Received: 21 May 2024 | Accepted: 23 September 2024

DOI: 10.1111/ejn.16563

RESEARCH REPORT

EJN European Journal of Neuroscience FENS WILEY

Differential modulations of theta and beta oscillations by audiovisual congruency in letter-speech sound integration

Dongyang Yan[1] | **Ayumi Seki**[2]

[1]Graduate School of Education, Hokkaido University, Sapporo, Japan

[2]Faculty of Education, Hokkaido University, Sapporo, Japan

Correspondence
Dongyang Yan, Graduate School of Education, Hokkaido University, Nishi-7, Kita-11, Kita-ku, Sapporo, Hokkaido, Japan.
Email: ydyxj@hotmail.com

Funding information
Japan Society for the Promotion of Science, Grant/Award Number: JP17H02713

Edited by: Ali Mazaheri

Abstract

The integration of visual letters and speech sounds is a crucial part of learning to read. Previous studies investigating this integration have revealed a modulation by audiovisual (AV) congruency, commonly known as the congruency effect. To investigate the cortical oscillations of the congruency effects across different oscillatory frequency bands, we conducted a Japanese priming task in which a visual letter was followed by a speech sound. We analyzed the power and phase properties of oscillatory activities in the theta and beta bands between congruent and incongruent letter-speech sound (L-SS) pairs. Our results revealed stronger theta-band (5–7 Hz) power in the congruent condition and cross-modal phase resetting within the auditory cortex, accompanied by enhanced inter-trial phase coherence (ITPC) in the auditory-related areas in response to the congruent condition. The observed congruency effect of theta-band power may reflect increased neural activities in the left auditory region during L-SS integration. Additionally, theta ITPC findings suggest that visual letters amplify neuronal responses to the following corresponding auditory stimulus, which may reflect the differential cross-modal influences in the primary auditory cortex. In contrast, decreased beta-band (20–35 Hz) oscillatory power was observed in the right centroparietal regions for the congruent condition. The reduced beta power seems to be unrelated to the processing of AV integration, but may be interpreted as the brain response to predicting auditory sounds during language processing. Our data provide valuable insights by indicating that oscillations in different frequency bands contribute to the disparate aspects of L-SS integration.

KEYWORDS
audiovisual integration, beta, inter-trial phase coherence, letter-speech sound integration, neural oscillations, reading, theta

Abbreviations: ANOVA, analysis of variance; AV, audiovisual; AVc, audiovisual congruent; AVi, audiovisual incongruent; EEG, electroencephalography; ERP, event-related potential; *dB*, decibel; Hz, Hertz; ITPC, inter-trial phase coherence; kΩ, kiloohm; kHz, kilohertz; L-SS, letter-speech sound; MEG, magnetoencephalography; ms, *millisecond*; *ROI*, regions of interest; SD, standard deviation; STC, superior temporal cortex.

1 | INTRODUCTION

Reading is an essential skill for modern life and education. Unlike spoken language, which has evolved biologically over time, reading is relatively recent cognitive achievement in human evolution, having existed for only a few thousand years (Liberman, 1992). Although audiovisual (AV) integration of visual letters with speech sounds and visual speech with speech sounds may share common neural circuitry (Calvert et al., 1999, 2000; Dehaene et al., 2015), research has shown that compared to visual speech, visual letters have weaker modulations on the sound processing (Stekelenburg et al., 2018). This is likely because the AV associations are arbitrary and require explicit learning and practice later in life. In contrast, lip reading relies on naturally developed associations between speech and visual information, which can be easily learnt in everyday conversation (Paulesu et al., 2003). Therefore, learning to read is a complex and challenging process that entails mastering both spoken and written languages. It is thought that the core of literacy acquisition lies in the integration of visual letters (or graphemes) and speech sounds (phonemes) (Ehri, 2005). The ability to establish AV associations between visual letters and corresponding speech sounds plays an important role in reading acquisition. Automatic activation of these associations is essential for achieving fluency in reading, failure of which is often linked to reading difficulties (Blomert, 2011).

The neural correlates underlying AV integration (e.g. McGurk illusion) have been widely explored in the literature by manipulating the congruency of different AV inputs (e.g. Roa Romero et al., 2015). Although letter-speech sound (L-SS) integration may involve different types of AV integration processing, early neuroimaging studies investigating L-SS integration using similar approaches have revealed a modulation by AV congruency, known as the congruency effect (Blau et al., 2008; Raij et al., 2000; van Atteveldt et al., 2004). These findings suggest that successful integration of letters and speech sounds would elicit significantly greater neural activities in response to the congruent condition compared to that to the incongruent condition. The increased neural responses were not only observed in multimodal areas, that is, the superior temporal cortex (STC), but also in sensory-specific cortices, including the auditory (i.e. Heschl's sulcus/planum temporale) and visual cortices. Furthermore, a comprehensive review of these previous neuroimaging studies (van Atteveldt et al., 2009) proposed the feedback hypothesis, suggesting that audiovisually presented visual letters and speech sounds (i.e. auditorily presented phonemes) are initially integrated within the STC, and the integrated information is subsequently fed back to the auditory cortex to modulate activities in that area. Moreover, the visual cortex (i.e. occipitotemporal cortex) has also been found to be influenced by the L-SS congruency, particularly when suboptimal visual stimuli are used (Blau et al., 2008) or when sufficiently automatized L-SS integration have not been acquired (Plewko et al., 2018; Wang et al., 2020).

Recent electrophysiological evidence (Du et al., 2022; Nash et al., 2017; Yan & Seki, 2024) has added that the modulation even generated by the congruent visual letter prior to the auditory onset, as reflected in event-related potentials (ERPs), is accompanied by amplitude increase of auditory N1 and P2 components compared with unmatching AV pairs. Specifically, the congruency effects occur within approximately 250 ms. The auditory N1/P2 waveform refers to an auditory ERP response, which is generally related to the physical attributes of an auditory stimulus, such as speech (Näätänen & Winkler, 1999; Tremblay et al., 2006). These studies on L-SS integration indicated that visual stimuli beginning before the speech sound onset can serve as a visual cue that potentially facilitates the auditory processing. However, all these results have revealed the congruency effects of L-SS integration in the time domain.

Recent research on multisensory processing and perception has directed attention to cortical oscillations in the power of either phase-locked ('evoked', i.e. ERP) or non-phase-locked ('induced') oscillations, as well as in the phase domain, as it may provide complementary information regarding the underlying mechanisms of multisensory processing. It has been suggested that oscillations in different frequency bands contribute to distributed cortical networks that support multisensory integration. Previous studies have proposed that power and phase properties of oscillatory neural responses in the theta and beta bands have been reported to be related to multisensory processing. Studies have shown that theta oscillations, mainly distributed in the left frontocentral area, tend to be related to working memory and long-term memory retrieval for cross-modal association (Michail et al., 2021; Sauseng et al., 2010; Simon & Wallace, 2018), and to cognitive control mechanisms, such as attention and predictions (Cavanagh & Frank, 2014; Herweg et al., 2020). Moreover, theta oscillations tend to reflect cross-modal influences on the unisensory cortices. For example, enhanced theta oscillations may reflect the neural pathways involved in AV speech integration, possibly suggesting the cross-modal interactions between higher and lower cortical regions (Keil et al., 2012; Lange et al., 2013; Lindborg et al., 2019). Moreover, AV incongruence may induce an increase in the theta-band activity, as suggested by Morís Fernández et al. (2018). They also proposed that this

increased theta-band power, similar to the response elicited by Stroop stimuli, may serve as an indicator of the brain response to conflict detection in the perception of the McGurk illusion. Regarding the beta band, empirical studies have proposed that it is related to important cognitive processing and that it might reflect the different stages of multisensory integration, particularly suggesting the involvement of a fronto-parieto-occipital beta-band network in AV perception (Hipp et al., 2011; Senkowski et al., 2008). Furthermore, an alternative view from the error monitoring framework suggests that the evaluation of deviant or mismatching information is accompanied by increased beta-band power in the supramarginal gyrus during AV speech processing (Lange et al., 2013). In another study, Roa Romero et al. (2015) compared congruent and incongruent AV speech syllables and observed a reduction in beta-band power for incongruent AV pairs, which were perceived as an illusory novel percept, during the early and late processing stages. This reduction in beta-band power was assumed to be associated with the early processing stage in primary sensory areas and the subsequent integration in the late processing stage respectively.

Additionally, the coherence and phase resetting of ongoing oscillations have been proposed to play important roles in multisensory integration. Phase-resetting of oscillatory responses can be quantified as inter-trial phase coherence (ITPC), as reflected increased ITPC values (Bauer et al., 2020). One early study in macaques has reported that the sensory input to the somatosensory cortex resets the phase of oscillations in the auditory cortex and amplifies neuronal responses to the auditory stimuli (Lakatos et al., 2007). In line with this observation, a later study in monkeys have shown that cross-modal phase resetting in low-frequency oscillations (~10 Hz) can significantly modulate the processing of auditory stimuli when visual input precedes the auditory input (Kayser et al., 2008). In humans, a recent magnetoencephalography (MEG) study of AV speech integration, using coherence analysis, revealed an increase of theta-band coherence in auditory-related areas elicited by the congruent stimuli, while decreased coherence for incongruent stimuli (Lange et al., 2013). Mercier et al. (2015) found that the visual inputs can reset the phase of the theta-band oscillations in the auditory activity and argued that synchronized low-frequency phases facilitate communication between cortical areas. Moreover, findings from another related electroencephalography (EEG) study support the role of the beta-band ITPC in non-speech AV simultaneity judgment processing (Kambe et al., 2015). The authors observed increased poststimulus beta-band ITPC, indicative of phase-resetting, in the visual and auditory-related electrodes that were linked to the subsequent stimulus, only when participants perceived AV stimuli as synchronous.

Although previous studies have highlighted the role of neural activities in different oscillatory frequency bands as supportive mechanisms for multisensory processing (for review, see e.g. Keil & Senkowski, 2018; van Atteveldt et al., 2014), understanding the ongoing oscillations in the human cortex during L-SS integration is still in its early stages. As yet, the only MEG neuroimaging study that directly investigated L-SS integration was conducted by Herdman et al. (2006), who explored the congruency effects in cortical oscillations during L-SS integration in Japanese participants, revealing that the visual congruency can modulate 2–10 Hz activity in the left auditory cortex (i.e. left Heschl's sulcus/planum temporale) within 250 ms time window. Despite these intriguing findings, the precise dynamics of the ongoing oscillations during L-SS integration are still not fully understood. Hence, we expect that verifying the congruency effects across different oscillatory frequency bands may offer valuable insights into identifying different network interactions underlying L-SS integration. As such, our aim was to investigate the congruency effects of L-SS integration and to evaluate the contribution of theta/beta-band powers and ITPC values. Additionally, we also sought evidence of cross-modal phase resetting. To achieve this, here we used EEG technique to measure a sample of Japanese participants, while they were passively exposed to Japanese letters and speech sounds (i.e. kana) in the context of an AV priming paradigm. Specifically, we compared the theta/beta-band powers and ITPC values between congruent and incongruent conditions involved in AV integration processing.

2 | MATERIALS AND METHODS

2.1 | Participants

This dataset was originally collected by Yan and Seki (2024). The experimental procedure was described thoroughly in the original study. In brief, the dataset contained data from 36 native Japanese-speaking university students (age range 18–27 years; 20 females). Data from two female participants who failed to complete the EEG task were excluded from the analysis. All participants were right-handed and had normal hearing, as well as either normal or corrected-to-normal vision. Additionally, they had no record of reading difficulties diagnoses or other neurological or psychiatric disorders. Written informed consent was obtained from all the participants in accordance with the Declaration of Helsinki.

2.2 | Stimuli and procedure

A priming paradigm in Japanese was performed by presenting visual letters or symbols followed by auditory sounds. The prime and target were matching and mismatching Japanese kana letters and sounds. The kana letters were presented at the uniform brightness in Arial font at a size of 28. Speech sounds were recorded by a male native Japanese speaker, and all of these audio files were subsequently down-sampled to 22.05 kHz with matched loudness at 70 dB using the PRAAT software (www.fon.hum.uva.nl/praat/; Boersma & Weenink, 2020). Auditory stimuli were presented binaurally through a pair of earphones. The stimulus material was categorized in audiovisual congruent (AVc), audiovisual incongruent (AVi), baseline (symbol-sound pairs) and two control conditions (see Table 1). The prime was either a kana letter or a symbol, while the auditory stimulus was either a congruent or an incongruent phoneme. For the control conditions, the auditory stimulus was a scrambled phoneme, edited from the corresponding speech sounds using the MATLAB code. Control conditions were included to balance button-press responses in the behavioural task, where participants were instructed to respond after hearing the auditory stimuli by deciding whether the sound was a 'real speech sound' (results are shown in Yan & Seki, 2024). To maintain consistency with the behavioural task, these control conditions were also included in the EEG task. Moreover, Hein et al. (2007) suggested that the congruent and baseline conditions might engage in different neural processing for adults. They found that unfamiliar incongruent pairs were integrated in the inferior frontal cortex, potentially reflecting the learning of novel AV associations. Comparing the congruent condition with the baseline condition could obscure the specific congruency effects we aimed to isolate. Therefore, data from the baseline and control conditions were not analyzed in the current study. All stimuli were positioned 70 cm in front of the participants' eyes and presented on a 17-in. monitor with a screen resolution of 1024 × 720 pixels.

The participants performed the task in a silent and dimly lit room, shielded from electrical interference. Each trial had a total duration of approximately 2000 ms (see Figure 1). Within each trial, a central fixation cross was presented ranging from 1000 to 1450 ms (150 ms per step) to mitigate expectancy effects. Subsequently, a visual prime presented as the black text on a white screen, was displayed for 200 ms, followed by the presentation of the auditory stimuli. Participants were not required to take any action in response to the AV pairs. To ensure that participants were attentive to the presented stimuli, an unrelated visual detection task was also administered, in which they were instructed to press a button when piano pictures were presented as the target condition. The Japanese experiment included nine blocks, and each block lasted for approximately 3 min. All conditions were presented in a randomized order, with a total of 108 trials for each condition, including 108 trials for the target condition.

TABLE 1 Examples of AV stimuli of each condition in the priming task.

Condition	Visual prime	Auditory stimuli
AVc	と	/to/
AVi	ね	/to/
Baseline	θ	/to/
Control 1	と	/tos/*
Control 2	θ	/tos/*

/tos/* represents a scrambled speech sound.

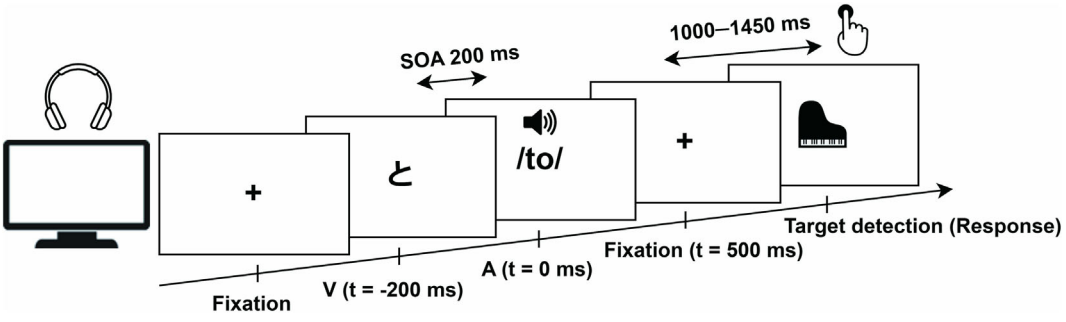

FIGURE 1 The procedure used for the priming paradigm in Japanese during EEG data collection. Speech sounds were presented 200 ms after the presentation of the visual letter. V (visual letter) and A (auditory sound) represent the onset of each stimulus. The 0-ms mark represents the onset of the speech sound. EEG, electroencephalogram.

2.3 | EEG recording, preprocessing and analyses

EEG data were recorded from 29 Ag/AgCl electrodes mounted on an electrode cap (Easycap GmbH, Herrsching, Germany), as specified by the International 10–20 System, and three electrooculogram electrodes. All the channels were referenced to POz. The recorded data were digitized at a sampling rate of 1000 Hz by SynAmps amplifier (NeuroScan, Sterling, VA, USA). Electrode impedance was maintained below 10 kΩ.

Preprocessing and offline EEG signal analyses were performed using EEGLAB 14.1.2b (Delorme & Makeig, 2004), FieldTrip (Oostenveld et al., 2011) and custom-made MATLAB scripts (The MathWorks, Natick, MA, USA). The EEG signal from each electrode was re-referenced to the common average and then down-sampled to 250 Hz. The continuous EEG data were first high-pass filtered (noncausal Butterworth filter, half-amplitude cutoff = .1 Hz, 12 dB/oct roll-off), and then segmented into epochs starting from 600 ms before to 800 ms after the auditory onset. Artifactual independent components including eye movements, blink artifacts, heartbeat and muscular activities were identified with the help of ICLabel (Pion-Tonachini et al., 2019). The channels with large artifacts were excluded, and the removed channels were spherically interpolated. Epochs with amplitudes ±100 μV at the electrooculogram and scalp electrodes were removed from analysis. One participant was excluded from further EEG analyses, due to having more than 30% of trials rejected. The average number [±SD] of epochs in each condition are shown as follows: AVc [90.12 (±12.55)] and AVi [92.03 (±11.72)].

To evaluate the oscillatory activities of each condition (i.e. congruent and incongruent), individual trials were then separately filtered into two frequency ranges using wavelet convolution, and then, power transform was computed: theta (1–20 Hz, wavelet cycles varied linearly from 3 to 7) and beta (10–40 Hz, wavelet cycles varied linearly from 5 to 10) with a 4-ms-step in the time domain and 1 Hz-step in the frequency domain. Averaged oscillatory activity was baseline corrected (decibel normalization) from −250 to −50 ms before the sound onset to focus on the auditory processing and subsequent integration processing, rather than potential impacts of visual letters.

To measure the properties of the phase, we computed the ITPCs for both congruent and incongruent conditions. ITPC values were obtained according to the following formula:

$$ITPC(f,t) = \frac{1}{N}\left|\sum_{k=1}^{N} e^{i\varphi^k(f,t)}\right|$$

where t stands for time, f for frequency, N for the number of trials, and ϕk is the local phase angle of the signal at time t and frequency f (van Diepen & Mazaheri, 2018). ITPC results are bound between zero and one, with zero indicating completely randomly distributed phase angles and one indicating completely identical phase angles (Cohen, 2014). Essentially, the ITPC value reflects the phase concentration among all trials, with a value of 1 indicating perfect phase locking and 0 indicating random phase distribution.

2.4 | Statistical analysis

2.4.1 | Power analysis

Statistical comparison of the oscillatory theta/beta-band activities during L-SS integration was performed on the time-frequency representations of congruent and incongruent trials. Based on the findings of prior research, we designated specific regions of interest (ROI) to select the time window and frequency values for the analysis of neural oscillations. Specifically, we determined the left frontocentral areas to explore the theta-band power (Herdman et al., 2006) and the fronto-parieto-occipital regions to examine beta-band power (Hipp et al., 2011; Senkowski et al., 2008) respectively. Following a visual inspection of the time course of the EEG power spectra on the ROI, we chose 5–7 Hz within the time window of 0–200 ms after the auditory onset in the theta band, and defined 20–35 Hz within the time window of 0–250 ms in the beta band as the intervals for analysing oscillation differences. After selecting the a priori time in our data, the following statistical comparison between congruent and incongruent conditions within the two specified time spans was conducted using cluster-based permutation tests with Monte Carlo randomization ($n = 2000$) (Maris & Oostenveld, 2007). Clusters with at least two adjacent electrodes were considered significant, if the probability of observing a cluster with a greater summed test statistic during data shuffling was less than 5%. Once a significant cluster was identified, a following repeated-measures analysis of variance (ANOVA) was conducted on the average power of the specified time window for each electrode within the cluster between the two conditions (condition × electrode). These ANOVA analyses were conducted by applying Greenhouse–Geisser corrections to the degrees of freedom with an alpha level of .05. Statistical analyses of ANOVAs were performed using the SPSS version 26.0 (SPSS, Inc., Chicago, IL, USA).

2.4.2 | ITPC analysis

To evaluate the phase properties, we analyzed the ITPC values of theta/beta band for each condition. To specifically focus on the activities in the auditory cortex, we selected the following electrodes based on previous studies (Kambe et al., 2015; Montoya-Martínez et al., 2021), which covered the frontocentral areas of the scalp (FC1, FC2, C3, C4, and Cz). To explore the ITPC modulations induced solely by the preceding visual stimulus, we conducted repeated-measures ANOVAs ($\alpha = .05$) comparing the time averaged prestimulus and poststimulus ITPC values (i.e. before and after the sound onset). We expected that the effects in the phase would be most prominent near the stimulus onset. Therefore, the prestimulus ITPCs was set at a time averaged from −250 to −50 ms, while the poststimulus ITPCs was determined at a time averaged from 0 to 200 ms. Furthermore, to investigate the effect of AV congruency on phase consistency of neural responses to auditory stimuli, we also compared poststimulus ITPCs between congruent and incongruent conditions using cluster-based permutation tests in both the theta and beta frequency ranges.

3 | RESULTS

3.1 | Theta and Beta power

The cluster-based permutation test in the theta-band power revealed one significant positive cluster in the time interval ranging from 0 to 200 ms. We observed the significant difference ($p = .0405$) between conditions ranging from 28 to 148 ms, with a stronger theta power after onset of auditory sounds for the congruent condition than for the incongruent condition at the left frontocentral electrodes. The topographic representation of theta power for each condition, averaged within a specific time window, is summarized in Figure 2(a). The time course of the time-frequency representation by time (x-axis) and frequency (y-axis) within this significant cluster is depicted in Figure 3(a). To further explore the effect of congruency within the identified cluster, a subsequent 2 (condition; AV congruent, AV incongruent) × 3 (electrode; Fz, FC1, Cz) ANOVA was conducted. The analysis revealed a significant main effect of condition [$F(1, 96) = 9.52$, $p = .003$, $\eta^2 = .09$], but there were no significant effect of electrode [$F(2, 96) = .17$, $p = .84$, $\eta^2 = .00$] or significant condition × electrode interaction [$F(2, 96) = .04$, $p = .96$, $\eta^2 = .00$]. The mean power of the theta band at the electrodes within the significant cluster, averaged over time for each condition, is shown in Figure 4(a). Our findings revealed the congruency effects in the theta band at left frontocentral electrodes.

Furthermore, the results of the cluster-based permutation test in the beta band (0–250 ms) showed one significant positive cluster. The significant difference ($p = .006$) was observed from approximately 40 to 248 ms, showing the higher spectral power for incongruent trials than for congruent trials at the right centroparietal electrodes. The topographic distribution of the grand time-averaged beta oscillation responses for congruent and incongruent conditions is summarized in Figure 2(b). The grand average of beta power by time (x-axis) and frequency (y-axis) within the identified right centroparietal sensors is shown in Figure 3(b). Subsequently, a repeated measures ANOVA was performed. The results of the condition × electrode (FC2, C4, CP1, CP2, Pz, P4, P8) ANOVA in the beta-band power yielded a significant main effect of condition [$F(1, 224) = 39.61$, $p < .001$, $\eta^2 = .15$]. However, the main effect of electrode was not significantly different [$F(6, 224) = .19$, $p = .98$, $\eta^2 = .01$], and there was no significant interaction between condition and electrode [$F(6, 224) = 0.04$, $p = 1.000$, $\eta^2 = .00$]. Figure 4(b) shows the average beta power of the electrodes within the significant cluster averaged over time for congruent and incongruent conditions. Thus, we found an incongruency effect in the beta band within the specified time window, rather than a congruency effect.

3.2 | Theta and Beta ITPC

The grand-average ITPC of the theta and beta bands is plotted by condition at the sensor level as shown by the topographical maps in Figures 5 and 6 respectively. The repeated-measures ANOVA for the theta-band ITPC over the selected auditory electrodes (FC1, FC2, C3, C4 and Cz) indicated that the prestimulus-poststimulus comparison reached a significant main effect for both the congruent [$F(1160) = 168.40$, $p < .001$, $\eta^2 = .51$] and incongruent [$F(1160) = 178.46$, $p < .001$, $\eta^2 = .53$] conditions, showing significant differences between time averaged ITPCs before and after the onset of auditory stimuli. However, no main effect of electrode and no interaction between the averaged prestimulus-poststimulus ITPC and electrode were found for either condition. Similarly, the results of the repeated-measures ANOVA in the beta-band ITPC showed a significant main effect for the prestimulus-poststimulus comparison in each condition [congruent: $F(1,160) = 152.88$, $p < .001$, $\eta^2 = .49$; incongruent: $F(1,160) = 122.22$, $p < .001$, $\eta^2 = .43$], but the main effect of electrode was not significantly different, nor a significant the prestimulus-poststimulus

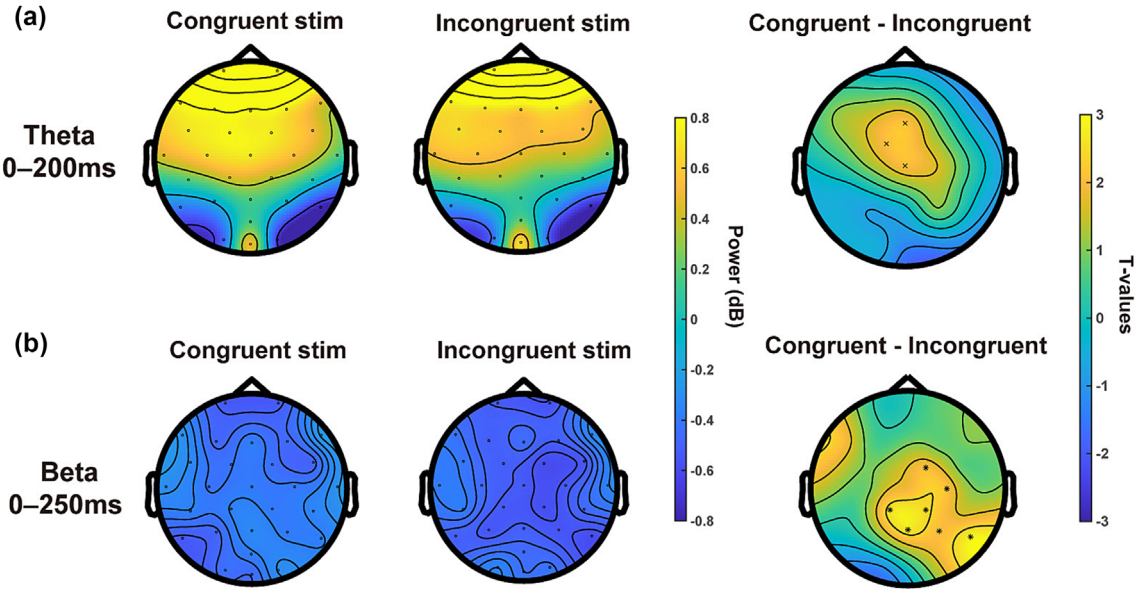

FIGURE 2 Topographic distribution of grand average power within the specific time windows for each AV condition and topography of the corresponding significant cluster between congruent and incongruent conditions in the theta band (a) and beta band (b). The right column displays permutation-based *T*-value distributions for statistical comparisons within conditions. Electrodes belonging to the significant cluster are shown as $p < 0.01$ (*), $p < 0.05$ (x), and $p < 0.1$ (.). Time 0 = auditory onset. AV, audiovisual.

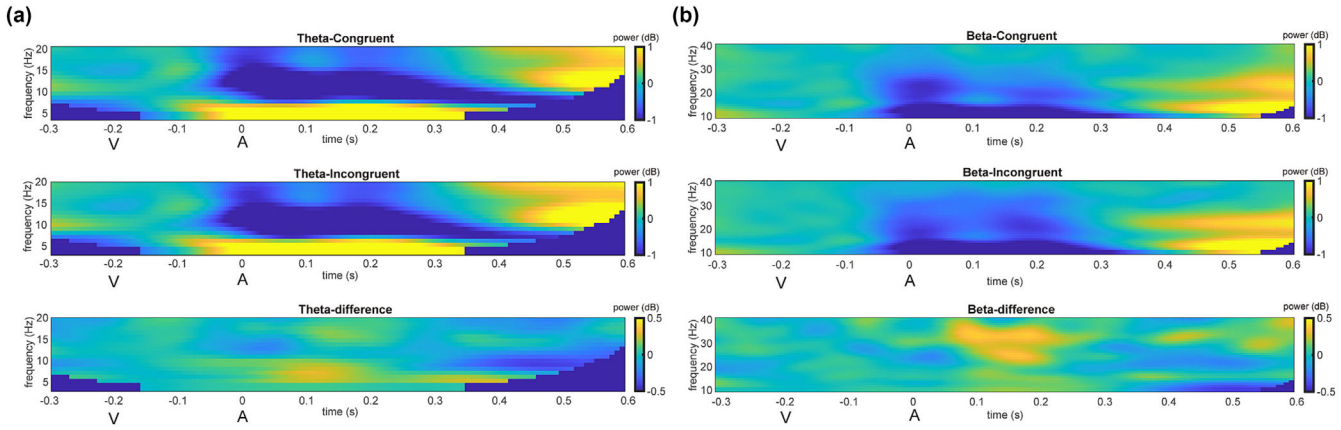

FIGURE 3 Grand average power in the theta band (a) at electrodes (Fz, FC1, and Cz) and in the beta band (b) at electrodes (FC2, C4, CP1, CP2, Pz, P4, and P8) within the positive significant clusters by time (*x*-axis) and frequency (*y*-axis) for congruent and incongruent conditions at the sensor level. Time 0 = auditory onset. Enhanced theta-band activity is observed from approximately 0 to 200 ms. Enhanced beta-band activity is observed from approximately 0 to 250 ms.

ITPC × electrode interaction. For a complete list of ANOVA statistics, see Table 2 for the details.

Furthermore, the cluster-based permutation test to verify differences of theta ITPC between conditions showed one significant positive cluster (see Figure 5). The positive cluster reflected larger ITPC values at the left centroparietal electrodes for congruent trials compared to incongruent trials in the time range from 36 to 164 ms ($p = .033$). The observed congruency effect of the theta-band ITPC in the left centroparietal regions differed from the findings related to the scalp distribution of congruency effects in theta power. Hence, the difference in ITPC between congruent and incongruent conditions may not be attributed to power differences. In contrast, no significant cluster in the beta ITPC was observed (see Figure 6); therefore, there was no significant difference in the beta-band ITPC between congruent and incongruent trials.

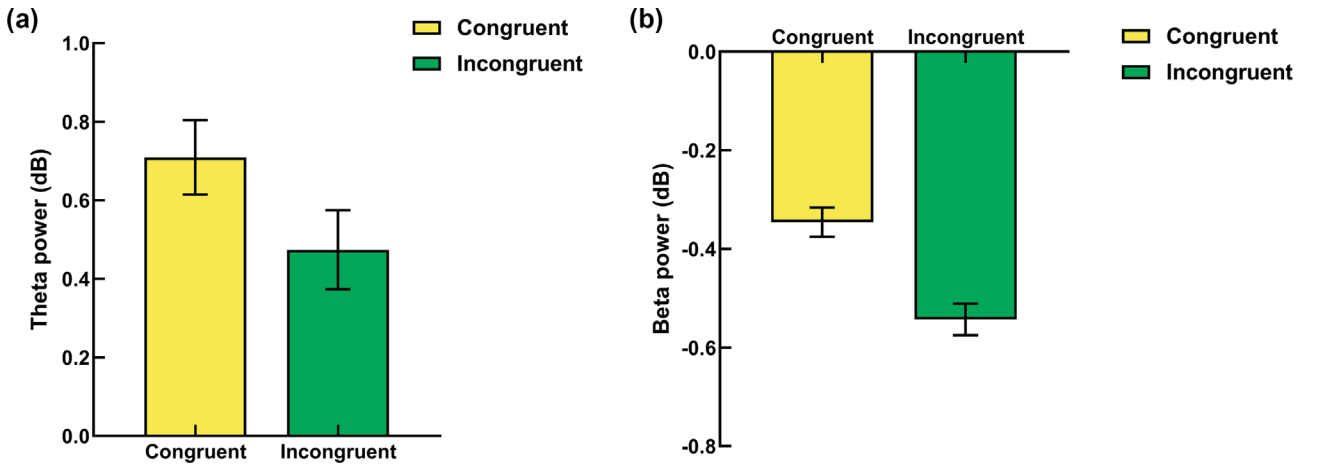

FIGURE 4 Mean power of electrodes within the significant cluster cross time in the theta band (a) and beta band (b) for each AV condition. Whiskers depict the standard error of the mean over participants. AV, audiovisual.

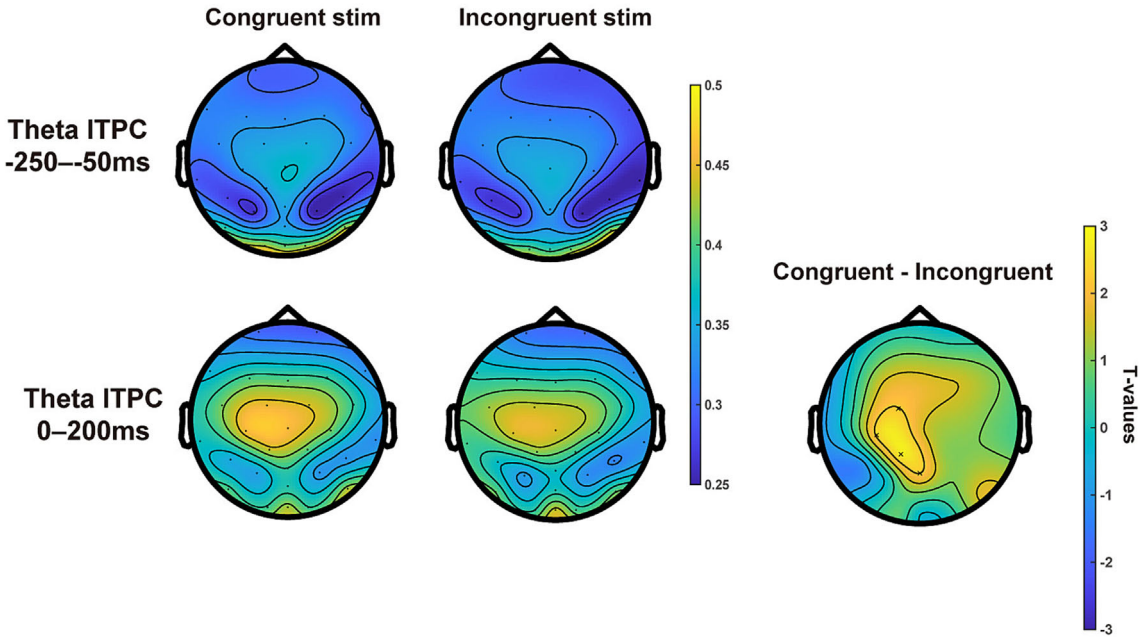

FIGURE 5 Topographies of the prestimulus and poststimulus ITPC and corresponding topography of the significant cluster between congruent and incongruent conditions in the theta band. The right column displays permutation-based *T*-value distributions for statistical comparisons within conditions. Electrodes belonging to the significant cluster are shown as $p < 0.01$ (*), $p < 0.05$ (x), and $p < 0.1$ (.). Time 0 = auditory onset. ITPC, inter-trial phase coherence.

4 | DISCUSSION

This study was set out to identify the electrophysiological signature of the congruency effects in L-SS integration across different oscillatory frequency bands. Additionally, we evaluated the contribution of the cortical oscillations to the processing of Japanese speech sounds influenced by preceding congruent or incongruent Japanese kana letters. Our focus was on exploring the power and phase properties of the theta and beta bands, along with examining the spatial distribution of cortical activity. To achieve this, we compared the theta/beta powers and ITPC values between the congruent and incongruent AV conditions.

4.1 | Power and phase properties in theta band

Our findings revealed increased left frontocentral theta oscillatory activities in response to the congruent trials within the 0–200-ms time window after the auditory

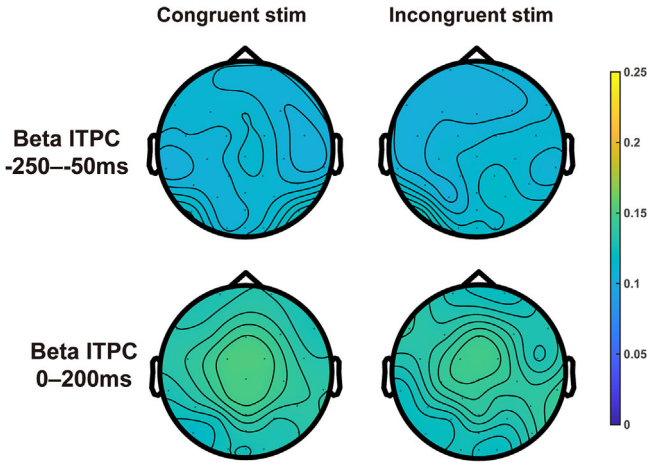

FIGURE 6 Topographies of the prestimulus and poststimulus ITPC in the beta band for each AV condition. Time 0 = auditory onset. AV, audiovisual; ITPC, inter-trial phase coherence.

onset. Previous studies on AV speech integration by Morís Fernández et al. (2018) demonstrated an increase of theta-band activity induced by the incongruent AV pairs in the late processing period, suggesting the conflicting mechanisms between congruent and incongruent conditions over the mid-frontal and mid-central electrodes for AV speech integration. The source of conflict-related activity in the theta band has been in the anterior cingulate cortex, and similar responses were also observed in the same participants during a Stroop task (Morís Fernández et al., 2017). However, if the conflict hypothesis holds true, we would expect lower theta power for the congruent AV condition in the late integration period, but not for incongruent trials, which contradicts our findings.

An alternative explanation could be that the enhanced theta-band activities in response to congruent trials may reflect the cross-modal influence between higher and primary cortical areas, which may suggest the feedback processing from higher association areas during AV integration (Keil et al., 2012; Keil & Senkowski, 2018; Lindborg et al., 2019). MEG research by Lange et al. (2013) investigated AV speech integration, reporting an early theta-band power difference between congruent and incongruent AV conditions after onset of auditory inputs (~250 ms) in the auditory cortex and the middle and superior temporal gyri. Moreover, a previous MEG study in typical Japanese adult readers during L-SS integration showed that visual congruency can modulate theta activities in the left primary auditory cortex (i.e. left Heschl's sulcus/planum temporale) within the 0–250-ms time window (Herdman et al., 2006). Their results suggest that increased theta activity for the congruent condition in the auditory

TABLE 2 Results of statistical analyses of ITPC for the auditory ROI from the repeated-measures ANOVA (prestimulus - poststimulus × electrodes).

	Theta						Beta					
	Congruent				Incongruent				Congruent				Incongruent			
Effect	*df*	*F*	*p*	η^2	*df*	*F*	*p*	η^2	*df*	*F*	*p*	η^2	*df*	*F*	*p*	η^2
Pre-post	1, 160	168.402	<.001	.513	1, 160	178.457	<.001	.527	1, 160	152.88	<.001	.489	1, 160	122.22	<.001	.433
Electrode	4, 160	1.124	.347	.027	4, 160	.899	.466	.022	4, 160	.618	.651	.015	4, 160	.101	.982	.003
Pre-post × electrode	4, 160	.511	.728	.013	4, 160	.191	.943	.005	4, 160	.253	.908	.006	4, 160	.421	.793	.01

Abbreviations: ITPC, inter-trial phase coherence; ROI, regions of interest.

cortex may be driven by feedback from multisensory areas during L-SS integration. While the spatial resolution of EEG data limits precise localization in our current study, the observed congruency effects in the theta-band power at the left frontocentral electrodes in our results may be explained by differential theta activities in the primary auditory cortex, driven by feedback projections. This explanation is further supported by an ERP study using a priming paradigm focused on L-SS integration. Nash et al. (2017) reported that the congruency effects observed in the auditory-evoked P1 component within the 50- to 125-ms time window, particularly in the left frontocentral scalp region, could be attributed to differential activities in the primary auditory cortex. Additionally, in our current study, visual letters were presented 200 ms before the speech sounds, allowing the visual stimulus to activate the corresponding L-SS association in the higher-order multisensory regions to prepare for the subsequent auditory input, potentially accelerating integration. This interpretation may support the automaticity and efficiency of L-SS integration in proficient readers. Thus, the early increased theta power in current study may indeed be indicative of feedback mechanisms, where higher-order areas (i.e. STC) send integrated information back to the primary auditory cortex, modulating its activity.

However, the early effects observed in our study may not be solely due to feedback projections. Although not definitive, the early stage of the time window during the congruency effects suggests that they may involve processes beyond direct feedback inputs from higher-order cortical regions. Foxe and Schroeder (2005) proposed that the early activities may reflect feedforward inputs to sensory cortices, which help the brain quickly determine the cross-sensory binding. In their study, where auditory and visual inputs were presented simultaneously, early AV multisensory interactions were detected as early as 46 ms post-stimulus onset (Molholm et al., 2002). They estimated that feedback inputs from the STC to early visual areas occurs at a minimal latency of about 67 ms after stimulus onset in humans, which may be too late to drive the integration effects found between 40 and 50 ms. Anatomical studies in monkeys proposed that nonauditory inputs may access auditory cortex via feedforward projections through these nonspecific thalamic systems (Hackett et al., 2007; Lakatos et al., 2007). Supporting this, an early ERP study demonstrated that visual facial input may affect auditory-evoked responses via subcortical areas as early as 11–30 ms post-auditory stimulation (Musacchia et al., 2006). In light of these previous studies, the early timing may imply that the early feedforward projections play a role in the observed effects. Hence, we conjecture that increased left frontocentral theta power in response to the congruent condition in the early stage may reflect these feedforward projections facilitating cross-modal integration and aiding the rapid combination of AV information. Nevertheless, further investigation is needed to validate this by confirming the distinction between predictive multisensory interactions that optimize feedforward encoding of auditory information and later feedback processes that alter auditory processing during L-SS integration.

Evoked oscillations (phase-locked) are closely related to the onset of an external stimulus, whereas induced oscillations (non-phase-locked) may arise from some distinct high-order processes (David et al., 2006). However, both evoked and induced oscillations can contribute to changes in power. To determine whether the observed congruency effect in theta power was related to an effect in the evoked ERP (Yan & Seki, 2024), we therefore performed an additional analysis to compare the individual theta powers of evoked potentials between conditions (i.e. the time-frequency transforms of the ERPs) by running a cluster-based permutation test. Unfortunately, we did not observe any significant differences between conditions. The lack of significant differences in the evoked potential theta powers suggests that the observed congruency effects of theta powers in our study may not be attributable to the neural responses elicited by the auditory stimuli. Instead, other factors, such as ongoing neural oscillations related to multisensory processing, may have contributed to the observed differences.

In addition, we identified a poststimulus increase in theta ITPC within the auditory cortex, showing a high phase coherence in the theta band. It has previously been shown that ITPC enhancements elicited by other sensory stimuli are often interpreted as phase resetting (Bauer et al., 2020; Kayser et al., 2008; Keil & Senkowski, 2018). Thus, it is plausible to consider that a preceding visual letter resets the phase of ongoing activities via theta oscillations in the auditory cortex and amplifies neuronal responses to prepare appropriately for the subsequent auditory stimulus. Meanwhile, we found an increase in ITPC for the congruent condition in the auditory related areas. The observed congruency effects in theta ITPC may be associated with the increased functional coupling between the cortices (Fries, 2015; Keil & Senkowski, 2018). Additionally, previous studies have also proposed that information on stimulus congruence is fed back from the multimodal areas to auditory cortical regions through theta-band phase coherence (Lange et al., 2013). Therefore, our findings of greater ITPC in the congruent condition reveal highly synchronous firing of assemblies of neurons oscillating in the multisensory network when L-SS integration occurs.

Taken together, we can interpret the observed congruency effects in the theta band as reflecting increased

neuronal activities in the auditory regions, which may possibly indicate the cross-modal influences in the primary auditory cortex. These findings are consistent with the rationale that low-frequency oscillations may facilitate communication between neural populations across long distances, such as the network involved in L-SS integration. Thus, cross-modal processes are likely to influence the primary sensory activity through low-frequency oscillatory activities.

4.2 | Power and phase properties in beta band

We found an increase in beta power in response to incongruent trials at the right centroparietal electrodes. The previous EEG study by Roa Romero et al. (2015) observed stronger oscillatory power in the beta band for successful McGurk fusions than for unsuccessful fusions at the left-lateralized central electrodes. A recent review summarized that the increased beta-band power in the congruent AV condition has been linked to multisensory integration processing (Keil & Senkowski, 2018). However, our results showed increased beta-band power for the incongruent trials, rather than for the congruent AV condition. Additionally, the spatial distribution of the incongruency effect was identified in the right centroparietal regions, which are potentially not responsible for the processing of multisensory integration. As such, our findings of increased beta-band power in the right centroparietal regions may not be generated by multisensory processing. Notably, the increased beta-band power for incongruent trials may reflect the error monitoring of incongruent, nonmatching stimuli (Lange et al., 2013), while the decrease of the beta oscillation in the congruent condition may be linked with the prediction update to prevent further error occurrence during the auditory processing (Chao et al., 2022). Furthermore, a related study proposed that the expectancy violations of the subsequent auditory stimuli may lead to an increase in beta power, but only in the absence of attention (Todorovic et al., 2015). Their findings are in line with the requirements of our task, which required participants to detect congruency in a passive paradigm. Additionally, changes in the beta-band power also have been associated with various aspects of language processing, including grapheme-to-phoneme conversion and word recognition (e.g. Brennan et al., 2014; Duncan Milne et al., 2003; Klimesch et al., 2001). Moreover, beta-band oscillations may be linked to the anticipation of upcoming stimuli during the language processing. For instance, a decrease in the beta-band power was found in the highly predictable condition during conversations in the mid-frontal and centroparietal regions (Magyari et al., 2014). Similar findings of reduced beta activity in the right temporo-parietal areas have also been observed while predicting upcoming words (Gastaldon et al., 2020). Accordingly, the decrease in beta-band power for the congruent condition in our study may be interpreted as the expectancy of the subsequent auditory input during language processing. Meanwhile, we found the increased poststimulus ITPC in the beta band within the auditory cortex for both congruent and incongruent conditions, considering as phase resetting. However, we did not observe any differences in the poststimulus beta ITPC values between conditions. These findings indicate that there is the lack of significant alteration in beta-band coherence within the auditory cortex despite the modulation of auditory stimulus processing by the preceding visual input. Therefore, we conjecture that the mechanisms underlying beta-band activity in the auditory cortex may not be directly influenced by cortical areas. To summarize, our findings reveal that the beta-band oscillation activity tends to be linked to the prediction of the following auditory sounds, as the visual letter can serve as a visual cue for the upcoming auditory stimuli, which is associated with the task design of our priming paradigm.

4.3 | Limitations and future study

While we have made efforts to conduct comprehensive analyses, it is important to acknowledge potential limitations that should be taken into consideration in future research. Notably, one limitation of our study is the use of EEG, which provides limited spatial resolution. This limitation makes it challenging to precisely identify the neural sources and pathways involved in the observed effects. Future studies employing advanced imaging techniques with higher spatial and temporal resolution such as magnetoencephalography combined with source localization algorithms, or multi-modal imaging methods, may shed light on more precisely delineating the contributions of feedforward and feedback processes for L-SS integration processing.

Furthermore, the automaticity and efficiency of L-SS integration observed in proficient readers may not fully extend to naturalistic speech scenarios, where the integration of continuous auditory and visual speech is required. However, in such contexts, factors like background noise and contextual cues may play significant roles in modulating the integration process. Therefore, future exploration of how the neural mechanisms identified in our study function in more ecologically valid contexts is warranted.

5 | CONCLUSIONS

To conclude, in this study, we investigated the ongoing oscillations that contribute to the processing of L-SS integration by utilizing a priming paradigm. The main findings of this study can be summarized as follows. Regarding theta band oscillation, our results revealed that the preceding visual congruency induced stronger theta-band power in the left auditory cortex, accompanied by cross-modal phase resetting at the theta frequency band within the auditory cortex. Meanwhile, we also found an enhancement of theta ITPC in response to the congruent AV condition in the auditory related areas. With respect to beta-band activity, an increase in beta-band oscillatory power was observed in the right centro-parietal regions for the incongruent trials. Therefore, these findings suggest that oscillations in different frequency bands are set for different aspects of integration that are crucial for L-SS integration. Specifically, theta-band activity tended to be closely related to the L-SS integration networks, while beta-band activity was more likely to be associated with the task design of our priming paradigm. These findings provide valuable insights in how oscillatory activity at different frequencies plays differential roles in supporting L-SS integration.

AUTHOR CONTRIBUTIONS

Dongyang Yan: conceptualization; investigation; data curation; formal analysis; visualization; writing—original draft; writing—review and editing. Ayumi Seki: conceptualization; supervision; resources; funding acquisition; writing—review and editing.

ACKNOWLEDGEMENTS

This work was supported by the Japan Society for the Promotion of Science, JSPS KAKENHI, Grant Number JP17H02713.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ejn.16563.

DATA AVAILABILITY STATEMENT

The datasets analysed during the current study are available from the corresponding author on reasonable request.

ORCID

Dongyang Yan https://orcid.org/0000-0003-3944-0508

REFERENCES

Bauer, A. R., Debener, S., & Nobre, A. C. (2020). Synchronisation of neural oscillations and cross-modal influences. *Trends in Cognitive Sciences*, *24*(6), 481–495. https://doi.org/10.1016/j.tics.2020.03.003

Blau, V., van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2008). Task-irrelevant visual letters interact with the processing of speech sounds in heteromodal and unimodal cortex. *European Journal of Neuroscience*, *28*(3), 500–509. https://doi.org/10.1111/j.1460-9568.2008.06350.x

Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *NeuroImage*, *57*(3), 695–703. https://doi.org/10.1016/j.neuroimage.2010.11.003

Boersma, P., & Weenink, D. (2020). Praat: Doing phonetics by computer (Version 4.1) [Computer program]. Retrieved June 6, 2020, from www.praat.org

Brennan, J., Lignos, C., Embick, D., & Roberts, T. P. (2014). Spectro-temporal correlates of lexical access during auditory lexical decision. *Brain and Language*, *133*, 39–46. https://doi.org/10.1016/j.bandl.2014.03.006

Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., & David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, *10*(12), 2619–2623. https://doi.org/10.1097/00001756-199908200-00033

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*(11), 649–657. https://doi.org/10.1016/s0960-9822(00)00513-3

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Chao, Z. C., Huang, Y. T., & Wu, C. T. (2022). A quantitative model reveals a frequency ordering of prediction and prediction-error signals in the human brain. *Communications Biology*, *5*(1), 1076. https://doi.org/10.1038/s42003-022-04049-6

Cohen, M. X. (2014). Analyzing neural time series data. In *Theory into practice* (pp. 241–258). MIT Press.

David, O., Kilner, J. M., & Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage*, *31*(4), 1580–1591. https://doi.org/10.1016/j.neuroimage.2006.02.034

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews. Neuroscience*, *16*(4), 234–244. https://doi.org/10.1038/nrn3924

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Du, Y. C., Li, Y. Z., Qin, L., & Bi, H. Y. (2022). The influence of temporal asynchrony on character-speech integration in Chinese children with and without dyslexia: An ERP study. *Brain and Language*, *233*, 105175. https://doi.org/10.1016/j.bandl.2022.105175

Duncan Milne, R. D., Hamm, J. P., Kirk, I. J., & Corballis, M. C. (2003). Anterior-posterior beta asymmetries in dyslexia during lexical decisions. *Brain and Language*, *84*(3), 309–317. https://doi.org/10.1016/s0093-934x(02)00557-6

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, *16*(5), 419–423. https://doi.org/10.1097/00001756-200504040-00001

Fries, P. (2015). Rhythms for cognition: Communication through coherence. *Neuron*, *88*(1), 220–235. https://doi.org/10.1016/j.neuron.2015.09.034

Gastaldon, S., Arcara, G., Navarrete, E., & Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. *Cortex*, *133*, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Hackett, T. A., De La Mothe, L. A., Ulbert, I., Karmos, G., Smiley, J., & Schroeder, C. E. (2007). Multisensory convergence in auditory cortex, II. Thalamocortical connections of the caudal superior temporal plane. *The Journal of Comparative Neurology*, *502*(6), 924–952. https://doi.org/10.1002/cne.21326

Hein, G., Doehrmann, O., Müller, N. G., Kaiser, J., Muckli, L., & Naumer, M. J. (2007). Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *The Journal of Neuroscience*, *27*(30), 7881–7887. https://doi.org/10.1523/JNEUROSCI.1740-07.2007

Herdman, A. T., Fujioka, T., Chau, W., Ross, B., Pantev, C., & Picton, T. W. (2006). Cortical oscillations related to processing congruent and incongruent grapheme-phoneme pairs. *Neuroscience Letters*, *399*(1–2), 61–66. https://doi.org/10.1016/j.neulet.2006.01.069

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta oscillations in human memory. *Trends in Cognitive Sciences*, *24*(3), 208–227. https://doi.org/10.1016/j.tics.2019.12.006

Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, *69*(2), 387–396. https://doi.org/10.1016/j.neuron.2010.12.027

Kambe, J., Kakimoto, Y., & Araki, O. (2015). Phase reset affects auditory-visual simultaneity judgment. *Cognitive Neurodynamics*, *9*(5), 487–493. https://doi.org/10.1007/s11571-015-9342-4

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, *18*(7), 1560–1574. https://doi.org/10.1093/cercor/bhm187

Keil, J., Müller, N., Ihssen, N., & Weisz, N. (2012). On the variability of the McGurk effect: Audiovisual integration depends on prestimulus brain states. *Cerebral Cortex*, *22*(1), 221–231. https://doi.org/10.1093/cercor/bhr125

Keil, J., & Senkowski, D. (2018). Neural oscillations orchestrate multisensory processing. *The Neuroscientist*, *24*(6), 609–626. https://doi.org/10.1177/1073858418755352

Klimesch, W., Doppelmayr, M., Wimmer, H., Gruber, W., Röhm, D., Schwaiger, J., & Hutzler, F. (2001). Alpha and beta band power changes in normal and dyslexic children. *Clinical Neurophysiology*, *112*(7), 1186–1195. https://doi.org/10.1016/S1388-2457(01)00543-0

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in the primary auditory cortex. *Neuron*, *53*(2), 279–292. https://doi.org/10.1016/j.neuron.2006.12.011

Lange, J., Christian, N., & Schnitzler, A. (2013). Audio-visual congruency alters power and coherence of oscillatory activity within and between cortical areas. *NeuroImage*, *79*, 111–120. https://doi.org/10.1016/j.neuroimage.2013.04.064

Liberman, A. M. (1992). The relation of speech to reading and writing. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 167–178). Elsevier Science Publishers B.V. https://doi.org/10.1016/S0166-4115(08)62794-6

Lindborg, A., Baart, M., Stekelenburg, J. J., Vroomen, J., & Andersen, T. S. (2019). Speech-specific audiovisual integration modulates induced theta-band oscillations. *PLoS ONE*, *14*(7), e0219744. https://doi.org/10.1371/journal.pone.0219744

Magyari, L., Bastiaansen, M. C., de Ruiter, J. P., & Levinson, S. C. (2014). Early anticipation lies behind the speed of response in conversation. *Journal of Cognitive Neuroscience*, *26*(11), 2530–2539. https://doi.org/10.1162/jocn_a_00673

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Mercier, M. R., Molholm, S., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H., & Foxe, J. J. (2015). Neuro-oscillatory phase alignment drives speeded multisensory response times: An electro-corticographic investigation. *Journal of Neuroscience*, *35*(22), 8546–8557. https://doi.org/10.1523/JNEUROSCI.4527-14.2015

Michail, G., Senkowski, D., Niedeggen, M., & Keil, J. (2021). Memory load alters perception-related neural oscillations during multisensory integration. *Journal of Neuroscience*, *41*(7), 1505–1515. https://doi.org/10.1523/JNEUROSCI.1397-20.2020

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002 Jun). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Brain Research. Cognitive Brain Research*, *14*(1), 115–128. https://doi.org/10.1016/s0926-6410(02)00066-6

Montoya-Martínez, J., Vanthornhout, J., Bertrand, A., & Francart, T. (2021). Effect of number and placement of EEG electrodes on measurement of neural tracking of speech. *PLoS ONE*, *16*(2), e0246769. https://doi.org/10.1371/journal.pone.0246769

Morís Fernández, L., Macaluso, E., & Soto-Faraco, S. (2017). Audiovisual integration as conflict resolution: The conflict of the McGurk illusion. *Human Brain Mapping*, *38*(11), 5691–5705. https://doi.org/10.1002/hbm.23758

Morís Fernández, L., Torralba, M., & Soto-Faraco, S. (2018). Theta oscillations reflect conflict processing in the perception of the McGurk illusion. *European Journal of Neuroscience*, *48*(7), 2630–2641. https://doi.org/10.1111/ejn.13804

Musacchia, G., Sams, M., Nicol, T., & Kraus, N. (2006). Seeing speech affects acoustic information processing in the human brainstem. *Experimental Brain Research*, *168*(1–2), 1–10. https://doi.org/10.1007/s00221-005-0071-5

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological*

Bulletin, *125*(6), 826–859. https://doi.org/10.1037/0033-2909.125.6.826

Nash, H. M., Gooch, D., Hulme, C., Mahajan, Y., McArthur, G., Steinmetzger, K., & Snowling, M. J. (2017). Are the literacy difficulties that characterize developmental dyslexia associated with a failure to integrate letters and speech sounds? *Developmental Science*, *20*(4), e12423. https://doi.org/10.1111/desc.12423

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869. https://doi.org/10.1155/2011/156869

Paulesu, E., Perani, D., Blasi, V., Silani, G., Borghese, N. A., De Giovanni, U., Sensolo, S., & Fazio, F. (2003). A functional-anatomical model for lipreading. *Journal of Neurophysiology*, *90*(3), 2005–2013. https://doi.org/10.1152/jn.00926.2002

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Plewko, J., Chyl, K., Bola, Ł., Łuniewska, M., Dębska, A., Banaszkiewicz, A., Wypych, M., Marchewka, A., van Atteveldt, N., & Jednoróg, K. (2018). Letter and speech sound association in emerging readers with familial risk of dyslexia. *Frontiers in Human Neuroscience*, *12*, 393. https://doi.org/10.3389/fnhum.2018.00393

Raij, T., Uutela, K., & Hari, R. (2000). Audiovisual integration of letters in the human brain. *Neuron*, *28*(2), 617–625. https://doi.org/10.1016/S0896-6273(00)00138-0

Roa Romero, Y., Senkowski, D., & Keil, J. (2015). Early and late beta-band power reflect audiovisual perception in the McGurk illusion. *Journal of Neurophysiology*, *113*(7), 2342–2350. https://doi.org/10.1152/jn.00783.2014

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1015–1022. https://doi.org/10.1016/j.neubiorev.2009.12.006

Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: Implications for multisensory processing. *Trends in Neurosciences*, *31*(8), 401–409. https://doi.org/10.1016/j.tins.2008.05.002

Simon, D. M., & Wallace, M. T. (2018). Integration and temporal processing of asynchronous audiovisual speech. *Journal of Cognitive Neuroscience*, *30*(3), 319–337. https://doi.org/10.1162/jocn_a_01205

Stekelenburg, J. J., Keetels, M., & Vroomen, J. (2018). Multisensory integration of speech sounds with letters vs. visual speech: Only visual speech induces the mismatch negativity. *The European Journal of Neuroscience*, *47*(9), 1135–1145. https://doi.org/10.1111/ejn.13908

Todorovic, A., Schoffelen, J. M., van Ede, F., Maris, E., & de Lange, F. P. (2015). Temporal expectation and attention jointly modulate auditory oscillatory activity in the beta band. *PLoS ONE*, *10*(3), e0120288. https://doi.org/10.1371/journal.pone.0120288

Tremblay, K. L., Billings, C. J., Friesen, L. M., & Souza, P. E. (2006). Neural representation of amplified speech sounds. *Ear and Hearing*, *27*(2), 93–103. https://doi.org/10.1097/01.aud.0000202288.21315.bd

van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, *43*(2), 271–282. https://doi.org/10.1016/j.neuron.2004.06.025

van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: Flexible use of general operations. *Neuron*, *81*(6), 1240–1253. https://doi.org/10.1016/j.neuron.2014.02.044

van Atteveldt, N., Roebroeck, A., & Goebel, R. (2009). Interaction of speech and script in human auditory cortex: Insights from neuro-imaging and effective connectivity. *Hearing Research*, *258*(1–2), 152–164. https://doi.org/10.1016/j.heares.2009.05.007

van Diepen, R. M., & Mazaheri, A. (2018). The caveats of observing inter-trial phase-coherence in cognitive neuroscience. *Scientific Reports*, *8*(1), 2990. https://doi.org/10.1038/s41598-018-20423-z

Wang, F., Karipidis, I. I., Pleisch, G., Fraga-González, G., & Brem, S. (2020). Development of print-speech integration in the brain of beginning readers with varying reading skills. *Frontiers in Human Neuroscience*, *14*, 289. https://doi.org/10.3389/fnhum.2020.00289

Yan, D., & Seki, A. (2024). The role of letter-speech sound integration in native and second language reading: A study in native Japanese readers learning English. *Journal of Cognitive Neuroscience*, *36*(6), 1123–1140. https://doi.org/10.1162/jocn_a_02137

How to cite this article: Yan, D., & Seki, A. (2024). Differential modulations of theta and beta oscillations by audiovisual congruency in letter-speech sound integration. *European Journal of Neuroscience*, 1–14. https://doi.org/10.1111/ejn.16563