Title of Thesis

Study on Attention and Audiovisual Integration by Event-related Potentials

Date of conferring

September 2013

Yulin Gao

The Graduate school of Natural Science and Technology

(Doctor's Course)

OKAYAMA UNIVERSITY

Abstract

In the real world, we are bombarded by information arriving via multiple sensory modalities. Humans constantly receive much information, including visual and auditory stimuli. Our attention system regulates the choice of task-relevant information, and neglecting irrelevant information. However, the irrelevant information is not completely neglected and can affect task completion. Despite the studies investigating the role of influences of attention on audiovisual integration, and studies investigating the influence of multisensory integration on the orienting of attention have been conducted more or less independently of each other, the relationship between attention and neural activity of multimodal audiovisual integration remains unclear.

In the present study, utilizing the high temporal resolution of event-related potentials (ERPs), we combined the relationship between attention and brain activity of audiovisual integration using four experiments. We primarily investigated the effects of spatial or temporal attention and the spatial characteristic of stimuli on audiovisual integration. The effects of audiovisual integration were observed by a simple model, in the difference between the ERPs to audiovisual stimulus and the sum of ERPs to the auditory and visual stimuli presented alone contrasting.

Our results showed that in cue-target paradigm with a task, the temporal attention had no effect on the early-stage of audiovisual sensory processing but the difference of late-stage of audiovisual cognitive processing was observed, and the differential modulation of late N2 in amplitude elicited by AV stimuli at right occipitotemporal area in spatial attention condition was greater than that in temporal attention condition. Moreover, these results that the spatial attention and temporal attention had a differential effect on audiovisual stimuli processing at right occipitotemporal area.

In addition, we also investigated the effects of the spatial characteristics of the audiovisual stimulus on audiovisual integration. Our results showed that the spatial source of the audiovisual stimulus also affected the processing of audiovisual integration.

According to the complexity between attention and brain activity of audiovisual integration, future studies will focus on separating the attention processing mechanism for audiovisual integration. For example, we will separately study the effect of spatial attention on audiovisual integration, the effect of temporal attention on audiovisual integration. Through studying the relationship between attention and brain activity of audiovisual integration, we hope that future studies to clarify its neural correlates.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Summary

This chapter introduces the concept of audiovisual integration and attention, related previous studies, event-related potentials (ERPs), and the method of event-related potential (ERP) analysis in audiovisual integration study. The aim and contents of the thesis are also briefly described.

1.1 Audiovisual integration in the brain

1.1.1 Audiovisual integration

In daily life, people perceive information from the outside world not by one single sensory stream, but by several sensory streams that encode and further process the information in a special area of brain. For example, when enjoying the color, aroma and taste of something delicious, this refers to the information process of the senses of sight, smell and taste. The effective integration and processing of sensory streams from various stimuli is named multisensory integration [1.1].

In early ethology research, the discovery of the McGurk effect is considered to be a landmark event in audiovisual integration research [1.2]. The study found that an integrated influencing effect existed between visual stimuli and auditory pronunciation. For example, when the visual stimuli is the face of a person who is pronouncing the syllable “ga” and the auditory stimuli is the syllable “ba”, the answer of participant is “da”. Since then, neural mechanisms of audiovisual integration have been extensively studied.

Attention is not the same as seeing or perceiving, but on deeper reflection, it is clear that attention involves something more than sensation and perception [1.3]. For example, when we are reading a magazine on a bus the sound of car engines, music from vehicles’ CD players, etc can be distracting. When those noises grab our attention, our eyes can inadvertently skip a line, or even stop reading.

1.1.2 Related studies with attention and audiovisual integration

Despite the studies investigating the role of influences of attention on audiovisual integration, and studies investigating the influence of multisensory integration on the orienting of attention have been conducted more or less independently of each other, the relationship between attention and neural activity of multimodal audiovisual integration remains unclear.

Detailed studies on the characteristics and properties of multisensory neurons in the cat superior colliculus led Stein and colleagues to define four “integration rules” [1.4]: (1) temporal coincidence; (2) spatial coincidence, according to which the largest integration effects are obtained when inputs from different modalities are in close temporal and spatial relationships, respectively; (3) “Magnitude or Inverse Effectiveness Rule,” stipulating that the less effective the unimodal stimuli are, the larger the magnitude of the enhancement they are capable of generating by combination; and (4) the receptive field preservation rule.

Many previous studies have also investigated multisensory audiovisual integration using a attention tasks in behavioral and event-related potential (ERP) measures in humans [1.5-1.8]. Behavioral results have shown that responses to audiovisual stimuli are more rapid and accurate than the responses to either a unimodal visual or auditory stimulus. The ERP results from those studies were limited in that their analysis of ERPs was of the first 200 ms after stimulus presentation, and few assessed the brain activity after this point. Previous neurological research showed that integration occurring during the first up to 200 ms post stimulus onset were mediated by early sensory-perceptual processing, whereas responses after 200ms was mediated by late cognitive processing [1.6, 1.9, 1.10]. Further, previous behavioral research have confirmed that audiovisual perceptual integration was greater when the auditory stimuli were presented in close spatial proximity with the visual stimuli [1.11, 1.12]. Talsma et al. (2007) studied the interactions between attention and audiovisual integration when auditory and visual stimuli were presented centrally. They reported that multisensory integration effects depended on the attention method; early audiovisual integration only occurred in divided-attention tasks. In visual attention tasks in which only visual information was task-relevant and required attention, the visual and auditory stimuli interacted with each other after 200 ms of presentation of the stimulus [1.13].

Moreover, previous studies also investigated effect of spatial characteristics of the audiovisual stimulus on audiovisual integration [1.14-1.16]. In the assumption of unity on multisensory integration, the more in common the properties of stimuli from various streams, the easier they are perceived as a whole [1.17, 1.18]. Macaluso et al. investigated the effect of spatial consistency on audiovisual integration by studying the spatial location relationship between visual stimuli (the face of a person who is pronouncing a word) and auditory stimuli (a word). Their research suggested that the activation of the posterior occipital was found under the condition of spatial inconsistency but not under the condition of spatial consistency. Furthermore, there was not a significant difference in the activation of the ventral occipital and the superior temporal sulcus [1.19]. Gondan et al.(2005)found in their research of ERP that in a waveform induced by audiovisual stimulus under the condition of both spatial inconsistency and spatial consistency, the original difference waves appeared in the component of the posterior parietal cortex from 150 ms to 180 ms [1.15]. Further research by Teder- Sälejärvil et al. in 2005 proved the effect of spatial consistency on audiovisual integration. It has been found that multisensory integration depends on a special neural process. The N260 component of the superior temporal cortex under the condition of spatial consistency is greater than that under the condition of spatial inconsistency from 260 ms to 280 ms. In addition, there existed an oscillatory 10 Hz component in the ventral occipito-temporal cortex by 120 – 300 ms. All the above studies showed that spatial consistency of visual stimuli and auditory stimuli affected neural mechanisms of audiovisual integration [1.14]. The previous research have investigated behavioral performance and neural responses with ERP recordings during audiovisual integration to unilateral AV pairings [1.14, 1.15]. However, effect of bilateral auditory stimuli on audiovisual integration remains unclear.

1.2 Event-related potential (ERP)

1.2.1 What is an event-related potential (ERP)?

The event-related potential (ERP) are voltage changes induced within the brain in response to a variety of sensory, cognitive, and motor processes.. The technique in cognitive neuroscience allows scientists to observe human brain activity that reflects specific cognitive processes [1.20]. Event-related potentials (ERPs) are very small voltages generated in the brain structures in response to specific events or stimuli [1.21]. ERPs in humans can be divided into 2 categories. The early waves, or components peaking roughly within the first 100 milliseconds after stimulus, are termed 'sensory' or 'exogenous' as they depend largely on the physical parameters of the stimulus. In contrast, ERPs generated in later parts reflect the manner in which the subject evaluates the stimulus and are termed 'cognitive' or 'endogenous' ERPs as they examine information processing [1.22].

The ERP consists of a sequence of positive and negative voltage fluctuations that are labeled components. Moreover, ERP components are useful as measures of covert information processing, as differences between conditions can be obtained in the absence of behavioral responding [1.23]. ERP researchers tend, for convenience sake, to identify the positive and negative fluctuations in the overall scalp ERP as the actual components themselves. Yet, this is, to some degree, misleading. Any given ERP waveform recorded at the scalp is actually the summation and cancellation of neural activity from a large number of neural generators from a number of different brain regions. Figure 1.1 shows an ERP experiment equipment.

In addition, ERP provide extremely high time resolution, in the range of one millisecond compared with techniques such as fMRI and PET and are capable of detecting changes in electrical activity in the brain on a time scale. Figure 1.2 shows the 10/20 international standard system, and Figure 1.3 shows recorded ERP data.

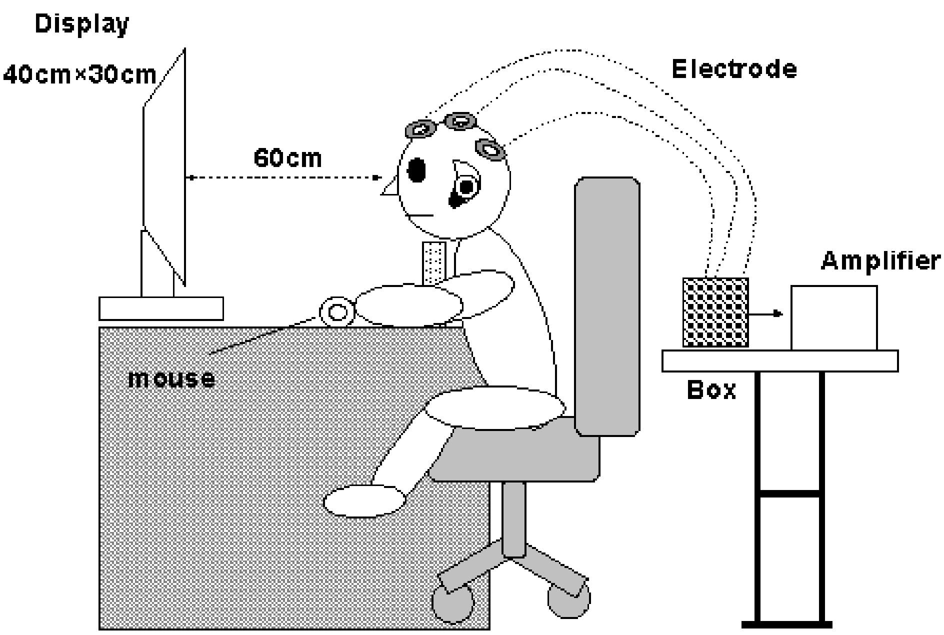

Figure 1.1 An ERPs experiment equipment

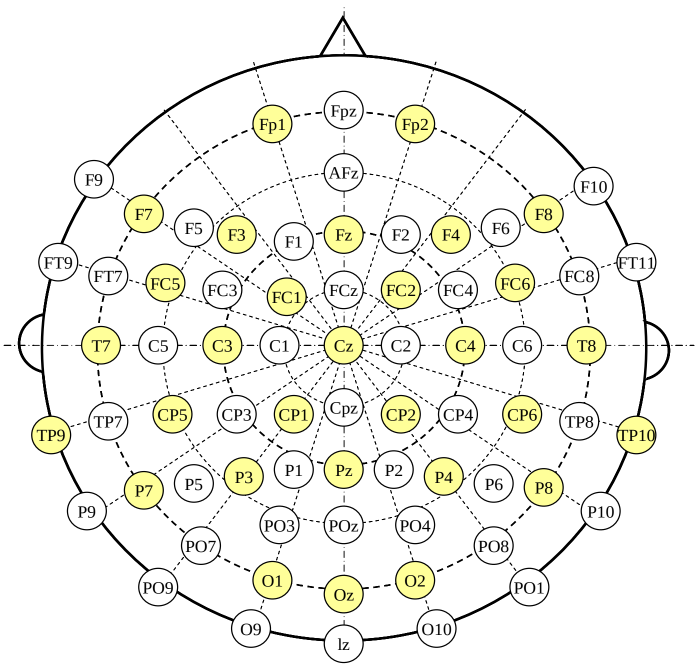

Figure 1.2 Electrode placements according to the 10/20 international standard system

1.2.2 Method of ERP analysis in the audiovisual integration study

In several ERP studies, auditory–visual integration was investigated by comparing the ERP to a bimodal audiovisual (AV) stimulus with the sum of the ERPs to the constituent auditory (A) and visual (V) stimuli [1.4, 1.6, 1.9]. Audiovisual interactions were revealed in the difference waveform formed by subtracting the sum of the ERPs to the individual A and V stimuli from the ERP to the bimodal AV stimulus. Therefore, Interaction components were observed in this AV - (A + V) difference wave.

We assumed the neural activities induced by the multimodal (AV) stimulus were equal to the sum of the neural activities induced separately by the auditory (A) and the visual (V) stimulus, in addition to the putative neuronal activities induced uniquely by multimodal stimulation (auditory-visual interactions). This assumption is valid only while the stimulus analysis is not "contaminated" by late activities related to target processing (P3 waves) or by activities related to the response selection or motor processes (all of these activities are common to all three stimulus types A, V, and AV). We may therefore use the summative model to estimate the AV interactions:

$$ERP(AV) = ERP(A) + ERP(V) + ERP(A \times V \quad Interactions)$$

This expression is valid regardless of the nature or configuration of the intracerebral generators and is based on the law of superposition of electric fields [1.24].

Figure 1.3 Electrode placements according to the 10/20 international standard system

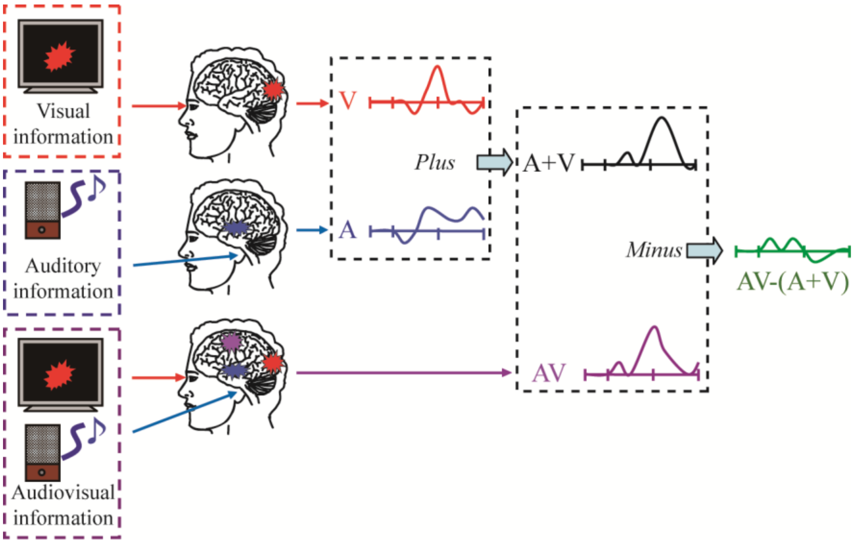

Figure 1.4 Method of ERP analysis in the audiovisual integration study

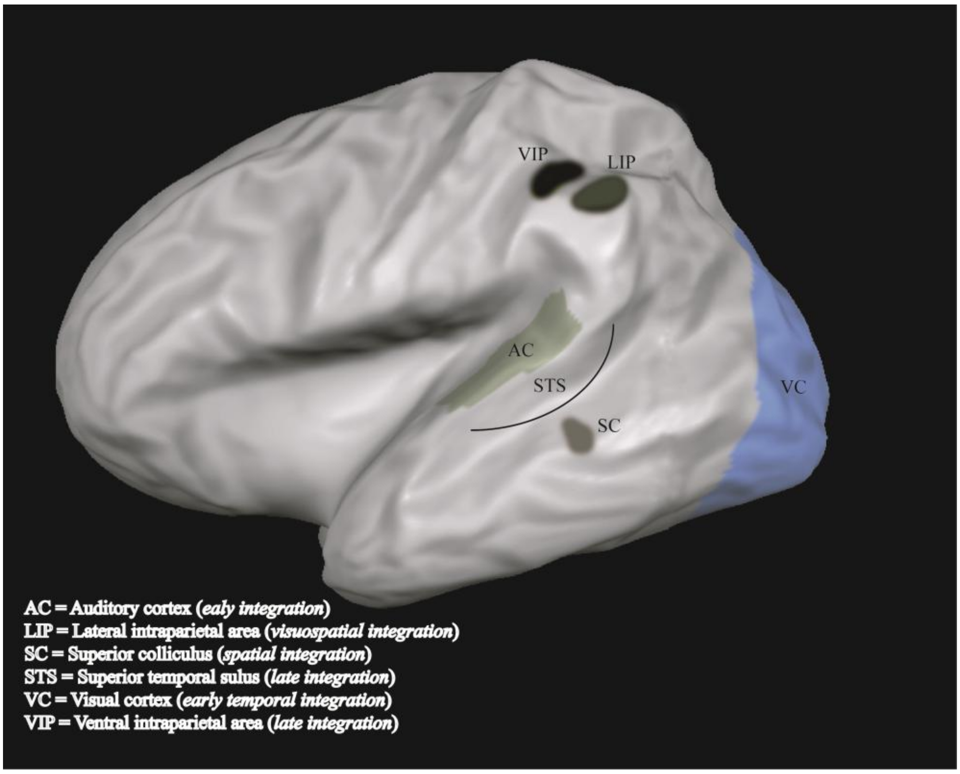

Figure 1.5 Multisensory audiovisual integration

1.3 The purpose of the present dissertation

The main aim of this thesis research was to investigate the brain activities of audiovisual integration using behavioral and ERP with high temporal resolution and to determine the neural mechanism of audiovisual integration in humans. To achieve these aims, four related experiments were conducted.

1.4 The contents of the dissertation

This dissertation mainly investigates the effects of attention and spatial characteristics of stimuli on audiovisual integration. Four experiments and a general discussion are briefly introduced below.

Chapter 1 introduces the concept of attention and audiovisual integration, related previous studies, event-related potentials (ERPs), and the method of event-related potential (ERP) analysis in audiovisual integration study.

Chapter 2 describes the first experiment, in which we applied an event-related potential measurement to investigate the effect of endogenous temporal attention on audiovisual stimuli processing.

Chapter 3 describes the second experiment, in this experiment, we investigated differential modulation of visually cued spatial and temporal attention on audiovisual stimuli processing using event-related potential measurements.

Chapter 4 describes the third experiment, in this experiment, we used behavioral and event-related potential measurements to investigate effect of ipsilateral visual and auditory spatial information interaction mechanisms in selective attention conditions.

Chapter 5 describes the fourth experiment. In this experiment, the effect of ipsilateral and bilateral auditory stimuli on audiovisual integration was investigated using behavioral and event-related potential measurements in a visual detection task.

Chapter 6 describes a general conclusion based on the findings of the four experiments and future challenges.

References

[1.1] Ernst MO, Bülthoff HH. Merging the senses into a robust percept. Trends in cognitive sciences 2004; 8: 162-169.

[1.2] McGurk H, MacDonald J. Hearing lips and seeing voices. 1976.

[1.3] Handy TC, Grafton ST, Shroff NM, Ketay S, Gazzaniga MS. Graspable objects grab attention when the potential for action is recognized. Nature neuroscience 2003; 6: 421-427.

[1.4] Giard M, Peronnet F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. Journal of cognitive neuroscience 1999; 11: 473-490.

[1.5] Fort A, Delpuech C, Pernier J, Giard M-H. Early auditory–visual interactions in human cortex during nonredundant target identification. Cognitive Brain Research 2002; 14: 20-30.

[1.6] Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ. Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. Cognitive Brain Research 2002; 14: 115-128.

[1.7] Teder-Sälejärvi W, McDonald J, Di Russo F, Hillyard S. An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. Cognitive Brain Research 2002; 14: 106-114.

[1.8] Vidal J, Giard M-H, Roux S, Barthelemy C, Bruneau N. Cross-modal processing of auditory–visual stimuli in a no-task paradigm: A topographic event-related potential study. Clinical Neurophysiology 2008; 119: 763-771.

[1.9] Talsma D, Woldorff MG. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. Journal of cognitive neuroscience 2005; 17: 1098-1114.

[1.10] Li Q, Wu J, Touge T. Audiovisual interaction enhances auditory detection in late stage: an event-related potential study. Neuroreport 2010; 21: 173.

[1.11] Lewald J, Guski R. Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. Cognitive Brain Research 2003; 16: 468-478.

[1.12] Frassinetti F, Bolognini N, Làdavas E. Enhancement of visual perception by crossmodal visuo-auditory interaction. Experimental Brain Research 2002; 147: 332-343.

[1.13] Talsma D, Doty TJ, Woldorff MG. Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? Cerebral Cortex 2007; 17: 679-690.

[1.14] Teder-Sälejärvi W, Russo FD, McDonald J, Hillyard SA. Effects of spatial congruity on audio-visual multimodal integration. Journal of cognitive neuroscience 2005; 17: 1396-1409.

[1.15] Gondan M, Niederhaus B, Rösler F, Röder B. Multisensory processing in the redundant-target effect: a behavioral and event-related potential study. Attention, Perception, & Psychophysics 2005; 67: 713-726.

[1.16] Harrington L, Peck C. Spatial disparity affects visual-auditory interactions in human sensorimotor processing. Experimental Brain Research 1998; 122: 247-252.

[1.17] Bertelson P. Ventriloquism: A case of crossmodal perceptual grouping. Advances in psychology 1999; 129: 347-362.

[1.18] Welch RB. Meaning, attention, and the “unity assumption” in the intersensory bias of spatial and temporal perceptions. Advances in psychology 1999; 129: 371-387.

[1.19] Macaluso E, George N, Dolan R, Spence C, Driver J. Spatial and temporal factors during processing of audiovisual speech: a PET study. Neuroimage 2004; 21: 725-732.

[1.20] Luck SJ. An introduction to the event-related potential technique (cognitive neuroscience). 2005.

[1.21] Blackwood D, Muir W. Cognitive brain potentials and their application. The British Journal of Psychiatry 1990.

[1.22] Sur S, Sinha V. Event-related potential: An overview. Industrial psychiatry journal 2009; 18: 70.

[1.23] Friedman D, Johnson R. Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. Microscopy research and technique 2000; 51: 6-28.

[1.24] Mozolic JL, Hugenschmidt CE, Peiffer AM, Laurienti PJ. Modality-specific selective attention attenuates multisensory integration. Experimental Brain Research 2008; 184: 39-52.

Chapter 2

Modulation of response to audiovisual stimuli presented peripherally by visually cued endogenous temporal attention

Summary

Previous studies have used the cue-target paradigm to study the effect of the endogenous temporal attention on only visual or auditory stimuli. Furthermore, some studies found that the visual and auditory stimuli is not processed in isolation but produce coherent cognition in the brain when the visual and auditory stimuli are simultaneously presented. However, the effect of endogenous temporal attention on audiovisual (AV) stimuli processing is unclear. Utilizing the high temporal resolution of event-related potentials (ERPs), we used a central cue that can predict the time point (600 ms or 1800 ms) of audiovisual target to investigate whether endogenous temporal attention could modulate AV stimuli processing. The results showed that the endogenous temporal attention could not change the amplitude of the early ERP component in conditions of either short (600 ms) or long (1800 ms) cue-target intervals, indicating that the endogenous temporal attention had no effect on the early-stage of AV stimuli processing. However, the late ERP component showed differences between the short (600 ms) and long (1800 ms) cue-target intervals, supporting a model in which endogenous temporal attention might determine the late stage of AV stimuli processing.

Keywords: Endogenous temporal attention, audiovisual stimuli, interstimulus interval (ISI), event-related potentials (ERPs)

2.1 Background

In the real world, we are bombarded by information arriving via multiple sensory modalities. Our attention system enables us to focus on task-relevant information and to ignore the irrelevant stimuli [2.1, 2.2]. Attention can be guided in exogenous and endogenous attention patterns [2.3]. Exogenous attention (stimulus-driven attention) can be triggered reflexively by a salient sensory event that demands attention, whereas endogenous attention (goal-driven attention) involves a more purposeful orienting process [2.4]. For example, when we study in the classroom, a man who breaks through the door suddenly can attract our attention reflexively, which is stimulus-driven (the intruder), the so-called exogenous attention. Yet, hearing the school bell, we will consciously pay attention to the door to wait for the teacher, which is goal-driven (the teacher comes to the classroom), the so-called endogenous attention.

One approach to studying endogenous attention is the visual cue-visual target paradigm, in which the central spatial cue (left or right arrow) or temporal cue (small or large circle) can predict the location or time point of the target stimuli [2.1]. This attention induced by a central spatial or temporal cue is usually called endogenous spatial or temporal attention. With the visual cue-visual target paradigm, we observed when stimulus processing was speeded up at attended locations or time points [2.1, 2.5-2.7]. Some neurophysiologic studies have investigated the neural mechanism of endogenous spatial attention orienting [2.8] that involves a more purposeful orienting process controlled by the attending person [2.4]. With the visual cue-visual target paradigm, it has been suggested that the time interval between the cue and the target stimulus (interstimulus interval, ISI) can modulate the attentional effect [2.9]. Several studies have reported that event-related potential (ERP) components elicited with relatively short ISIs have smaller amplitudes than those obtained with longer ISIs [2.10-2.12]. In addition, some studies have used a similar paradigm to investigate the effect of visually cued endogenous temporal attention on visual stimuli processing presented peripherally [2.6]. In a case study, the production of endogenous temporal attention is through two concentric circles (rather than an arrow) at the center of the visual field, where the two concentric circles provide a cue to the time interval (short or long) between the cue and target stimuli. Specifically, when the inner small circle was presented, it indicated that the target would appear within a short time interval, whereas when the outer large circle was presented, it indicated that the target would appear after a longer time interval [2.13]. The imaging result showed that the frontoparietal network (FPN) is involved in endogenous temporal attention [2.13]. Moreover, Li et al. used the cue-target task to determine the effect of visually cued temporal attention on auditory stimulus processing and found that the specific activations related to temporal cognition were confirmed within the superior occipital gyrus, tegmentum, motor area, thalamus and putamen [2.14]. Furthermore, previous studies have indicated that both behavioral and electrophysiological responses to visual and auditory stimuli are improved when those stimuli are accompanied by a stimulus in a different modality and that bimodal audiovisual (AV) stimuli are detected and discriminated more accurately than either visual (V) or auditory (A) unimodal stimuli presented alone [2.15, 2.16]. Thus, we predicted that the visually cued endogenous temporal attention modulation of response to audiovisual stimuli presented peripherally would have different mechanisms than that of response to visual or auditory stimuli presented alone. However, how the visually cued endogenous temporal attention modulates the response to audiovisual stimuli presented peripherally is still unclear.

The aim of the present study is to investigate whether different ERPs elicited by AV stimuli can be observed in the short or long ISI conditions of endogenous temporal attention. Using the high temporal resolution of the ERP technique, we can observe the general processing of AV stimuli, which are preceded by a central visual temporal cue predicting the visual cue and AV target interval. In the context of the previous results mentioned above, we predicted that dissociate interaction could be found between the short and long ISI, either in earlier ERP components or in later ERP components, or even both.

2.2 Methods

2.2.1 Subjects

Behavioral experiment

Eleven undergraduate students (age range: 23-30 years; mean age: 25.3 years; 9 male) were recruited for the behavioral experiment as paid volunteers.

ERP experiment

Eleven undergraduate students (age range: 21-25 years; mean age: 23.2 years; 7 male) were recruited for the ERP experiment as paid volunteers.

All participants were right-handed and had normal or corrected-to-normal vision and had no history of neurological or psychiatric disorders. They had not participated in similar experiments during the past year. The experimental protocol was approved by the Ethics Committee of Okayama University.

2.2.2 Stimuli and procedure

The experiment was conducted in a dimly lit, sound-attenuated, electrically shielded room (laboratory room, Okayama University, Japan). The procedure was totally consistent between the behavioral and ERP experiments.

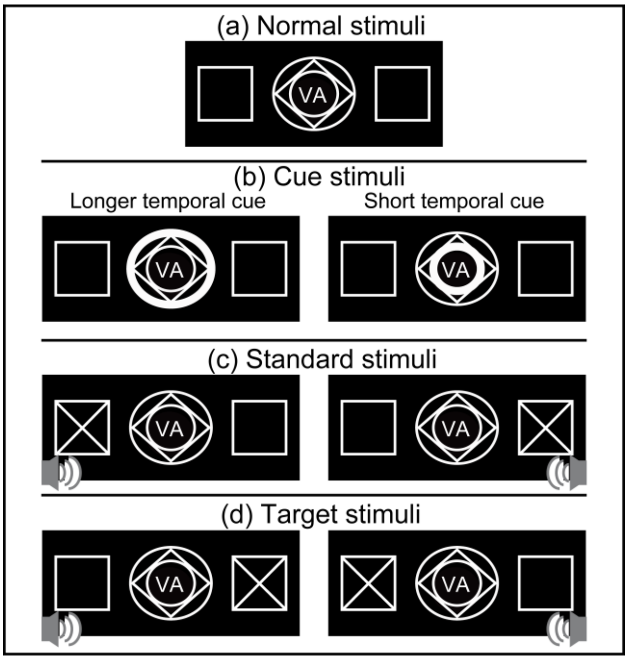

Figure 2.1 Attentional cues used to direct subjects' attention to a stimulus-onset time. The temporal cue directs attention to a short or long stimulus-onset time. The standard stimuli were simultaneous AV stimuli presented in the same locations, and the target stimuli were simultaneous AV stimuli presented in opposite locations. (a) Normal stimuli; (b) cue stimuli; (c) standard stimuli; (d) target stimuli.

As shown in Fig 2.1, a stream comprised normal stimuli, cue stimuli, standard AV stimuli, and target AV stimuli. The normal stimuli are shown in Fig 2.1a and were composed of two peripheral left and right boxes (2°×2°, centers 7° from the center of the monitor) and a fixation stimulus (2°× 2°); the fixation stimulus was a compound stimulus consisting of a diamond (2 × 2 cm, subtending a visual angle of approximately 2°) and two concentric circles (large circle, 2°× 2°; small circle, 1.4°× 1.4°). The temporal cue was reported in Fig 2.1b, which could completely predict (100%) the cue-target intervals. One part of the cue was highlighted to inform the subject whether to attend to the time of the target (600 or 1800 ms from cue presentation): a brightening of the inner small circle indicated that the target would appear within a short time interval (600 ms), whereas a brightening of the outer large circle represented a longer time interval (1800 ms). There were 2 types of AV stimuli: standard and target stimuli. As shown in Fig 2.1c, standard stimuli refer to the visual stimulus (was the letter X, 2° × 2°) and the auditory stimulus (single pure tone, 1600 Hz, played at an intensity of 65 dB) being presented at the same location (left or right boxes, centers 7° from the center of the monitor), and these stimuli, to which no response was required. The target stimuli, as shown in Fig. 2.1d, are defined as the condition in which the visual stimulus and the auditory stimulus were presented at opposite locations; these stimuli, which required responses. In this study, the purpose of the targets was to ensure that participants attended to the stimuli during the entire experiment [2.17].

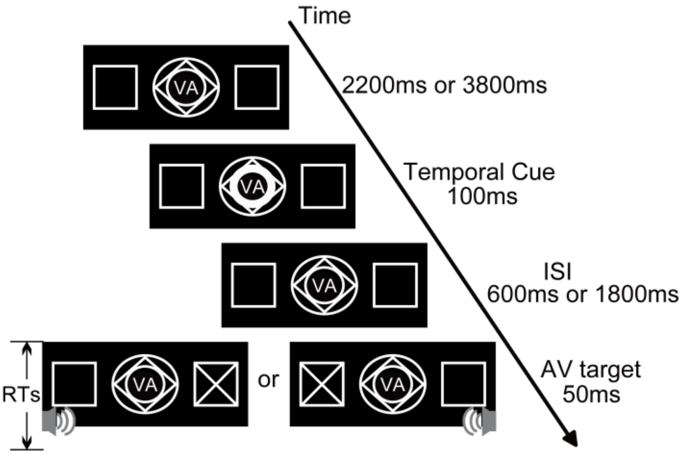

Figure 2.2 A typical trial, which, in this example, directs the subjects' attention to a stimulus-onset time with no information about the target's location. The attentional cue is on for 100 ms, the cue-target interval is either 600 or 1800 ms (short/long cue), and the AV target appears for 50 ms, which was the visual stimulus and the auditory stimulus were presented at opposite locations.

As shown in Fig 2.2, streams of stimuli were randomly presented against a black background. During the experimental task, the participants were asked to fix their eyes on a central location and to pay attention to the temporal cue for one of two temporal interval lengths (600 or 1800 ms). The participants were seated 60 cm from the center of the monitor. At the beginning of each trial, the normal stimuli were presented in the center of the monitor. After 2200 ms or 3800 ms, one of the two temporal cues was presented for 100 ms. During the temporal cue, the participants were instructed to estimate when the AV stimuli event would occur; the interval length between presentation of the cue and the AV stimulus event was either 600 or 1800 ms, which occurred with equal probability (50%). Following the interval after cue-stimulus presentation, the AV stimuli appeared for 50 ms. There were two types of AV stimuli: standard stimuli, presented at the same location (left or right), and target stimuli, presented at the opposite location. Standard stimuli were used 77% of all stimuli, and target stimuli were used 23% of all stimuli. In the last, the normal stimuli appeared for 1450 ms to allow participants to make corresponding responses. For AV target stimuli, the response was required through a reaction key as accurately as possible. Throughout the experiment, subjects were required to fix their eyes on a centrally presented fixation point on a screen.

For the standard stimuli, there were 200 trials, including 100 trials each for the ISIs of 600 ms and 1800 ms, and the probability of the stimulus appearing at the left or right location was equivalent. For the target stimuli, there were 60 trials, and the distribution was the same as for the standard target. In each block, there were 40 standard trials and 12 target trials. Participants were allowed to take a 5-min break between 5 sessions. At the beginning of the normal experiment, there was a practice block for participants to teach them the experimental task.

2.2.3 Apparatus and software

Stimulus presentation was controlled by a personal computer running Presentation software (Neurobehavioral Systems, Albany, CA). An EEG system (BrainAmp MR plus, Gilching, Germany) was used to record EEG signals through 32 electrodes mounted on an electrode cap (Easy cap, Herrsching Breitbrunn, Germany), as specified by the International 10-20 System. All signals were referenced to the bilateral earlobe electrodes. Horizontal eye movements were recorded from the outer canthus of the right eye; eye blinks and vertical eye movements were recorded from the vEOG electrode. The impedance of all of the electrodes was kept below 5 kΩ. Raw signals were digitized with a sample frequency of 500 Hz with a 60 Hz notch filter and stored continuously on a compatible computer for offline analysis. The event-related potential (ERP) analysis was carried out using Brain Vision Analyzer software (version 1.05, Brain Products GmbH, Munich, Bavaria, Germany).

2.2.4 Data analysis

Behavioral results analysis

Reaction times (RTs), hit rates (HRs) and detection sensitivity (d') were analyzed separately for each ISI condition. Hit rates were the number of correct responses to target stimuli divided by the total number of target stimuli. Response time data were analyzed for correct responses. All data from eleven participants was normally distributed (Shapiro-Wilk test, p>.05). Therefore, RTs, HRs and d' for different ISI conditions were subjected to a repeated-measures analysis of variance (ANOVA) with the factor of ISI (long and short) and a significance level of 0.05.

ERP results analysis

Only the ERPs elicited by the AV standard stimuli were analyzed to remove the response movement effect. The EEG and EOG signals were amplified and band-pass filtered with an analog filter of 0.01-100 Hz at a sampling rate of 500 Hz. Continuous EEG and EOG signals were divided into epochs from 100 ms before AV stimuli onset to 800 ms after stimulus onset. Baseline corrections were made against -100 ms-0 ms. We automatically rejected trials with vertical eye movements and eye blinks (vertical EOG amplitudes exceeding ±100 μV), horizontal eyeball movements (horizontal EOG amplitudes exceeding ±25 μV), or other artifacts (a voltage exceeding ±75 μV at any electrode location relative to baseline). Responses associated with false alarms were also rejected from the analysis. The data were then averaged for each stimulus type, following digital filtering with a band-pass filter of 0.01-30 Hz. The grand-averaged data were obtained across all participants for each stimulus type. Because no significant lateralization effect of AV stimuli has been found, ERP data from the left and right hemispaces were combined to improve the signal-to-noise ratio of the ERPs [2.18].

The ERP wave was measured as the mean amplitude for the short ISI condition and the long ISI condition from each electrode (32 channels). All electrodes (except vEOGs & hEOGs) were selected for statistical analysis. The ERP result for each condition was derived by averaging 90 valid trials after the removal of null data. Based on an inspection of the grand averages, we selected five time windows within the following parentheses (P1: 70-90 ms; N1: 140-160 ms; a positivity: 220-260 ms; a late negativity: 320-340 ms; a late positivity: 400-500 ms). The mean amplitudes were entered into a repeated-measures analysis of variance with factors of ISI (600 ms vs. 1800 ms) and electrodes (P1: Fz, F3, F4, FC1, FC2, Cz; N1: Fz, F3, F4, FC1, FC2, Cz; a positivity: Fz, F3, F4, FC1, FC2, Cz, C3, C4, CP1, CP2, CP5, CP6, PZ, P3, P4; a late negativity: Fz, F3, F4, FC1, FC2, Cz, C3, C4; a late positivity: Cz, C3, C4, CP1, CP2, CP5, CP6, Pz, P3, P4) separately. We also tested the vEOGs & hEOGs for each condition to control for any differences in ERPs caused by micro-movements of the eye or by slight winking. The Greenhouse-Geisser epsilon correction was used for non-sphericity when appropriate. Statistical significance was set at 0.05. In addition, we applied the Bonferroni correction to post-hoc comparisons. All statistical analyses were carried out using SPSS, version 16.0 (SPSS, Tokyo, Japan).

2.3 Results

2.3.1 Behavioral results

A significant difference between the long and short ISI conditions is shown in Table 2.1. RT to the AV target stimuli in the short ISI condition were significantly faster [$F(1,10) = 10.206$, $p < 0.01$] than those to AV target stimuli in the long ISI condition. However, there were no differences in HR [$F(1,10) = 1.000$, $p = 0.341$] or d' [$F(1,10) = 0.149$, $p = 0.707$] associated with the long and short ISI conditions. All subjects responded to the task with accuracy above 99%.

Table 2.1 Mean reaction times (RTs), hit rates (HRs) and detection sensitivity (d') in each ISI condition

ISI condition	Long	Short
RTs (ms)	697 (45)	672 (40)
HRs (%)	98.9 (0.7)	99.3 (0.4)
d′	3.3 (0.2)	3.4 (0.1)

Standard deviations are given in parentheses.

2.3.2 Event-related potential results

Fig 2.3 shows the ERPs elicited by AV standard stimuli. ERPs to AV standard stimuli were characterized by a small P1 wave (peaking approximately 80 ms post-stimulus), which was followed by a N1 component (peaking approximately 150 ms); in the time window of 220-260 ms, a positivity component developed. Then, a late negativity component (320-340 ms) was induced. Lastly, a late positivity component (400-500 ms) developed.

The Effects of ERP Components before 300 ms Post-Stimulus

As shown in Fig 2.3, first, the mean amplitude of P1 (70-90 ms) at the Fz, F3, F4, FC1, FC2, and Cz electrodes was analyzed by ANOVA. The main effect of the ISI was not significant [$F(1, 10) = 0.32$, $p > 0.05$]. The main effect of the electrode was significant [$F(5, 50) = 3.265$, $p < 0.05$]. The interaction between the electrode and the ISI was also not significant [$F(5, 50) = 0.413$, $p > 0.8$]. Topographic differences between the short ISI and the long ISI were observed. Paired-sample t-tests were not significantly different on the vEOG or the hEOG during the epoch of 70-90 ms ($p > 0.05$).

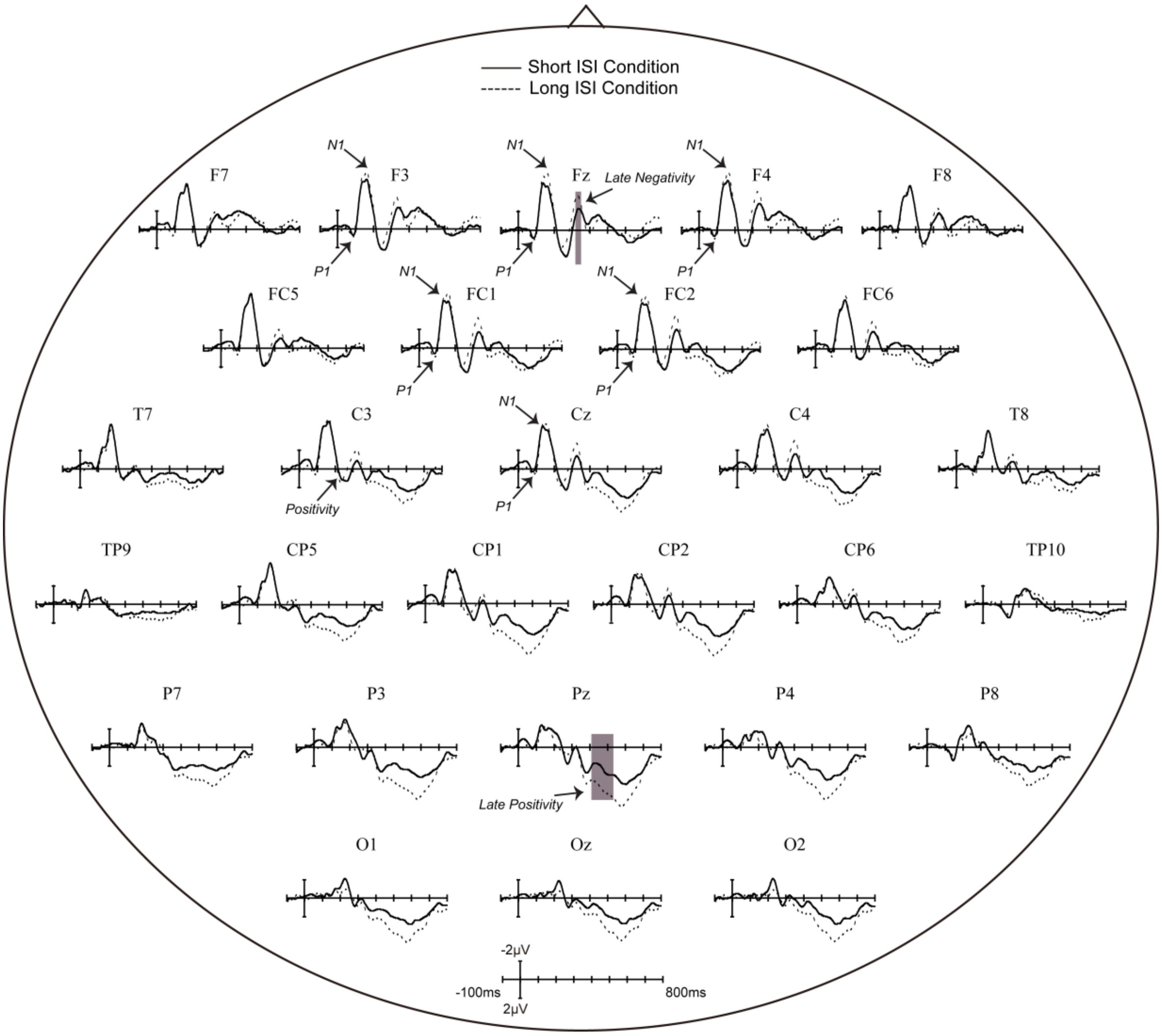

Fig 2.3 Grand average ERPs elicited by AV stimuli in short ISI (black solid line) and long ISI (black dotted line) conditions. Electrode distribution is also illustrated.

Second, the mean amplitude of N1 (140-160 ms) at the Fz, F3, F4, FC1, FC2, and Cz electrodes was analyzed by ANOVA. The main effect of the ISI was not significant [$F(1, 10) = 0.271$, $p > 0.6$]. The main effect of the electrode was significant [$F(5, 50) = 4.691$, $p < 0.01$]. The interaction between the electrode and the ISI was also not significant [$F(5, 50) = 0.22$, $p > 0.9$]. Topographic differences between the short ISI and the long ISI were observed. Paired-sample t-tests were not significantly different on the vEOG or the hEOG during the epoch of 140-160 ms ($p > 0.05$).

Last, the mean amplitude of the positivity component (220-260 ms) at the Fz, F3, F4, FC1, FC2, Cz, C3, C4, CP1, CP2, CP5, CP6, Pz, P3, and P4 electrodes was analyzed by ANOVA. The main effect of the ISI was not significant [$F(1, 10) = 1.371$, $p > 0.4$]. The main effect of the electrode was not significant [F(14, 140) = 1.12, p > 0.05]. The interaction between the electrode and the ISI was also not significant [F(14, 140) = 0.4, p > 0.08]. Here, no topographic differences between the short ISI and the long ISI were observed. Paired-sample t-tests were not significantly different on the vEOG or the hEOG during the epoch of 220-260 ms (p > 0.05).

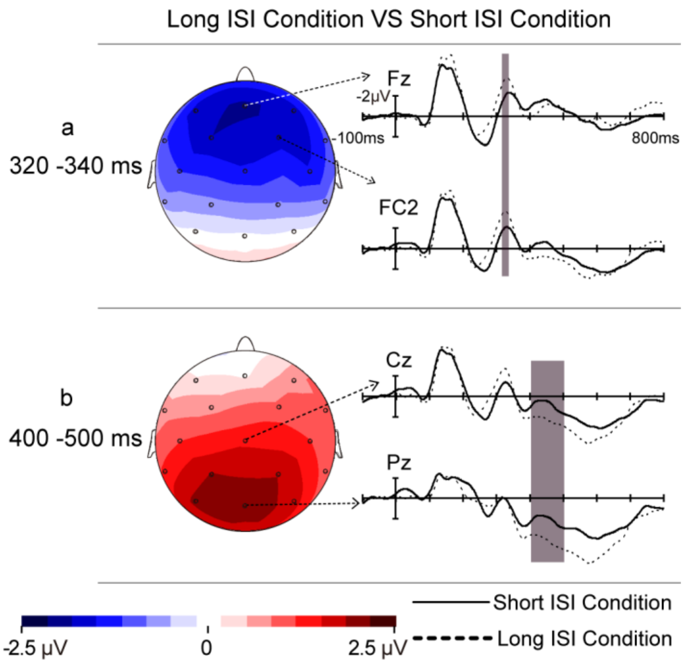

Fig 2.4 Two-dimensional (2D) parameter images of the long ISI condition minus the short ISI condition during each epoch after the onset of the AV stimuli are shown on the left. Grand average ERPs elicited by AV standard stimuli in the short ISI condition (black solid line) and the long ISI condition (black dotted line) in the most activated electrode are shown on the right. (a) The time epoch used for analyzing the positivity during the epoch of 320-340 ms is shaded grey; the ERPs displayed are from the Fz and FC2 electrodes. (b) The time epoch used for analyzing the late negativity (400-500 ms) is shaded grey; the ERPs displayed are from the Cz and Pz electrodes.

The Effects of the ERP Component after 300 ms Post-Stimulus

The mean amplitude of the late negativity component at the Fz, F3, F4, FC1, FC2, Cz, C3, and C4 electrodes during the epoch of 320-340 ms was analyzed by ANOVA. The main effect of the ISI was significant [F(1, 10) = 8.23, p < 0.05], and the main effect of the electrode was significant [F(7, 70) = 3.32, p < 0.01]. Although the interaction between the electrodes and the ISI was also significant [F(7, 70) = 5.11, p < 0.01], post-hoc analysis showed that the greatest difference between the short and long ISI conditions was found at the Fz (p < 0.05) and FC2 (p < 0.05) electrodes (see

Fig 2.3 and Fig 2.4a), indicating that the mean amplitude of the late negativity component in the long ISI condition (-3.97 μV in Fz, -3.68 μV in FC2) is larger than that in the short ISI condition (-2.27 μV in Fz, -2.13 μV in FC2). Paired-sample t-tests were not significantly different on the vEOG or the hEOG during the epoch of 300-320 ms ($p > 0.05$), indicating that the modulatory effects of ISI were not caused by small movements or blinking. The electrode distribution over the fronto-central areas is shown in Fig 2.4a.

The mean amplitude of the late positivity component at the Cz, C3, C4, CP1, CP2, CP5, CP6, Pz, P3, and P4 electrodes during the epoch of 400-500 ms was analyzed by ANOVA. The main effect of the ISI was significant [$F(1, 10) = 7.15$, $p < 0.05$], and the main effect of the electrode was significant [$F(9, 90) = 2.65$, $p < 0.01$]. Although the interaction between the electrodes and the ISI was not significant [$F(9, 90) = 1.47$, $p > 0.05$], post-hoc analysis showed that the greatest difference between the short and long ISI conditions was found at the Cz ($p < 0.05$) and Pz ($p < 0.01$) electrodes (see Fig 2.3 and Fig 2.4b), indicating that the mean amplitude of the late positivity component in the long ISI condition (2.22 μV in Cz, 4.46 μV in Pz) is larger than that in the short ISI condition (0.80 μV in Cz, 2.32 μV in Pz). Paired-sample t-tests showed no significant differences in the vEOG or the hEOG during the epoch of 400-500 ms ($p>0.05$), indicating that the modulatory effects of ISI were not caused by small movements or blinking. The electrode distribution over the centro-parietal areas is shown in Fig 2.4b.

2.4 Discussion

The present study aimed to investigate whether the ISI affected the processing of AV stimuli by endogenous temporal attention. The behavioral results showed that the responses to the AV target were faster in the short ISI condition than in the long ISI condition. Previous studies indicated that, apart from this orienting effect, cues are also thought to enhance alertness, as reflected in faster responses in conditions with cues compared to conditions without cues [2.19]. Through our research, the enhancement of alertness was shown to be stronger with a short ISI than with a long ISI. Moreover, the d’ parameter reflecting the subject’s accuracy to discern a sensory event from its background (perceptual level) was investigated. However, in this study, we found no differences in detection sensitivity (d') between the long and short ISI conditions. And no enhancement of d’ for AV target detection was found when the auditory stimulus was presented simultaneously different or opposite location [2.20]. Therefore, the present d' results suggested that this effect of endogenous temporal attention is not mediated by early-stage sensory processing.

The ERP results showed that the short ISI and the long ISI had the same effect on early ERP components, which are the P1 component (70-90 ms), the N1 component (140-160 ms) and a positivity component (220-260 ms) elicited by AV stimuli. However, from 300 ms after the stimulus onset to approximately 500 ms, we observed distinct effects of the short and long ISIs. Specifically, the mean amplitudes of the positivity component during the epochs of 320-340 ms and 400-500 ms were larger in the long ISI condition than in the short ISI condition. The ERP results were consistent with the behavioral results, in which the effect of endogenous temporal attention originates from late-stage cognitive processing rather than early-stage sensory processing.

The ERP Components before 300 ms Post-Stimulus

We found that the short and long ISIs had the same effect on audiovisual stimuli processing before 300 ms stimulus onset (see Fig 2.3), such as the P1 and N1 components and the positivity component, which are related to early-stage sensory processing. While previous studies found that the auditory N1 and the visual P1 can be enhanced in the long ISI rather than in the short ISI [2.21]. the audiovisual interaction might account for this difference, it has been suggested that the auditory/visual stimulus can enhance the simultaneous visual/auditory stimulus detection, and this enhancement effect was mediated by early-stage sensory processing [2.22]. In the present study, the auditory stimulus was presented simultaneously with the visual stimulus; consequently, the early ERPs elicited by the audiovisual stimuli might be enhanced. Therefore, a similar modulation effect of the short ISI and the long ISI was observed here, indicating that the modulatory mechanism of the ISI was affected by the audiovisual interaction.

The Late Components after 300 ms Post-Stimulus

We found that, beginning at 300 ms after stimulus onset, there were distinctions between the short ISI and the long ISI (see Fig 2.4). Specifically, the negativity component (320-340 ms) and the positivity components (400-500 ms) were larger in the long ISI condition than in the short ISI condition. The late positivity ERP components are P3-like components [2.23, 2.24]. The P3-like component elicited by standard stimuli with temporally endogenous attention could be affect by the ISI, which was similar but not identical to results from previous studies. We found that the mean amplitude of the late negativity and the positivity in the long ISI condition at fronto-central areas was greater than in the short ISI condition. These findings are in accordance with those of previous studies [2.25, 2.26]. The previous studies investigated the effects on endogenous spatial attention by long ISI and short ISI conditions in tactile events, and they found substantial activation of the late negativity component at fronto-central sites in the long ISI condition [2.25]. The frontal sites are related to the orienting of attention [2.14]. Moreover, previous studies on audiovisual interaction have reported that the spread of spatial attention can cross modalities in a multisensory object [2.27, 2.28]. These findings may indicate that the endogenous spatial attention can modulate multisensory audiovisual events. Furthermore, the frontal cortices were also activated during endogenous spatial attention and endogenous temporal attention tasks [2.14]. Therefore, we speculated that the endogenous temporal attention can also modulate responses to audiovisual stimuli.

Lastly, the change in amplitude of P3 with different cue-target intervals is related to the generation mechanism of the late positivity component. Gonsalvez et al. (2002) proposed that target-target interval (TTI) was important to P3 [2.26]. In our experiment, we controlled the interval between cue and target, which means that the interval between target and target also changed, and the amplitude of the late positivity component increased with the cue-target interval. Moreover, it has been found that the cue-target interval (ISI) had effects on target stimulus processing. Typically, a facilitation effect was found in short ISI condition, while an inhibition effect was observed in long ISI condition [2.29, 2.30]. Therefore, the exogenous attention orienting effect would change with differences in the cue-target interval. Similar with exogenous, although still be controversial, the central cue also could induce early facilitation and inhibit the endogenous attention effect with saliency-driven stimuli [2.31]. In summary, our results show that endogenous temporal attention can also affect audiovisual processing.

2.5 Conclusion

In this study, we investigated whether peripherally presented AV stimuli could be modulated by endogenous temporal attention when the central visual cue could completely predict the cue-target interval (600 ms or 1800 ms). Specifically, the modulation effect appeared in the late ERP components from 300 ms post-stimuli, such as the late negativity component (320-340 ms) and the late positivity component (400-500 ms), rather than in the early ERP components (P1 & N1). The ERP results suggested that endogenous temporal attention can modulate the response to AV stimuli reflected in the late components but cannot affect the early ERP components.

References

[2.1] Posner MI. Orienting of attention. Quarterly journal of experimental psychology 1980; 32(1): 3-25.

[2.2] Carrasco M, Ling S, Read S. Attention alters appearance. Nature neuroscience 2004; 7(3): 308-313.

[2.3] Theeuwes J. Exogenous and endogenous control of attention: The effect of visual onsets and offsets. Perception & psychophysics 1991; 49(1): 83-90.

[2.4] Brosch T, Pourtois G, Sander D, et al. Additive effects of emotional, endogenous, and exogenous attention: behavioral and electrophysiological evidence. Neuropsychologia 2011; 49(7): 1779-1787.

[2.5] Coull JT, Nobre AC. Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. The Journal of Neuroscience 1998; 18(18): 7426-7435.

[2.6] Griffin IC, Miniussi C, Nobre AC. Multiple mechanisms of selective attention: differential modulation of stimulus processing by attention to space or time. Neuropsychologia 2002; 40(13): 2325-2340.

[2.7] Xiaoyu Tang CL, Qi Li, Yulin Gao, Weiping Yang, Jingjing Yang, Soushirou Ishikawa, Jinglong Wu. Modulation of auditory stimulus processing by visual spatial or temporal cue: an event-related potentials study. Neuroscience Letters 2013: in press.

[2.8] Peelen MV, Heslenfeld DJ, Theeuwes J. Endogenous and exogenous attention shifts are mediated by the same large-scale neural network. Neuroimage 2004; 22(2): 822-830.

[2.9] Doallo S, Lorenzo-Lopez L, Vizoso C, et al. The time course of the effects of central and peripheral cues on visual processing: an event-related potentials study. Clinical Neurophysiology 2004; 115(1): 199-210.

[2.10] Fitzgerald PG, Picton TW. Temporal and sequential probability in evoked potential studies. Canadian journal of psychology 1981; 35(2): 188.

[2.11] Woods DL, Courchesne E. The recovery functions of auditory event-related potentials during split-second discriminations. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section 1986; 65(4): 304-315.

[2.12] Woods DL, Courchesne E, Hillyard SA, et al. Recovery cycles of event-related potentials in multiple detection tasks. Electroencephalography and clinical neurophysiology 1980; 50(5): 335-347.

[2.13] Coull JT, Vidal F, Nazarian B, et al. Functional anatomy of the attentional modulation of time estimation. Science 2004; 303(5663): 1506-1508.

[2.14] Li C, Chen K, Han H, et al. An fMRI Study of the Neural Systems Involved in Visually Cued Auditory Top-Down Spatial and Temporal Attention. PLoS One 2012; 7(11): e49948.

[2.15] Li Q, Wu J, Touge T. Audiovisual interaction enhances auditory detection in late stage: an event-related potential study. Neuroreport 2010; 21(3): 173-178.

[2.16] Wu J, Li Q, Bai O, et al. Multisensory interactions elicited by audiovisual stimuli presented peripherally in a visual attention task: a behavioral and event-related potential study in humans. Journal of Clinical Neurophysiology 2009; 26(6): 407-413.

[2.17] Senkowski D, Saint-Amour D, Kelly SP, et al. Multisensory processing of naturalistic objects in motion: a high-density electrical mapping and source estimation study. Neuroimage 2007; 36(3): 877-888.

[2.18] Talsma D, Woldorff MG. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. Journal of cognitive neuroscience 2005; 17(7): 1098-1114.

[2.19] Van der Lubbe RH, Keuss PJ, Stoffels E-J. Threefold effect of peripheral precues: alertness, orienting, and response tendencies. Acta Psychologica 1996; 94(3): 319-337.

[2.20] Frassinetti F, Bolognini N, Làdavas E. Enhancement of visual perception by crossmodal visuo-auditory interaction. Experimental brain research 2002; 147(3): 332-43.

[2.21] Coch D, Skendzel W, Neville HJ. Auditory and visual refractory period effects in children and adults: An ERP study. Clinical Neurophysiology 2005; 116(9): 2184-2203.

[2.22] Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. Cognitive Brain Research 2003; 17(2): 447-453.

[2.23] Verleger R. Event-related potentials and cognition: A critique of the context updating hypothesis and an alternative interpretation of P3. Behavioral and brain sciences 1988; 11(3): 343-456.

[2.24] Lange K, Rösler F, Röder B. Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: An event ‐ related potential study. Psychophysiology 2003; 40(5): 806-817.

[2.25] Velzen J, Forster B, Eimer M. Temporal dynamics of lateralized ERP components elicited during endogenous attentional shifts to relevant tactile events. Psychophysiology 2002; 39(6): 874-878.

[2.26] Gonsalvez CJ, Polich J. P300 amplitude is determined by target ‐ to ‐ target interval. Psychophysiology 2002; 39(3): 388-396.

[2.27] Busse L, Roberts KC, Crist RE, et al. The spread of attention across modalities and space in a multisensory object. Proceedings of the National Academy of Sciences of the United

States of America 2005; 102(51): 18751-18756.

[2.28] Donohue SE, Roberts KC, Grent T, et al. The cross-modal spread of attention reveals differential constraints for the temporal and spatial linking of visual and auditory stimulus events. The Journal of Neuroscience 2011; 31(22): 7982-7990.

[2.29] Posner MI, Cohen Y. Components of visual orienting. Attention and performance X: Control of language processes 1984; 32: 531-556.

[2.30] Ming Zhang XTaJW. Blocking the Link between Stimulus and Response at Previously Attended Locations: Evidence for Inhibitory Tagging Mechanism. Neuroscience and Biomedical Engineering 2013; 1(1): 13-21.

[2.31] Henderickx D, Maetens K, Soetens E. The involvement of bottom-up saliency processing in endogenous inhibition of return. Attention, Perception, & Psychophysics 2012; 74(2): 285-299.

Chapter 3

Effect of visually cued spatial and temporal attention on audiovisual stimuli processing

Summary

Previous studies investigate the effect of the spatial and temporal attention on visual or auditory stimuli processing. Furthermore, the visual and auditory information simultaneously received from the different modalities must be integrated by several systems to produce coherent cognition in the brain. However, how the spatial and temporal attention modulates the audiovisual (AV) stimulus processing is still unclear. The aim of this study was to compared the modulatory effects of spatial attention versus temporal attention on audiovisual stimuli processing using event-related potentials (ERPs) with high temporal resolution. The spatial attention was triggered by a visual spatial cue (usually an arrow) and the temporal attention was triggered by a visual temporal cue (two concentric circles). Behavioral responses to audiovisual stimuli in spatial attention condition were faster than those in temporal attention condition, and the false alarm rate in spatial attention condition was lower than that in temporal attention condition. The ERPs results show that the amplitude of N2 elicited by AV stimuli at right occipitotemporal area in spatial attention condition was greater than that in temporal attention condition. Those results indicated that the spatial attention and temporal attention had a differential effect on AV stimuli processing at right occipitotemporal area.

Keywords: Spatial attention, temporal attention, audiovisual stimuli, interstimulus interval (ISI), event-related potentials (ERPs)

3.1 Background

In real world, there is a large amount of information bombarding us. Attention can helps us to selectively concentrating on one aspect of the useful information in the environment while ignore others. Attention is distributed across time as well as space, the spatial attention manipulate attention to spatial location of target presented [3.1] while the temporal attention manipulate attention toward when the target could appear [3.2, 3.3].

A previous neurophysiologic PET & fMRI study used the visual cue-visual target paradigm to investigate effects of spatial and temporal attention on visual stimulus processing [3.4], these author found that a partial overlap between neural systems involved in the spatial and temporal attention. Specifically, the hemispheric asymmetries revealed preferential right for spatial attention and left parietal activation for temporal attention respectively, and the parietal cortex was activated bilaterally by attending to both dimensions simultaneously. Moreover, using the same paradigm, an ERPs study found that spatial attention affected the amplitude of early visual components. While modulation by temporal attention started later, and mainly affected late stages of processing related to decisions and responses [3.5].

With the visual cue-auditory target paradigm, Smith et al. reported that the dorsal frontoparietal network involved in modulation of auditory spatial attention [3.6], and a recent fMRI studies suggested that the specific activations related to temporal attention were confirmed within the superior occipital gyrus, tegmentum, motor area, thalamus and putamen [3.7]. Previous ERPs study shown that the spatial attention could modulate amplitude of auditory component [3.8]. The temporal attention could affect detection [3.9] and early perceptual processing of auditory stimulus [3.10]. Furthermore, several studies indicated that responses to bimodal audiovisual stimuli (AV) are more rapid and accurate than are responses to either unimodal visual (V) or unimodal auditory stimuli (A) [3.11, 3.12]. This facilitative effect is called "audiovisual integration". Numerous research have investigated the effect of visually cued spatial and temporal attention on unimodal visual or unimodal auditory stimulus. However, how the visually cued spatial and temporal attention modulates the audiovisual stimuli processing is still unclear.

The aim of our study was to investigate how visually cued spatial and temporal attention modulate ERPs corresponding to audiovisual stimulus processing. Visual spatial cue predicted the spatial location (left, right) of visual stimuli and visual temporal cue predicted the short or long time interval for cue-target interval. We utilizing the high temporal resolution of ERPs to observe that stages of audiovisual stimuli processing are effected by visually cued spatial or temporal attention.

3.2 Methods

3.2.1 Subjects

Eleven healthy students (age range: 21-25 years; mean age: 23.2 years; 7 male) from Okayama University participated in the experiment. All subjects were right-handed, and had normal or corrected-to-normal vision and normal hearing ability. In addition, they had no history of neurological or psychiatric disorders. The individuals provided written informed consent for participation in this study, and the experimental protocol was approved by the Ethics Committee of Okayama University

3.2.2 Stimuli and procedure

The experiment was conducted in a dimly lit, sound-attenuated, electrically shielded room. Stimuli streams were randomly presented in black background. The participants were seated 60 cm from the center of the monitor. At the beginning of each trial, the normal stimuli were presented in the center of the monitor, which was comprised of two peripheral left and right boxes (2°×2°, centers 7° from the center of the monitor) and fixation stimuli (2°×2°) (see Fig 3.1). After 2200 ms or 3800 ms, the cue was presented for 100ms. The spatial cue predicted the spatial location (left, right) of the visual stimuli, gave no time information for cue-target interval, and was consisted of brightening of either the left or right border of the diamond to inform the subject to attend to the presented location of the stimuli (left or right) (Fig 3.1a). The temporal cue predicted the cue-target interval, but not provide information for location, and was a compound stimulus consisting of a diamond and two concentric circles highlighted to inform the subject to attend to the time of the cue-target interval (600 or 1800 ms from cue presentation) (Fig 3.1b), If either the inner or outer border brightened, the interstimulus interval (ISI, the time interval between visual cue offset and audiovisual stimuli onset) was 600ms or 1800ms, respectively. Before the AV stimuli that appeared, the inter stimulus interval (ISI) lasted for 600ms or 1800ms. After that, the AV stimuli will be presented for 50ms. There were 2 kinds of AV stimuli: no-go and go stimuli. The no-go stimuli refer that the visual stimulus (×, 2°×2°) and the auditory stimulus (single pure tone, 1600 Hz, played at an intensity of 65dB) were presented at the same location (left or right) and they are comprised 77% of AV stimuli, to which no response is required. While, the go-stimuli refer that the visual stimulus and the auditory stimulus were presented at different locations and they are comprised 23% of AV stimuli, to which the response is required through a reaction key as accurately as possible. For each spatial cue trial, the ISI was set to either 600 ms or 1800 ms randomly; for the temporal cue trials, stimuli appeared in the left or right at random.

The cue type (spatial or temporal) was blocked to evoke spatial or temporal attention independently, but the sequence was counterbalanced across participants. Participants were required to fix their eyes on a centrally presented fixation point on a screen and were allowed to take a 5-min break between blocks. There were 5 blocks per cue type. In each block, there were 40 no-go trials and 12 go trials. The total experiment time was approximately 2 hours.

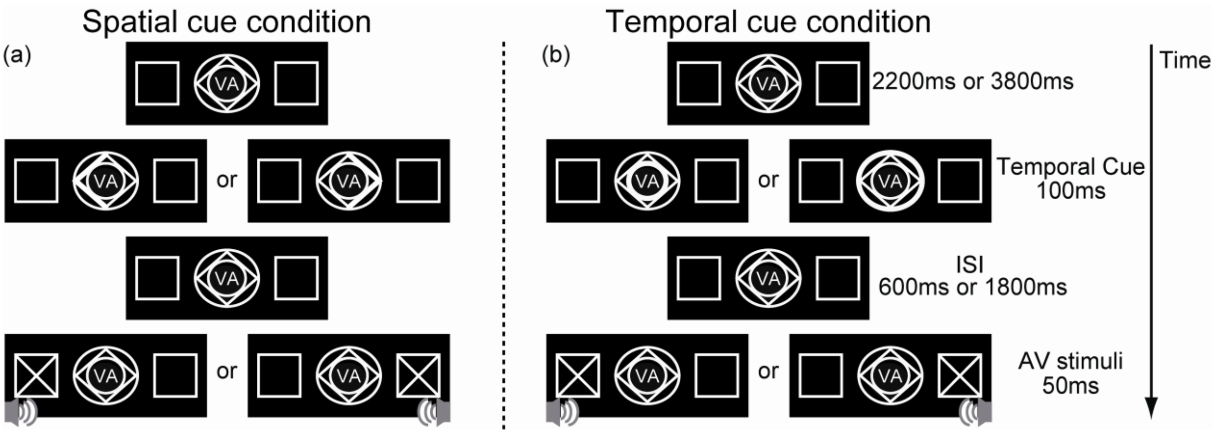

Figure 3.1 An example showing the sequence of trial events. (a) the visual spatial cue indicates spatial information but provides no information about the cue–target interval, and (b) the visual temporal cue: the light of inner circle indicates the target came within a short cue–target interval and the light of outer circle indicates the target came within a long cue–target interval.

3.2.3 Apparatus and software

Stimulus presentation was controlled by a personal computer running Presentation software (Neurobehavioral Systems, Albany, CA). An EEG system (BrainAmp MR plus, Gilching, Germany) was used to record EEG signals through 32 electrodes mounted on an electrode cap (Easy cap, Herrsching Breitbrunn, Germany), as specified by the International 10-20 System. All signals were referenced to the bilateral earlobe electrodes. Horizontal eye movements were recorded from the outer canthus of the right eye; eye blinks and vertical eye movements were recorded from the vEOG electrode. The impedance of all of the electrodes was kept below 5 kΩ. Raw signals were digitized with a sample frequency of 500 Hz with a 60 Hz notch filter and stored continuously on a compatible computer for offline analysis. The event-related potential (ERP) analysis was carried out using Brain Vision Analyzer software (version 1.05, Brain Products GmbH, Munich, Bavaria, Germany).

3.2.4 Data analysis

Behavioral analysis

Reaction times (RTs) for the correct detection of go stimuli, hit rates (HRs), and the false alarm rates (FARs) for incorrect responses to no-go stimuli from each participant were analyzed separately for each stimulus type. HRs were the number of correct responses to go stimuli divided by the total number of go stimuli. FARs were the number of incorrect responses to no-go stimuli divided by the total number of no-go stimuli. RTs, HRs and FARs different for two cue type were subjected to a repeated-measures analysis of variance (ANOVA) using the factors of spatial cue and temporal cue conditions and a significance level of 0.05.

Event-related potential analysis

Only the ERPs elicited by the AV no-go stimuli were analyzed to remove the response. The EEG and EOG signals were amplified and band-pass filtered with an analog filter of 0.01-100 Hz at a sampling rate of 500 Hz. Continuous EEG and EOG signals were divided into epochs from 100 ms before stimulus onset to 500 ms after stimulus onset. Baseline corrections were made against -100 ms - 0 ms. Responses associated with false alarms were also rejected from the analysis. The data were then averaged for each stimulus type, following digital filtering with a band-pass filter of 0.01-30 Hz. The grand-averaged data were obtained across all participants for each stimulus type. Because no significant lateralization effect of AV stimuli has been found, ERP data from the left and right hemisphere were combined to improve the signal-to-noise ratio of the ERPs [3.13]. Based on analysis results, the amplitudes of the N2 components were measured as the mean voltages within the intervals 300–340 ms post stimulus. The mean amplitudes were using a repeated-measures ANOVA with factors of cue type (spatial vs. temporal) and electrodes (CP2, CP6, Pz, P4, P8, O2, T8) separately. We automatically rejected trials with vertical eye movements and eye blinks (vertical EOG amplitudes exceeding $\pm$100 μV), horizontal eyeball movements (horizontal EOG amplitudes exceeding ±25 μV), or other artifacts (a voltage exceeding ±75 μV at any electrode location relative to baseline). The Greenhouse-Geisser epsilon correction was used for non-sphericity when appropriate. Statistical significance was set at 0.05 level. All statistical analyses were carried out using SPSS, version 16.0 (SPSS, Tokyo, Japan).

3.3 Results

3.3.1 Behavioral results

Table 3.1 Mean reaction times (RTs), hit rates (HRs) and false alarm rates (FARs) in each cue condition

Cue type	Spatial	Temporal
RTs (ms)	571 (38)	683 (42)
HRs (%)	95.6 (0.9)	99.1 (0.6)
FARs (%)	9.5 (2.1)	16.8 (3.2)
Standard deviations are given in parentheses.		

The RTs, HRs, and FARs for each cue type condition are shown in Table 3.1. The significant difference between the spatial and temporal cue conditions was found in RTs and FARs. The RTs in response to go AV stimuli in the spatial cue condition were significantly faster [$F(1,10) = 19.033$, $p < 0.001$] than those to go AV stimuli in the temporal cue condition. Furthermore, The false alarm rate for spatial cue condition (9.5%) was significantly lower [$F(1,10) = 7.574$, $p < 0.05$] than that for temporal cue condition (16.8%). However, there were no differences in HRs [$F(1,10) = 1.267$, $p = 0.287$] associated with the spatial and temporal cue conditions. All subjects responded to the task with accuracy above 95%.

3.3.2 Event-related potential results

As shown in Fig 3.2, the mean amplitude of N2 (300-340 ms) at the CP2, CP6, Pz, P4, P8, O2 and T8 electrodes was analyzed by ANOVA. The main effect of the cue type was significant [$F(1, 10) = 11.23$, $p < 0.01$], the result showed that the amplitude of N2 in spatial cue condition was significant strong than those in temporal cue condition (Fig 3.2). In addition, the main effect of the electrode was significant [$F(6, 60) = 2.98$, $p < 0.05$] and the interaction between the electrode and the cue type was no significant [$F(6, 60) = 1.21$, $p > 0.05$]. Post-hoc analysis showed significant differences in amplitude at all of the calculated electrodes ($p < 0.05$) and the maximum difference was found at P4 (-1.108 μV, $p < 0.001$). The scalp distribution of audiovisual N2 component (300-340 ms) for spatial cue, temporal cue and the difference between two cue types were observed in Fig 3.3.

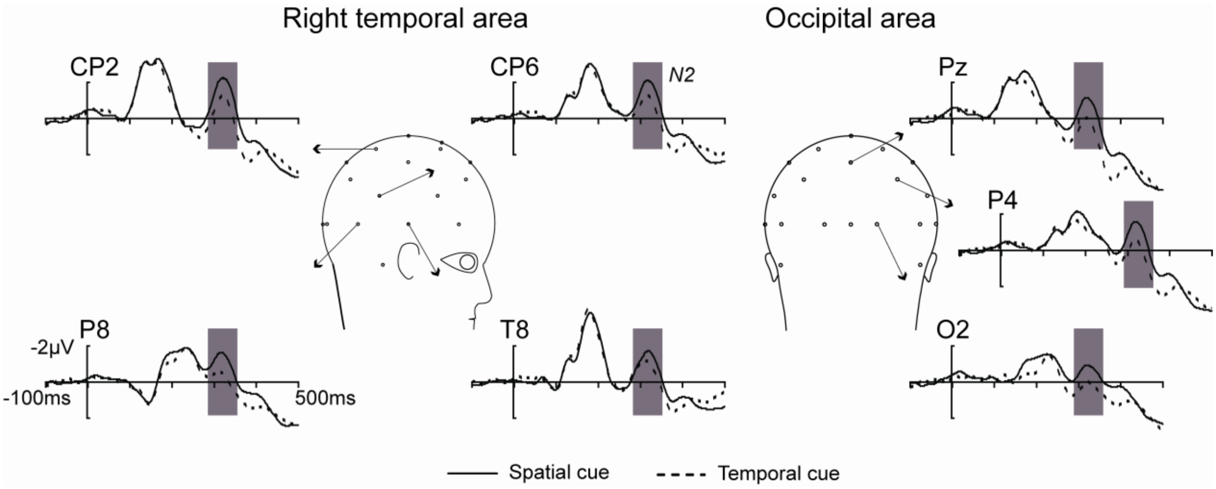

Figure 3.2 The overlap of grand average ERPs elicited by no-go AV stimuli from the right temporal and occipital electrodes in spatial cue (black solid line) and temporal cue (black dotted line) conditions.

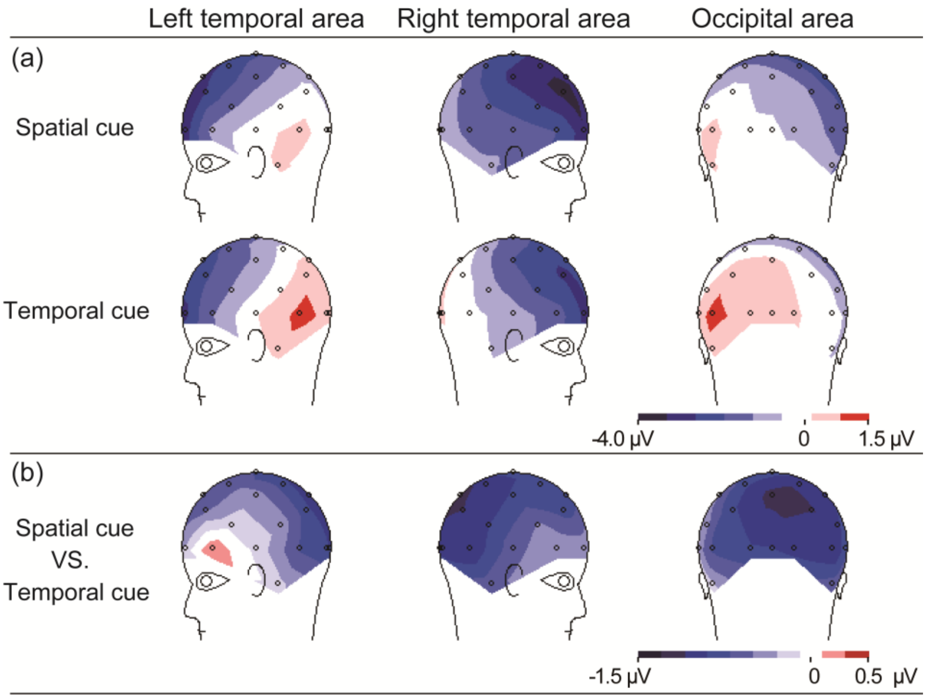

Figure 3.3 The scalp topographies of N2 activity (300-340 ms) elicited by AV stimuli at left and right temporal, occipital area related to (a) the effect of spatial or temporal cue, and (b) effect in spatial cue minus temporal cue.

3.4 Discussion

Behavioral data

The present study aimed to investigate differential modulation of audiovisual stimuli processing by spatial and temporal attention. We found that behavioral responses to AV stimuli in spatial attention condition were faster than those in temporal attention condition, which is consistent with those of previous studies in which responses to auditory stimulus [3.14]. Previous behavioral studies have shown that behavioral responses to bimodal audiovisual (AV) stimuli are improved than auditory stimuli presented alone [3.15, 3.16]. Our findings suggest that visual and auditory stimuli interact to facilitate the behavioral response in spatial attention are better than that in temporal attention.

ERPs data

We observed that the stronger N2 amplitude at right occipitotemporal area is associated with low false alarm rates in the spatial attention condition compared with the temporal attention condition (table 3.1 and Fig 3.3b), which indicate that the greater inhibit responses were induced by spatial attention. Previous studies found that the N2 enhancement has been interpreted as reflecting response inhibition [3.17, 3.18], in which task require subjects to respond to go stimulus and inhibit responses to the no-go stimulus with enhanced N2 potentials being found to no-go stimuli.

Several neurobiological studies have investigated the effect of spatial attention [3.19] or temporal attention [3.20], or differential effect of spatial and temporal attention [3.4, 3.5] on visual stimulus processing. ERPs study found that spatial attention enhanced the amplitude of the late N2 component which was larger when attention was directed to the location of the stimulus position, and the dN2 (difference N2 wave between that under attended condition and under unattended condition) was always larger in the right hemisphere than in the left regardless of the stimulus position, showed the effect of spatial attention to be maximal at the right occipitotemporal area [3.19]. Moreover, ERPs study of temporal attention observed that the N2 component was enhanced only when subjects attended the short interval [3.5], which demonstrate the N2 component can be modulated by temporal attention. By comparing to the both studies, we found that the modulation of spatial attention on stimulus processing at right occipitotemporal area was comprehensive in which spatial attention enhanced the N2 activity regardless of the cue direction(left or right), whereas N2 enhancement only occurred in a short interval during temporal attention. Thus, we believe the visual N2 component in amplitude during spatial attention was more enhanced than that during temporal attention. However, the difference of auditory N2 component in amplitude between spatial attention and temporal attention was not found [3.14].

Furthermore, the visual and auditory information simultaneously received from the different modalities must be integrated by several systems to produce coherent cognition in the brain [3.21, 3.22]. Talsma and Woldorff reported that spatial attention could modulate AV integration processes [3.13, 3.22, 3.23], is that audiovisual integration under attended condition was greater than unattended condition [3.23]. In addition, previous studies proposed that the right occipitotemporal cortex play a important role on audiovisual interaction, which was the integration of audiovisual information into coherent boundaries [3.24-3.27]. Therefore, the present results may be interpreted in the differential modulation of audiovisual N2 component by visually cued spatial and temporal attention originates from the typical cortex of audiovisual integration in brain.

3.5 Conclusion

This study provides behavioral evidence that detection of an audiovisual stimuli in spatial attention is faster than those in temporal attention. Additionally, the ERP results showed that a enhanced N2 amplitude at right occipitotemporal area is associated with low false alarm rates in the spatial attention condition compared with the temporal attention condition. Our ERP results suggest that the differential modulation of audiovisual N2 component by visually cued spatial and temporal attention was generated by a audiovisual integration region of the brain.

References

[3.1] Posner MI. Orienting of attention. Quarterly journal of experimental psychology 1980; 32(1): 3-25.

[3.2] Miniussi C, Wilding E, Coull J, et al. Orienting attention in time modulation of brain potentials. Brain 1999; 122(8): 1507-1518.

[3.3] Griffin IC, Miniussi C, Nobre AC. Orienting attention in time. Frontiers in Bioscience 2001; 6: 660-671.

[3.4] Coull JT, Nobre AC. Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. The Journal of Neuroscience 1998; 18(18): 7426-7435.

[3.5] Griffin IC, Miniussi C, Nobre AC. Multiple mechanisms of selective attention: differential modulation of stimulus processing by attention to space or time. Neuropsychologia 2002; 40(13): 2325-2340.

[3.6] Smith DV, Davis B, Niu K, et al. Spatial attention evokes similar activation patterns for visual and auditory stimuli. Journal of cognitive neuroscience 2010; 22(2): 347-361.

[3.7] Li C, Chen K, Han H, et al. An fMRI Study of the Neural Systems Involved in Visually Cued Auditory Top-Down Spatial and Temporal Attention. PLoS One 2012; 7(11): e49948.

[3.8] Choi I, Rajaram S, Varghese LA, et al. Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. Frontiers in human neuroscience 2013; 7.

[3.9] Rimmele J, Jolsvai H, Sussman E. Auditory target detection is affected by implicit temporal and spatial expectations. Journal of cognitive neuroscience 2011; 23(5): 1136-47.

[3.10] Sanders LD, Astheimer LB. Temporally selective attention modulates early perceptual processing: Event-related potential evidence. Perception & psychophysics 2008; 70(4): 732-742.

[3.11] Wu J, Li Q, Bai O, et al. Multisensory interactions elicited by audiovisual stimuli presented peripherally in a visual attention task: a behavioral and event-related potential study in humans. Journal of Clinical Neurophysiology 2009; 26(6): 407.

[3.12] Li Q, Wu J, Touge T. Audiovisual interaction enhances auditory detection in late stage: an event-related potential study. Neuroreport 2010; 21(3): 173-178.

[3.13] Talsma D, Woldorff MG. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. Journal of cognitive neuroscience 2005; 17(7): 1098-1114.

[3.14] Tang X, Li C, Li Q, et al. Modulation of Auditory Stimulus Processing by Visual Spatial or

Temporal Cue: An Event-related Potentials Study. Neuroscience letters 2013; in press.

[3.15] Odgaard EC, Arieh Y, Marks LE. Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. Cognitive, Affective, & Behavioral Neuroscience 2004; 4(2): 127-132.

[3.16] Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. Cognitive Brain Research 2003; 17(2): 447-453.

[3.17] Jodo E, Kayama Y. Relation of a negative ERP component to response inhibition in a Go/No-go task. Electroencephalography and clinical neurophysiology 1992; 82(6): 477-482.

[3.18] Kok A. Effects of degradation of visual stimuli on components of the event-related potential (ERP) in go/nogo reaction tasks. Biological psychology 1986; 23(1): 21-38.

[3.19] Wang J, Jin Y, Xiao F, et al. Attention-sensitive visual event-related potentials elicited by kinetic forms. Clinical Neurophysiology 1999; 110(2): 329-341.

[3.20] Coull JT, Frith C, Büchel C, et al. Orienting attention in time: behavioural and neuroanatomical distinction between exogenous and endogenous shifts. Neuropsychologia 2000; 38(6): 808-819.

[3.21] Spence C. Audiovisual multisensory integration. Acoustical science and technology 2007; 28(2): 61-70.

[3.22] Talsma D, Doty TJ, Woldorff MG. Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? Cerebral Cortex 2007; 17(3): 679-690.

[3.23] Talsma D, Senkowski D, Soto-Faraco S, et al. The multifaceted interplay between attention and multisensory integration. Trends in cognitive sciences 2010; 14(9): 400-410.

[3.24] Giard M, Peronnet F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. Journal of cognitive neuroscience 1999; 11(5): 473-490.

[3.25] Fort A, Delpuech C, Pernier J, et al. Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. Cerebral Cortex 2002; 12(10): 1031-1039.

[3.26] Yang W, Li Q, Ochi T, et al. Effects of Auditory Stimuli in the Horizontal Plane on Audiovisual Integration: An Event-Related Potential Study. PLoS One 2013; 8(6): e66402.

[3.27] Paulesu E, Harrison J, Baron-Cohen S, et al. The physiology of coloured hearing A PET activation study of colour-word synaesthesia. Brain 1995; 118(3): 661-676.

Chapter 4

Ipsilateral visual and auditory spatial information interaction mechanisms in selective attention conditions

Summary

We used event-related potentials (ERPs) to evaluate the neural mechanism of which auditory spatial information affect audiovisual integration in visual attention task. Some previous studies showed that the ERPs components elicited by audiovisual integration existed in the nonspecific cortices as well as specific visual and auditory cortex. It was widely considered that there weren't significant difference in ERPs of multi-sensory interaction on the left and right. In this study, we compared the integration which elicited by left and right stimuli. The significant differences were found between left and right, which were supported by the following results: (1)There are significant difference between left and right visual and auditory stimuli in early integration. There was integration effect in frontal area around 100-120ms when visual and auditory stimuli appeared in the right. But there was none in the left. (2)There are also significant differences between left and right visual and auditory stimuli in late integration. It showed difference in left and right audiovisual integration in frontal area and centro-medial. Furthermore, it indicated that the processing of multi-sensory information was different between left and right brain.

Keywords: Spatial information, Audiovisual interaction, Event-related potential

4.1 Background

When orienting attention to a certain location in space, the responses to visual stimuli appearing at the location are faster and more accurate than to that at other locations [4.1-4.3]. The effects may affected by at least two different attention mechanisms. One is a voluntary mechanism that is activated by expectancy about which location a relevant visual stimulus will appear in, the other one is an involuntary mechanism that is activated by suddenly occurring stimuli anywhere in the visual field. Furthermore, the same behavioral results have been found in auditory [4.4, 4.5] and somatosensory [4.6] stimuli, suggesting that spatial attention mechanisms might be shared among the various spatial senses.

Lots of previous ERP studies have solved the problem of modality specificity of voluntary attention orienting [4.7, 4.8]. Such studies indicated that responses to visual events appearing in a particular location in space affects the processing of auditory events, and vice versa. The effects demonstrated that stimuli still exist even on the irrelevant-modality, though smaller than the effects of attention on the relevant-modality stimuli. These effects included modulations of early components arising from modality-specific cortex. The results indicated that the mechanisms of voluntary attention are not complete modality specific.

Previous studies have focused on the behavioral responses to auditory (A) and visual (V) stimuli when accompanied by a task-irrelevant stimulus in a different modality. It is indicated that bimodal audiovisual (AV) stimuli are considered faster and more accurately than either visual or auditory unimodal stimuli presented alone [4.9-4.14]. However, these researches never come to an agreement. Some suggested that this effect is regulated by early-stage sensory processing [4.11-4.13]. Meanwhile, the studies suggested that the facilitation effect resulted from the interaction of visual and auditory information at the early stage of sensory processing was a bottom–up effect. On the contrary, other studies presented that the improvement in detecting multimodal stimuli was mediated by late-stage cognitive processing and was a top–down effect [4.14, 4.15]. Now, many studies have been conducted to identify brain activities related to this phenomenon by means of behavioral or neuroimaging methods.

Previous studies indicated that the audiovisual integration exists on the processing of visual and auditory information. Many researches were conducted on unilateral integration test. This paper compared the left and right audiovisual integration respectively, putting forward a new perspective of separately comparison of left and right.

4.2 Methods

4.2.1 Subjects

Fourteen healthy students from Okayama University (seven females) participated in the experiment. All subjects pos-sessed normal or corrected-to-normal vision and normal hearing capability. The experimental protocol was approved by the ethics committee of Kagawa University. After receiving a full explanation of the purpose and risks of the research, subjects provided written informed consent for all studies as per the protocol approved by the institutional research review board. The whole experiment was conducted in soundproof and dark circumstance.

4.2.2 Stimuli

The experiment comprised of three stimuli, that is visual stimulus, auditory stimulus and audiovisual stimulus. Each stimulus is divided into two type, standard one and target one. So, in the whole experiment, the stimuli are as follows: visual standard stimulus, visual target stimulus, auditory standard stimulus, auditory target stimulus, audiovisual standard stimulus and audiovisual target stimulus. The target stimulus of experiment is visual stimulus and the stimuli that participants are requested to response to are visual target stimulus and audiovisual target stimulus. In the central of display, a cross with visual angle of 5 degree appeared as fixation. During the whole experiment, the participants was asked to neglect any sound stimuli.

Visual standard stimuli used a picture of white and black squares with sides of 5.2 cm (visual angle is 5 degree), which appeared on the low left or low right. Visual target stimulus is adding a grey square in the black area of standard stimuli. The cue method is the same as visual standard stimuli and the cue time is 40ms.

Auditory standard stimulus is a 1200Hz simple tone. Cue time is 25 ms and the stimulus was cued in left ear or right ear by earphone, respectively. Audiovisual standard stimulus is visual standard stimulus and auditory standard stimulus appearing simultaneously on the left or right. Cue time of visual standard stimulus is 40ms and that of auditory standard stimulus is 25ms. Audiovisual target stimulus is visual target stimulus and auditory target stimulus appearing simultaneously on the left or right. Cue time of visual target stimulus is the same as that of audiovisual standard stimulus.

4.2.3 Task and procedure

An experiment procedure is as seen in Figure 4.1.

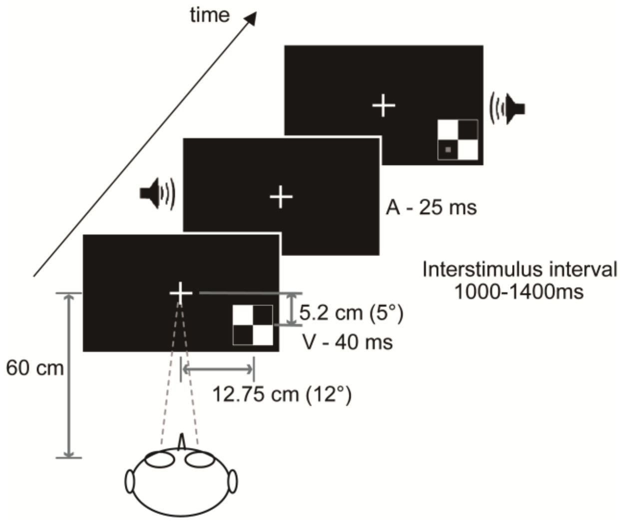

Figure 4.1 Method of stimulus presentation. “V”, “A” and “AV” represent visual stimulus, auditory stimulus, audiovisual stimulus.

The test began with a cross fixation appeared in the central of the display. The central fixation always existed and participants should watch it during the whole experiment. One of the five stimuli appeared randomly. The time interval between two trials which is about 1000 to 1400 ms was set randomly. In the experiment, participants were asked to neglect the auditory stimuli and detect the visual stimuli. They should reflect to target stimuli as quickly as possible. The participants were instructed to press the left button of a computer mouse when the target stimuli appeared on the lower left, and press the right button of a computer mouse when the target stimuli appeared on the lower right, with their right hand. In the whole experiment, stimuli with response are visual target stimulus, audiovisual target stimulus. The whole experiment lasted for 40 minutes and was divided into six sessions.

4.2.4 Apparatus

According to the 10/20 International Standard System, an EEG system (BrainAmp MR Plus, Germany) was applied to record EEG signals via 32 electrodes mounted on an electrode cap (Easy Cap, Germany). All signals were referenced to the bilateral earlobe electrodes. Horizontal eye movements were recorded from the outer canthus of the right eye; eye blinks and vertical eye movements were recorded from the vEOG electrode. Electrode impedance was maintained below 5kΩ except for the electrooculogram (EOG) electrode, which was maintained below 13kΩ. Raw signals were digitized with a sample frequency of 500 Hz with a 60-Hz notch filter; all data were stored digitally for off-line analysis.

4.2.5 Data analysis

Analysis of behavioral results

Reaction times (RTs) for the correct detection of target stimuli and accuracy were computed separately for each stimulus type and each hemispace. RTs and accuracy for different stimulus types were subjected to repeated-measures analysis of variance (ANOVA) using the subject factors of stimulus modality (visual, audiovisual) and stimulus location (left hemispace, right hemispace).

ERP results off-line analysis

Only the ERPs elicited by standard stimuli were analyzed. Continuous EEG signals were divided into epochs from 100 ms before stimulus onset to 500 ms after stimulus onset. After baseline corrections were performed using the period between –100 to 0 ms as the baseline, an artifact criterion over ± 80μV was used at all channels to reject trials with noise transients; responses associated with false alarms or omissions were also rejected from the analysis. The data were then averaged for each stimulus type following digital filtering using a band-pass filter of .01–30 Hz. Final grand-averaged data were obtained across all subjects for each stimulus type.

4.3 Results

4.3.1 Behavioral results

The results can be seen in Figure 4.2. The results analyzed with analysis of variance (ANOVA) showed that the response time to audiovisual target stimuli was much shorter than that to visual target stimuli [$F(1,13)=19.437$, $P<0.001$], which can be explained as the result of Redundant Signal Effect (RSE). RSE means that response to bi-channel sensory information (such as audiovisual information) can process quicker comparing to response time of single channel sensory information. Furthermore, there wasn't a significant main effect of stimuli position and there wasn't a significant interaction effect between stimuli type and position. Datum of response time to target stimuli less than 100ms or more than 1000ms and wrong response datum were excluded from the analyses.

In addition, ANOVA of accuracy rate shows that there wasn't a significant main effect of stimuli type and stimuli position and there wasn't a significant interaction effect between stimuli type and position (Figure 4.3).

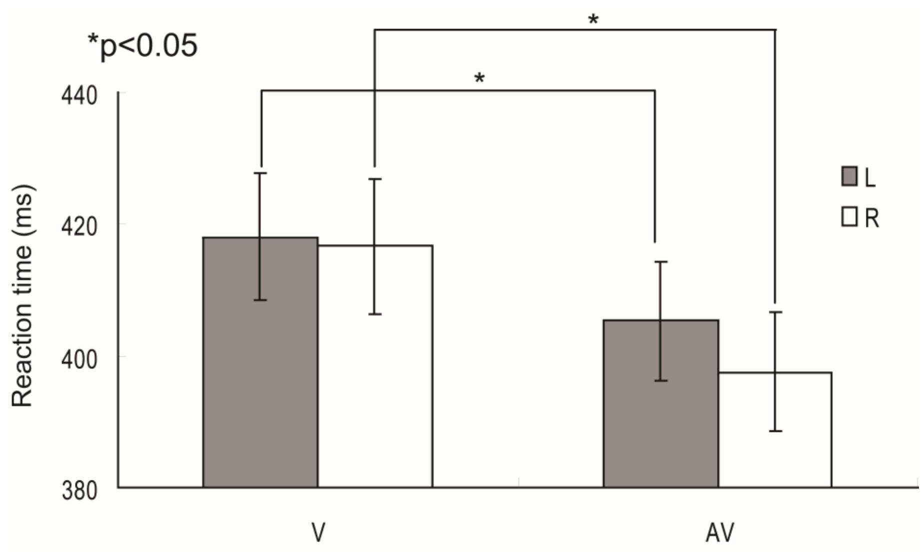

Figure 4.2 Averaged reaction times of spatial auditory task

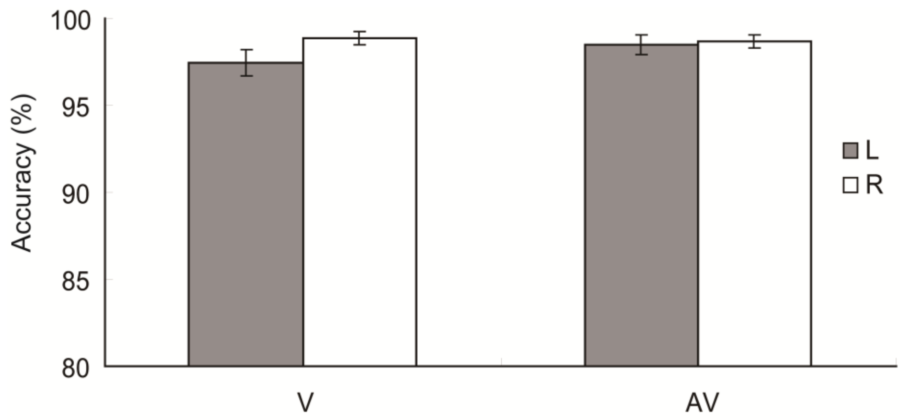

Figure 4.3 The correct response of spatial auditory task

4.3.2 ERP results

The results of audiovisual integration are illustrated in Figure 4.4 (early integration) and Figure 4.5 (late integration), which were shown as follows: (1) at early stage, left audiovisual integration appeared 80-100ms after the stimuli, located in the frontal and central areas .Right audiovisual integration also appeared 80-100ms after the stimuli. But in the period of 100-120ms, the right integration, not left one appeared in frontal area. In the period of 140-180ms, left integration was strong in frontal area, weak in the central area. The similar effect appeared on the right and lasted to

200ms. (2) at later stage, 260-320ms, when stimuli appeared on the left, stronger integration effect was found in the front area and weaker one was in the top head. But when stimuli appeared on the right, the apparent integration existed both in the front lobe and the top head. The results indicated that although there are some similar characteristics, the mechanism were different between the left and right stimuli.

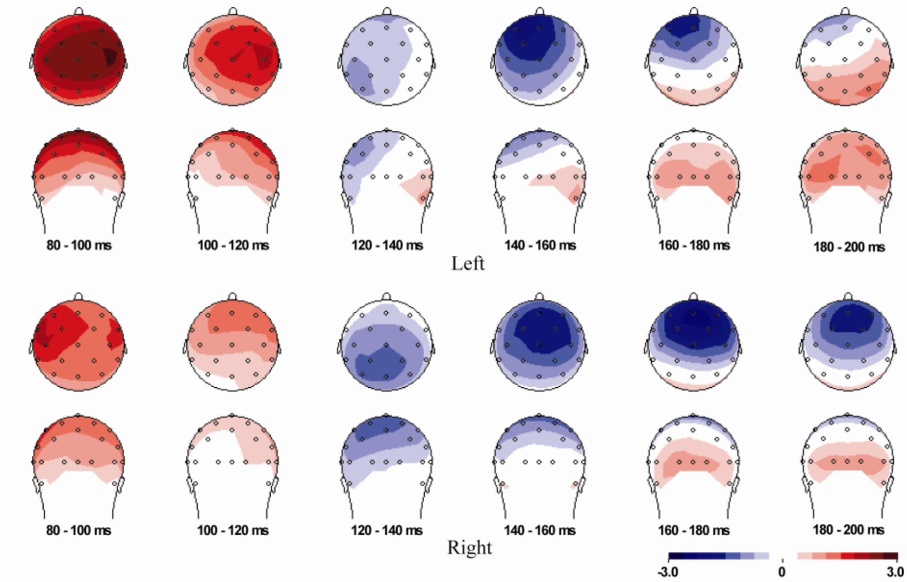

Figure 4.4 Scalp topographies of [AV-(A+V)] related to the audiovisual interaction elicited at early

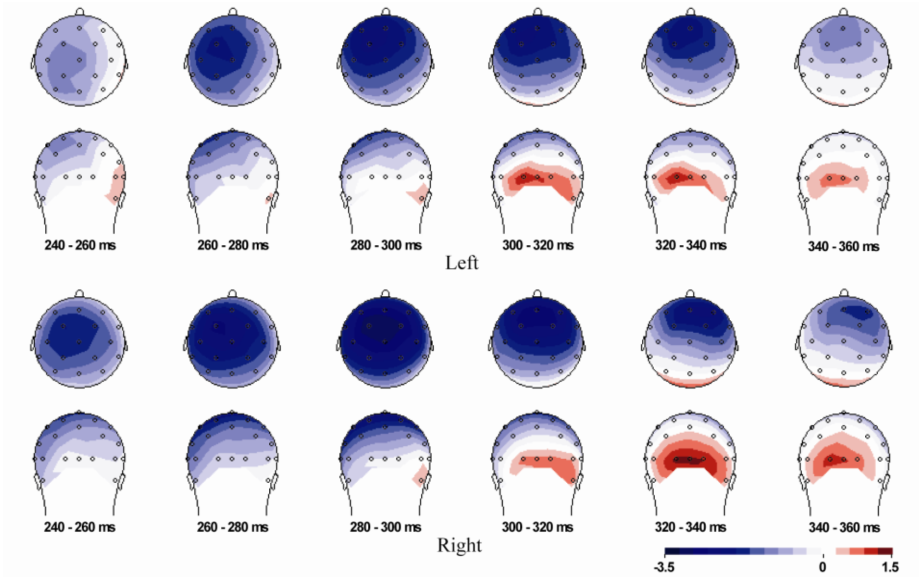

Figure 4.5 Scalp topographies of [AV-(A+V)] related to the audiovisual interaction elicited at later.

4.4 Discussion

In this study, we designed a visual attention task, in which audiovisual stimuli presented on the left and right hemispace, and researched the multimodal integration. Behavioral results revealed that it were faster and more accurate of the responses to multimodal stimuli than those to unimodal visual or auditory stimuli[4.16-4.19]. The results of this study on the audiovisual integration were consistent with the results of previous studies.

In ERP results, we observed the significant audiovisual integration effect elicite by the left stimuli at frontal-central areas around 80-100ms. The result was similar to previous studies[4.20]. The latency of this integration effects corresponded to that auditory N1 wave in unimodal responses. According to the brain topography, the interaction effects could decreased amplitude of the auditory N1 wave, which indicated weaker activation of its generators in the auditory cortex for bimodal stimuli. Furthermore, it was observed that another audiovisual integration elicite by the right stimuli effects at frontal-central areas around 280-300 ms. The latency of the results were similar to the N260 in previous studies [4.13, 4.21]. It was considered that N260 was generally localized to auditory-specific temporal lobe. The possible neural path-ways were a direct, feed-forward projection from the primary visual cortex or the visual association cortex to the auditory cortex; or an indirect feed-forward projection in which input from the visual cortex projects first to the superior temporal poly sensory region where the AV multisensory information converges and then to the auditory cortex [4.12].

Some previous studies limited their analysis time window to within the first 200ms after stimulus presentation. Few analyzed the integration condition after 200ms following presentation of the stimulus. The integration effect results suggested that the activities of visual cortex were increased for bimodal stimuli after 200ms [4.20]. Responses occurring during the first 200ms after presentation of the stimulus were thought to be mediated by early sensory processing, whereas activity after 200ms was thought to be mediated by late cognitive processing [4.22]. Some studies have suggested that this effect is mediated by early-stage sensory processing and was a bottom–up effect [4.16, 4.23]. Yet, other studies suggest that the improvement in detecting multimodal stimuli was mediated by late-stage cognitive processing and was a top–down effect [4.24, 4.25].

It has been indicated in this paper that the effect of integration also existed in unilateral audiovisual information processing. the mechanism of audiovisual integration of human being has been further probed. the latency of the ERP activity was later than that found in the previous studies [4.22]. That may result from effects of modality select attention, which was only auditory stimuli were task-relevant and required to be attended in this study. However, the previous studies required participants to attend both visual and auditory stimuli. Talsma and Woldorff reported that audio-visual interaction occurred somewhat earlier when stimuli were attended to than when stimuli were not attended to. Therefore, we speculated that the difference of modality select attention resulted in the postponement of audiovisual interaction. However, further studies are needed to confirm and elucidate those details.

4.5 Conclusion

The results indicated that left and right audiovisual integration occurring in both early-stage and late-stage in the tasks of visual selective attention. It came to some conclusion as follows: (1) There was integration effect in front area around 100-120ms when visual and auditory stimuli appeared in the right. But there was none in the left. (2) There are also significant difference between left and right visual and auditory stimuli in late integration. It showed difference in left and right audiovisual integration in front area and top head. Furthermore, it indicated that the processing of multi-sensory information was different between left and right brain. Furthermore, the study put forward a new perspective of separately comparison of left and right.

References

[4.1] Eriksen, C.W., & Yeh, Y.Y. (1985). Allocation of attention in the visual field. Journal of Experimental Psychology: Human Perception and Performance, 11, 583–597

[4.2] Hawkins, H.L., Hillyard, S.A., Luck, S.J., Mouloua, M., Downing, C.J., & Woodward, D.P. (1990). Visual attention modulates signal detectability. Journal of Experimen-tal Psychology: Human Perception and Performance, 16, 802–811.

[4.3] Posner, M.I. (1980). Orienting of attention. Quarterly Journal of Experimental Psychology, 32, 3–25.

[4.4] McDonald, J.J., & Ward, L.M. (1999b). Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial orienting. Journal of Experimental Psychology: Human Perception and Performance, 25, 1–19.

[4.5] Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial orienting. Journal of Experimental Psychology: Human Perception and Performance, 22, 1005–1030.

[4.6] Kilgard, M.P., & Merzenich, M.M. (1995). Anticipated stimuli across skin. Nature, 373-663.

[4.7] Eimer, M., & Schro ¨ger, E. (1998). Effects of intermodal attention and cross-modal links in spatial attention. Psychophysiology, 35, 313–327.

[4.8] Teder-Sa ¨leja ¨rvi, W.A., Mu ¨nte, T.F., Sperlich, F.-J., & Hillyard, S.A. (1999). Intra-modal and cross-modal spatial attention to auditory and visual stimuli: An event-related brain potential (ERP) study. Cognitive Brain Research, 8, 327–343

[4.9] Ferreira A, Celeste WC, Cheein FA, Bastos-Filho TF, Sarcinelli-Filho M, Carelli R. Human-machine interfaces based on EMG and EEG applied to robotic systems. J Neuroeng Rehabil., vol. 26, pp. 5-10, 2008

[4.10] Ushiba J. Brain-machine interface--current status and future prospects, Brain Nerve., vol. 62, pp. 101-111, 2010

[4.11] A. Fort, C. Delpuech, J. Pernier, and M. H. Giard, Early auditory-visual interactions in human cortex during nonredundant target identification, Brain Res Cogn Brain Res, vol. 14, pp. 20-30, 2002.

[4.12] S. Molholm, W. Ritter, M. M. Murray, D. C. Javitt, C. E. Schroeder, and J. J. Foxe, Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. Brain Res Cogn Brain Res, vol. 14, pp. 115-128, 2002.

[4.13] W. A. Teder-Salejarvi, J. J. McDonald, F. Di Russo, and S. A.Hillyard, An analysis of audio-visual crossmodal integration by means of eventrelated potential (ERP) recordings. Brain Res Cogn Brain Res, vol. 14, pp. 106-114, 2002.

[4.14] Stein BE, Meredith MA. The Merging of the Senses. Cambridge: MIT Press, 1993a; 123-156.

[4.15] M. H. Giard and F. Peronnet, Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. J Cogn Neurosci, vol. 11, pp. 473-490, 1999.

[4.16] Frassinetti F, Bolognini N, Ladavas E., Enhancement of visual perception by crossmodal visuo-auditory interaction. Exp Brain Res., 147 (2002), 332-343.

[4.17] Giard MH, Peronnet F., Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. J Cogn Neurosci., 11 (1999), 473-490

[4.18] Schroger E, Widmann A., Speeded responses to audiovisual signal changes result from bimodal integration. Psychophysiology, 35 (1998), 755-759.

[4.19] Watkins S, Shams L, Tanaka S, Haynes JD, Rees G., Sound alters activity in human V1 in association with illusory visual perception. Neuroimage, 31 (2006), 1247-1256.

[4.20] Vidal J, Giard MH, Roux S, Barthelemy C, Bruneau N., Cross-modal processing of auditory-visual stimuli in a no-task paradigm: a topographic event-related potential study. Clin Neurophysiol., 119 (2008), 763-771.

[4.21] Teder-Salejarvi W. A., Russo F. D, McDonald J. J., and Hillyard S. A., Effects of spatial congruity on audio-visual multimodal integration, J Cogn Neurosci, 17 (2005), 1396-1409

[4.22] Talsma D, Woldorff MG. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. J Cogn Neurosci, 17 (2005), 1098–1114.

[4.23] Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. Brain Res Cogn Brain Res, 17 (2003), 447–453.

[4.24] Lippert M, Logothetis NK, Kayser C. Improvement of visual contrast detection by a simultaneous sound. Brain Res, (2007) 1173:102–109.

[4.25] Marks LE, Ben-Artzi E, Lakatos S. Cross-modal interactions in auditory and visual discrimination. Int J Psychophysiol (2003) 50:125–145.

Chapter 5

Effects of ipsilateral and bilateral auditory stimuli on audiovisual integration

Summary

We used event-related potential (ERP) measures to compare the effects of ipsilateral and bilateral auditory stimuli on audiovisual (AV) integration. Behavioral results showed that responses to visual stimuli with either type of auditory stimulus were faster than those to visual stimuli only, and an enhancement of perceptual sensitivity (d') for visual detection was found for visual stimuli with ipsilateral auditory stimuli. Furthermore, ERP components related to AV integrations were identified over the occipital areas at approximately 180 to 200 ms during early-stage sensory processing via the effect of ipsilateral auditory stimuli and over the fronto-central areas at approximately 300 to 320 ms during late-stage cognitive processing via the effect of ipsilateral and bilateral auditory stimuli. Our results confirmed that audiovisual integration was also elicited, despite the effect of bilateral auditory stimuli, and only occurred at later stages of cognitive processing in response to a visual detection task. Furthermore, integration from early-stage sensory processing was observed via the effect of ipsilateral auditory stimuli, suggesting that the integration of audiovisual information in the human brain might be particularly sensitive to ipsilaterally presented audiovisual stimuli.

Keywords: ipsilateral auditory stimuli, bilateral auditory stimuli, visual detection, audiovisual integration, event-related potential.

5.1 Background

In the real world, humans constantly receive a large amount of information, including visual and auditory stimuli. Our attention system regulates the choice of task-relevant information and the neglect of irrelevant information [5.1]. However, the irrelevant information is not completely neglected and can affect task completion. It has been shown that detection and responses to unimodal visual (V) stimuli can be improved when accompanied by a task-irrelevant auditory (A) stimulus [5.2, 5.3], a crossmodal facilitating phenomenon known as audiovisual (AV) integration.

Previous studies have suggested that AV integration is dependent on the spatial relationship between visual and auditory stimuli [5.4, 5.5]. On the basis of the visual detection task, several studies have found that the detection of V stimuli was improved when A stimuli were simultaneously presented in the same location. In contrast, when the V and A stimuli were presented in different locations, little or no enhancement was observed [5.6-5.8]. Furthermore, in a study conducted by Teder-Sälejärvi et al. (2005), the auditory and visual stimuli were presented unilaterally; for example, either at an ipsilateral spatial location (i.e., 30° to the left or right of the central fixation point) or at a contralateral spatial location (i.e., V to the left and A to the right of the fixation point, 60° apart). Teder-Sälejärvi and colleagues found different ERP interactions with both AV pairings in the brain. Specifically, an amplitude modulation of activity localized to the superior temporal region was observed for ipsilateral AV pairings, and a phase shift of visual-evoked activity localized to the ventral occipito-temporal cortex was observed for contralateral AV pairings [5.9]. In addition, the enhancement of the audiovisual stimuli was optimal when the visual stimulus was presented with an ipsilateral, spatially congruent auditory stimulus [5.4].

Sound lateralization also had a major influence on the magnitude and extent of the activations in the auditory cortex. For example, unilateral sounds produced significantly larger activations in the hemisphere contralateral to the stimulation [5.10, 5.11]. Moreover, the amplitude of the bilateral responses differed from the amplitude of the responses to unilateral sounds in the auditory cortex [5.12]. Previous studies have investigated behavioral performances and neural responses using ERP recordings during audiovisual integration in response to unilateral AV pairings [5.9, 5.13]. However, the effect of bilateral auditory stimuli on audiovisual integration still remains unclear.

The aims of the present study were to investigate the effect of bilateral auditory stimuli on audiovisual integration and to compare this effect to that of ipsilateral auditory stimuli. We designed a visual detection task and performed behavioral and ERP measurements to determine the neural basis of audiovisual integration elicited by this task. In this task, the audiovisual stimuli consisted of unilateral visual stimuli accompanied with simultaneous ipsilateral or bilateral auditory stimuli.

5.2 Methods

5.2.1 Subjects

Fourteen healthy students (age: 22 to 33 years, mean: 24.2 years, 5 females) from Okayama University participated in this experiment. All of the subjects were right-handed, possessed normal or corrected-to-normal vision and demonstrated normal hearing ability. They had no history of neurological or psychiatric disorders. The experimental protocol was approved by the Ethics Committee of Okayama University.

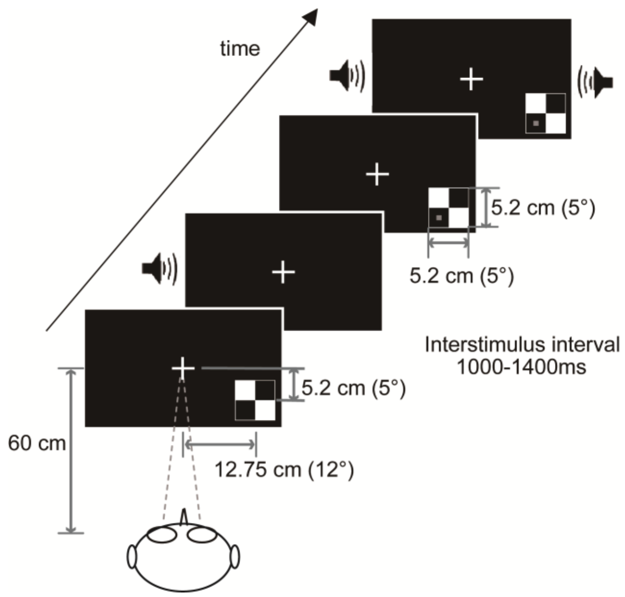

Figure 5.1 Method of stimulus presentation during sequential trials in the experiment.

5.2.2 Stimuli and design

The experiment was performed in a dimly lit, sound-attenuated, electrically shielded room. Presentation software (Neurobehavioral Systems Inc., Albany, California, USA) was used to present the stimuli and collect the responses. Stimulus streams consisting of unimodal V stimuli, unimodal A stimuli and bimodal AV stimuli (simultaneous visual and auditory events) were randomly presented in the experiment. The unimodal V stimulus consisted of a checkerboard image (duration 40 ms, 5.2×5.2 cm, subtending a visual angle of approximately 5°, Fig 5.1). These V stimuli were presented peripherally to lateral locations on either the left or right side of the display at an angle of approximately 12° from a centrally presented fixation point on a black background; moreover, they were positioned 60 cm from the participant's eyes in the lower visual fields (approximately 5° below central fixation). The V target stimuli were checkerboards containing 1 dot. Those with no dots were referred to as "standard stimuli." The A stimulus consisted of a tone pip (25 ms in duration, 1200 Hz, 60 dB SPL, 5 ms rise and fall periods), which was presented through an earphone either unilaterally into one ear (A_{uni}: left or right) or bilaterally into both ears (A_{bi}: left and right). This A stimulus was a task-irrelevant and ignored event. The bimodal AV target stimulus consisted of simultaneous V target stimuli and task-irrelevant A stimuli. The bimodal AV standard stimulus consisted of simultaneous V standard stimuli and task-irrelevant A stimuli, which were presented in two multisensory conditions: visual with ipsilateral auditory stimulus ($A_{ipsi}V$) and visual with bilateral auditory stimulus ($A_{bi}V$). The subjects were required to detect the V or AV target stimuli.

Each subject participated in twelve blocks, of which each block lasted for approximately 4 min. The participants were allowed to take a 2-min rest between blocks. Each block consisted of 40 unimodal V-only stimuli (32 standards, 8 targets), 40 bimodal AipsiV stimuli (32 standards, 8 targets), 40 bimodal AbiV stimuli (32 standards, 8 targets), 32 unimodal Auni stimuli and 32 unimodal Abi stimuli. All of the stimuli were randomly presented. Before the experiment, a display without stimuli appeared on the screen for 4000 ms, followed by the experimental stimulus. The interstimulus interval (ISI) varied randomly between 1000 and 1400 ms. The participants were asked to neglect the auditory stimuli and to detect the visual target stimuli. They were instructed to press the left button of a mouse when the visual target stimulus appeared on the lower left and to press the right button of a mouse when the visual target stimulus appeared on the lower right, using the index or middle finger of their right hand as quickly and accurately as possible.

5.2.3 Apparatus

Consistent with the 10/20 International Standard System, an EEG system (BrainAmp MR Plus, Germany) was used to record the EEG signals of 32 electrodes mounted on an electrode cap (Easy Cap, Germany). All of the signals were referenced to the bilateral ear lobe electrodes. Horizontal eye movements were recorded from the outer canthus of the left eye; eye blinks and vertical eye movements were recorded from the vEOG electrode, which was placed 1.5 cm below the left eye. The impedance of all electrodes was below 5 kΩ. The raw signals were digitized with a sample frequency of 500 Hz with a 60-Hz notch filter; all of the data were stored digitally for off-line analysis.

5.2.5 Data analysis

Analysis of behavioral results

The reaction times (RTs) for the correct detection of target stimuli, hit rates (HRs), and false alarm rates (FARs) for incorrect responses to standard stimuli from each participant were separately computed for each stimulus type. The perceptual sensitivity (d'), which reflected the subject's accuracy in discriminating a sensory event from its background, was computed using the z-transformed ratio [5.14]. All of the results were analyzed using repeated-measures analysis of variance (ANOVA) with the factor of stimulus modality (visual-only, AipsiV and AbiV) and a significance level of 0.05.

Off-line analysis of ERP results

We used the Brain Vision Analyzer software (version 1.05, Brain Products GmbH, Munich, Bavaria, Germany) to analyze the off-line ERPs elicited by standard stimuli. Continuous EEG signals were divided into epochs from -100 ms before stimulus onset to 500 ms after stimulus onset. After baseline corrections were performed using the period between -100 and 0 ms, an artifact criterion of ±80 μV was used in all channels to exclude trials with noise transients; responses associated with false alarms or omissions were also excluded from the analysis. The data were then averaged for each stimulus type after digital filtering with a band-pass filter of 0.01 to 30 Hz. The grand-averaged data were obtained across all participants for each stimulus type. Audiovisual integration was quantified as the difference wave [AV-(A+V)], which was obtained by subtracting the sum of the responses to the unimodal stimuli from that of the bimodal stimuli [5.15]. The mean amplitudes were then averaged for all of the electrodes at consecutive windows of 20 ms each between the stimulus onset and 500 ms after presentation of the stimulus. The mean amplitude data were analyzed using repeated-measures ANOVA with the within-subjects factors of modality (AV, A+V), time-window and electrodes [5.16]. In addition, we did not observe a significant effect between left and right lateralization (all $p>0.05$), and the data from the left and right hemispace were combined to improve the signal-to-noise ratio of the ERPs [5.17]. The Greenhouse-Geisser procedure was applied to all F-tests to correct for potential violations of the sphericity assumption.

5.3 Results

5.3.1 Behavioral results

The RTs, HRs, and FARs for each stimulus condition are shown in Table 5.1. The RTs in response to unimodal visual and bimodal audiovisual stimuli ($A_{ipsi}V$ and $A_{bi}V$) differed significantly [F (2, 26)= 4.880; p= 0.016]. Furthermore, post-hoc comparisons revealed that responses to the $A_{ipsi}V$ (p =0.031) and $A_{bi}V$ (p= 0.032) stimuli were significantly faster than those to the V stimuli. However, no significant differences were observed between the $A_{ipsi}V$ and $A_{bi}V$ stimuli (p= 0.912).

Table 5.1 Mean reaction times (RTs), hit rates (HRs), false alarm rates (FARs) and perceptual sensitivity (d') for each stimulus type.

Stimulus type	V	$A_{ispi}V$	$A_{bi}V$
RTs (ms)	508 (16)	488 (15)	496 (14)
HRs (%)	94.4 (2.1)	94.8 (2.5)	96.8 (1.6)
FARs (%)	0.8 (0.3)	0.6 (0.2)	1.0 (0.3)
d'	4.4 (0.9)	4.7 (0.9)	4.6 (0.8)

V, visual stimulus; $A_{ispi}V$, visual with ipsilateral auditory stimulus; $A_{bi}V$, visual with bilateral auditory stimulus; Standard errors are shown in parentheses.

The HRs showed a main effect of the stimulus type [F (2, 26)= 6.692; p= 0.005], and the maximal hit rate was observed in the $A_{bi}V$ condition (96.8%). The FARs for the stimulus type showed a significant effect [F (2, 26)= 4.658; p= 0.019], and the maximal false alarm rate was also observed in the $A_{bi}V$ condition (1.0%). Taken together, these results indicated that if the hit rate for the $A_{bi}V$ condition was high, then the false alarm rate for this stimulus type was also high.

The hit and false alarm rates were then used to compute the signal detection measure d' (Table 5.1). There was a significant effect of stimulus type [F (2, 26)= 4.150; p= 0.027]. In addition, post-hoc comparisons revealed significant differences between V and $A_{ipsi}V$ (p =0.004). Moreover, no other significant effect of stimulus type on d' was found.

5.3.2 Event-related potential results

The ERPs elicited by task irrelevant unimodal auditory standard stimuli are shown in Figure 5.2. For the auditory stimuli from the left or right or simultaneous bilateral stimuli (left and right), a negative N1 wave peaked at approximately 110 ms post-stimulus over the fronto-central sites (peak: -4.36 μV, -3.69 μV, -3.31 μV at Fz). Importantly, the topographies showed that the amplitude of the auditory N1 component was greater over the contralateral scalp sites in response to the stimulus presentation.

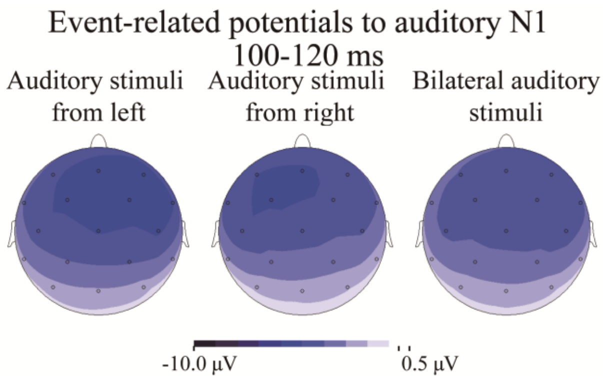

Figure 5.2 Auditory ERP elicited by auditory standard stimuli. Topographic distributions of the N1 (100 –120 ms) sensory components are shown separately for auditory stimuli presented on the left, on the right or bilaterally. The auditory N1 activity showed a clear contralateral effect

As shown in Fig. 3, the effect of audiovisual integration was investigated by comparing the ERPs for the AV stimulus to the sum of the ERPs for the A and V stimuli presented alone. Significant differences in amplitude were observed at 180–200 ms post-stimulus over the parietal-occipital area in response to the $A_{ipsi}V$ stimuli (Figure 5.3a) and at 300–320 ms post-stimulus over the fronto-central area in response to both the $A_{ipsi}V$ (Figure 5.3a) and $A_{bi}V$ stimuli (Figure 5.3b).

ERPs over the occipital areas (180 to 200 ms).

When visual stimuli were presented with ipsilateral auditory stimuli, significant differences in the amplitudes between the AV and (A+V) conditions [F (1, 13)= 10.824, p=0.006] were observed at a latency of 180–200 ms post-stimulus over the occipital electrodes. The difference in amplitude was found at P3, P4, Oz, O1, and O2 (all p<0.05) and approached significance at Pz (p=0.053). The difference in the mean amplitude between AV and A+V was maximal at P4 (1.072 μV) (Figure 5.3a). However, neither a significant effect of electrode type nor significant interactions between electrode and modality were found.

However, within the same time window (180-200 ms), there were no significant differences between AV and (A+V) in the occipital areas when visual stimuli were presented with bilateral auditory stimuli [F (1, 13)=3.633, p=0.079].

As shown in Figure 5.4a, the topography of [AV-(A+V)] revealed the effects of the ipsilateral and bilateral auditory stimuli. Moreover, a reduced visual N1 response was generated by the addition of an auditory stimulus for the bimodal AV stimuli, which was shown in Figure 5.4a (first, fourth and fifth columns).

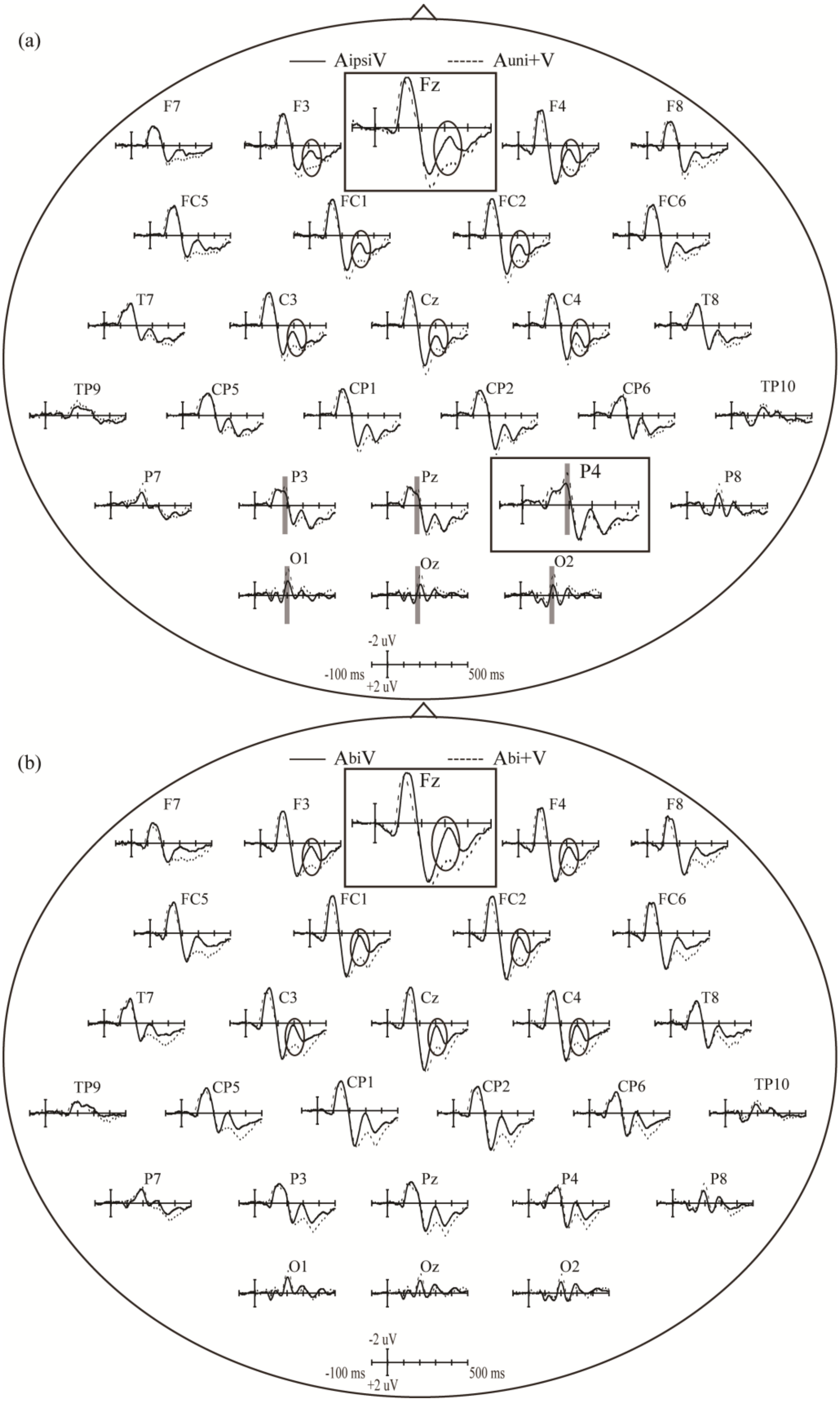
(a)
AipsiV
Auni+V
F7
F3
Fz
F4
F8
FC5
FC1
FC2
FC6
T7
C3
Cz
C4
T8
TP9
CP5
CP1
CP2
CP6
TP10
P7
P3
Pz
P4
P8
O1
Oz
O2
-2 uV
-100 ms
500 ms
+2 uV
(b)
AbiV
Abi+V
F7
F3
Fz
F4
F8
FC5
FC1
FC2
FC6
T7
C3
Cz
C4
T8
TP9
CP5
CP1
CP2
CP6
TP10
P7
P3
Pz
P4
P8
O1
Oz
O2
-2 uV
-100 ms
500 ms
+2 uV

Figure 5.3 Overlap of the grand averaged ERPs elicited by audiovisual (AV) and auditory (A) plus visual (V) stimuli. (a) Auditory stimuli were presented unilaterally. (b) Auditory stimuli were presented bilaterally. The rectangular and oval areas indicate the time periods when the AV response significantly differed from the A plus V response (p<0.05). (A_{uni}, unilateral auditory stimulus from left or right; A_{bi}, bilateral auditory stimulus from left and right; $A_{ipsi}V$, visual with ipsilateral auditory stimulus; $A_{bi}V$, visual with bilateral auditory stimulus).

ERPs over the fronto-central areas (300 to 320 ms).

Analyses of the mean amplitudes revealed main effects of modality [F(1,13)=13.395, P=0.003] and a significant interaction between modality and electrode [F(7,91)=3.819, P=0.020] at 300–320 ms over the fronto-central electrodes (F3, F4, Fz, FC1, FC2, Cz, C3, C4) when visual stimuli were presented with ipsilateral auditory stimuli (Figure 3a). Post-hoc analysis showed significant differences in amplitude at all of the examined electrodes (p<0.05). The difference in amplitude between $A_{ipsi}V$ and $A_{ipsi}+V$ was maximal at Fz (-2.708 μV, p=0.001) (Figure 5.3a).

When visual stimuli were presented with bilateral auditory stimuli, similar effects were found. A main effect of modality with a 300-320 ms latency was found at the fronto-central electrodes [F(1,13)=18.133, P=0.001] (Figure 5.3b); however, neither a significant effect of electrode type nor significant interactions between electrode and modality were found. The difference in the mean amplitudes over F3-F4-Fz-FC1-FC2-Cz-C3-C4 at 300-320 ms were -3.031 μV for AV-(A+V). The difference in amplitude between $A_{bi}V$ and $A_{bi}+V$ was also maximal at Fz (-3.33 μV, p<0.001) (Figure 5.3b).

The scalp distribution over the fronto-central areas is shown in Fig 5.4b.

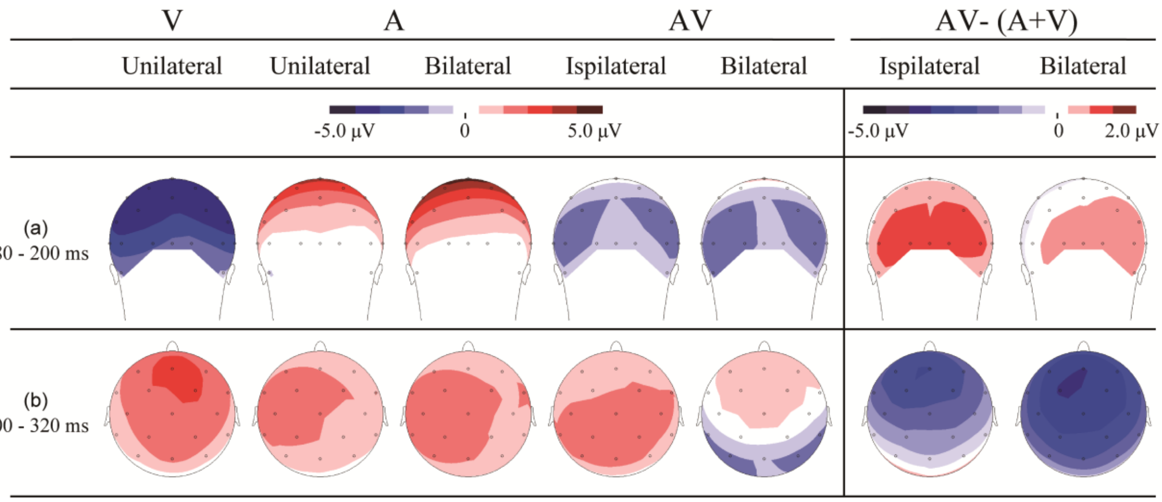

Figure 5.4 The scalp topographies of visual (V), auditory (A), audiovisual (AV) and AV – (A + V) stimuli related to the audiovisual integration effects (a) at 180–200 ms post-stimulus over the occipital area and (b) at 300–320 ms post-stimulus over the fronto-central area.

5.4 Discussion

Behavioral data

In this study, we used a visual detection task to examine the behavioral and neural mechanisms of visual stimuli when simultaneously presented with ipsilateral or bilateral auditory stimuli. The behavioral responses to both $A_{ispi}V$ and $A_{bi}V$ stimuli were faster than those to unimodal V stimuli, which was consistent with the results of previous behavioral performances of AV interaction [5.18]. In addition, a previous behavioral study had shown an enhancement of sensitivity (d') for visual detection when audiovisual stimuli were simultaneously presented at a consistent ipsilateral position [5.2]. Similarly, we observed that the sensitivity (d') to bimodal audiovisual stimuli in which the auditory stimuli were presented ipsilaterally was significantly larger than that to unimodal visual stimuli. In our study, an increased d' might indicate that the integration of ipsilateral audiovisual stimuli was mediated by early-stage sensory processing at the behavioral level.

ERP data

Investigation of the ERP components related to audiovisual integration revealed significant differences between AV and A+V in early-stage processing (before 200 ms) over the occipital area when visual stimuli were presented with ipsilateral auditory stimuli and in late-stage processing (after 200 ms) over the fronto-central areas when visual stimuli were presented with both ipsilateral or bilateral auditory stimuli (Fig 5.4).

Audiovisual integration in the occipital areas (180 to 200 ms).

The initial difference between AV and A+V was observed at 180–200 ms over the occipital areas with ipsilateral auditory stimuli (Fig. 4a), which was consistent with previous studies [5.19]. However, the latency of the ERP activity in the current study was slightly delayed compared to those reported in a previous study [5.15]. This effect may have been due to the location of the presented stimuli. Our study presented both visual and auditory stimuli peripherally, whereas previous studies [5.15] presented both stimuli centrally. Several studies have confirmed that early integration induced by peripherally presented stimuli occurred later compared to early integration induced by centrally presented stimuli [5.16]. Thus, our results reflected the phenomenon of a postponement of audiovisual interaction.

In this study, we observed that both the peak latency and topography of this interaction pattern corresponded to those of the visual N1 components in the unimodal condition. Moreover, we found that a reduced N1 amplitude was associated with faster reaction times in the bimodal AV condition compared to the unimodal V condition (Table 5.1 and Fig 5.4a), which was consistent with findings obtained in previous studies [5.15, 5.20]. At the neuronal level of audiovisual integration, the reduced N1 response may reflect the requirement of visual processing with simultaneous auditory stimuli for less energy from the visual system compared to visual processing without auditory stimuli [5.15].

However, in the present study, an enhanced perceptual sensitivity (d') and an audiovisual integration of the occipital visual areas at 180-200 ms were not found after bilateral auditory stimulus presentation (Table 5.1 and Fig 5.4a). Previous neurological studies have shown that integrations occurring during the first up to 200 ms post-stimulus onset were mediated by early sensory-perceptual processing, whereas responses after 200 ms were mediated by late cognitive processing [5.20, 5.21]. Furthermore, previous behavioral studies have confirmed that audiovisual perceptual integration was greater when the auditory stimuli were presented in close spatial proximity to the visual stimuli [5.2, 5.22]. Auditory imaging studies have demonstrated that when the auditory stimulus is presented simultaneously to both of the participant's ears, the subject will hear the stimulus as being centrally presented [5.23], whereas visual stimuli were presented laterally in our study. Thus, we speculate that the spatial incongruity between the visual and auditory stimuli resulted in an absence of audiovisual integration at early-stage processing and a weak perceptual sensitivity (d') via the effect of bilateral auditory stimuli. However, further electrophysiological studies are required to confirm and elaborate on these results.

Audiovisual integration in the fronto-central areas (300 to 320 ms).

A significant audiovisual integration was found at 300–320 ms over the fronto-central areas with ipsilateral and bilateral auditory stimuli (Fig 5.4b). The scalp distribution of the ERP activity in this study was very similar to that reported in previous studies [5.3, 5.15, 5.16, 5.20, 5.24]. Some studies had found that fronto-central multisensory ERP activity was elicited during early-stage processing (before 200 ms) when both the auditory and visual stimuli were task relevant and the participants were required to attend to both the visual and auditory stimuli [5.15, 5.20, 5.24]. However, other studies indicated that the audiovisual integration over the fronto-central areas were elicited during late stage processing (after 200 ms) when the auditory stimuli were task irrelevant and the participants were only required to attend to visual stimuli [5.3, 5.16]. A comparison of these studies revealed that the difference in latency in AV integration over the fronto-central areas might be due to the effects of selective attention. Talsma and Woldorff reported that audiovisual integration is dependent on the stimuli being fully attended to, which occurred earlier when the participants attended to both modalities compared to when the participants attended to only one modality or did not attend to either modality [5.16, 5.17, 5.25]. In our study, the auditory stimuli were task irrelevant, and only visual stimuli were required to be attended to. Thus, a fronto-central AV integration elicited during late-stage processing (300–320 ms) was observed in this study, which was mostly consistent with previous studies [5.3, 5.16]. In addition, the frontal negative activity elicited by the AV interaction, which was thought to attend to the visual modality, resulted in an association of the visual and auditory stimulus features and a spread of enhanced processing to the auditory components of the stimulus [5.16]. In the present study, enhanced auditory processing via audiovisual integration was found during late-stage processing via the effects of both the ipsilateral and bilateral auditory stimuli (Fig. 4b), which may indicate that ipsilateral and bilateral auditory stimuli play similar roles in fronto-central audiovisual integration.

5.5 Conclusion

This study provides behavioral evidence that visual detection was enhanced by ipsilateral or bilateral auditory stimuli, whereas enhanced perceptual sensitivity (d') was only found in visual detection with ipsilateral auditory stimuli. The ERP results show that behavioral performance was enhanced by audiovisual integration at approximately 180-200 ms in occipital areas via the effect of ipsilateral auditory stimuli and at approximately 300-320 ms in fronto-central areas via the effect of ipsilateral or bilateral auditory stimuli. Our ERP results indicate that the effect of ipsilateral auditory stimuli on audiovisual integration originated from both early-stage sensory and late-stage cognitive processing, while the effect of bilateral auditory stimuli on audiovisual integration originated from late-stage cognitive processing rather than early-stage sensory processing. Taken together, these results demonstrate overlapping, yet distinct patterns of audiovisual integration for ipsilateral and bilateral auditory stimuli.

References

[5.1] Carrasco M, Ling S, Read S. Attention alters appearance. Nature neuroscience 2004; 7: 308-313.

[5.2] Frassinetti F, Bolognini N, Làdavas E. Enhancement of visual perception by crossmodal visuo-auditory interaction. Experimental Brain Research 2002; 147: 332-343.

[5.3] Wu J, Li Q, Bai O, Touge T. Multisensory interactions elicited by audiovisual stimuli presented peripherally in a visual attention task: a behavioral and event-related potential study in humans. Journal of Clinical Neurophysiology 2009; 26: 407.

[5.4] Frens M, Van Opstal A, Van der Willigen R. Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. Attention, Perception, & Psychophysics 1995; 57: 802-816.

[5.5] Stein BE, Meredith MA. The merging of the senses. The MIT Press; 1993.

[5.6] Harrington L, Peck C. Spatial disparity affects visual-auditory interactions in human sensorimotor processing. Experimental Brain Research 1998; 122: 247-252.

[5.7] Lippert M, Logothetis NK, Kayser C. Improvement of visual contrast detection by a simultaneous sound. Brain research 2007; 1173: 102-109.

[5.8] Turatto M, Mazza V, Umilta C. Crossmodal object-based attention: Auditory objects affect visual processing. Cognition 2005; 96: B55-B64.

[5.9] Teder-Sälejärvi W, Russo FD, McDonald J, Hillyard SA. Effects of spatial congruity on audio-visual multimodal integration. Journal of cognitive neuroscience 2005; 17: 1396-1409.

[5.10] Woldorff MG, Tempelmann C, Fell J, Tegeler C, Gaschler-Markefski B, Hinrichs H et al. Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. Human Brain Mapping 1999; 7: 49-66.

[5.11] Schönwiesner M, Krumbholz K, Rübsamen R, Fink GR, von Cramon DY. Hemispheric asymmetry for auditory processing in the human auditory brain stem, thalamus, and cortex. Cerebral Cortex 2007; 17: 492-499.

[5.12] Behne N, Scheich H, Brechmann A. Contralateral white noise selectively changes right human auditory cortex activity caused by a FM-direction task. Journal of neurophysiology 2005; 93: 414-423.

[5.13] Gondan M, Niederhaus B, Rösler F, Röder B. Multisensory processing in the redundant-target effect: a behavioral and event-related potential study. Perception & psychophysics 2005; 67: 713-726.

[5.14] Stanislaw H, Todorov N. Calculation of signal detection theory measures. Behavior research methods, instruments, & computers 1999; 31: 137-149.

[5.15] Giard M, Peronnet F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. Journal of cognitive neuroscience 1999; 11: 473-490.

[5.16] Talsma D, Doty TJ, Woldorff MG. Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? Cerebral Cortex 2007; 17: 679-690.

[5.17] Talsma D, Woldorff MG. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. Journal of cognitive neuroscience 2005; 17: 1098-1114.

[5.18] Schröger E, Widmann A. Speeded responses to audiovisual signal changes result from bimodal integration. Psychophysiology 1998; 35: 755-759.

[5.19] Yang W, Li Q, Ochi T, Yang J, Gao Y, Tang X et al. Effects of Auditory Stimuli in the Horizontal Plane on Audiovisual Integration: An Event-Related Potential Study. PLoS One 2013; 8: e66402.

[5.20] Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ. Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. Cognitive Brain Research 2002; 14: 115-128.

[5.21] Li Q, Wu J, Touge T. Audiovisual interaction enhances auditory detection in late stage: an event-related potential study. Neuroreport 2010; 21: 173.

[5.22] Lewald J, Guski R. Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. Cognitive Brain Research 2003; 16: 468-478.

[5.23] Goupell MJ, Kan A, Litovsky RY. Mapping procedures can produce non-centered auditory images in bilateral cochlear implantees. The Journal of the Acoustical Society of America 2013; 133: EL101-EL107.

[5.24] Fort A, Delpuech C, Pernier J, Giard M-H. Early auditory–visual interactions in human cortex during nonredundant target identification. Cognitive Brain Research 2002; 14: 20-30.

[5.25] Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG. The multifaceted interplay between attention and multisensory integration. Trends in cognitive sciences 2010; 14: 400-410.

Chapter 6

General Conclusion and Future Challenges

Summary

This dissertation has investigated the neural mechanisms and correlates of audiovisual integration in human by event-related potential experiments. Particularly, we discussed the effects of crossmodal attention and the spatial characteristics of stimuli on audiovisual integration. The findings are generalized below. In addition, we describe directions for future research.

6.1 General conclusions

The dissertation is composed of four experiments. The first and second experiments investigated the effect of spatial and temporal attention on bimodal audiovisual stimuli processing. The third and fourth experiments investigated the spatial characteristics of stimuli in human audiovisual integration.

Chapter 2 describes the first experiment, utilizing the high temporal resolution of event-related potentials (ERPs), we used a central cue that can predict the time point (600 ms or 1800 ms) of audiovisual target to investigate whether endogenous temporal attention could modulate audiovisual stimuli processing. The results showed that the endogenous temporal attention could not change the amplitude of the early ERP component in conditions of either short (600 ms) or long (1800 ms) cue-target intervals, indicating that the endogenous temporal attention had no effect on the early-stage of AV stimuli processing. However, the late ERP component showed differences between the short (600 ms) and long (1800 ms) cue-target intervals, supporting a model in which endogenous temporal attention might determine the late stage of audiovisual stimuli processing.

Chapter 3 introduces the second experiment, in which we to compared the modulatory effects of spatial attention versus temporal attention on audiovisual stimuli processing using event-related potentials (ERPs) with high temporal resolution. The spatial attention was triggered by a visual spatial cue (usually an arrow) and the temporal attention was triggered by a visual temporal cue (two concentric circles). Behavioral responses to audiovisual stimuli in spatial attention condition were faster than those in temporal attention condition, and the false alarm rate in spatial attention condition was lower than that in temporal attention condition. The ERPs results show that the amplitude of N2 elicited by audiovisual stimuli at right occipitotemporal area in spatial attention condition was greater than that in temporal attention condition. Those results indicated that the spatial attention and temporal attention had a differential effect on audiovisual stimuli processing at right occipitotemporal area.

Chapter 4 describes the third experiment, in which we used event-related potentials (ERPs) to evaluate the neural mechanism of which auditory spatial information affect audiovisual integration in visual attention task, and compared the integration which elicited by left and right stimuli. The significant differences were found between left and right, which were supported by the following results: (1)There are significant difference between left and right visual and auditory stimuli in early integration. There was integration effect in frontal area around 100-120ms when visual and auditory stimuli appeared in the right. But there was none in the left. (2)There are also significant differences between ipsilateral left or right visual and auditory stimuli in late integration. It showed difference in ipsilateral left or right audiovisual integration in frontal area and centro-medial. Furthermore, it indicated that the processing of multisensory information was different between left and right brain.

Chapter 5 describes the fourth experiment. In this experiment, we compared the effect of ipsilateral and bilateral auditory stimuli on audiovisual integration using behavioral and ERPs measures in humans. Behavioral results showed that responses to visual stimuli with both types of auditory stimuli were faster than those to visual stimuli only, and an enhancement of perceptual sensitivity (d') for visual detection was found when the visual stimuli with ipsilateral auditory stimuli. Further, ERP components related to audiovisual integrations were identified over the occipital areas, approximately 180 to 200 ms of early-stage sensory processing via the effect of ipsilateral auditory stimuli, and over the fronto-central areas, approximately 300 to 320 ms of late-stage cognitive processing via the effect of ipsilateral and bilateral auditory stimuli. Our results confirmed that audiovisual integration was also elicited, even though effect of bilateral auditory stimuli, only occurring at later stages of cognitive processing in response to a visual detection task. Further, the integration from early-stage sensory processing was observed via the effect of ipsilateral auditory stimuli, which suggesting the integration of audiovisual information in human brain might be particularly sensitive to ipsilaterally presented audiovisual stimuli.

Effects of the spatial and temporal attention on audiovisual stimuli processing.

Activity of integration that occurring during the first up to 200 ms post stimulus onset were mediated by early sensory-perceptual processing, whereas responses after 200ms was mediated by late cognitive processing. In the first experiment, we investigated whether peripherally presented audiovisual stimuli could be modulated by endogenous temporal attention when the central visual cue could completely predict the cue-target interval (600 ms or 1800 ms). Specifically, the modulation effect appeared in the late ERP components from 300 ms post-stimuli, such as the late negativity component (320-340 ms) and the late positivity component (400-500 ms), rather than in the early ERP components (P1 & N1). In our second experiments, the ERP results showed that a enhanced N2 amplitude at an audiovisual integration region of the brain (right occipitotemporal area) is associated with low false alarm rates in the spatial attention condition compared with the temporal attention condition. Therefore, we concluded that the attention affected audiovisual stimuli processing in humans.

Effects of the spatial characteristics of stimuli on human audiovisual integration

The effects of the spatial location of stimuli on audiovisual integration were investigated in our third and fourth experiment. The results through third experiment showed that siginificant differece between ipsilateral left or right visual and auditory stimuli in audiovisual integration generated by frontal and center areas. Furthermore, it indicated that the processing of multisensory information was different between left and right brain. In the fourth experiment, ERPs results indicate that the effect of ipsilateral auditory stimuli on audiovisual integration originates from both early-stage sensory and late-stage cognitive processing, while the effect of blateral auditory stimuli on audiovisual integration originates from late-stage cognitive processing rather than early-stage sensory processing. These results demonstrate overlapping but distinctive patterns of audiovisual integration for ipsilateral and bilateral auditory stimuli.

6.2 Future challenges

The neural mechanism of audiovisual integration is very complex. Previous studies have shown that an audiovisual stimulus with spatial and temporal coincidence can generate the largest integration effects. In future studies, it is considered that researchers should apply an integrated pattern of behavioral methods combined with neurological methods to investigate audiovisual integration. As to the neurological methods, different methods can be used to investigate the same problem, such as using both ERP and fMRI. Due to virtual reality technology not being applied in neuroimaging studies, research in laboratories loses the ecological effect in reality. Furthermore, the relationship among multisensory cortices and that between multisensory cortices and single-sensory cortices should be further investigated in future.

Appendix-- Simple Introduction of EEG Apparatus

The BrainAmp MR *plus* was manufactured by BrainProduct Inc., Germany. This amplifier is a compact solution for neurophysiology research that can be combined with other units within the same product family to cover a vast range of possible application areas. This fully portable solution can be used for standard EEG/ERP recordings and can also be placed inside of the MRI bore for simultaneous EEG/fMRI acquisitions.

Thanks to its 5 kHz sampling rate per channel, the BrainAmp can be used to record EEG, EOG, and EMG signals as well as evoked potentials with a frequency up to 1 kHz. The 16-bit TTL trigger input allows the detection of a large number of markers from visual, acoustic, electrical, magnetic or other stimulation modalities. The BrainAmps can be used both with passive and active electrodes offering a great degree of flexibility.

The 32 channel units can be stacked to expand the number of channels up to 256 and combined with the BrainAmp ExG to record EEG, EOG, EMG, ECG, GSR (Galvanic Skin Response) and many other types of bipolar and auxiliary signals.

Figure A1. EEG amplifier of BrainAmp MR *plus*

Table A1. Technical specifications of BrainAmp MR *plus*

Number of Channels per unit	32
Max. Number of channels	128
Reference Type	unipolar

MR-compatibility	Yes (for scanners up to 4 Tesla)
Bandwidth [Hz]	DC - 1000
High Pass Filter [Hz]	0.016 / 10 s AC or DC switchable
Low Pass Filter [Hz]	1000 / 250 switchable
Input Noise [μVpp]	≤ 1
Input Impedance [MΩ]	10 / 10000
Input Measurement Ground / Reference	Yes
A/D-C [bit]	16
A/D-Rate [Hz]	5000
Max. Sampling Frequency [Hz]	5000
Offset Compatibility [mV]	± 300
Operating Range [mV]	selectable: ±3.2768; ±16.384; ±327.68
Resolution [μV]	selectable: 0.1; 0.5; 10.0
CMRR [dB]	≥ 110
TTL Trigger Input [bit]	16
Synchronized Digital Trigger Input [bit]	up to 16
Max. Power Consumption [mA]	160
Power Supply	rechargeable Battery
Signal Transmission	optical
PC Interface	PCI, USB 2.0
Deblocking Function	Yes
Blocking of Unused Channels	Yes
Safety	Twin Fiber optical Transmission Protection Class II, Type BF IEC EN 60601 EMC tested, electrically safe
Classification to MDD 93/42/EEC	Class IIa

Dimensions H x W x D [mm]	68 x 160 x 187
Weight [kg]	1.1

Publications

Journal papers:

1) Yulin Gao, Chunlin Li, Xiaoyu Tang, Qi Li, Weiping Yang, Jingjing Yang, Ishikawa Soushirou and Jinglong Wu. Modulation of response to audiovisual stimuli presented peripherally by visually cued endogenous temporal attention. *Neuroscience and Biomedical Engineering*, Accept. (2013).

2) Jinglong Wu, Weiping Yang, Yulin Gao and Takahiro Kimura. Age-related multisensory integration elicited by peripherally presented audiovisual stimuli. *NeuroReport*, 2012, 23(10), 616-620.

3) Weiping Yang, Qi Li, Tatsuya Ochi, Jingjing Yang, Yulin Gao, Xiaoyu Tang, Satoshi Takahashi, Jinglong Wu. Effects of auditory stimuli in the horizontal plane on audiovisual integration: an event-related potential study. *Plos One*, Accept. (2013).

4) Qi Li, Jingjing Yang, Yulin Gao and Jinglong Wu. Unpredictable auditory spatial information modulates spatial-congruent audiovisual integration. *Journal of Information*. Accept. (2013).

5) Xiaoyu Tang, Chunlin Li, Qi Li, Yulin Gao, Weiping Yang, Jingjing Yang, Soushirou Ishikawa and Jinglong Wu. Modulation of Auditory Stimulus Processing by Visual Spatial or Temporal Cue: An Event-related Potentials Study. *Neuroscience Letters*. Accept. (2013).

International conference papers

1) Yulin Gao, Jinglong Wu and Ming Zhang. Investigation on Human Characteristic of Relationship between Spatial Working Memory and Controlled Attention for Education. *IEEE/BMEI International Conference on BioMedical Engineering and Informatics (BMEI'10)*. (2010) pp.1927-1931.

2) Yulin Gao, Jingjing Yang, Qi Li, Ryota Morikawa and Jinglong Wu. A Basic Study of Event-Related Potentials (ERPs) on Human Audiovisual Spatial Integration for Human-machine Interface. *IEEE/ROBIO International Conference on Robotics and Biomimetics*. (2010) pp.1244-1249.

3) Jingjing Yang, Qi Li, Yulin Gao and Jinglong Wu. Task-irrelevant Auditory Stimuli Affect Audiovisual Integration in a Visual Attention Task: Evidence from Event-related Potentials.

IEEE/CME International Conference on Complex Medical Engineering. (2011) pp.248-253.

4) Yulin Gao, Qi Li, Jingjing Yang and Jinglong Wu. Ipsilateral visual and auditory spatial information interaction mechanisms in selective attention conditions: Behavioral and ERP study. *IEEE/ICMA International Conference on Mechatronics and Automation.* (2011)pp.1560-1565.

5) Yulin Gao, Xiaoyu Tang, Chunlin Li, Jingjing Yang, Weiping Yang, Satoshi Takahashi and Jinglong Wu. Effect of cue-target interval on temporally endogenous attention of No-Go target: Event-related potential evidence. *IEEE/ICMA International Conference on Mechatronics and Automation.* (2012) pp.1405-1410.

6) Xingjuan Liu, Yulin Gao, Ming Zhang. The Left to Right in Inhibition of Return in discrimination tasks. *IEEE/ICMA International Conference on Mechatronics and Automation.* (2012). pp.2579 - 2583

7) Weiping Yang, Yulin Gao, Jingjing Yang, Satoshi Takahashi, Takahiro Kimura and Jinglong Wu. Multisensory integration of peripherally presented audiovisual stimuli in older adults: An event-related potential study. *IEEE/ICMA International Conference on Mechatronics and Automation.* (2012) pp.563-568.

8) Jingjing Yang, Tatsuya Ochi, Qi Li, Yulin Gao, Weiping Yang, Xiaoyu Tang, Satoshi Takahashi and Jinglong Wu. The Effect of Spatial Information of Auditory Stimuli on Audiovisual Interaction: Evidence from Event-related Potentials. *IEEE/CME International Conference on Complex Medical Engineering,* (2012) pp.699-704.

9) Xiaoyu Tang, Ming Zhang, Yulin Gao, Weiping Yang, Chunlin Li, Jingjing Yang, Ishikawa Soushirou and Jinglong Wu. Effect of cue-target interval on endogenous attention in Go/No-Go task: evidence from an event-related potentials study. *IEEE/CME International Conference on Complex Medical Engineering,* (2012) pp.669-672.

10) Ming Zhang, Xingjuan Liu, Yulin Gao. From Left to Right in Inhibition of Return of Congenitally Deaf People Depending on Task Type. *IEEE/CME International Conference on Complex Medical Engineering,* (2012) pp.271-275.

11) Xiaoyu Tang, Chunlin Li, Qi Li, Yulin Gao, Weiping Yang, Jingjing Yang, Ishikawa Soushirou, Satoshi Takahashi and Jinglong Wu. Effect of Cue-target Interval on Audiovisual Stimuli Processing in Endogenous Spatial Attention: An Event-related Potentials Study. *IEEE/CME International Conference on Complex Medical Engineering,* (2013) pp.656-661.

12) Weiping Yang, Qi Li, Tatsuya Ochi, Jingjing Yang, Yulin Gao, Xiaoyu Tang, Satoshi Takahashi and Jinglong Wu. Influences of auditory stimuli in front and rear space on visual detection: an event-related potential study. *IEEE/CME International Conference on Complex Medical Engineering*, (2013) pp.7-12.

Book chapters

1) Yulin Gao, Weiping Yang, Jingjing Yang, Satoshi Takahashi and Jinglong Wu. Neural Mechanisms of Audiovisual Integration in Integrated Processing for Verbal Perception and Spatial Factors. *Chapter 34, (2012) IGI Global*. pp.327-336.

2) Weiping Yang, Yulin Gao and Jinglong Wu. Effects of Selective and Divided Attention on Audiovisual Interaction. *Chapter 32, (2012) IGI Global*. pp.311-319.

3) Jingjing Yang, Qi Li, Yulin Gao and Jinglong Wu. Temporal Dependency of Multisensory Audiovisual Integration. *Chapter 33, (2012) IGI Global.* pp.320-326.

4) Xiaoyu Tang, Yulin Gao, Weiping Yang, Ming Zhang and Jinglong Wu. Audiovisual Integration of Natural Auditory and Visual Stimuli in the Real-World Situation. *Chapter 35, (2012) IGI Globa*l. pp.337-344.

Acknowledgements

First and foremost, I would like to express my heartfelt thanks and gratitude to my supervisor, Professor Jinglong Wu, who has provided valuable guidance for me during the doctor course and the writing of this dissertation. Without his enlightening instruction, impressive kindness and patience, I could not have completed my study of doctor course and this dissertation. His study theory led me to the study of cognitive neuroscience. His diligence, will power, courage, and kindness enlightens me not only at present but also at my future work. Beyond the university, Professor Wu has also kindly helped me, a foreigner in Japan, with everyday life.

I additionally extend my great appreciation to Professor Shinichiro Yanase, Professor Eiji Tomita, Professor Akihiko Horibe, Associate Professor Satoshi Takahashi and Assistant Professor Jiajia Yang, who provided valuable advice about writing the dissertation and oral presentation.

I also express my sincere thanks to Associate Professor Qi Li of Changchun University of Science and Technology and Master Weiping Yang, Xiaoyu Tang in Wu lab, who provided valuable advice about my research and kindly manuscript review. In addition, I also thank the students and staff in Wu lab. Without their cooperation, I cannot imagine how I could have finished my experiments. I sincerely thank all those who contributed to my experiment, my paper, and my dissertation.

Finally, I extend my deep appreciation and dedicate this dissertation to my family, especially to my mother, who brought me up and have always supported and encouraged me.