

Different Effects of Attentional Mechanisms between Visual and Auditory Cueing

Yasuhiro Takeshima¹ & Jiro Gyoba²

¹ Department of Psychology, Faculty of Human Studies, Bunkyo Gakuin University, Japan

² Department of Psychology, Graduate School of Arts & Letters, Tohoku University, Japan

Correspondence: Yasuhiro Takeshima, Department of Psychology, Faculty of Human Studies, Bunkyo Gakuin University, 1196 Kamekubo, Fujimino, Saitama, 356-8533, Japan. Tel: 81-49-261-6663. E-mail: yasuhito.takeshima@gmail.com

Received: June 23, 2015 Accepted: August 12, 2015 Online Published: August 14, 2015

doi:10.5539/ijps.v7n3p176

URL: <http://dx.doi.org/10.5539/ijps.v7n3p176>

Abstract

Audio-visual integration interacts with attentional mechanisms. Additionally, salient auditory stimuli automatically draw attention to an audio-visual event, while spatial attention can modulate audio-visual integration. Attention induced by auditory inputs (sound-driven attention) facilitates visual perception. Similarly, visual attention improves performance on a visual task. However, the difference between attention driven by auditory and visual cues is not clear. When visual attention facilitates visual perception, there is a trade-off between spatial and temporal resolution. In contrast, audition has superior temporal resolution to vision. In the present study, we investigated the difference between auditory and visual cue-driven attention with respect to this trade-off. The results indicated that visual cueing increased spatial resolution but decreased temporal resolution. On the other hand, auditory cueing affected the efficiency of visual processing (i.e., response time) for temporal gap detection. These findings suggest that auditory cueing capitalizes on resources available for visual processing. In contrast, visual cueing may increase activation of the spatial channel instead of inhibiting the temporal channel, as proposed in previous study. Overall, there appear to be clear differences between mechanisms involved in auditory and visual cues-driven attention.

Keywords: multisensory integration, sound-driven attention, visual attention, temporal/spatial resolution

1. Introduction

Individuals perceive the external environment through sensory information, which can involve multisensory information. In using multisensory information, each sense compensates for ambiguity in perception in relation to other senses. Such multisensory integration is fundamental for producing stable and efficient perception.

Multisensory interactions interface with attentional mechanisms during perceptual processing. Synchronous sound presentation with visual stimuli drives attention and affects visual processing due to two types of attentional mechanism: bottom-up and top-down (see Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010, for a review). Regarding bottom-up mechanisms, attention is driven automatically by salient stimuli, and is captured by multisensory events. For example, when a brief sound is presented concurrently with a target color or orientation change, search efficiency improves dramatically even among difficult visual search displays (Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008; Van der Burg, Talsma, Olivers, Hickey, & Theeuwes, 2011). A salient sensory stimulus elicits a strong neural response that is automatically linked to a stimulus in another modality (Talsma et al., 2010). In addition, auditory stimuli can speed up attentional shifts toward visual targets (Keetels & Vroomen, 2011).

For top-down mechanisms, spatial attention induced by various cues modulates multisensory integration processing. Furthermore, top-down attention directed toward one sensory stimulus spreads to stimuli in other modalities that occur at the same time (Busse, Roberts, Crist, Weissman, & Woldorff, 2005). When attention is directed to visual lip movements that match a spoken auditory sentence, activity increases in multiple brain areas (Fairhall & Macaluso, 2009). In contrast, allocating spatial attention towards irrelevant lip movements interferes with recognition of audio-visual speech signals (Senkowski, Saint-Amour, Gruber, & Foxe, 2008). When multisensory stimuli compete for processing resources (i.e., the saliency of an individual sensory signal is low),

top-down selective attention is likely to be necessary for multisensory integration (e.g., Talsma, Doty, & Woldorff, 2007).

However, the differences in attentional mechanisms between attention induced by auditory and visual cues have not been investigated sufficiently. In the auditory domain, attention induced by auditory input (sound-driven attention) also increases the sensory input signal in audio-visual processing because auditory inputs induce feedback projections from the early auditory cortex to the early visual cortex during audio-visual interactions (Watkins, Shams, Josephs, & Rees, 2007; Watkins, Shams, Tanaka, Haynes, & Rees, 2006). Visual attention also increases the sensory input signal via visual cueing (e.g., Carrasco, Williams, & Yeshurun, 2002). However, visual attention does not simply improve processing but also reduces temporal channel performance, while increasing spatial channel performance (Yeshurun & Levy, 2003). Therefore, we examined the differences between attention induced by auditory and visual cueing using temporal and spatial gap detection tasks. We hypothesized that response speed would increase with auditory cueing according to an increasing sensory input signal. On the other hand, temporal gap detection sensitivity would decrease with visual cueing instead of increasing spatial gap detection sensitivity.

2. Experiment 1

Experimental 1 examined the differences in the attentional mechanisms underlying auditory and visual cues-driven attention. Here, we examined effects on temporal resolution by measuring a two-flash fusion threshold. The experimental paradigm was adopted from a previous study (see Yeshurun & Levy, 2003). We hypothesized that the sensitivity would decrease by visual cue compared with neutral cue condition. On the other hand, response time (RT) would be shorter in the auditory cue condition than in the neutral cue condition.

2.1 Participants

A group of 11 Tohoku University graduate and undergraduate students (8 women and 3 men) participated in Experiment 1. All participants reported normal or corrected-to-normal vision and audition. None had been informed as to the purpose of the experiment. They were awarded experimental credit for their participation.

2.2 Materials

The visual target was a white (43.5 cd/m^2) disk (1.0 deg in diameter), and the fixation was a white cross (about 1.4 deg). The target and fixation were displayed on a black (1.6 cd/m^2) background. The target was presented 10 deg to the left or right of the fixation. The visual cue was a red (15.3 cd/m^2) frame (1.5×1.5 deg). In the neutral and auditory cue conditions, the red frame was presented at the location of the fixation (i.e., presented in the center of the display). In the visual cue condition, the red frame appeared at the same location where the target would be presented (i.e., presented at either the left or right side of the fixation according to the location of the target). The red frame remained during the target offset period. The auditory cue was a pure tone, with a frequency of 1000 Hz and sound pressure level of 80 dB. The auditory cue was presented simultaneously with target onset. The auditory cue duration was only 50 ms.

2.3 Apparatus

The stimuli were generated and controlled by custom-made MATLAB scripts (MathWorks, Inc.), Cogent Graphics and 2000 toolbox (www.vislab.ucl.ac.uk/cogent.php), and a PC (XPS720, Dell; OS: Windows Vista, Microsoft). Visual stimuli were displayed on a CRT-display (Trinitron GDM-F520, Sony; resolution: 1024×768 pixels; refresh rate: 60 Hz). The auditory stimuli were conveyed through an audio interface (Edirol FA-66, Roland) and headphones (HDA200, Sennheiser). The synchrony of the visual and auditory stimuli was confirmed using a digital oscilloscope (TS-80600, Iwatsu). The experiment was conducted in a dark room with 43.6 dB (A) of background noise. The participants viewed the monitor binocularly at a distance of 60.0 cm with their heads stabilized using a chin rest.

2.4 Procedure

Each trial began with the participant pressing the “0” key. The fixation cross was presented for 1000 ms followed by the visual cue (Figure 1). The target was presented 100 ms after the onset of the visual cue. For the auditory cue condition, the pure tone was presented simultaneously with target onset. In the case of a double flash, two disks appeared, each for 50 ms, separated by one of three inter-stimulus intervals (ISIs; 17, 33, or 50 ms). In the case of a single flash, a single disk was presented for one of three durations (117, 133, or 150 ms). All participants completed 20 trials for each 3 (Cueing; neutral, auditory cue, or visual cue) \times 3 (ISI; 17, 33, or 50 ms) \times 2 (Flash; single or double) condition for a total of 360 trials. Participants were asked to discriminate the number of flashes (one or two). Accuracy and RT were recorded.

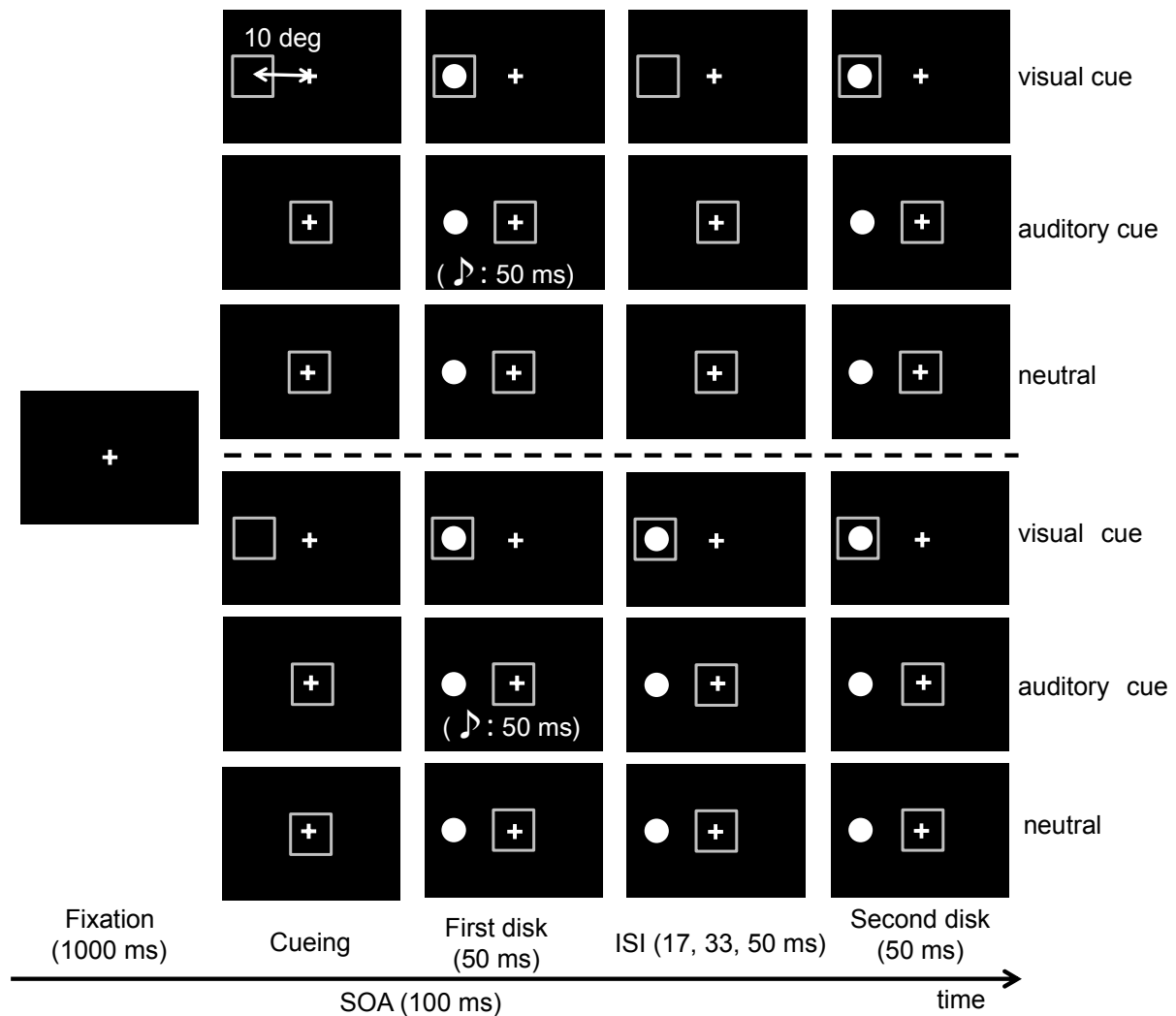


Figure 1. Schematic representation of the procedure in Experiment 1. The upper three and lower three streams indicate the double disk and single disk conditions, respectively

2.5 Results and Discussion

Sensitivity (d') and criterion (c) scores for single vs. double flash discriminations were computed using signal detection theory (Yeshurun & Levy, 2003). Additionally, the average RT was computed using correct response RT data. The results are shown in Figure 2. A two-way analysis of variance (ANOVA) with Cueing (3) \times ISI (3) was conducted for d' scores. The main effect of Cueing was significant ($F(2, 20) = 10.97, p < .001, \eta_p^2 = .52$). Multiple comparisons (Ryan's method) indicated that sensitivity was higher in the neutral and auditory cue conditions compared to the visual cue condition ($ps < .005$). No significant difference in sensitivity between the neutral and auditory cue conditions was observed ($p = .68$). Moreover, the main effect of ISI was also significant ($F(2, 20) = 31.67, p < .001, \eta_p^2 = .76$). Multiple comparisons indicated that sensitivity was higher in the 33 and 50 ms ISI conditions than in the 17 ms ISI condition ($ps < .001$). The difference in sensitivity between the 33 ms and 50 ms ISI conditions was not significant ($p = .26$). The interaction between Cueing and ISI was not significant ($F(4, 40) = 0.76, p = .56, \eta_p^2 = .07$). Differences in criterion score between cueing conditions were small (neutral: $c = 0.05$; auditory cue: $c = 0.05$; visual cue: $c = -0.09$). Finally, a two-way ANOVA with the Cueing (3) \times ISI (3) was conducted for the RT data. The main effect of Cueing was significant ($F(2, 20) = 4.78, p < .05, \eta_p^2 = .32$). Multiple comparisons indicated that RTs were shorter in the auditory cue condition as compared to the neutral condition ($p < .01$). The main effect of ISI ($F(2, 20) = 0.68, p = .52, \eta_p^2 = .06$) and the two-way interaction ($F(4, 40) = 0.80, p = .54, \eta_p^2 = .07$) were not significant.

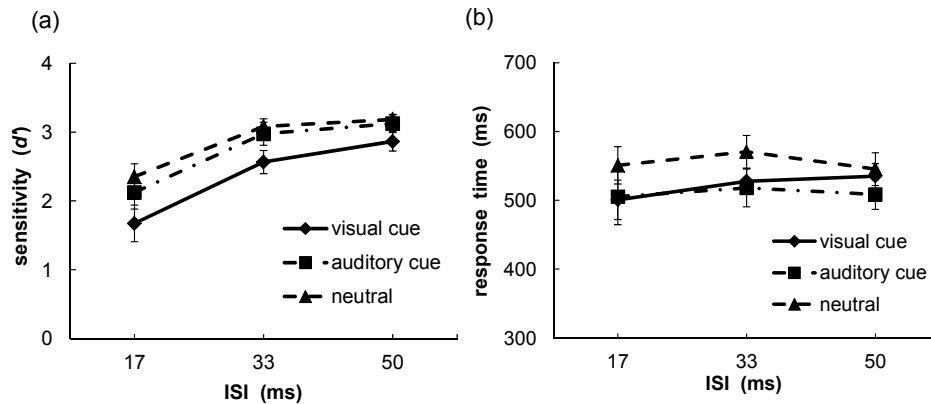


Figure 2. Temporal gap detection sensitivity and response time in Experiment 1. (a) Sensitivity for each ISI and Cueing condition. (b) Response time for each ISI and Cueing condition. Error bars represent standard errors of the mean ($n = 11$)

In this experiment, visual cueing reduced the participants' ability to detect a temporal gap, replicating Yeshurun and Levy's (2003) findings. In contrast, a co-occurring tone did not influence temporal sensitivity. The results of the RT analyses revealed that the auditory cue accelerated temporal judgments. Therefore, the auditory cue did not impair vision's temporal resolution but rather improved visual temporal judgment speed. Yeshurun and Levy (2003) have proposed that spatial visual cues improve spatial resolution instead of impairing temporal resolution. In Experiment 2, we investigated the difference in spatial gap detection sensitivity between visual and auditory cueing.

3. Experiment 2

In Experiment 1, a difference in temporal gap detection performance was observed between auditory and visual cue-driven attention. This difference may be induced by the trade-off between temporal and spatial resolution in attention. In Experiment 2, we attempted to confirm this hypothesis by examining differences in spatial resolution performance between visual and auditory cueing.

3.1 Participants

The participants were the same as in Experiment 1.

3.2 Materials

The visual target was a white circle (1.0 deg in diameter), with a 0.2 deg gap on the left or right side. The visual target was presented on the left or right side of the fixation. The distance between the fixation and target was 5.0 deg. The duration of the visual target was one of three types (67, 85, or 100 ms). The other stimuli and conditions were the same as in Experiment 1.

3.3 Procedure

Each trial began by the participant pressing the "0" key. The fixation was presented for 1000 ms followed by the attentional cue (Figure 3). The cueing conditions were the same as in Experiment 1. The target was presented for one of three durations, 100 ms after cue onset. All participants completed 20 trials for each 3 (Cueing; neutral, auditory cue, or visual cue) \times 3 (Duration; 67, 85, or 100 ms) \times 2 (Gap location; left or right) for a total of 360 trials. Participants were asked to discriminate the location of the gap (left or right).

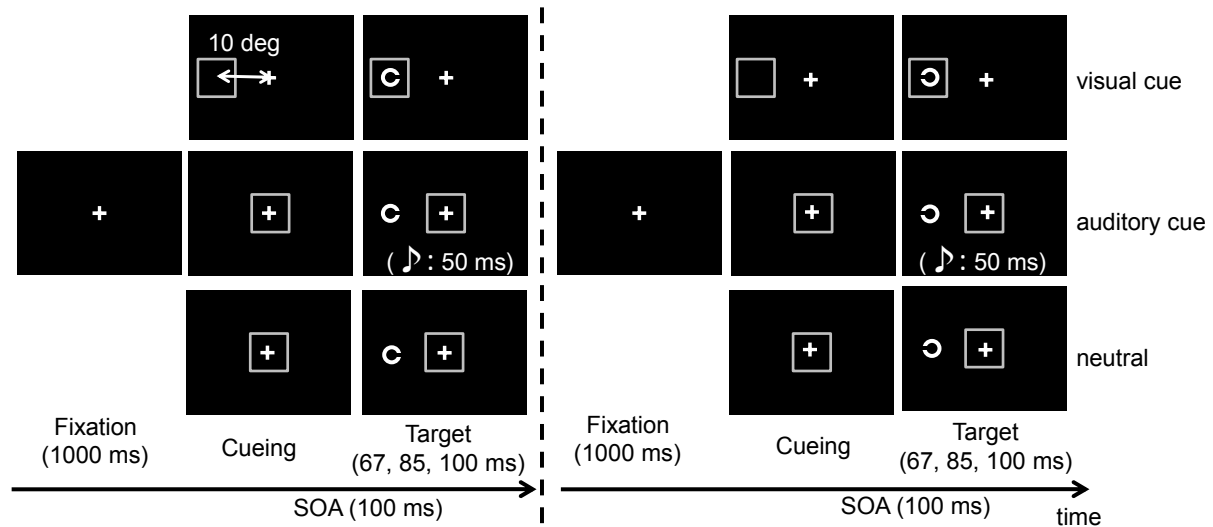


Figure 3. Schematic representation of the procedure in Experiment 2. The left three and right three streams indicate the right-side gap and left-side gap conditions, respectively

3.4 Results and Discussion

Sensitivity (d') and criterion (c) scores for gap location discrimination were computed according to signal detection theory (Yeshurun & Levy, 2003). Additionally, the average RT was computed using correct response RT data. The results are shown in Figure 4. A two-way ANOVA with Cueing (3) \times Duration (3) as factors was conducted for d' scores. The main effect of Cueing was significant ($F(2, 20) = 16.46, p < .001, \eta_p^2 = .62$). Multiple comparisons indicated that sensitivity was higher in the visual cue condition than in the neutral and auditory cue conditions ($ps < .001$). No significant difference in sensitivity was observed between the neutral and auditory cue conditions ($p = .70$). The main effect of Duration was also significant ($F(2, 20) = 4.34, p < .05, \eta_p^2 = .30$). Moreover, the interaction between Condition and Duration was significant ($F(4, 40) = 4.74, p < .005, \eta_p^2 = .32$). The simple main effect of Cueing was significant at the 85 and 100 ms durations ($F(2, 60) = 13.92, p < .001, \eta_p^2 = .32$; $F(2, 60) = 17.01, p < .001, \eta_p^2 = .36$, respectively). Multiple comparisons indicated that performance was higher in the visual cue condition than in the neutral and auditory cue conditions ($ps < .001$). The simple main effect of Duration was also significant in the visual cue condition ($F(2, 60) = 13.24, p < .001, \eta_p^2 = .31$). Multiple comparisons indicated that sensitivity was higher for durations of 85 ms and 100 ms than for the 67 ms duration condition when a visual cue was presented ($ps < .001$). Differences in criterion scores between cueing conditions were small (neutral: $c = 0.08$; auditory cue: $c = 0.16$; visual cue: $c = 0.10$). Finally, a two-way ANOVA with the Cueing (3) \times ISI (3) was conducted for the RT data. The main effect of Cueing ($F(2, 20) = 1.66, p = .22, \eta_p^2 = .14$) and ISI ($F(2, 20) = 0.02, p = .99, \eta_p^2 = .00$) were not significant. On the other hand, the interaction between Cueing and ISI was significant ($F(4, 40) = 4.11, p < .01, \eta_p^2 = .29$). The simple main effect of cueing was significant for the 100 ms duration conditions ($F(2, 60) = 5.53, p < .01, \eta_p^2 = .15$). Multiple comparisons indicated that the RT was longer in the auditory cue condition than in the other two cueing conditions ($ps < .01$).

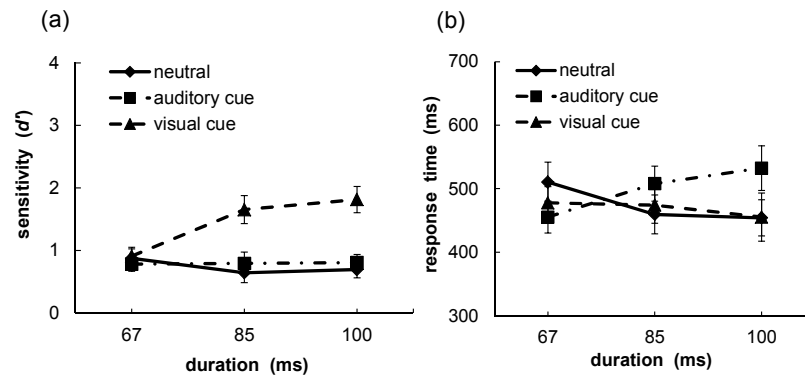


Figure 4. Spatial gap detection sensitivity and response time in Experiment 2. (a) Sensitivity for each duration and Cueing condition. (b) Response time for each duration and Cueing condition. Error bars represent standard errors of the mean ($n = 11$)

The results of Experiment 2 showed that visual cueing increased spatial gap detection sensitivity, consistent with Yeshurun and Levy (2003). In contrast, a simultaneous tone did not affect spatial gap detection sensitivity and RT. In Experiment 1, a reduction in temporal resolution was not observed in the auditory cue condition. Therefore, differences do emerge between auditory and visual cue-driven attention. Visual attention increases spatial resolution but decreased temporal resolution. In contrast, sound-driven attention did not modulate spatial or temporal resolution in the present experiments. Furthermore, response speed decreased in the auditory cue condition. Thus, an auditory tone could interrupt spatial gap detection at a decisional level.

4. General Discussion

The current study examined differences in the mechanisms underlying auditory and visual cues-driven attention. Visual cueing increased spatial gap detection sensitivity and decreased temporal detection sensitivity. In contrast, auditory cueing did not influence the sensitivity to detect either spatial or temporal gaps. Instead, auditory cueing partially increased processing efficiency (i.e., led to faster RTs). Therefore, we observed a clear difference in attention between the visual and auditory cueing conditions.

Facilitation due to visual cueing revealed a trade-off between spatial and temporal resolution. The same trade-off has been found by Yeshurun and Levy (2003). Other studies have also shown that visual attention affects spatial resolution (Gobell & Carrasco, 2005), contrast sensitivity (Carrasco, Ling, & Read, 2004), and perceived size (Anton-Erxleben, Henrich, & Treue, 2007). Visual attention increases perceptual performance via two distinct mechanisms: stimulus enhancement and external-noise exclusion (Doshier & Lu, 2000). Additionally, a valid cue decreases discrimination uncertainty by selecting and restricting the target location and thereby, reducing the decision loads (Davis, Kramer, & Graham, 1983). Yeshurun and Levy (2003) have proposed that facilitation for a visual target can be attributed to activation of a trade-off between the spatial and temporal channels of vision. In other words, visual attention increases activation of the spatial channel and facilitates perceptual performance while decreasing the temporal channel activation. This characteristic of visual attention was replicated in the present study. In addition, visual cueing hardly influenced the RT. In a previous study using a similar paradigm (Yeshurun & Levy, 2003), the facilitation of discrimination speed was not observed. In the present study, while a visual cue controls spatial attention, the task involves temporal or spatial *gap* discrimination of a visual target. Therefore, the controlling of spatial attention by a visual cue would make it hard to affect processing efficiency during a gap detection task. However, further research may be needed to better validate this possibility.

On the other hand, auditory cueing facilitated processing speed in the temporal gap detection task. Other studies have shown that simultaneous sound improves visual processing speed (Molholm, Ritter, Murray, Javitt, Schroeder, & Foxe, 2002; Schroger & Widmann, 1998). Synchronous tones automatically direct attention to an audio-visual event and capitalize on available processing resources (Van der Burg et al., 2008). A simultaneous sound improves visual detection or identification performance (e.g., Bolognini, Frassinetti, Sereno, & Ladavas, 2005; Frassinetti, Bolognini, & Ladavas, 2002; Stein, London, Wilkinson, & Price, 1996); an auditory cue increases not only processing speed but also accuracy. Synchronous sound facilitates visual object representations (Chen & Spence, 2011). Salient audio-visual events can attract attention, which enhances later processing (Busse et al., 2005; Talsma et al., 2007; Van der Burg et al., 2011). It has been proposed that

attentional resources are not shared between visual and auditory sensations (Arrighi, Lunardi, & Burr, 2011). Therefore, auditory cueing increases the efficiency of visual processing by capitalizing on different attentional resources from visual attentional resources.

In the auditory modality, attentional capture is related to activation in a more dorsal network comprising the left precentral gyrus, right superior parietal gyrus, and right intraparietal sulcus (Watkins, Dalton, Lavie, & Rees, 2007). In the visual modality, attentional capture is associated with activation in the left prefrontal and bilateral superior parietal cortices (de Fockert, Rees, Frith, & Lavie, 2004). These brain regions are very close to each other, consistent with a common supra-modal network for stimulus-driven attentional shifts (Downer, Grawley, Mikulis, & Davis, 2002). In visual attention, signal enhancement is related to thalamic activation, and noise exclusion is associated with a network spanning the pulvinar to V4 (Posner & Raichle, 1994). Audio-visual integration correlates with activation in the superior colliculus and superior temporal sulcus (Fairhall & Macaluso, 2009). Therefore, the neural mechanisms of attentional capture are very similar in vision and audition. However, there are different mechanisms underlying how auditory and visual cues guide attention and facilitate visual processing.

In previous studies, it has been shown that a simultaneous auditory stimulus can affect the perception of the number of presented visual stimuli, which is well known as the fission and fusion illusions (e.g., Andersen, Tiippana, & Sams, 2004; Shams, Kamitani, & Shimojo, 2000). In the fusion illusion, only a single flash is perceived even if two flashes are presented when a simultaneous single beep is presented. This phenomenon affects the temporal gap detection sensitivity of visual stimuli in the present study. However, an auditory cue did not affect the temporal gap detection sensitivity of visual stimuli in Experiment 1. Moreover, the fission and fusion illusions occurred when visual stimuli were presented at shorter intervals (e.g., 17 ms). In the present study, the duration of the visual stimuli was longer, for 50 ms. Thus, the effects of the fission and fusion illusions would be weak in this study.

The present study indicated the differences between auditory and visual cue-driven attention. Both auditory and visual cue-driven attention facilitates visual processing in different fashions. Moreover, the underlying mechanisms of the attentional facilitation effects are clearly different. Therefore, the functions of auditory and visual cue-driven attention should be distinguished and discussed.

Acknowledgments

We are most grateful to the participants. This research was supported by a Japan Society for the Promotion of Science KAKENHI (Grant-in-Aid for JSPS Fellows: No. 24-4354).

References

- Andersen, T., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*, 301-308. <http://dx.doi.org/10.1016/j.cogbrainres.2004.06.004>
- Anton-Erxleben, K., Henrich, C., & Treue, S. (2007). Attention changes perceived size of moving visual patterns. *Journal of Vision*, *7*, 1-9. <http://dx.doi.org/10.1167/7.11.5>
- Arrighi, R., Lunardi, R., & Burr, D. (2011). Vision and Audition Do Not Share Attentional Resources in Sustained Tasks. *Frontiers in Psychology*, *2*, 56. <http://dx.doi.org/10.3389/fpsyg.2011.00056>
- Bolognini, N., Frassinetti, F., Serino, A., & Ladavas, E. (2005). "Acoustical vision" of below threshold stimuli: Integration among spatially converging audiovisual inputs. *Experimental Brain Research*, *160*, 273-282. <http://dx.doi.org/10.1007/s00221-004-2005-z>
- Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the of National Academy of Sciences*, *102*, 18751-18756. <http://dx.doi.org/10.1073/pnas.0507704102>
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*, 308-313. <http://dx.doi.org/10.1038/nn1194>
- Carrasco, M., Williams, P. E., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masks: Support for signal enhancement. *Journal of Vision*, *2*, 467-479. <http://dx.doi.org/10.1167/2.6.4>
- Chen, Y.-C., & Spence, C. (2011). The crossmodal facilitation of visual object representations by sound: Evidence from the backward masking paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1784-1802.

- Davis, E. T., Kramer, P., & Graham, N. (1983). Uncertainty about spatial frequency, spatial position, or contrast of visual patterns. *Perception & Psychophysics*, *33*, 20-28. <http://dx.doi.org/10.3758/BF03205862>
- De Fockert, J., Rees, G., Frith, C., & Lavie, N. (2004). Neural correlates of attentional capture in visual search. *Journal of Cognitive Neuroscience*, *16*, 751-759. <http://dx.doi.org/10.1162/089892904970762>
- Doshier, B. A., & Lu, Z.-L. (2000). Noise exclusion in spatial attention. *Psychological Science*, *11*, 139-146. <http://dx.doi.org/10.1111/1467-9280.00229>
- Downer, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2002). A cortical network sensitive to stimulus salience in a neutral behavioral context across multiple sensory modalities. *Journal of Neurophysiology*, *87*, 615-620.
- Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *European Journal of Neuroscience*, *29*, 1247-1257. <http://dx.doi.org/10.1111/j.1460-9568.2009.06688.x>
- Frassinetti, F., Bolognini, N., & Ladavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, *147*, 332-343. <http://dx.doi.org/10.1007/s00221-002-1262-y>
- Gobell, J., & Carrasco, M. (2005). Attention alters the appearance of spatial frequency and gap size. *Psychological Science*, *16*, 644-651. <http://dx.doi.org/10.1111/j.1467-9280.2005.01588.x>
- Keetels, M., & Vroomen, J. (2011). Sound affects the speed of visual processing. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 699-708. <http://dx.doi.org/10.1037/a0020564>
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, *14*, 115-128. [http://dx.doi.org/10.1016/S0926-6410\(02\)00066-6](http://dx.doi.org/10.1016/S0926-6410(02)00066-6)
- Posner, M. I., & Raichle, M. E. (1994). *Images of mind*. New York, USA: Scientific American Library.
- Schroger, E., & Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*, *35*, 755-759. <http://dx.doi.org/10.1111/1469-8986.3560755>
- Senkowski, D., Saint-Amour, D., Gruber, T., & Foxe, J. J. (2008). Look who's talking: The deployment of visuo-spatial attention during multisensory speech processing under noisy environmental conditions. *NeuroImage*, *43*, 379-387. <http://dx.doi.org/10.1016/j.neuroimage.2008.06.046>
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, *408*, 788. <http://dx.doi.org/10.1038/35048669>
- Stein, B. E., London, N., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience*, *8*, 497-506. <http://dx.doi.org/10.1162/jocn.1996.8.6.497>
- Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: Is attending to both modalities a prerequisite for early integration? *Cerebral Cortex*, *17*, 679-690.
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Science*, *14*, 400-410. <http://dx.doi.org/10.1016/j.tics.2010.06.008>
- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1053-1065.
- Van der Burg, E., Talsma, D., Olivers, C. N. L., Hicker, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *NeuroImage*, *55*, 1208-1218. <http://dx.doi.org/10.1016/j.neuroimage.2010.12.068>
- Watkins, S., Dalton, P., Lavie, N., & Rees, G. (2007). Brain mechanisms mediating auditory attentional capture in humans. *Cerebral Cortex*, *17*, 1694-1700. <http://dx.doi.org/10.1093/cercor/bhl080>
- Watkins, S., Shams, L., Josephs, O., & Rees, G. (2007). Activity in human V1 follows multisensory perception. *NeuroImage*, *37*, 572-578. <http://dx.doi.org/10.1016/j.neuroimage.2007.05.027>
- Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage*, *31*, 1247-1256.

<http://dx.doi.org/10.1016/j.neuroimage.2006.01.016>

Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science*, *14*, 225-231. <http://dx.doi.org/10.1111/1467-9280.02436>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).