:89:453–479.

Hari R, Puce A. 2017. Auditory responses. In: MEG-EEG Primer, Oxford University Press; 2017. p. 200–213.

Hirata M, Ikeda T, Kikuchi M, Kimura T, Hiraishi H, Yoshimura Y, Asada M. Hyperscanning MEG for understanding mother-child cerebral interactions. Front Hum Neurosci. 2014:8: Article 118.

Hirsch J, Noah JA, Zhang X, Dravida S, Ono Y. A cross-brain neural mechanism for human-to-human verbal communication. Soc Cogn Affect Neurosci. 2018:13:907–920.

Jiang J, Zheng L, Lu C. A hierarchical model for interpersonal verbal communication. Soc Cogn Affect Neurosci. 2021:16:246–255.

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. 2013:3:Article number 1692.

Kelsen BA, Sumich A, Kasabov N, Liang SHY, Wang GY. What has social neuroscience learned from hyperscanning studies of spoken communication? A systematic review. Neurosci Biobehav Rev. 2022:132:1249–1262.

Kennedy DP, Adolphs R. The social brain in psychiatric and neurological disorders. Trends Cogn Sci. 2012:16: 559–572.

Klein A, Ghosh SS, Bao FS, Giard J, Häme Y, Stavsky E, Lee N, Rossa B, Reuter M, Chaibub Neto E, Keshavan A. Mindboggling morphometry of human brains. PLoS Comput Biol. 2017:13: e1005350.

Kuhlen AK, Bogler C, Brennan SE, Haynes JD. Brains in dialogue: Decoding neural preparation of speaking to a conversational partner. Soc Cogn Affect Neurosci. 2017:12:871–880.

Kurita T, Id S, Kitawaki N. Effects of transmission delay in audiovisual communication, Electronics and Communications in Japan, Part 1, 1994:77:331–339.

Liu D, Liu S, Liu X, Zhang C, Li A, Jin C, Chen Y, Wang H, Zhang X. Interactive brain activity: Review and progress on EEG-based hyperscanning in social interactions. Front Psychol. 2018: 9:Article 1862.

Lotter LD, Kohl SH, Gerloff C, Bell L, Niephaus A, Kruppa JA, Dukart J, Schulte-Rüther M, Reindl V, Konrad K. Revealing the neurobiology underlying interpersonal neural synchronization with multimodal data fusion. Neurosci Biobehav Rev. 2023:146:105042.

Mckiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. 2003:15:394–408.

Miyazaki T, Motoyama H, Hishitani S. Emotional values for nouns and epithets: Standardization on the positive-negative dimension (in Japanese). Japanese Journal of Mental Imagery. 2001:1: 48–59.

Montague PR, Berns GS, Cohen JD, McClure SM, Pagnoni G, Dhamala M, Wiest MC, Karpov I, King RD, Apple N, Fisher RE. Hyperscanning: Simultaneous fMRI during linked social interactions. Neuroimage. 2002:16:1159–1164.

Pérez A, Carreiras M, Duñabeitia JA. 2017. Brain-To-brain entrainment: EEG interbrain synchronization while speaking and listening. Sci Rep. 7.

Pérez A, Davis MH. Speaking and listening to inter-brain relationships. Cortex. 2023:159:54–63.

Redcay E, Schilbach L. Using second-person neuroscience to elucidate the mechanisms of social interaction. Nat Rev Neurosci. 2019:20:495–505.

Sassa Y, Sugiura M, Jeong H, Horie K, Sato S, Kawashima R. Cortical mechanism of communicative speech production. Neuroimage. 2007:37:985–992.

Schurz M, Radua J, Tholen MG, Maliske L, Margulies Daniel S, Mars RB, Sallet J, Kanske P. Toward a hierarchical model of social cognition—A neuroimaging meta-analysis and integrative review of empathy and theory of mind—. Psychol Bull. 2021:147:293–327.

Sebanz N, Bekkering H, Knoblich G. Joint action: Bodies and minds moving together. Trends Cogn Sci. 2006:10:70–76.

Stolk A, D'Imperio D, di Pellegrino G, Toni I. Altered communicative decisions following ventromedial prefrontal lesions. Current Biology. 2015:25:1469–1474.

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM. Brainstorm: A user-friendly application for MEG/EEG analysis. Comput Intell Neurosci 2011. 2011:879716.

Watanabe H, Shimojo A, Yagyu K, Sonehara T, Takano K, Boasen J, Shiraishi H, Yokosawa K, Saito T. Construction of a fiber-optically connected MEG hyperscanning system for recording brain activity during real-time communication. PLoS One. 2022:17:e0270090.

Table 1. Brain regions with significant differences in the normalized alpha-band rhythm amplitude between conditions (L: Left, R: Right, Bonferroni correction).

Alpha-band (Patterned > Natural)

	Region	p-value		Region	p-value
1	caudalanteriorcingulate L	0.026	22	medialorbitofrontal L	0.048
2	caudalanteriorcingulate R	0.043	23	middletemporal L	0.010
3	cuneus L	＜0.001	24	middletemporal R	0.048
4	cuneus R	＜0.001	25	paracentral L	0.001
5	entorhinal L	0.004	26	paracentral R	0.002
6	entorhinal R	0.014	27	parahippocampal L	0.002
7	fusiform L	＜0.001	28	parahippocampal R	0.005
8	fusiform R	0.010	29	pericalcarine L	＜0.001
9	inferiorparietal L	0.002	30	pericalcarine R	＜0.001
10	inferiorparietal R	0.003	31	posteriorcingulate L	0.005
11	inferiortemporal L	0.004	32	posteriorcingulate R	0.002
12	inferiortemporal R	0.045	33	precuneus L	0.002
13	insula L	0.026	34	precuneus R	0.002
14	insula R	0.037	35	rostralanteriorcingulate L	0.027
15	isthmuscingulate L	0.002	36	superiorparietal L	0.003
16	isthmuscingulate R	0.001	37	superiorparietal R	0.002
17	lateraloccipital L	＜0.001	38	superiortemporal L	0.024
18	lateraloccipital R	＜0.001	39	superiortemporal R	0.034
19	lateralorbitofrontal L	0.031	40	supramarginal L	0.040
20	lingual L	0.001	41	supramarginal R	0.042
21	lingual R	＜0.001	42	transversetemporal R	0.039

Table 2. Brain regions with significant differences in the normalized alpha-band rhythm amplitude between states (L: Left, R: Right, Bonferroni correction).

Alpha-band (Think > Wait)

	Region	p-value		Region	p-value
1	caudalanteriorcingulate R	0.034	21	medialorbitofrontal L	0.009
2	cuneus L	＜0.001	22	medialorbitofrontal R	0.008
3	cuneus R	＜0.001	23	middletemporal L	0.015
4	entorhinal L	0.005	24	middletemporal R	0.010
5	entorhinal R	0.002	25	paracentral L	0.008
6	fusiform L	0.000	26	paracentral R	0.012
7	fusiform R	0.001	27	parahippocampal L	0.001
8	inferiorparietal L	＜0.001	28	parahippocampal R	0.003
9	inferiorparietal R	＜0.001	29	pericalcarine L	＜0.001
10	inferiortemporal L	0.001	30	pericalcarine R	＜0.001
11	inferiortemporal R	0.002	31	posteriorcingulate L	0.005
12	insula R	0.039	32	posteriorcingulate R	0.007
13	isthmuscingulate L	0.001	33	precuneus L	＜0.001
14	isthmuscingulate R	0.001	34	precuneus R	＜0.001
15	lateraloccipital L	0.001	35	rostralanteriorcingulate L	0.016
16	lateraloccipital R	0.001	36	rostralanteriorcingulate R	0.011
17	lateralorbitofrontal L	0.018	37	superiorparietal L	＜0.001
18	lateralorbitofrontal R	0.006	38	superiorparietal R	0.001
19	lingual L	0.001	39	superiortemporal R	0.027
20	lingual R	0.001	40	supramarginal L	0.019
			41	supramarginal R	0.006

Table 3. Brain regions with significant differences in the normalized beta-band rhythm amplitude between conditions (L: Left, R: Right, Bonferroni correction).

Beta-band

	Region (Patterned > Natural)	p-value
1	cuneus L	0.008
2	cuneus R	0.003
3	pericalcarine L	0.002
4	pericalcarine R	0.007

	Region (Natural > Patterned)	p-value
1	inferiortemporal R	0.025
2	insula R	0.020
3	parsopecularis R	0.037
4	parsorbitalis R	0.021
5	superiortemporal R	0.025
6	transversetemporal R	0.037

Appendix B. List of semantic words (left two columns) and non-semantic syllables (right two columns) presented as prompts. Japanese words and syllables were used in the experiment (a), and English translations (b) are shown.

(a)

Natural condition		Patterned condition	
For the task			
はさみ	かめら	あいう	にぬね
ろうか	びーる	いうえ	ねのは
ふくろ	ばたー	えおか	のはひ
けがわ	こーひー	かきく	あいうえ
きつね	えんぴつ	きくけ	うえおか
とこや	しんぶん	けこさ	えおかき
でんわ	やまごや	こさし	おかきく
さかな	はんかち	さしす	かきくけ
おかね	すぽんじ	しすせ	くけこさ
いんく	へあぴん	せそた	けこさし
らじお	ほうたい	そたち	こさしす
ちーず	げんかん	たちつ	さしすせ
かざん	かーてん	ちつて	すせそた
きのこ	くちびる	てとな	せそたち
たまご	とーすと	となに	そたちつ
くるま	らいおん	なにぬ	たちつて
For the training			
ごるふ	ひろば	いうえ	たちつて
ふくろう	たくしー	なにぬね	そたちつ
		あいう	けこさし
		こさし	かきくけ

(b)

Natural condition		Patterned condition	
For the task			
scissors	camera	a-i-u	ni-nu-ne
hallway	beer	i-u-e	ne-no-ha
bag	butter	e-o-ka	no-ha-hi
fur	coffee	ka-ki-ku	a-i-u-e
fox	pencil	ki-ku-ke	u-e-o-ka
barber	newspaper	ke-ko-sa	e-o-ka-ki
phone	mountain hut	ko-sa-shi	o-ka-ki-ku
fish	handkerchief	sa-shi-su	ka-ki-ku-ke
money	sponge	shi-su-se	ku-ke-ko-sa
ink	hairpin	se-so-ta	ke-ko-sa-shi
radio	bandage	so-ta-chi	ko-sa-shi-su
cheese	entranceway	ta-chi-tsu	sa-shi-su-se
volcano	curtain	chi-tsu-te	su-se-so-ta
mushroom	lip	te-to-na	se-so-ta-chi
egg	toast	to-na-ni	so-ta-chi-tsu
car	lion	na-ni-nu	ta-chi-tsu-te
For the training			
golf	plaza	i-u-e	ta-chi-tsu-te
owl	taxi	na-ni-nu-ne	so-ta-chi-tsu
		a-i-u	ke-ko-sa-shi
		ko-sa-shi	ka-ki-ku-ke

Captions of figures

Figure 1. Experimental setup and the time course of one trial when participant A was instructed to speak first. Two participants spoke alternately, triggered by a grey fixation cross that appeared for 5 s, until a green cross appeared. The color (red/blue) of the banner behind the prompt indicated who would be the first speaker. The time windows for analysis were set as 2 seconds before speaking (Think state) and 2 seconds before the participant heard their partner's utterance (Wait state). Magnetoencephalogram of two participants were recorded simultaneously. See Appendix A for direct-eye-contact system.

Figure 2. Illustration of brain regions shown in Table 1 and 2. In all 42 brain regions shown in blue, normalization amplitudes were significantly larger in the Patterned condition than in the Natural condition ((a); see Table 1). In all 41 brain regions shown in blue, normalization amplitudes were significantly larger in the Think state than in the Wait state ((b); see Table 2). L: Left, R: Right, A: Anterior, P: Posterior.

Figure 3. Several examples of time courses of normalized alpha-band rhythm amplitudes, mean ± SE among participants (see Table 1 and 2; L: Left, R: Right). In all these brain regions, significant differences were observed in both the condition factor and state factor.

Figure 4. Illustration of brain regions shown in Table 3. Normalized amplitudes were larger in the Patterned condition than in the Natural condition in the 4 brain regions indicated in blue, and vice versa for the 6 brain regions indicated in red. L: Left, R: Right, A: Anterior, P: Posterior.

Figure 5. Several examples of time courses of normalized beta-band rhythm amplitudes, mean ± SE among participants. The normalized amplitude was larger for the Patterned or Natural condition, depending on the brain region (see Table 3; L: Left, R: Right).

Appendix A. The original setting (a) and improved setting for direct-eye-contact adopted in this study (b). In both settings, a video image of the partner's face was projected onto a back-projection screen by a projector located outside the magnetically shielded room. In the original setting, the camera was located at the top of the screen, hence, it looked down on the participant's face. Thus, the gazes of the participant and the partner were mismatched. To match the gazes, a half-mirror (c) was inserted into the setup. The image of the partner's face was transmitted directly to the half mirror, while the participant's own face image was reflected by 90 degrees and captured by the camera. Additionally, in the original setting at site A, the video image of the partner's face was projected onto a lower position ((a) left). To correct and match the gazes, two mirrors were inserted to function like a periscope ((b) left and (d)). In this setting, all mirrors were front surface reflective to ensure image reflection accuracy. Figure (c) and (d) show the detailed sizes of the half mirror and the "periscope" together with their black containers. This setting made face-to-face communication more natural.

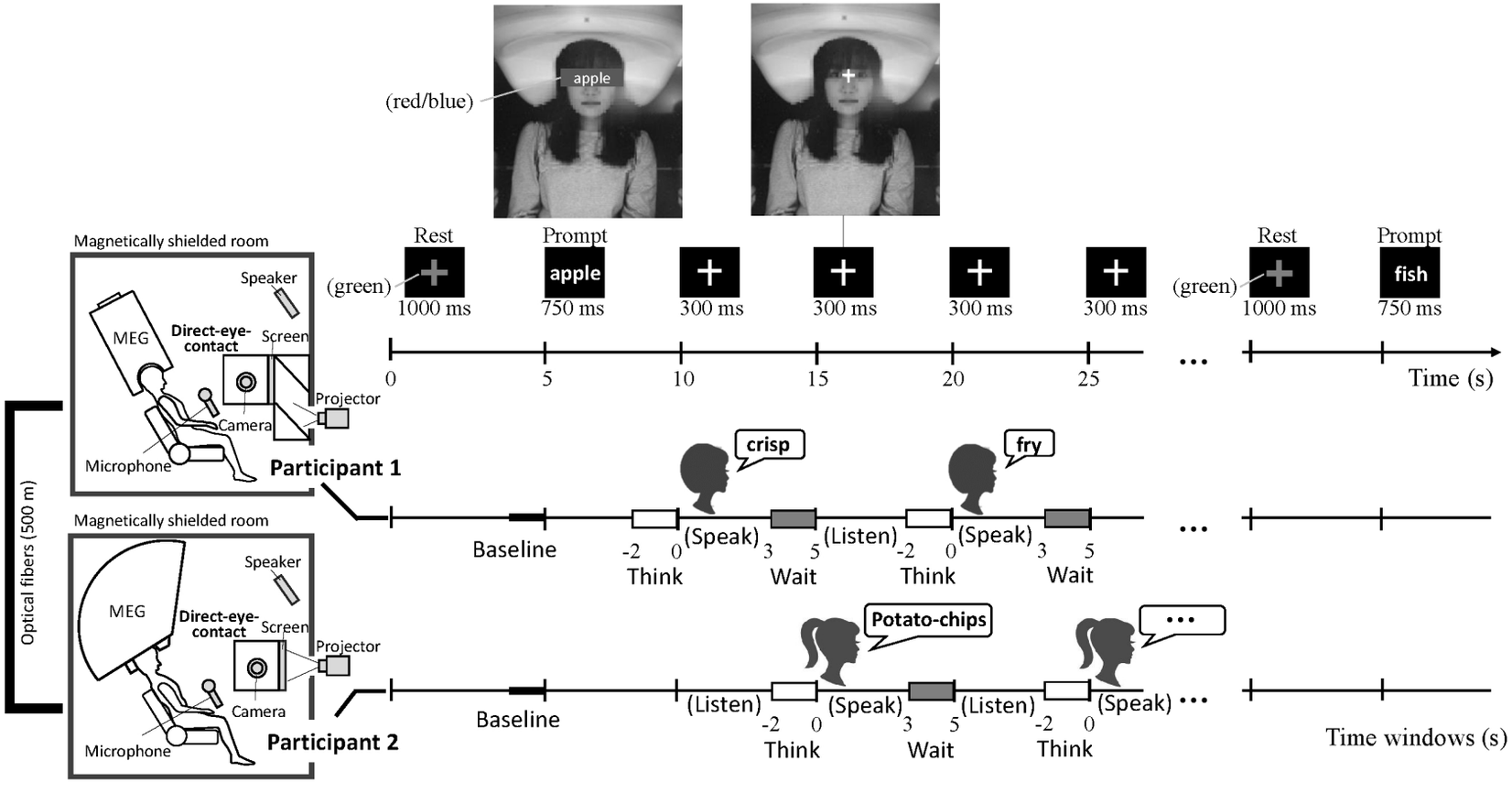

(red/blue)
apple
Rest
Prompt
(green)
apple
1000 ms
750 ms
300 ms
300 ms
300 ms
300 ms
(green)
Rest
Prompt
fish
1000 ms
750 ms
0
5
10
15
20
25
Time (s)
Magnetically shielded room
Speaker
MEG
Direct-eye-contact
Screen
Projector
Camera
Microphone
Participant 1
Optical fibers (500 m)
Magnetically shielded room
Speaker
MEG
Direct-eye-contact
Screen
Projector
Camera
Microphone
Participant 2
crisp
fry
Baseline
-2
0
(Speak)
3
5
(Listen)
-2
0
(Speak)
3
5
Think
Wait
Think
Wait
Potato-chips
Baseline
(Listen)
(Speak)
(Listen)
(Speak)
-2
0
3
5
-2
0
Think
Wait
Think
Time windows (s)

Fig. 1

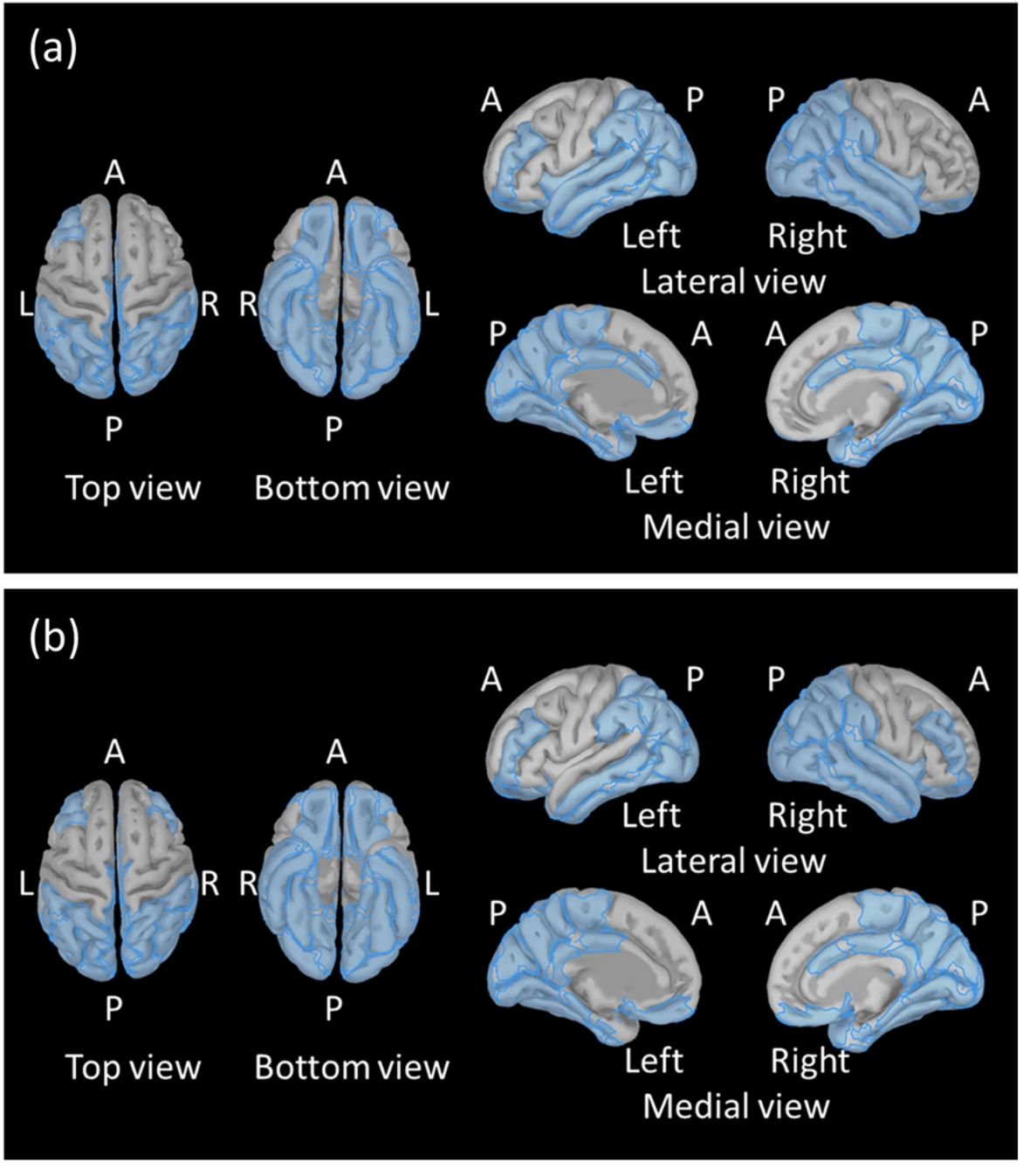

(a)
A
L
R
P
Top view
A
R
L
P
Bottom view
A
P
Left
P
A
Right
Lateral view
P
A
Left
A
P
Right
Medial view
(b)
A
L
R
P
Top view
A
R
L
P
Bottom view
A
P
Left
P
A
Right
Lateral view
P
A
Left
A
P
Right
Medial view

Fig. 2

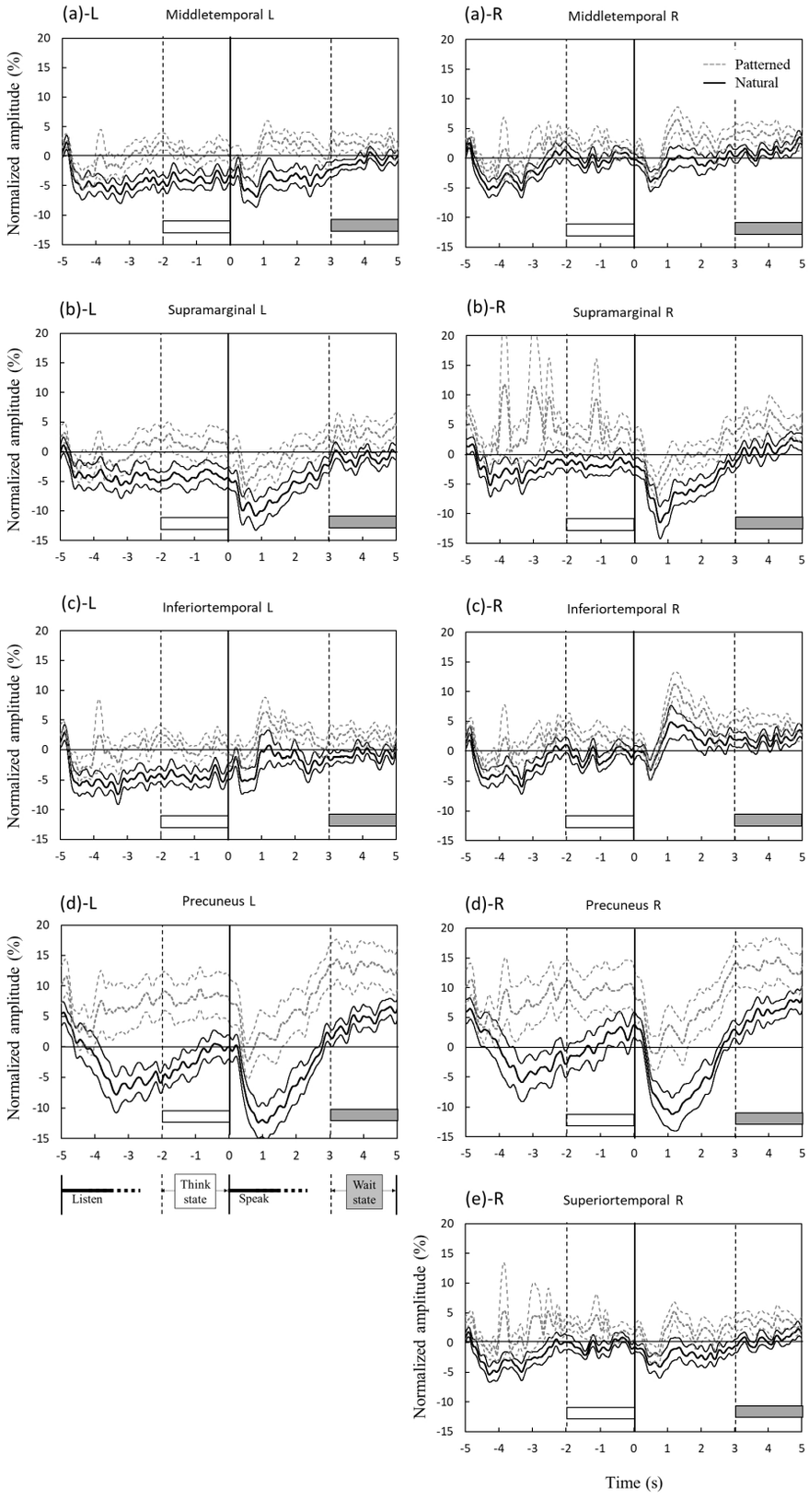

(a)-L
Middletemporal L
(a)-R
Middletemporal R
Patterned
Natural
(b)-L
Supramarginal L
(b)-R
Supramarginal R
(c)-L
Inferiortemporal L
(c)-R
Inferiortemporal R
(d)-L
Precuneus L
(d)-R
Precuneus R
(e)-R
Superiortemporal R
Normalized amplitude (%)
Time (s)
Listen
Think state
Speak
Wait state

Fig. 3

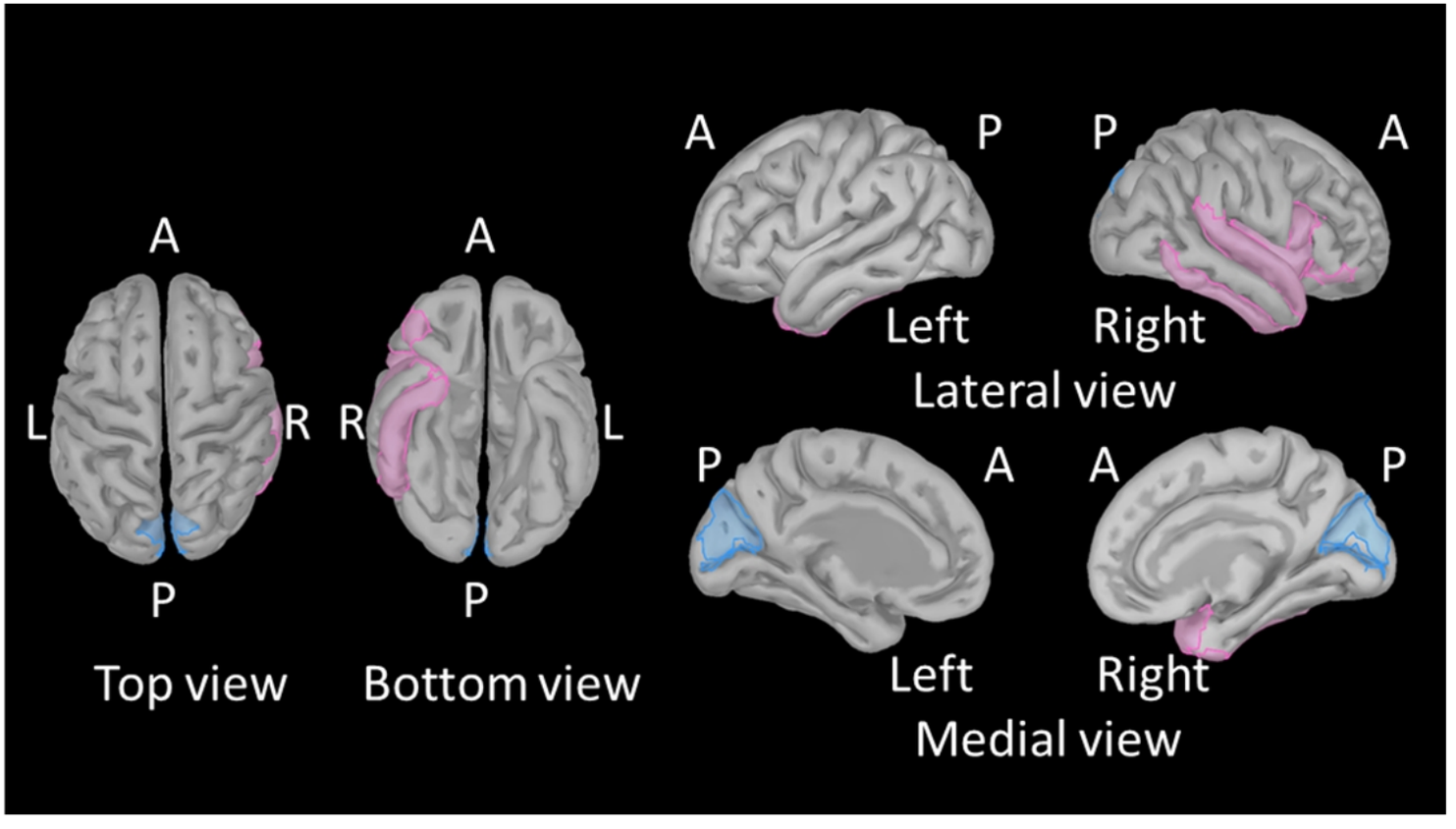

A
L
R
P
Top view
A
R
L
P
Bottom view
A
P
Left
P
A
Right
Lateral view
P
A
Left
A
P
Right
Medial view

Fig. 4

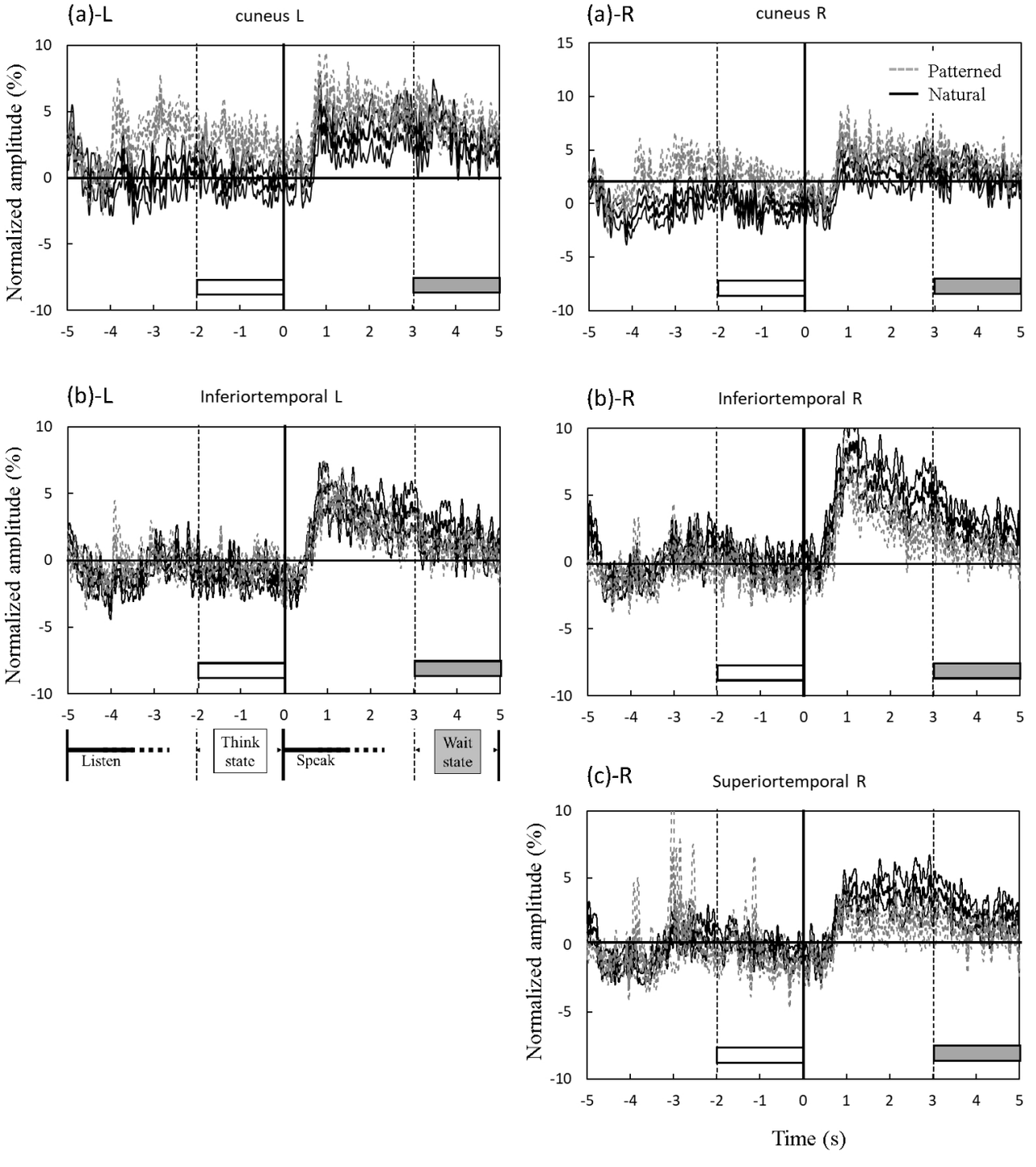
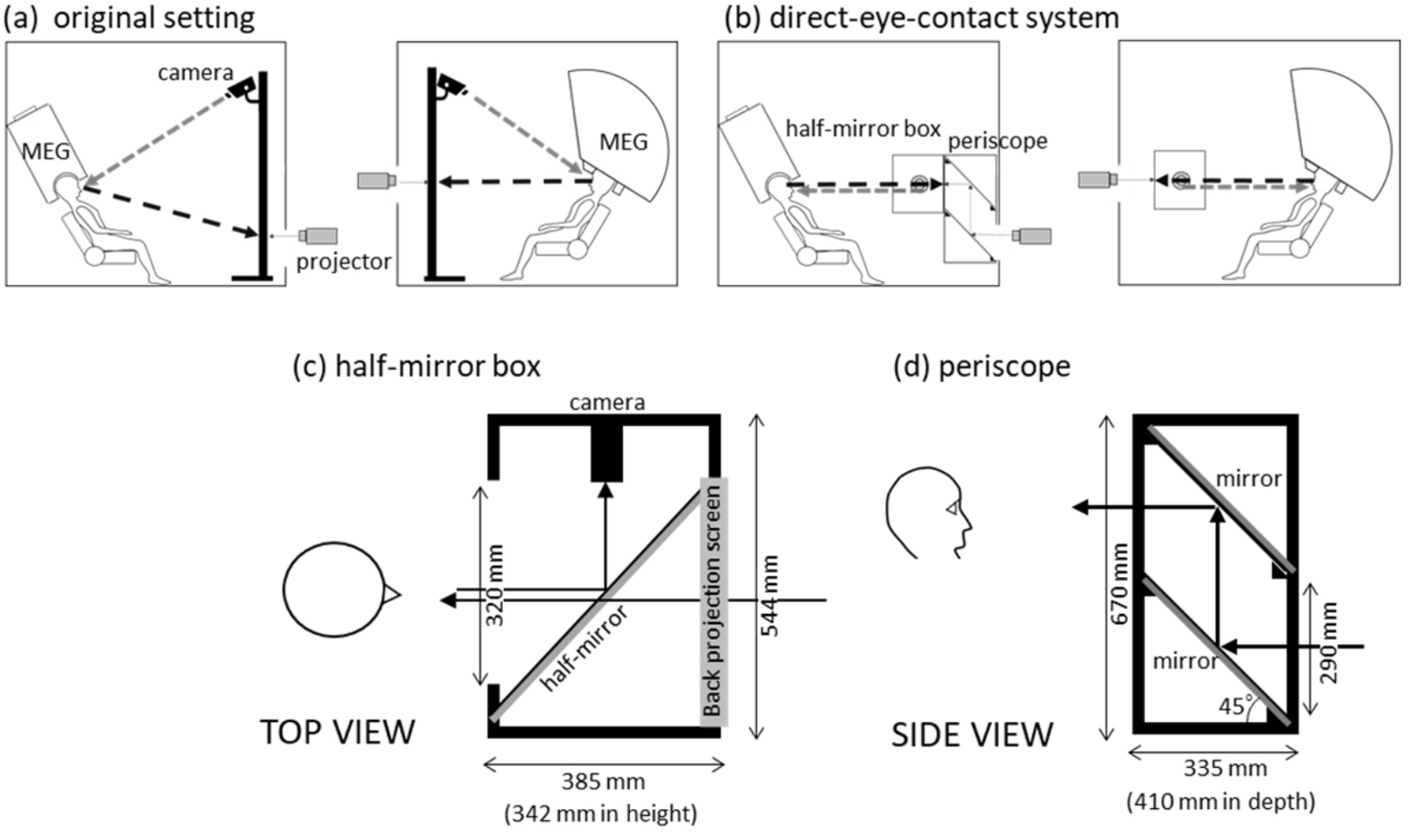

(a)-L
cuneus L
(a)-R
cuneus R
Patterned
Natural
Normalized amplitude (%)
(b)-L
Inferiortemporal L
(b)-R
Inferiortemporal R
Listen
Think state
Speak
Wait state
(c)-R
Superiortemporal R
Time (s)

Fig. 5


(a) original setting
camera
MEG
projector
MEG
(b) direct-eye-contact system
half-mirror box
periscope
(c) half-mirror box
camera
320 mm
half-mirror
Back projection screen
544 mm
TOP VIEW
385 mm
(342 mm in height)
(d) periscope
mirror
670 mm
mirror
290 mm
45°
SIDE VIEW
335 mm
(410 mm in depth)

Appendix A