Pergamon

PII: S0042-6989(96)00313-6

Vision Res., Vol. 37, No. 12, pp. 1575–1580, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Visual Motion Sensation Yielded by Non-visually Driven Attention

SHINSUKE SHIMOJO,*¶ SATORU MIYAUCHI,† OKIHIDE HIKOSAKA‡

Received 1 April 1996; in revised form 25 October 1996

When a visual stimulus (the "cue") is presented and followed by a line, the line is perceived to grow rapidly from the cued side even when it is presented physically simultaneously (the "line-motion effect"). We now report that the same line motion can be observed when the cue is presented in a non-visual modality, such as auditory or somatosensory. A beep sound was presented either from the left or the right speaker as an auditory cue, or an electric pulse was applied to a finger put on the left or the right side of a CRT display as a somatosensory cue. A line probe was then presented between the two possible cue positions. Both the auditory and the somatosensory cues led to line motion, thus the line motion could not be interpreted as a variation of within-modality effects, such as visual apparent motion. When the cue lead time was manipulated, the obtained time courses of the effects were similar across the three cue modalities (Experiment 1). The minor differences could be explained simply in terms of latency of detection, according to results of another experiment (Experiment 2). Finally, the line-motion task was compared with a task of temporal order judgment, where two targets were presented simultaneously at the cued and the uncued sides, and the subject was asked to judge which of the targets had appeared first. As a result, similar dependencies on cue lead time were obtained between the two tasks within subjects (Experiment 3). Thus, the non-visual cue seems to facilitate "prior entry" of a visual stimulus nearby in the spatial representation, much the same way as a visual cue does. These effects should be attributed to modality non-specific spatial attention, i.e., a "gradient" of information processing efficiency across various locations. © 1997 Elsevier Science Ltd.

Attention   Motion   Cross-modal representation   Space perception   Auditory cue
Somatosensory stimulation   Line motion

INTRODUCTION

When a visual stimulus (the "cue") is presented and followed by a line, the line is perceived to grow or elongate from the cued side, even when it is presented physically simultaneously (the "line-motion effect"; Hikosaka *et al.*, 1993a,b,c; the effect is indicated by the arrow in Fig. 1).

Our previous studies indicate that the direction of this illusory motion sensation is the same, regardless of the nature of cue: whether the cue is stimulus-onset or -offset (Hikosaka *et al.*, 1993a). Moreover, the line motion could be induced without a cue, only by voluntary attention (Hikosaka *et al.*, 1993b; also see Schmidt *et al.*, 1997) or anticipation of visual events based on memory without actual visual stimulus (von Gruenau & Faubert, 1992; Shimojo, 1995) at a particular location in the visual field. We argued that the line-motion effect is induced by local facilitation of visual information processing. That is, the cue (or the observer's voluntary effort or anticipation) drives attention, thus locally facilitating visual processing, which results in prior entry for input from the cued side. Here, we use the term "prior entry" to indicate *earlier entry* of input from the cued side, relative to that from the uncued, *into* the mechanism of motion detection. This hypothesis was supported by another set

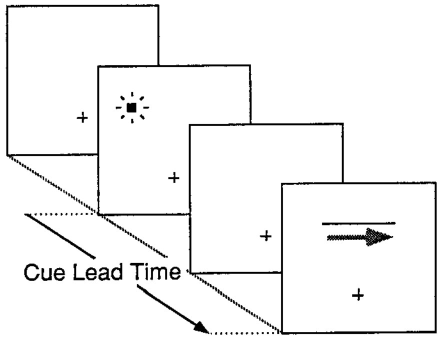

FIGURE 1. Line-motion task applied to visual, auditory, and somatosensory cues. The figure illustrates the stimulus sequence in the visual cue condition as an example.

*Computational Neuval Systems, Division of Biology, California Institute of Technology, Pasadena, CA91125, U.S.A.
†Communication Research Laboratory, Koganei, Tokyo 184, Japan.
‡Department of Physiology, College of Medicine, Jyuntendo University, Ochyanomizu, Tokyo 113, Japan.
§To whom all correspondence should be addressed [*Fax* +1-818-844 4514; *Email* sshimojo@cns.caltech.edu].

of experiments in which we employed the paradigm of temporal order judgment (Hikosaka *et al.*, 1993a). In this experiment, a stimulus on the cued side tended to be judged as prior to the other stimulus on the uncued side. Moreover, this effect of illusory temporal order depended upon the cue lead time in the way expected from the line-motion data and the hypothesis. Thus, the line-motion effect seems to reflect a spatial gradient in efficiency of information processing. It has been widely used as a sensitive psychophysical tool to measure visual attention.

In the example above, a visual stimulus was used to attract attention to its location. Space is super-modal, however; a real object occupies a location in space, and we perceive it by hearing or touching, as well as by seeing. If the object moves, our attention would be attracted to and follow the object to maximize *multi-modal* processing of it, regardless of whether the cue for orientation is given only in one sensory modality or in different modalities altogether. How these sensory modalities prime each other and become integrated in a spatially selective manner is an intriguing question (Alho *et al.*, 1992; Buchtel & Butter, 1988; Farah *et al.*, 1989; Woods *et al.*, 1992; Stein & Meredith, 1993). For example, would an auditory or somatosensory stimulus induce visual attention by local facilitation of visual processing, much the same way as a visual stimulus does? If so, the same illusory visual motion should be perceived by presenting the non-visual cues.

Another reason for employing non-visual cues was to test the possibility of line motion as a purely visual artifact, totally unrelated to attention. According to Downing & Treisman (1995, 1997), effects of spatial–temporal parameters on the line motion were so similar to effects of those on classical apparent motion that the line motion, in particular the stimulus-driven version of it, should be interpreted as a variation of apparent motion. If they are correct, however, then an auditory or a somatosensory cue could not trigger the line-motion effect.

The first experiment tests these predictions.

EXPERIMENT 1: THE LINE-MOTION EFFECT INDUCED BY NON-VISUAL CUES

Method

Subjects. Six subjects (three naïve, three non-naïve) participated in the experiment with visual cue, six subjects (three naïve, three non-naïve) in the experiment with auditory cue, and five subjects (two naïve, three non-naïve) in the experiment with somatosensory cue. Four of the five subjects participated also in the cross-arm condition of the somatosensory cue experiment.

Materials and procedures. The experiment consisted of three sessions, each of which employed a cue stimulus in each of three cue modalities: visual, auditory and somatosensory. The visual cue condition is illustrated as an example in Fig. 1. A visual fixation point was presented on the display first. It was then followed by a brief cue stimulus on the left or the right, which was presented in one of the three modalities. Finally, a line was visually presented physically simultaneously. Its length and location were such that its terminators spatially overlapped the two possible locations of visual cue. The subject's task was to judge from which side the line appeared to grow, and push one of the two mouse buttons accordingly (a two-alternative, forced-choice judgment). No specific instruction was given as for attention: they were asked just to observe the stimuli passively.

Duration of the fixation and that of the line were 1000 and 500 msec, respectively. The cue lead time varied in 13 steps, from −534 to 534 msec (15 steps from −204 to 1020 msec in the case of somatosensory cue), and randomized across trials. A negative value of cue lead time indicates that the line was presented prior to the cue. The position of the cue (left/right) was pseudo-randomized so that the cue stimulus was presented on the same side in no more than three successive trials.

The visual fixation was a cross of 24 × 30 min (luminance: 7.9 cd/m^2), and the line extended 13.5 deg × 6 min(7.9 cd/m^2). The luminance of background was 0.1 cd/m^2. The distance from the fixation point to the visual cue was 18.5 deg, that to the auditory cue (speaker) was 25 deg, and that to the somatosensory cue (the electrode attached to the index finger) was 20 deg. All the visual stimuli were presented on a CRT display (Commodore 1950-B) controlled by a personal computer (Commodore AMIGA 3000 in the visual auditory cue conditions; Mitsubishi XC1498 in the somatosensory condition).

The visual cue was rectangular (21 × 30 min at the observation distance of 57 cm, 34 cd/m^2), whose duration was approximately 17 msec. The auditory cue was a burst sound generated by a computer and presented through one of the two speakers which were located on the left and the right sides of the CRT. The waveform of the sound was a sine wave in an amplitude envelope whose duration was approximately 17 msec. The peak frequency of the spectrum varied from 100 to 1000 Hz, and was randomized across trials, a procedure to avoid habituation. The somatosensory cue was a single electric pulse of 1 msec duration, which was generated by a physiological electric stimulus generator (Nihon Koden SEN-7103) and applied to the subject's index finger. For this, electrodes were attached to the subject's index fingers of both hands, and the subject positioned the fingers on the left and the right edges of the CRT. In some sessions, subjects were asked to cross their arms so that the left finger was positioned on the right edge of the CRT, and *vice versa*. (We plan to publish this pert of the experiment elsewhere, so will not describe further details in the present paper.) In a preliminary session, the somatosensory threshold of detection was first measured in each hand of each subject, and then the voltage was doubled for each. Finally, a minor readjustment was made between the two according to the subject's verbal report so that the subjective strength of the stimulus was equal between the hands. The voltages obtained through this procedure were then employed for the main experimental session. The voltages which were actually employed were in a range of 40–80 V.

Thirty trials were conducted in the cases of visual and auditory cues, whereas 20 were conducted in the case of the somatosensory cue, for each cue lead time in each individual subject. Thus, altogether, 2 (positions) × 13 (cue lead times) × 15 = 390 trials were conducted each for the visual and the auditory cue conditions, whereas 2 × 15 × 10 = 300 trials were conducted for the somatosensory cue condition.

Results and discussion

Results are shown in Fig. 2, where the proportion of trials in which line motion was perceived away from the cued side was plotted as a function of cue lead time for each of the three cue modality conditions. Each curve is for each individual subject. See results in the visual cue condition first [Fig. 2(a)]. The strongest effect of line motion was obtained at cue lead times of 0–300 msec. This essentially duplicated our previous data (Hikosaka *et al.*, 1993a,b). In several subjects, there were effects in the opposite direction (line unfolding from the uncued side) at small values of negative cue lead times, ranging from −150 to −17 msec. This may be attributed to backward masking; that is, visibility of one end of the line might have been reduced by the cue which was presented later.

Similar results were obtained in the auditory cue [Fig. 2(b)] and in the somatosensory cue [Fig. 2(c)] conditions. Again, the strongest effect of line motion was obtained at cue lead times between 0 and 300 msec. The results were statistically significant within this range of cue lead time in virtually all subjects in all the three cue modality conditions ($P < 0.01$ by chi-square test). There was no systematic difference between the data obtained from the naïve subjects and those from non-naïve ones. Thus, auditory and somatosensory cues, as well as a visual cue, can give rise to the visual line motion. These results suggest that the perceptual system performed modulation while preserving the spatial coherence between the visual and non-visual signals (the results in cross-arm condition were also consistent with a multimodal representation of environmental space; Miyauchi *et al.*, 1993).

One could still argue that the results could be explained by so-called Type II errors, or a cognitive bias. That is, the percentage of illusory line motion from the cued side might have been grossly inflated because the subjects could rely on the cue location to make the line-motion judgment, even when they were completely unsure about the effect. We do think this is highly unlikely, however, considering the following two facts:

1. The subjects indeed did not always judge that the line appeared from the cued side. Instead, they changed the percentage of the illusion *system-*

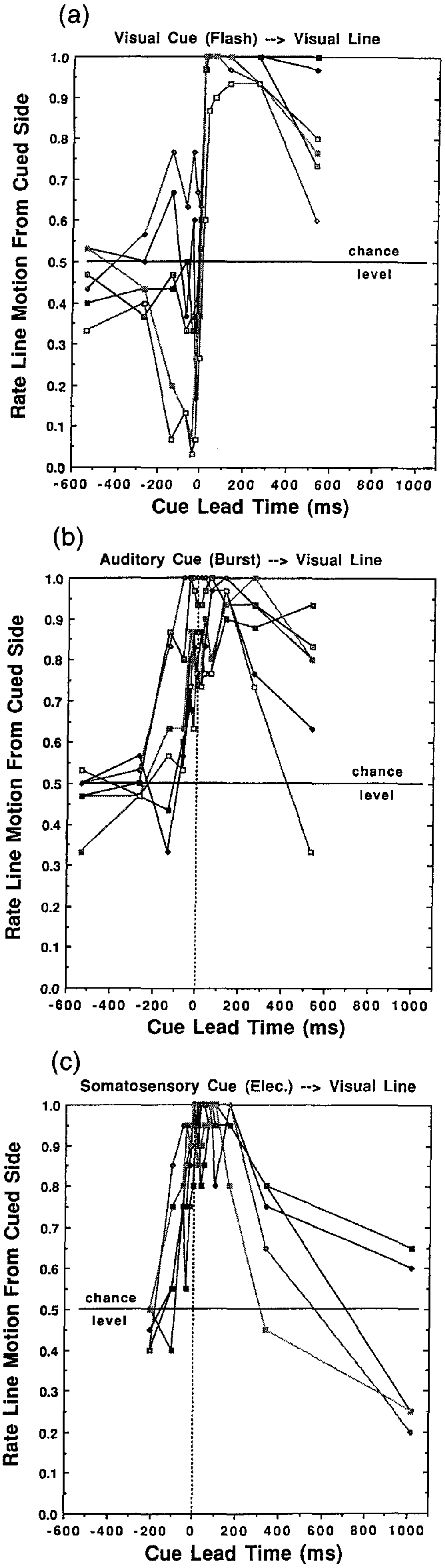

FIGURE 2. Effects of visual and the non-visual stimuli on illusory line-motion sensation. Rate of line motion perceived from the cue side is plotted as a function of cue lead time. (a) Group results in the visual cue condition. Each curve corresponds to each subject. (b) Group results in the auditory cue condition. (c) Group results in the somatosensory cue condition.

atically across cue lead time, and it was consistent across modalities (Fig. 2). Moreover, the naïve subjects are not different from the non-naïve ones in this regard. These are difficult to explain solely by a cognitive bias.

2. To examine this possibility more closely, we conducted an additional experiment. This experiment was basically a duplication of the auditory cue experiment, except that we randomly mixed 90% of experimental trials with 10% of catch trials, where the line was physically drawn *against* the illusory line motion. The physical time delay between the two ends of line was carefully adjusted (60–80 msec) such that it would overcome the illusory effect and clearly appear as growing *towards* the cue, yet not very noticeably unless the subjects pay attention to the direction of motion perception. The line-motion results obtained from the experimental trials in two subjects (one non-naïve, one new naïve) were virtually the same as those in the main experiment. Further, the percentage of correct judgment in the catch trials ("towards the cued side") was over 95% in both of them, suggesting that the subjects in fact had paid attention to the direction of line motion and responded accordingly in each trial. This is quite the opposite to what would be expected from the "cognitive bias" account.

There still might have been a minor bias caused by a cognitive bias in the data of the main experiments, but it could hardly explain their overall tendency.

There was yet another unexpected finding in the main results. The line-motion effect produced by auditory and somatosensory cues had a faster onset, already reaching statistically significant levels at small negative values of cue lead time; the line-motion sensation was produced even when the non-visual cues appeared after the line [compare Fig. 2(a, b and c)]. Within-subject comparisons of the three cue-modality conditions confirmed this qualitatively, although there were some quantitative variations across subjects.

A simple explanation would be that the non-visual signals reach the central nervous system, particularly the area responsible for spatial attention, faster than the visual signals. If so, the reaction time to detected auditory and somatosensory stimuli should also be faster than that to detect a visual stimulus, as suggested in the literature (Luce, 1986; Farah *et al.*, 1989; Pascual-Leone *et al.*, 1992; Stein & Meredith, 1993). To see if this is the case with the type of stimuli that we used in Experiment 1, we conducted the second experiment.

EXPERIMENT 2: REACTION TIMES TO VISUAL, AUDITORY AND SOMATOSENSORY STIMULI

Method

Subjects. Four subjects (one naïve, three non-naïve) participated in the experiment.

Materials and procedures. The stimulus parameters were the same as those employed in Experiment 1, except that this time there was no line probe presented. Thus, in each trial a visual fixation was presented first, then followed by a stimulus (the "cue" in the previous experiment) either on the left or the right side of the screen in one of the three modalities. The subject's task was to respond by a mouse button as soon as he/she detected the cue, regardless of whether it was on the left or the right. One hundred trials were conducted for each of the three cue modalities in each subject, while the position of the cue stimulus was pseudo-randomized across trials. Mean reaction time (RT) was then calculated for each stimulus modality for each subject.

Results and discussion

The results were highly consistent across subjects, as shown in Table 1. The mean RTs for the auditory stimulus were the shortest, ranging from 146 to 185 msec; those for the somatosensory were the next, ranging from 186 to 211 msec; and those for the visual stimulus were the longest, ranging from 216 to 272 msec. The difference between the auditory and the visual conditions ranged from 46 to 87 msec, while the difference between the somatosensory and the visual conditions ranged from 30 to 61 msec across the subjects.

The results are, in general, consistent with the literature, where the RT delay in visual detection has been reported in the order of 40–60 msec when compared with auditory, and somewhat less when compared with somatosensory detection (Luce, 1986; Farah *et al.*, 1989; Pascual-Leone *et al.*, 1992; Stein & Meredith, 1993). Although one would expect the exact values of cross-modal RT difference to vary depending on the stimulus intensities and conditions, our pilot study and the literature suggest that the direction and amount of difference would not change drastically, as long as the intensity of each stimulus is well beyond the detection threshold.

The RT results were also highly consistent with the onset order of line-motion effect found in the first experiment. To examine it more closely, we defined the "onset time" of line motion as 75% cutting point in each cue condition for each subject in the results shown in Fig. 2(a–c). We then obtained rank orders across cue conditions and subjects in the onset time of line motion,

TABLE 1. Results of the detection RT experiments

Subjects	Target's modality		
	Visual	Auditory	Somatosensory
SS	216	146	186
	(3.8)	(3.5)	(5.0)
OH	226	180	193
	(3.6)	(4.0)	(4.2)
SM	233	166	202
	(3.5)	(2.3)	(4.7)
HO	272	185	211
	(3.8)	(4.3)	(5.1)

The mean RTs for visual, auditory and somatosensory targets were shown for each of the four subjects. Values in parentheses are standard errors.

as well as in the RT data (Table 1). If our interpretation is correct, then the correlation between these two rank orders should be highly positive. Spearman's r_{rank} which was calculated from the actual data sets was 0.81 (and when one extreme deviant, i.e. subject HO's onset time of line motion, was eliminated from the data, r_{rank} was 0.95). Thus, the non-visual cues seem to have the same facilitatory effect as the visual cue in the early visual pathway, yielding the same motion sensation. Yet, the auditory or the somatosensory signal reaches the locus of visual facilitation faster than the visual cue, and sometimes even faster than the input from the line probe when the cue lead time is negative but small.

With all these consistent results, it has not been examined directly whether the "hypothesis of the local facilitation" which we have proposed to account for the line motion yielded by visual cues (Hikosaka *et al.*, 1993b) could apply to the cases with non-visual cue as well. To test this, we conducted the third experiment. Here, we employed the visual temporal-order task with the auditory and the somatosensory cues, making comparisons between this and the line-motion tasks within each cue modality. We wondered whether the time profiles of the effect developing and decaying as a function of cue lead time would be similar between the two tasks, even when the cue modality is not visual.

EXPERIMENT 3: TEMPORAL ORDER JUDGMENT

Method

Subjects. Five subjects (two naïve, three non-naïve) participated in the experiment.

Materials and procedures. The stimulus configuration, parameters and the design of the experiment were identical to those in the first experiment, except that instead of a visual line probe, a pair of dots was presented simultaneously, one on the left and the other on the right side of the screen. Thus, there was a fixation point at the beginning of each trial, followed by a cue in one of the three modalities, and further followed by a pair of dots. Since our interest here was a direct comparison between the temporal-order and the line-motion tasks, the two dots were always presented simultaneously in order to be more comparable with the line probe in the other task. The subject's task was to judge which of these two dots appeared first, and to press one of the two mouse buttons accordingly.

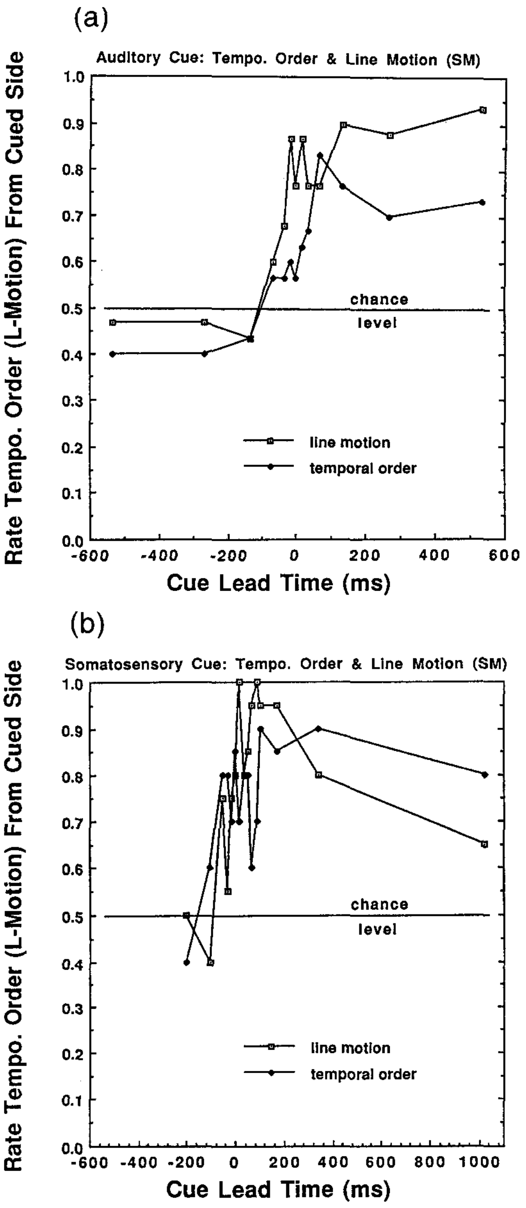

FIGURE 3. Effects of visual and non-visual stimuli on temporal order judgment. The procedure and stimuli of this experiment were similar to those in the first experiment, except that instead of a line, a pair of dot probes (12*30 min of size) was presented after the cue, and the subject had to judge which of them appeared first (two-alternative, forced-choice). Examples of within-subject comparison between the temporal order and the line motion tasks are shown. Rate of the stimulus on the cued side judged as prior is plotted against cue lead time (msec), together with the result in the line-motion experiment. (a) Comparison between the two tasks in the auditory cue condition. (b) Comparison in the somatosensory cue condition.

Results and discussion

Some examples of the results are shown in Fig. 3, where the results in the two tasks with either auditory cue [Fig. 3(a)] or somatosensory cue [Fig. 3(b)] are compared within subjects. As is obvious from the figures, the functions obtained for the line-motion and the temporal-order tasks were highly comparable, thus compatible with our hypothesis that a cue in various cue modalities could give rise to the same effect of local facilitation in a relatively early stage of visual pathway. The local facilitation would make the target at the cued location appear earlier in the temporal-order task. By the same token, it results in the illusory line motion away from the cued side, owing to the "prior entry" at the cued end.

If there were any differences, the results in the line-motion task seemed to be slightly more sensitive and robust than those in the temporal-order task, a difference which was more obvious in subject SM. The remarkable sensitivity of line motion as a tool to detect the spatial gradient of attention has been emphasized in our previous studies as well (Hikosaka *et al.*, 1993a,b,c).

GENERAL DISCUSSION

To summarize, we found that the visual line-motion effect could be induced by auditory and somatosensory cues, as well as a visual cue.

There are other lines of evidence that the kind of attention gradient on which the line motion is based is not restricted to local visual cueing. For example, the line-motion paradigm can reveal local facilitation which is induced by purely top-down voluntary attention (Hikosaka *et al.*, 1993a), or by "odd ball pop-out" (Shimojo *et al.*, 1992). Also, the peak of efficiency gradient, from which the line appears to grow, can move in the visual field by object-bound attention, not restricted to a retinotopic location of a transient visual event (Hikosaka *et al.*, 1993a; Kanwisher & Driver, 1992). Taken altogether, the line motion could not be attributed to a purely visual artifact, but rather reflects a more central and modality-non-specific gradient of spatial attention, i.e., a spatial gradient in processing efficiency of sensory signals. This is also consistent with the reaction time studies which have indicated that attention-related effects could be found in auditory, as well as visual modalities (Luce, 1986; Spence & Driver, 1994, 1996).

Even though the attentional modulation can be induced by non-visual cues, the facilitatory effect itself seems to be executed in the vision-specific pathway. To be more specific, the current results support the hypothesis that stimulus-driven attention locally facilitates visual information processing prior to, or at the same level as, the neural locus for motion perception. Closely related to this interpretation is the recent finding which underscores the role of attention on the ambiguity solving process for motion perception (Cavanagh, 1992).

REFERENCES

Alho, K., Woods, D. L., Algazi, A. & Naatanen, R. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography & Clinical Neurophysiology, 82*, 356–368.

Buchtel, H. A. & Butter, C. M. (1988). Spatial attentional shift: implications for the role of polysensory mechanisms. *Neuropsychologia, 26*, 499–509.

Cavanagh, P. (1992). Attention-based motion perception. *Science, 257*, 1563–1565.

Downing, P. & Treisman, A. (1995). The shooting line-motion: attention or apparent motion? *Investigative Ophthalmology & Visual Science, 36, Annual Meeting Abstract Issue, 4*, 856.

Downing, P. & Treisman, A. (1997). The line-motion illusion: attention or impletion? *Journal of Experimental Psychology: Human Perception and Performance*, in press.

Farah, M. J., Wong, A. B., Monheit, M. A. & Morrow, L. A. (1989). Parietal lobe mechanisms of spatial attention: modality specific or supramodal. *Neuropsychologia, 27*, 461–470.

Hikosaka, O., Miyauchi, S. & Shimojo, S. (1993a) Voluntary and stimulus-induced attention detected as motion sensation. *Perception, 22*, 517–526.

Hikosaka, O., Miyauchi, S. & Shimojo, S. (1993b) Focal visual attention produces illusory temporal order and motion sensation. *Vision Research, 33*, 1219–1240.

Hikosaka, O., Miyauchi, S. & Shimojo, S. (1993c) Visual attention revealed by an illusion of motion. *Neuroscience Research, 18*, 11–18.

Kanwisher, N. & Driver, J. (1992). Objects, attributes, and visual attention: which, what and where. *Current Directions in Psychological Science, 1*, 26–31.

Luce, D. R. (1986). *Response time: their role in inferring elementary mental organization*. New York: Oxford University Press.

Miyauchi, S., Hikosaka, O., Shimojo, S. & Okamura, H. (1993). Spatial attention is crossmodal – an evoked potential study. *Investigative Ophthalmology & Visual Science, 34, Annual Meeting Abstract Issue, 4*, 1234.

Pascual-Leone, A., Brasil-Neto, J. P., Valls-Sole, J., Cohen, L. G. & Hallet, M. (1992). Simple reaction time to focal transcranial magnetic stimulation. *Brain, 115*, 109–122.

Schmidt, W. C., Fisher, B. D. & Pylyshyn, Z. W. Multiple onset stimuli elicit illusory line motion. *Journal of Experimental Psychology: Human Perception and Performance*, in press.

Shimojo, S. (1995). Line-motion effect yielded by anticipation based on associative memory. A paper presented at the Annual Meeting of the Japanese Psychological Association (Okinawa).

Shimojo, S., Miyauchi, S. & Hikosaka, O. (1992). Voluntary and involuntary attention detected by the line-motion effect. *Perception, 212*, 12.

Spence, C. J. & Driver, J. (1994). Covert spatial orienting in audition: exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 555–574.

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

Stein, B. E. & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.

von Gruenau, M. W. & Faubert, J. (1992). Interactive effects in motion induction. *Perception, 212*, 12.

Woods, D. L., Alho, K. & Algazi, A. (1992). Intermodal selective attention. I. Effects on event-related potentials to lateralized f attentional auditory and visual stimuli. *Electroencephalography & Clinical Neurophysiology, 82*, 341–355.

Acknowledgement—Supported in part by a Grant-in-Aid for Scientific Research on Priority Areas, from MESC, Japan.