The Different Effects of Selective Attention and Divided Attention on Audiovisual Temporal Integration

Jingjing Yang[*], Qi Li and Xiujun Li

Changchun University of Science and Technology, 7089 Weixing Road, Changchun , 130022, China

yangjingjing@cust.edu.cn

Keywords: selective attention, divided attention, audiovisual temporal integration

Abstract. Human brains integrate various kinds of information from different modalities to perceive our complex environment. The purpose of this study was to investigate the effect of attention on the processing of audiovisual temporal information. Fourteen subjects participated in our experiment. Visual, auditory and audiovisual information was presented by the display and speaker, and the volunteers were asked to judge the target stimuli. In present study, audiovisual stimuli onset asynchrony (SOA= ±250 ms, ±200 ms, ±150 ms, ±100 ms, ±50 ms, 0 ms). Volunteers completed two experiments: first, only the auditory stimulus was attended and determined the auditory target stimuli. Second, pay attention to the visual and auditory channel at the same time and judge the target stimuli. The results showed that no matter what kind of attention, multimodal audiovisual stimuli improved the judgment. Under the SOA=-50 and SOA=0 conditions, the effect of selective attention and divided attention was different, the promotion effect under the divided attention is stronger. Our results can be helpful for the information processing of artificial intelligence.

Introduction

Artificial intelligence (AI) is using artificial methods and techniques to make the machines or software with the human-like intelligence. The information processing mechanism of human brain is the basis of artificial intelligence research. Human brain can effectively integrate information from multisensory. Previous studies have shown that reaction is more quickly and accurately to bimodal audiovisual stimuli compared with either unimodal visual or auditory stimuli [1-3]. Spatially and temporally coincidence are basic conditions that elicited the audiovisual integration [4]. Perception of time is necessary for human life, but no specific sense organs for temporal perception in human body. Many previous researchers have investigated the perceptual processes of audiovisual integration when visual and auditory stimuli are presented simultaneously [5].

In daily life, the phenomenon of audiovisual stimuli onset asynchrony is ubiquitous. For example, we always see the lightning before we hear the thunder. When the input of the auditory and visual stimuli are asynchronously, the perception of the visual channel would be affected by the auditory channel. [6]. For example, temporal disparity was easier in arousing the illusion if the visual stimulus was presented before the auditory stimulus by about 50-100ms [7-8]. Time window of integration is about 250 ms, in which the visual and audiotory stimuli tend to be integrated into a unified percept [9-12].

Many studies have shown that temporal asynchrony of visual-auditory stimuli can influence the audiovisual integration, however, the effect of attention on the processing of audiovisual temporal information are not well understood. The aim of our study is to investigate the effect of attention on audiovisual temporal integration.

In present study, the auditory and visual stimuli onset asynchrony (SOA = ±250 ms, ±200 ms, ±150 ms, ±100 ms, ±50 ms, 0 ms),subjects attended to auditory modality or attended to auditory-visual modality. Unimodal auditory, unimodal visual and audiovisual (11 types) stimuli were presented randomly. We investigated the effects of selective attention and divided attention on audiovisual temporal integration .We hope that this research may be helpful to the development of artificial intelligence.

Methods

Subjects. Fourteen collage students from the Changchun University of Science and Technology participated in this study(5 females,18 to 22 years, mean: 18.93 years). All subjects were healthy, with normal vision or corrected normal visual and normal hearing. Before the experiment, the experiment instruction was carefully read, and the informed consent was signed.

Stimuli. Unimodal visual (V) stimuli , unimodal auditory (A) stimuli, and multimodal audiovisual (AV) stimuli were contained in this experiment. Each type had two subtypes of standard stimuli and target stimuli.

The unimodal V standard stimulus was a white, horizontal, square wave grating (subtending a visual angle of approximately 5.8°(the spatial frequency was 0.78 cycles/degree, Fig.1). The unimodal A standard stimulus was 1600 Hz sinusoidal tone, with linear rise and fall times of 5ms and the intensity was 75 dB. The time duration was 105ms.

The multimodal AV stimulus consisted of both unimodal A and unimodal V stimuli onset asynchrony (SOA). The SOA were ±100 ms, ±50 ms, 0ms. The "−"means auditory stimulus was presented later than visual grating. "SOA = 0" means visual and auditory stimuli onset synchronously. The AV standard stimulus consisted of both A standard stimulus and V standard stimulus.

The target stimuli of V and A were very similar to the V and A standard stimuli, but contained a distance of no stimulation exhibition time halfway through the duration of the stimulus, which subjectively made the stimulus appear to flicker (V target stimulus) or stutter (A target stimulus). The no stimulation exhibition time halfway were 35ms. The multimodal AV target stimulus consisted of both unimodal A and V target stimulus. The target stimulus was presented with 50% frequency.
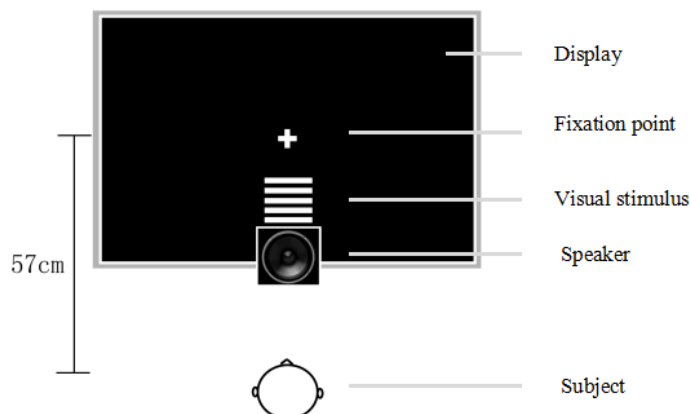

Fig.1 Experimental arrangement

The distance from subject to display is 57cm.The Visual stimulus was presented 4º below the plus. The auditory stimuli were presented through two speakers that was placed on the left and right of the display. Both the auditory and visual stimuli were attended. The subject's task was to detect the target stimuli as quickly and accurately as possible and to respond by pressing the left button of a computer mouse. The inter stimulus interval (ISI) is 800 to 1200ms.

Task and Procedure. Experiments were carried out in a dark and sound-attenuating room. A 17-in display was used to present the visual stimuli. The distance is 57cm from display to subject. One speaker was used to present A stimuli. The speaker was placed before the display at 0º (Fig. 1). The experiment consisted of two parts. First, selective attention part, only the auditory stimulus was attended and subjects were asked to response to auditory target stimuli. Second, divided attention part, both attention to the visual and auditory channel and judge the target stimuli. The probability of each stimulus type was equal and all stimuli were randomly presented. Subjects placed his right hand on a computer mouse. Pressing the left key of a computer mouse when the target stimuli was detected. The response of subjects were required as accurately and quickly as possible.

Data Analysis. The Presentation (Version 0.71) experimental software were used to present the stimulus and acquire the data. Only the data of targets were acquired and analysized. Response times (RTs) and hit rates (HRs) to target stimuli of each stimulus type were computed separately. Hit rates were the number of correct responses to target stimuli divided by the total number of target stimuli.

Response time were the average of correct responses to targets. We used T-test to assessment the reaction times and hit rates for the different types of target stimuli.

Results

Reaction time were recorded from stimulus onset to the button press. All stimulus types of reaction time and hit rates were recorded.
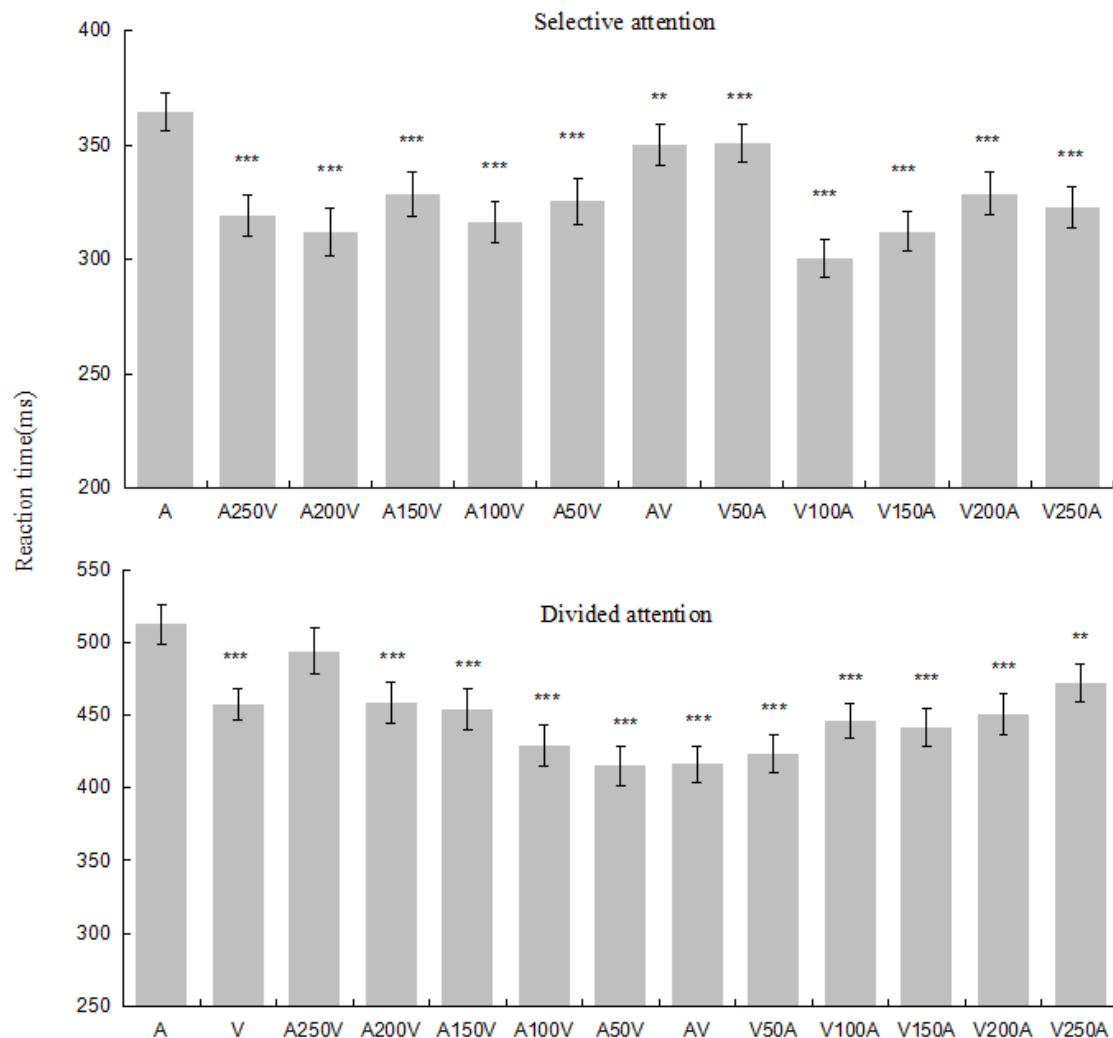

Fig.2 Average response times (RTs) in the detection of each stimulus type

The above figure shows the results from selective attention, below shows the divided attention. SOA = 0ms: visual and auditory stimuli onset synchronously; A250V: SOA = -250ms, visual stimulus delayed by 250ms relative to the onset of the auditory; A100V: SOA = -100ms, visual stimulus delayed by 100ms relative to the onset of the auditory; V250A: SOA = 250ms, auditory stimulus delayed by 250ms relative to the visual. Asterisks (*) denote conditions in which RTs are significantly(**;P<0.01, ***;P<0.001)faster than for the auditory stimulus alone (i.e. A).Error bars represent the group SEM.

In selective attention part, analysis of the factor stimulus type confirmed that reaction times to unimodal auditory and multimodal audiovisual stimuli (11 types of SOA) differed significantly [F (10, 130)=546.12; p<0.001]. In addition, post-hoc comparisons found that the reaction times to the multimodal audiovisual stimuli in which auditory stimuli were presented after the visual stimuli 100ms (p<0.001) were significantly faster than those to the unimodal auditory stimuli. Although visual stimuli did not require attention, but the multimodal audiovisual stimuli still contributed to the detection. These results indicate that a synergistic effect occurred when the auditory stimuli were aligned.

Fig.3 shows the hit rates, the black line was the result of selective attention , no significant differences were found between unimodal visual stimuli and bimodal audiovisual stimuli. In selective attention part, the visual stimuli were not attended. Responses were the fastest when the visual stimuli preceded A stimuli about 100ms (SOA = -100ms). The visual stimuli which were presented earlier than auditory stimuli were regarded as a visual cue. The temporal relationship of visual cues to auditory targets can speed the auditory detection [13-14]. The preceding visual cue enhanced attention to the following auditory stimuli, even though the visual stimuli were task-irrelevant. Task-irrelevant information cannot be completely ignored. The attention could enhance or modulate multisensory integration.

In divided attention part, reaction times to AV target stimuli (excepted A250V condition)were significantly shorter than that to unimodal auditory condition (p<0.01). The Reaction times to A100V, A50V, AV, V50A were significantly shorter than that to unimodal visual condition(p=0.007, p<0.001, p<0.001, p<0.001). Analysis of the stimulus type factor confirmed that the Reaction times to unimodal auditory and multimodal AV stimuli (11 types of SOA) differed significantly [F (10, 130)=546.12; p<0.001]. These results indicate that the alignment of audiovisual stimuli elicited synergistic effect. Fig. 3 shows the Hit rates, the gray line was the result of selective attention , the responses to temporal asynchronous audiovisual stimuli were more accurate than unimodal auditory stimuli. Hit rates to AV target stimuli (excepted A250V condition) were more accurate than that to unimodal visual condition. In divided attention part, both the visual stimuli and the auditory were attended. The enhancement was the most prominent when the visual and auditory stimuli were approached. So we propose that the visual temporal information were used effectively.

Compare the results of two experiments found that no matter what kind of attention, multimodal audiovisual stimuli improved the judgment. Under the SOA=-50 and SOA=0 conditions, the effect of selective attention and divided attention was different. The promotion effect under the divided attention is stronger.
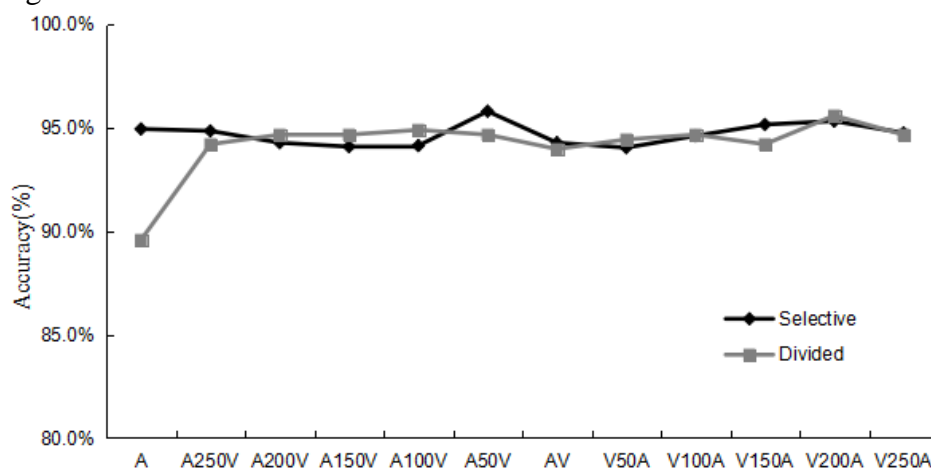

Fig.3 Average hit rates (HRs) in the detection of each stimulus type. The black line shows the result of selective attention, the gray line shows the result of divided attention.

Conclusion

The aim of this study was to investigate the effect of attention on the processing of audiovisual temporal information. No matter what kind of attention, multimodal audiovisual stimuli improved the judgment. In other words, the temporal asynchronous AV stimuli enhanced the detection [8,12]. Under the SOA=-50 and SOA=0 conditions, the effect of selective attention and divided attention was different, the promotion effect under the divided attention is stronger. We expect that these results could helpful to process the audiovisual information of artificial intelligence.

References

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

[2] M.Lippert, N.K.Logothetis, C.Kayser, Improvement of visual contrast detection by a simultaneous sound, Brain research, 2007,1173: pp.102-109.

[3] Qi Li, Jinglong Wu and Tetsuo Touge, Audiovisual interaction enhances auditory detection in late stage: an event-related potential study. Neuroreport 2010; 21:173–178.

[4] Hershenson, M., Reaction time as a measure of intersensory facilitation. J. Exp. Psychol. 1962, 63: pp.289–293.

[5] Posner, M.I., Nissen, M.J., Klein, R.M.,. Visual dominance: an information-processing account of its origins and significance. Psychol. Rev. 1976,83: pp.157–171.

[6] Daniel A.Slutsky and Gregg H. Recanzone, Temporal and spatial dependency of the ventriloquism effect, NeuroReport,2001,12:pp.7-10.

[7] Jinglong Wu, Qi Li, Ou Bai, and Tetsuo Touge, 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 : 407-413.

[8] S.Morein-Zamir, S.Soto-Faraco and A. Kingstone, Auditory capture of vision: Examining temporal ventriloquisum, Cognit. Brain Research,2003, 17: pp.154-163.

[9] C.Spence and J.Driver, Attracting attention to the illusory location of a sound reflexive crossmodal orienting and ventriloquism. NeuroReport, 2000,11:2057-2061.

[10] D.Raab, Statistical facilitation of simple reaction times. Transactions of the New York Academy of Sciences, 1962,24:pp. 574-590.

[11] M. Eimer, Attending to quadrants and ring-shaped regions: ERP effects of visual attention in different spatial selection tasks. Psychophysiology, 36, 491-503, (1999).

[12] J. Lewald and S. Getzmann, Horizontal and vertical effects of eye-position on sound localization, Hear Res, 2006,213:pp.99-106.

[13] Simon R, Sylvain D, Pascal D, Sabine DD, Jacques H and Henrique S. Peripherally presented emotional scenes: A spatiotemporal analysis of early ERP response. Brain Topogr; 2008,20: pp.216-223.

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