

## Research Report

## Head Up, Foot Down

## Object Words Orient Attention to the Objects' Typical Location

Zachary Estes,<sup>1</sup> Michelle Verges,<sup>2</sup> and Lawrence W. Barsalou<sup>3</sup><sup>1</sup>University of Warwick; <sup>2</sup>Indiana University, South Bend; and <sup>3</sup>Emory University

---

**ABSTRACT**—Many objects typically occur in particular locations, and object words encode these spatial associations. We tested whether such object words (e.g., head, foot) orient attention toward the location where the denoted object typically occurs (i.e., up, down). Because object words elicit perceptual simulations of the denoted objects (i.e., the representations acquired during actual perception are reactivated), we predicted that an object word would interfere with identification of an unrelated visual target subsequently presented in the object's typical location. Consistent with this prediction, three experiments demonstrated that words denoting objects that typically occur high in the visual field hindered identification of targets appearing at the top of the display, whereas words denoting low objects hindered target identification at the bottom of the display. Thus, object words oriented attention to and activated perceptual simulations in the objects' typical locations. These results shed new light on how language affects perception.

---

Attention is often guided by environmental cues (see Berger, Henik, & Rafal, 2005). Here we focus on cues that orient attention away from themselves. For example, when preceded by an arrow pointing leftward (Posner, Snyder, & Davidson, 1980), the word *left* (Hommel, Pratt, Colzato, & Godijn, 2001), a head facing leftward (Langton, Watt, & Bruce, 2000), or eyes gazing leftward (Kingstone, Smilek, Ristic, Friesen, & Eastwood, 2003), visual targets are identified faster on the left than on the right. These directional cues orient attention even when the target is no more likely to occur at the cued location than at an uncued location. Thus, social and symbolic cues can reflexively orient attention to an implied location.

---

Address correspondence to Zachary Estes, Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom, e-mail: z.estes@warwick.ac.uk.

Do object words such as *head* and *foot* similarly direct attention to specific locations? Words that denote objects are, after all, both symbolic and social. They are verbal symbols for interpersonal communication. Some objects, such as apples and books, occur in diverse locations and hence have no particular spatial connotation. Other objects, however, typically occur in particular locations. For instance, branches and clouds are typically overhead, whereas roots and puddles are typically underfoot. Indeed, object words are judged faster when presented on a computer screen in the objects' canonical locations (Zwaan & Yaxley, 2003). For instance, *eagle* is judged faster when presented at the top rather than the bottom of a display, whereas *snake* is judged faster at the bottom (Šetić & Domijan, 2007). Given that many object words encode spatial associations, we hypothesized that an object word directs attention toward the location where its referent typically occurs.

Whereas directional cues such as *left* simply point to a location, object words such as *bird* also evoke a perceptual simulation of the denoted object. *Perceptual simulation* is the activation of perceptual representations that were acquired during the actual perception of a stimulus. More specifically, it is the reactivation of neural pathways that have become associated with perceiving a particular stimulus (see Barsalou, 1999, in press; Martin, 2007). For example, the word *bird* and an image of a bird activate highly overlapping cortical networks (Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996; see Pulvermüller, 2001). Such perceptual simulations emerge automatically during language comprehension, often without conscious awareness. They may also be intentionally generated and consciously inspected, as in the case of mental imagery.

Perceptual simulation of word meaning may have significant implications for attention and perception. If the word *bird* activates the neural mechanisms involved in the perception of a bird, then perception of another visual stimulus that requires these same mechanisms should be delayed. Indeed, this may explain why mental imagery hinders visual perception (Craver-Lemley & Reeves, 1992), particularly when the mental image

and the physical stimulus overlap spatially (Craver-Lemley & Arterberry, 2001). Generalizing to more implicit forms of perceptual simulation, we hypothesized that object words might also hinder visual perception. More specifically, if an object word activates a perceptual simulation in the object's typical location, then that word should hinder perception of a visual target in that location. Essentially, because the perceptual mechanisms required for target identification are engaged in the simulation of the denoted object at its typical location, identification of a target at that location should be delayed. In contrast, when the visual target and the perceptual simulation do not overlap spatially, interference should not occur (cf. Craver-Lemley & Arterberry, 2001). The experiments reported here assessed this hypothesis.

### EXPERIMENT 1

In this experiment, object words associated with an upper or lower location served as cues. To ensure that the object words unambiguously denoted an upper or lower location, we presented a context word before each object word. Thus, on each trial, a context word (e.g., *cowboy*) was followed by an upper-location (e.g., *hat*) or lower-location (e.g., *boot*) cue. Context and cue words were both presented centrally. A target letter (*X* or *O*) then appeared at the top or bottom of the display, and participants identified the target as quickly as possible. The target was equally likely to appear at the top or bottom location, regardless of the cue object's typical location. Thus, the cue did not predict the target's location. If an object word activates a perceptual simulation of the denoted object in its typical location, then target identification should be hindered at that location.

#### Method

The experimental stimuli were 30 context words, each paired with one upper-location and one lower-location cue word (i.e., 60 spatial cues). We also included 60 filler stimuli consisting of 30 context words, each paired with 2 nonspatial cue words (e.g., *chocolate powder*, *chocolate shavings*). Participants initiated each trial by pressing the space bar, which triggered a central fixation cross that appeared for 250 ms. The context word then appeared centrally for 500 ms, replaced immediately by the cue word for 250 ms. After a 50-ms delay, a target letter subtending approximately 1° of visual angle appeared at the top or bottom of the screen. The "top" and "bottom" locations were centered horizontally approximately 8° vertically from the center of the display. Participants were instructed to identify each target letter as quickly and accurately as possible by pressing the appropriate key. Location cue (upper, lower), target location (top, bottom), and target letter (*X*, *O*) were fully crossed and balanced, such that each target letter was equiprobable at each target location, and each target location was equiprobable within each cue condition. Ten practice trials preceded the 120 experimental trials. Eighteen undergraduates participated for course credit.

### Results and Discussion

Data were coded according to whether the target letter appeared in the location associated with the object word. The *typical* condition included upper-location cues followed by top targets and lower-location cues followed by bottom targets, whereas the *atypical* condition included upper-location cues followed by bottom targets and lower-location cues followed by top targets. Response times on incorrect trials were removed from all analyses. Response times greater than 1,000 ms were also removed, resulting in the exclusion of 1% to 4% of trials (across experiments). Data were analyzed using analysis of variance (ANOVA) across participants ( $F_1$ ) and items ( $F_2$ ).

Targets were identified more slowly and less accurately when they appeared in the typical location of the object denoted by the preceding word than when they appeared in the opposite location (see Fig. 1, left-most graphs). In the response time data, this effect was significant,  $F_1(1, 17) = 40.19$ ,  $p_{\text{rep}} \approx 1.00$ , and  $F_2(1, 58) = 41.54$ ,  $p_{\text{rep}} \approx 1.00$ , and was robust (37 ms,  $\eta^2 = .70$ ). In the error-rate data, this effect was also significant,  $F_1(1, 17) = 9.33$ ,  $p_{\text{rep}} = .97$ , and  $F_2(1, 58) = 18.96$ ,  $p_{\text{rep}} \approx 1.00$ , and robust (4.83%,  $\eta^2 = .35$ ). As predicted, the cue word evoked a perceptual simulation in the object's typical location, thus hindering perception of a target letter at that location.

### EXPERIMENT 2

If the interference observed in Experiment 1 was due to perceptual simulation, then disrupting that simulation should attenuate the effect. In Experiment 2, we tested this prediction by replicating the procedure of Experiment 1, but including a condition in which both target locations were visually masked prior to presentation of the target. Another possibility is that the observed effect reflected *inhibition of return* (IOR), whereby the perception of a stimulus at a recently attended location is temporarily inhibited (Posner & Cohen, 1984). The spatial cue may have elicited an attentional shift to the implied location, thereby triggering IOR and inhibiting target identification in that location. Indeed, in Experiment 1, the target appeared 300 ms after the spatial cue, a delay that is well within the typical onset of IOR around 225 ms poststimulus (see Klein, 2000). To test this explanation, we reduced the cue-target asynchrony to 150 ms in Experiment 2. Any interference at this brief delay could not be attributed to IOR.

#### Method

Fifty-nine undergraduates were randomly assigned to one of two conditions: unmasked or masked. In the unmasked condition, the stimuli and procedure were the same as in Experiment 1, except that the context and cue words appeared for only 150 and 100 ms, respectively. With the 50-ms delay preceding the target, the cue-target asynchrony was 150 ms. The masked condition was identical, except that a visual mask appeared during the 50-ms delay. The mask—three contiguous rows of eight ampersands

(approximately  $3^\circ \times 5.25^\circ$ )—appeared simultaneously at the two target locations on each trial.

## Results and Discussion

Data were analyzed as in Experiment 1. Six participants (3 from each group) were excluded because of overall latencies or accuracies more than 2.5 standard deviations from the group mean. Data were initially analyzed via 2 (target location: typical, atypical)  $\times$  2 (mask: unmasked, masked) mixed ANOVAs.

As Figure 1 illustrates, the perceptual interference effect on both response times and error rates was replicated in the unmasked condition. In the masked condition, however, these effects were attenuated. In the response time data, the main effect of target location was significant,  $F_1(1, 51) = 18.28$ ,  $p_{\text{rep}} \approx 1.00$ , and  $F_2(1, 58) = 155.18$ ,  $p_{\text{rep}} \approx 1.00$ ; target identification was again slower in the typical location of the preceding object word. The main effect of mask was also significant,  $F_1(1, 51) = 16.54$ ,  $p_{\text{rep}} = .99$ , and  $F_2(1, 58) = 509.41$ ,  $p_{\text{rep}} \approx 1.00$ ; visual masking slowed target identification. Most important, however, target

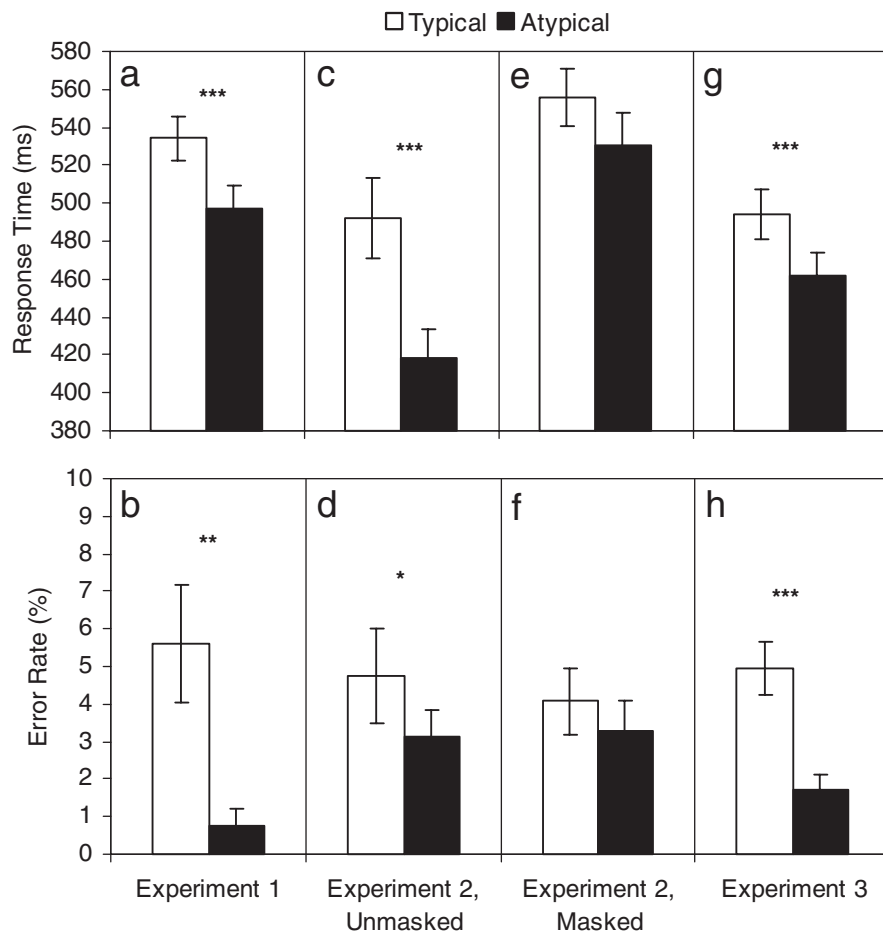
location and mask interacted: The interference effect was larger in the unmasked condition than in the masked condition,  $F_1(1, 51) = 4.31$ ,  $p_{\text{rep}} = .89$ , and  $F_2(1, 58) = 32.70$ ,  $p_{\text{rep}} \approx 1.00$ . In the error-rate data, only the main effect of target location was significant,  $F_1(1, 51) = 8.01$ ,  $p_{\text{rep}} = .96$ , and  $F_2(1, 58) = 13.21$ ,  $p_{\text{rep}} = .99$ . The unmasked and masked conditions were also analyzed separately, and we discuss the results of those analyses next.

### Unmasked Condition

In the response time data for the unmasked condition, perceptual interference was significant,  $F_1(1, 25) = 19.65$ ,  $p_{\text{rep}} = .99$ , and  $F_2(1, 58) = 176.11$ ,  $p_{\text{rep}} \approx 1.00$ , and was robust (74 ms,  $\eta^2 = .44$ ). In the error-rate data, interference was also significant,  $F_1(1, 25) = 7.24$ ,  $p_{\text{rep}} = .94$ , and  $F_2(1, 58) = 11.49$ ,  $p_{\text{rep}} = .98$ , and robust (1.60%,  $\eta^2 = .23$ ).

### Masked Condition

Response times in the masked condition yielded only mixed evidence of perceptual interference,  $F_1(1, 26) = 2.48$ ,  $p_{\text{rep}} =$



**Fig. 1.** Mean response times (top row) and error rates (bottom row) as a function of target location (typical, atypical). From left to right, the graphs present results for Experiment 1, the unmasked condition of Experiment 2, the masked condition of Experiment 2, and Experiment 3. Error bars show standard errors. Asterisks indicate a significant difference between results for typical and atypical locations, \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

.79, and  $F_2(1, 58) = 27.90$ ,  $p_{\text{rep}} \approx 1.00$ . The error-rate data for the masked condition did not show significant interference.

### Summary

The perceptual interference observed in Experiment 1 was replicated in the unmasked condition, but was attenuated in the masked condition. Although the same qualitative pattern was observed in the two conditions, the significant interactive effect of target location and mask on response times indicates that these patterns differed quantitatively. Furthermore, error rates showed the interference effect in the unmasked condition, but not in the masked condition. Given the brief cue-target asynchrony (150 ms), perceptual interference cannot be explained as IOR, which occurs only at longer delays (Klein, 2000). Thus, an object word appears to elicit a perceptual simulation in the object's typical location, and this perceptual simulation interferes with perceiving a target in that location. A visual mask in that location, however, disrupts the simulation and hence attenuates its interfering effect.

## EXPERIMENT 3

In the preceding experiments, the location cue (e.g., *hat*) was preceded by a context word (e.g., *cowboy*). To ensure that interference was unrelated to the context words, we presented only the location cues in Experiment 3.

### Method

The stimuli and procedure were identical to those in the unmasked condition of Experiment 2, except that only cue words were presented. Thirty undergraduates participated.

### Results and Discussion

Data were analyzed as in Experiment 1. Three participants with outlying data (2.5 standard deviations beyond the mean) were excluded. The response times (see Fig. 1, upper right graph) showed an interference effect that was significant,  $F_1(1, 26) = 22.59$ ,  $p_{\text{rep}} \approx 1.00$ , and  $F_2(1, 58) = 25.15$ ,  $p_{\text{rep}} \approx 1.00$ , and robust (32 ms,  $\eta^2 = .47$ ). The error rates (see Fig. 1, lower right graph) also showed an interference effect that was significant,  $F_1(1, 26) = 12.77$ ,  $p_{\text{rep}} = .98$ , and  $F_2(1, 58) = 13.04$ ,  $p_{\text{rep}} = .99$ , and robust (3.21%,  $\eta^2 = .33$ ). Thus, an object word orients attention to and evokes a perceptual simulation in the denoted object's typical location, thereby hindering perception in that location.

## GENERAL DISCUSSION

Words denoting objects that typically occur in high places (e.g., *hat*, *cloud*) hindered identification of targets appearing at the top of the display, whereas words denoting objects that typically occur in low places (e.g., *boot*, *puddle*) hindered identification of targets at the bottom. This perceptual interference is attribut-

able to attentional orienting and perceptual simulation. First, an object word orients attention toward the object's typical location. For example, because birds typically are seen in high places, they become associated with the upper visual field (cf. Richardson & Spivey, 2000), and the word *bird* elicits an upward shift of attention. Second, the object word activates a perceptual simulation of the denoted object (Barsalou, in press; Martin, 2007). That is, the neural systems associated with the actual perception of the denoted object are activated. In the experiments reported here, object words were followed by a visual target at either the top or the bottom of a display, such that the target appeared in either the denoted object's typical location or the opposite location. Because the simulated object (e.g., a bird) and the perceptual target (e.g., *X*) shared few features, perception of the target required inhibition of the neural circuits activated during object simulation, thereby slowing target identification at the denoted object's typical location. At the opposite (atypical) location, the perceptual mechanisms were not engaged in object simulation, and hence target identification proceeded without interference.

Attentional orienting and perceptual simulation together explain why some linguistic cues hinder perception of a physical object whereas others facilitate perception of the object. When the linguistic cue activates a perceptual representation that shares few features with the physical stimulus, as in the present experiments, perception is delayed. Richardson, Spivey, Barsalou, and McRae (2003) presented sentences describing either a vertically oriented event (e.g., *X* respected *Y*) or a horizontally oriented event (e.g., *X* argued with *Y*), followed by an unrelated visual target (i.e., a square or circle) at the top, bottom, left, or right of the display. Targets on the axis associated with the preceding sentence (e.g., a vertical event followed by a top or bottom target) were identified more slowly than targets associated with the other axis (e.g., a vertical event followed by a left or right target). Kaschak et al. (2005) reported an analogous pattern of interference between visual depictions of motion and unrelated sentential descriptions of motion. The present results suggest that participants in all these studies attended to the location, axis, or direction implied by the linguistic cue and simulated the described object or activity. Interference occurred because the simulation and the visual stimulus activated different perceptual representations.

In contrast, when the linguistic cue activates many perceptual features of the physical stimulus, perception is facilitated. Stanfield and Zwaan (2001) presented prime sentences that described an object in one orientation or another (e.g., "He hammered the nail into the *wall/floor*"), and each sentence was followed by an image of the target object in one of these orientations (e.g., a nail oriented horizontally or vertically). Judgments of the target object were faster when it matched the orientation implied by the sentence. Similarly, Zwaan, Madden, Yaxley, and Aveyard (2004) found that a ball in motion was processed faster when preceded by a sentence describing a ball

moving in the same direction rather than the opposite direction. In both cases, facilitation occurred because the simulation and the visual stimulus activated similar perceptual representations.

These results have implications not only for perception, but also for attention. When a central cue orients attention to a peripheral location, despite being nonpredictive of the actual location of the target, that orienting is said to be *reflexive*, or automatic. For example, symbolic cues such as arrows, head orientation, and eye-gaze direction elicit reflexive orienting (Hommel et al., 2001; Kingstone et al., 2003; Langton et al., 2000; Posner et al., 1980). Notably, because the object words in the present experiments did not predict the location of the target, our results indicate that orienting from object words is reflexive. And given the ubiquity of object words, these results indicate that reflexive orienting from symbolic cues is more general than previously known.

## REFERENCES

- Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou, L.W. (in press). Grounded cognition. *Annual Review of Psychology*.
- Berger, A., Henik, A., & Rafal, R. (2005). Competition between endogenous and exogenous orienting of visual attention. *Journal of Experimental Psychology: General*, 134, 207–221.
- Craver-Lemley, C., & Arterberry, M.E. (2001). Imagery-induced interference on a visual detection task. *Spatial Vision*, 14, 101–119.
- Craver-Lemley, C., & Reeves, A. (1992). How visual imagery interferes with vision. *Psychological Review*, 99, 633–649.
- Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science*, 12, 360–365.
- Kaschak, M.P., Madden, C.J., Theriault, D.J., Yaxley, R.H., Aveyard, M., Blanchard, A.A., & Zwaan, R.A. (2005). Perception of motion affects language processing. *Cognition*, 94, B79–B89.
- Kingstone, A., Smilek, D., Ristic, J., Friesen, C.K., & Eastwood, J.D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science*, 12, 176–180.
- Klein, R.M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, 4, 138–147.
- Langton, S.R.H., Watt, R.J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, 4, 50–59.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.
- Posner, M.I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531–556). Hillsdale, NJ: Erlbaum.
- Posner, M.I., Snyder, C.R.R., & Davidson, B. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160–174.
- Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, 5, 517–524.
- Richardson, D.C., & Spivey, M.J. (2000). Representation, space, and Hollywood Squares: Looking at things that aren't there anymore. *Cognition*, 76, 269–295.
- Richardson, D.C., Spivey, M.J., Barsalou, L.W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, 27, 767–780.
- Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes*, 22, 297–312.
- Stanfield, R.A., & Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153–156.
- Vandenbergh, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R.S.J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, 383, 254–256.
- Zwaan, R.A., Madden, C.J., Yaxley, R.H., & Aveyard, M.E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science*, 28, 611–619.
- Zwaan, R.A., & Yaxley, R.H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, 10, 954–958.

(RECEIVED 5/2/07; REVISION ACCEPTED 8/2/07)