Object-based spatial attention when objects have sufficient depth cues

Ryuji Takeya
Graduate School of Medicine, Hokkaido University, Sapporo, Japan

Tetsuko Kasai
Faculty of Education, Hokkaido University, Sapporo, Japan

Attention directed to a part of an object tends to obligatorily spread over all of the spatial regions that belong to the object, which may be critical for rapid object-recognition in cluttered visual scenes. Previous studies have generally used simple rectangles as objects and have shown that attention spreading is reflected by amplitude modulation in the posterior N1 component (150–200 ms poststimulus) of event-related potentials, while other interpretations (i.e., rectangular holes) may arise implicitly in early visual processing stages. By using modified Kanizsa-type stimuli that provided less ambiguity of depth ordering, the present study examined early event-related potential spatial-attention effects for connected and separated objects, both of which were perceived in front of (Experiment 1) and in back of (Experiment 2) the surroundings. Typically, the P1 (100–140 ms) and N1 (150–220 ms) attention effects of ERP in response to unilateral probes were observed in both experiments. Importantly, the P1 attention effect was decreased for connected objects compared to separated objects only in Experiment 1, and the typical object-based modulations of N1 were not observed in either experiment. These results suggest that spatial attention spreads over a figural object at earlier stages of processing than previously indicated, in three-dimensional visual scenes with multiple depth cues.

Introduction

In everyday life, we easily interact with objects in three-dimensional (3-D) space, although retinal images are two-dimensional (2-D) and fragmented. This may partly be achieved through object-based attention: Attention selects the products of perceptual organization processes as the most fundamental units. Object-based attention has generally been evidenced by using simple geometric shapes on a 2-D display. For example, better behavioral performance is achieved when participants divide their attention within a rectangle or grouped lines, rather than between two rectangles (e.g., Duncan, 1984; Marino & Scholl, 2005). Furthermore, when a participant pays attention to part of a rectangle, the detection of stimuli that appear at other parts of the rectangle is facilitated (Egly, Driver, & Rafal, 1994). Such object-based effects may occur because attention tends to involuntarily spread over the entire spatial region and over features within a group or an object (Richard, Lee, & Vecera, 2008; Wegener, Galashan, Aurich, & Kreiter, 2014; Zhao, Kong, & Wang, 2013). On the other hand, when the regions of a rectangle have binocular disparities and are perceived in back of the surroundings, like “holes,” object-based effects have not been observed (Albrecht, List, & Robertson, 2008). This suggests that objects selected by attention are structures with depth in 3-D space, rather than 2-D shapes.

Since selective attention involves multiple stages of processing (e.g., Luck & Hillyard, 1999), it may be important to clarify the processing stages or the timing of object-based attention. Previous studies using event-related potentials (ERPs) have provided insights regarding this issue. Through the use of two superimposed surfaces with moving random dots, it has been shown that the P1 (90–120 ms poststimulus) and N1 (160–240 ms) components were greater in response to changes in the direction of motion at attended surfaces compared to those at unattended surfaces (Valdes-Sosa, Bobes, Rodrigue, & Pinilla, 1998). Moreover, when a surface was cued exogenously, an earlier component (C1, 75–110 ms) was also modulated (Khoe, Mitchell, Reynolds, & Hillyard, 2005), which may indicate that space-invariant object-based attention can modulate very early visual processing. On the other hand, ERP studies of spatial attention (i.e., in the left or right visual field) with simple geometric shapes.
have repeatedly shown that the N1 component is object-based, both for probes that appeared on sustained objects (He, Fan, Zhou, & Chen, 2004; Martinez, Ramanathan, Foxe, Javitt, & Hillyard, 2007; Martinez, Teder-Salejarui, & Hillyard, 2007; Martinez et al., 2006) and for the onset of objects themselves (Kasai, 2010; Kasai, Moriya, & Hirano, 2011; Kasai & Takeya, 2012; Takeya & Kasai, 2014). Since N1 modulation is a typical spatial attention effect of ERP, the previous studies suggest that object-based attention shares a single common mechanism with spatial attention. The object-based N1 attention effects were estimated to originate at the lateral occipital cortex (LOC), which contributes to object or figure perception (Flevaris, Martinez, & Hillyard, 2013; Martinez, Ramanathan, et al., 2007; Martinez, Teder-Salejarui, & Hillyard, 2007; Martinez et al., 2006). However, it remains unclear why earlier stages of processing are not involved in the case of spatial selection for objects.

While previous ERP studies have indicated that P1 and N1 spatial attention effects reflect gain-control mechanisms of sensory inputs (e.g., Hillyard, Vogel, & Luck, 1998), some functional differences between them have been addressed (for a review, Luck & Kappenman, 2012). However, less attention may have been paid to the nature and processes that construct perceptual 3-D objects from fragmented 2-D visual images. The present study focused on this issue. Although two overlapping surfaces are generally used to examine space-invariant object attention, they may be perceived to be separated in the depth direction. In contrast, previous studies of object-based spatial attention generally used simple rectangles or Kanizsa-like figures with subjective contours. Although the stimuli were defined to be “objects” by researchers, they may not necessarily have had sufficient depth cues that define figural objects. In particular, as shown in Figure 1, rectangles in a 2-D display may be objects according to closure or size cues, while they can also be interpreted to be mere 2-D patterns or holes, and Kanizsa stimuli may be interpreted as Pac-men rather than a completed figure. Thus, even if a spatial region is subjectively perceived as a figure or an object, other interpretations may arise implicitly in early stages of cortical processing (for a review, see Long & Toppino, 2004).

To clarify depth ordering of objects and the surroundings, the present study used modified Kanizsa stimuli, which were combinations of rectangles and Pac-men: Although they were on a 2-D display, their relative depths were made relatively explicit by occlusion cues (Figure 2). The objects in Experiment 1 were perceived to be located in front of the surroundings; those in Experiment 2 had similar stimulus configurations but were perceived to be located in back of the surroundings. A typical ERP paradigm of sustained spatial attention was used. When attending to the left or right sides of stimuli/objects that were continuously presented throughout a block, participants detected infrequent target probes at the attended sides. If attention spreads over objects, the typical ERP attention effects (i.e., ERPs in response to frequent standard probes at attended locations have greater amplitudes than those in response to such probes at unattended locations) should be decreased when bilateral stimuli are connected or form an “object.”

### Experiment 1

The purpose of this experiment was to examine ERP spatial attention-spreading effects for objects that were perceived to be “figures” in front of the background with less ambiguity than prior studies (Figure 2). If spatial object selection occurs regardless of the clarity of 3-D interpretation, N1 attention effects should be decreased in the connected condition compared to the
We also tested whether attention spreads homogeneously over the whole object by comparing the effects of connectedness on two unattended locations (central, peripheral). However, ERPs in response to center probes did not cause any differences regarding connectedness, which was very likely because spatial attention has an insignificant effect on early cortical responses to visual stimuli at foveal location (Handy & Khoe, 2005). Thus, data regarding central probes are not reported.

**Method**

**Participants**

Twelve volunteers (six women), aged 19 to 24 years, participated in this study. All had normal or corrected-to-normal vision and were right-handed. Written informed consent was obtained from all participants after the nature of the study was fully explained. The experiment was conducted following the guidelines laid down in the Helsinki Declaration and was approved by the local ethics committee.

**Stimuli**

Stimuli were displayed on a Hitachi CRT monitor (Hitachi, Tokyo, Japan), at a viewing distance of 70 cm, and controlled by E-Prime version 1.0 (Psychology Software Tools, Inc., Sharpsburg, PA). A central fixation cross that extended across a visual angle of 0.4° × 0.4° was presented on a gray background throughout the experiment. The display included a single rounded rectangle (9.4° × 1.1°) at 1.3° above the fixation point (to the inner edge) in the connected condition, or three short rectangles with a spacing of 1.0° in the separated condition (Figure 2). The rectangles were white in color and partly occluded three black circles (2.4° × 2.4°), which were arranged horizontally at a spacing of 1.6°. Probe stimuli showed indentations in the rectangles and extended 0.9° horizontally and 0.2° vertically for standards and 0.1° for targets. They appeared at the upper and lower edges together and either at the left side, right side (3.2° to the inner edges), or center of the rectangles.

**Procedure**

During a run, the rectangle(s) and occluded circles were presented continuously and the probes were presented briefly (150 ms) 40 times at intervals of 400–600 ms (five steps, rectangular distribution). The standard probes were presented frequently (left, \( p = 0.33 \); right, \( p = 0.33 \); center, \( p = 0.17 \)). The target probes were presented infrequently on either the left or right side (\( p = 0.08 \) each). Each participant was seated in a reclining chair in a sound- and electric-shielded room and instructed to attend to either the left or right side of the rectangles during the runs, and to press a button with the right thumb in response to a target presented in the attended location as accurately and quickly as possible. The visual fields to be attended were indicated by arrows that were presented at the beginning of runs. Attended sides in a run were changed alternately, and connected and separated conditions were changed every two runs, in the order ABBA. Overall, each subject participated in 96 runs, and the attended side and the connectedness of the rectangles in the first run were counterbalanced across the participants.

**Recording and analysis**

Behavioral performance was measured, including the percentage of correct target detections (hits) and RTs for hits. Responses were scored as correct if they occurred within 300–1200 ms after a target was presented in the attended location. Responses to other stimuli were classified as false alarms (FAs). The hit rate and response time (RT) were subjected to repeated-measures analysis of variance (ANOVA) with factors of stimulus visual field (VF: left, right) and connectedness (connected, separated). FAs were ana-
Experiment 2

eliminate epochs that were contaminated above 75
baseline. Automatic artifact rejection was applied to
correcting for differences in the 200-ms prestimulus
stimulus and ending 800 ms poststimulus, while
1000 ms, starting 200 ms before the onset of the
location, and object context. Averaging epochs were
separately for each stimulus location, to-be-attended

was kept below 10 k Ohm. EEGs were filtered with a
eye (vertical EOG). The impedance of the electrodes
electrooculogram [EOG]) and Fp1 and below the left
electrodes at the outer canthi of the eyes (horizontal
and horizontal eye movements were monitored with
20 System), which were referenced to the nose. Blinks
POz, PO4, and PO8 according to the International 10–
C4, T4, T5, P3, Pz, P4, T6, O1, Oz, O2, PO7, PO3,
electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz,
C4, T4, T5, P3, Pz, P4, T6, O1, Oz, O2, PO7, PO3,
POz, PO4, and PO8 according to the International 10–
20 System), which were referenced to the nose. Blinks
and horizontal eye movements were monitored with
electrodes at the outer canthi of the eyes (horizontal
electrooculogram [EOG]) and Fp1 and below the left
eye (vertical EOG). The impedance of the electrodes
was kept below 10 k Ohm. EEGs were filtered with a
bandpass of 0.1–30 Hz and sampled at 500 Hz.

ERPs in response to standards were averaged
separately for each stimulus location, to-be-attended
location, and object context. Averaging epochs were
1000 ms, starting 200 ms before the onset of the
stimulus and ending 800 ms poststimulus, while
correcting for differences in the 200-ms prestimulus
baseline. Automatic artifact rejection was applied to
eliminate epochs that were contaminated above 75 μV,
and those epochs or the next epochs with either
responses or target stimuli were excluded.

The P1 (100–140 ms poststimulus) and N1 (150–220
ms) components of the visual ERP were quantified in
terms of mean amplitudes averaged over the same
cluster of five posterior electrode sites in each
hemisphere (O1/O2, PO3/PO4, PO7/PO8, P3/P4, T5/
T6). The measurements were subjected to repeated-
measures ANOVA: The factors considered were
stimulus VF (left, right), attention to stimulus location
(attended, unattended), connectedness (connected,
separated), and hemisphere (left, right).

### Results

#### Behavioral data

Table 1 summarizes the behavioral data. Hit rates
for the separated condition were higher than those for
the connected condition, which was reflected by a main
effect of connectedness, $F(1, 11) = 4.9, p < 0.05$. Effects
that involved connectedness were not statistically
significant with respect to RTs ($ps > 0.57$) or FA rates
($ps > 0.17$).

#### ERP data

In grand-averaged ERPs in response to standards,
spatial attention effects were observed as differences in
amplitude between ERPs for attended and unattended
stimuli, which were distributed most prominently at the
occipital temporal sites (Figure 3). Table 2 summarizes
the $p$ values of omnibus ANOVAs for the clustered
ERPs.

As a typical effect of unilateral probes, P1 (100–140
ms) at sites contralateral to stimuli had greater
amplitudes, which was indicated by the significant
interaction of VF × Hemisphere, $F(1, 11) = 11.1, p =
0.01$. P1 also had a greater amplitude for attended
versus unattended probes, which was reflected by the
main effect of attention, $F(1, 11) = 9.1, p = 0.01$.

Importantly, the P1 attention effect was modulated
by connectedness, as shown in Figure 4. There was
significant interaction among attention, connectedness,
and hemisphere in the P1 latency range, $F(1, 11) = 4.9,
p = 0.05$. To simplify further analyses, the attended
minus unattended subtraction ERPs (attention effects)
were calculated and tested. The interaction of
connectedness and hemisphere was significant, $F(1, 11) =
4.9, p = 0.05$. In posthoc tests, the attention effects in
the separated condition were greater than those in the
connected condition at both hemispheres: left hemi-
sphere, $F(1, 11) = 8.0, p = 0.02$, right hemisphere, $F(1,
11) = 6.3, p = 0.03$. The attenuation of the attention
effects by connectedness were greater at the left
hemisphere compared to those at the right hemisphere,
which was shown by further tests for double subtrac-
tions (the attention effects in the connected condition
minus from those in the separated condition), $F(1,
11) = 4.9, p = 0.05$.

N1 also had a greater amplitude for attended versus
unattended probes, as reflected by the main effect of
attention, $F(1, 11) = 15.3, p = 0.002$, whereas effects that
involved Attention × Connectedness were not signifi-
cant ($ps > 0.12$).
Discussion

In this experiment, we used objects that were more explicitly located in front of the surroundings than prior studies, and observed typical spatial attention effects of the P1 and N1 components: Amplitudes were enlarged in response to probes at attended visual fields. However, we did not observe object-based N1 modulation, which has robustly been found in previous ERP studies (He et al., 2004; Kasai, 2010; Kasai et al., 2011; Kasai & Takeya, 2012; Martinez, Ramanathan et al., 2007; Martinez, Teder-Salejarui et al., 2007; Martinez et al., 2006; Takeya & Kasai, 2014). Instead, we observed an object-based spatial attention effect of P1: The P1 spatial attention effects were decreased when probes appeared at connected objects compared to when they appeared at separated objects. The result suggests that object-based spatial attention spreading occurred at an earlier stage of visual processing than previously indicated.

We also observed worse behavioral performance for connected objects as an object-based attention effect, which is consistent with the findings in behavioral flanker tasks: Due to attention spreading, features at task-irrelevant locations interfere with the discrimination of targets at task-relevant locations (e.g., Richard et al., 2008; Zhao et al., 2013). However, such a behavioral object-based effect was not observed in a previous ERP study that used a similar spatial-attention paradigm with stable object(s) and unilateral probes (Martinez, Ramanathan et al., 2007). These findings suggest that the objects used in the present experiment influenced spatial attention more strongly or in a different manner.

Figure 3. (a) Grand-averaged ERPs at the occipito-temporal electrodes (PO7, PO8) in response to standard stimuli in the left and right visual fields in Experiment 1. (b) Scalp distributions of spatial attention effects for each condition: ERPs in the unattended condition were subtracted from those in the attended condition. While P1 and N1 attention effects were clearly observed, the P1 effect was decreased in the connected condition.
Taken together, the present results suggest, for the first time, that sufficiency of depth cues that define objects or “figures” is a critical factor in determining whether earlier stages of visual processing are involved in spatial selection for objects. However, it is possible that the unique stimuli or stimulus configurations, rather than the clarity of figures, in the present experiment incidentally caused these results. For example, in the present experiment, the rectangles were rounded and were presented only in the upper visual field, and these conditions have not been used in previous ERP studies of spatial attention over stable objects. This possibility was tested in the next experiment.

**Experiment 2**

This experiment used the same experimental procedure as Experiment 1, except that Pac-men were aligned such that objects were perceived at the back of the surroundings (Figure 2). Although the stimuli consisted of three discontinuous parts, a rectangle was completed and visible through three circle-shaped windows in the connected condition. Since attention selects perceptual objects after modal completion as well as amodal completion of partly occluded portions, rather than 2-D features (e.g., Albrecht, List, & Robertson, 2008; Kasai & Takeya, 2012; Moore, Yantis, & Vaughan, 1998), one can expect that object-based attention spreading also occurs in Experiment 2. However, it is unclear whether spatial regions behind windows or holes have the same advantage as figures. Since 2-D stimulus configurations or features were similar in Experiments 1 and 2, if the results in Experiment 1 were a side effect of the stimuli used, the same pattern of results should be observed in Experiment 2.

**Method**

The methods were the same as those in Experiment 1, except as noted below. Twelve volunteers (five women), aged 21 to 26 years, participated in this experiment. The inner edges of white regions fit exactly within the circles to be perceived behind the gray surrounding surface (Figure 2). Stimuli in the separated condition were made by swapping the left and right circles in the connected condition.

**Results and discussion**

In behavioral data (Table 1), effects involving connectedness were not statistically significant for all of
the indices examined (hit rates, $p > 0.09$; RTs, $p > 0.13$; FA rates, $p > 0.08$). Thus, inconsistent with Experiment 1, behavioral object-based effects were not found.

Attention effects of ERP were most prominent at occipital temporal sites (Figures 5 and 6). In the clustered ERPs, unilateral standard probes elicited greater P1 (100–140 ms) amplitudes at sites contralateral to stimuli, which were indicated by the significant interaction of VF $\times$ Hemisphere, $F(1, 11) = 6.7, p = 0.03$. The P1 amplitude for attended probes was more positive than that for unattended probes, which was reflected by the main effect of attention, $F(1, 11) = 85.1, p = 0.000002$. N1 (150–220 ms) also showed amplitude enhancement for attended probes, $F(1, 11) = 9.7, p = 0.01$. For both P1 and N1, there were no significant effects that involved the interactions of attention and connectedness ($p > 0.14$; Table 3).

The results in Experiment 2 did not show any object-based attention effects. This indicates that the P1 attention-spreading effect in Experiment 1 was not due to the 2-D stimulus configurations and supports the view that attention spreads over figural objects in front of the surroundings.

**General discussion**

The purpose of this study was to examine spatial selection processes for objects that had more depth cues and were more explicit regarding depth ordering than prior studies. Experiment 1 showed that the P1 spatial
attention effect was associated with object-based attention spreading, which was confirmed not to be a confounding of 2D features in Experiment 2. This finding is clearly in contrast with previous studies that have shown object-based N1 spatial attention effects (He et al., 2004; Kasai, 2010; Kasai et al., 2011; Kasai & Takeya, 2012; Martinez, Ramanathan et al., 2007; Martinez, Teder-Salejarui et al., 2007; Takeya & Kasai, 2014). The present study suggests that spatial selection for objects can occur at earlier stages of processing when the objects are less ambiguously perceived in front of the surrounding.

Although P1 attention effects have generally been interpreted in a 2-D display as a gain control mechanism of sensory inputs (Hillyard et al., 1998), depth information in 3-D space may also be involved. In a study that compared 2-D attention effects (i.e., left vs. right visual fields) in front of fixation from those in back by using a stereoscopic 3-D display (Kasai, Morotomi, Katayama, & Kumada, 2003), the typical P1 attention effect was found only for stimuli in the frontal plane. This suggests that the P1 attention effect is associated with the selection of early cortical representation of figural objects, since segregating borders belong to frontal surfaces. The present object-based P1 spatial attention effect supports the notion that clearly defined figural objects can play a role in the early stages of spatial selection. Such early selection of figural objects is possible because neurons in V1 or V2 are sensitive to border-ownership (De Yoe & Van Essen, 1988; Lamme, 1995; Qiu & von der Heydt, 2005). On the other hand, it is unclear how objects without a “figural advantage” in the back of the surroundings are represented in visual cortices.

In the present results, the P1 object-based attention effect was greater at the left hemisphere. The result may associate with the left-hemisphere dominance for the control of object-based spatial attention (Shomstein & Behrmann, 2006; Wilson, Woldorff, & Mangun, 2005). However, there are considerable inconsistencies. In neuropsychological studies, patients with left-hemisphere lesions experienced impaired object-based effects, while those with right-hemisphere lesions did not show such impairments (Egly et al., 1994), whereas transcranial magnetic stimulation of the left and right parietal cortex did not cause differences in object-based effects in healthy participants (Du, Chen, & Zhou, 2012). The previous studies that showed object-based N1 modulations also did not report hemispheric differences (He et al., 2004; Martinez, Ramanathan et al., 2007; Martinez, Teder-Salejarui et al., 2007). Thus, further studies will be needed to clarify the functional roles of the hemispheres in object-based spatial attention.

A growing body of evidence has suggested that object-based N1 modulation reflects figural enhancement for early cortical representations of visual objects (Kasai, Takeya, & Tanaka, 2015; Martinez, Ramanathan et al., 2007; Martinez, Teder-Salejarui et al., 2007). However, the absence of this effect in the present study indicates that the figural enhancement reflected

<table>
<thead>
<tr>
<th>P1 (100–140 ms)</th>
<th>N1 (150–220 ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>F</strong></td>
<td><strong>p</strong></td>
</tr>
<tr>
<td>Attention</td>
<td>85.00</td>
</tr>
<tr>
<td>Connectedness</td>
<td>3.30</td>
</tr>
<tr>
<td>Hemisphere*</td>
<td>6.70</td>
</tr>
<tr>
<td>Attention × Connectedness</td>
<td>0.40</td>
</tr>
<tr>
<td>Attention × Hemisphere*</td>
<td>0.00</td>
</tr>
<tr>
<td>Connectedness × Hemisphere</td>
<td>0.40</td>
</tr>
<tr>
<td>Attention × Connectedness × Hemisphere</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 3. Summary of statistical results for clustered ERPs in Experiment 2. Note: All significant p values for omnibus ANOVAs are shown. *Indicates interaction with the stimulus visual field.
by N1 is an outcome that resolves the ambiguity in 3-D interpretation. This notion is consistent with the fact that N1 modulation effects and the originating LOC are associated with illusory figures (Murray, Foxe, Javitt, & Foxe, 2004; Murray et al., 2002), although the LOC has generally been linked with object recognition (for a review, see Grill-Spector, Kourtzi, & Kanwisher, 2001). However, we cannot exclude the possibility that a prolonged P1 object-based effect canceled out N1 object-based effects in the present study.

## Conclusion

The present study showed that the spatial attention effect of P1 is involved in the selection of figural objects with sufficient depth cues. Together with previous studies of object-based attention, the present results suggest that clarity of depth ordering is a critical factor in determining whether earlier cortical stages of processing are involved in spatial selection for objects. However, in the present study, we presented objects in a 2-D display: The depth ordering may still be somewhat ambiguous, because binocular depth cues indicate that the objects and the background exist in the same depth plane. Although pictorial or monocular depth cues can modulate early selection processes (Parks & Corballis, 2006), depth cues must be manipulated systematically to clarify the selection processes of objects with which we interact every day.

**Keywords:** spatial attention, object, figure, event-related potential

## Acknowledgments

This research was supported by Japan Society for the Promotion of Science KAKENHI Grant Number 13J00809 to Ryuji Takeya and Ministry of Education, Culture, Sports, Science and Technology KAKENHI Grant Number 23500320 to Tetsuko Kasai.

Commercial relationships: none.

Corresponding author: Ryuji Takeya.

Email: takeryu200@med.hokudai.ac.jp.

Address: Graduate School of Medicine, Hokkaido University, Sapporo, Japan.

## References


