Author(s): Zachary Estes , Michelle Verges and Lawrence W. Barsalou
Article Title: Head Up, Foot Down: Object Words Orient Attention to the Objects' Typical Location
Year of publication: 2008
Link to published version: http://dx.doi.org/ 10.1111/j.1467-9280.2008.02051.x
Publisher statement: The definitive version is available at www.blackwell-synergy.com

IN PRESS, Psychological Science

Head up, foot down: Object words orient attention to the object's typical location

Zachary Estes
University of Warwick

Michelle Verges
Indiana University, South Bend

Lawrence W. Barsalou
Emory University

Word Count: 2500

Correspondence to:

Zachary Estes

Department of Psychology

University of Warwick

Coventry, CV4 7AL

United Kingdom

Email: z.estes@warwick.ac.uk

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 a perceptual simulation of the denoted object (i.e., the representations acquired during actual perception are reactivated), they were predicted to 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 a perceptual simulation in the object’s typical location. These results shed new light on how language affects perception.

KEYWORDS: language and perception; mental imagery; perceptual simulation; reflexive orienting; spatial attention.

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, visual targets are identified faster on the left 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). These directional cues orient attention even when the target is no more likely to occur at the cued location (i.e., left) than at an uncued location (e.g., right). Thus, social and symbolic cues can reflexively orient attention to an implied location.

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 communicating with others about a given object. 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 positions (Zwaan & Yaxley, 2003). For instance, "eagle" is judged faster when presented at the top 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 object words direct attention toward the location where their 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, 2008; 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). Critically, however, perceptual interference only occurs when the mental image and the physical stimulus overlap spatially (Craver-Lemley & Arterberry, 2001). Generalizing to more implicit forms of perceptual simulation, then, object words might also hinder visual perception. More specifically, if object words activate perceptual simulations in the object’s typical location, then they should hinder perception of a visual target in that location. Essentially, because the perceptual mechanisms required for identification of the target are engaged in the simulation of the denoted object at its typical location, target identification should be delayed at that location. In contrast, when the visual target and the perceptual simulation do not overlap spatially, interference should not occur (cf. Craver-Lemley & Arterberry, 2001). The present experiments assessed this hypothesis.

Experiment 1

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, each was preceded by a context word. Thus, each trial consisted of a context word (e.g., “cowboy”) followed by an upper (i.e., “hat”) or lower (i.e., “boot”) location cue, 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. Targets were equally likely to appear at the top or bottom location, regardless of the cue object’s typical location. Thus, the cue did not predict target location. If object words activate a perceptual simulation in the object’s typical location, then target identification should be hindered at that location.

Method

Experimental stimuli were 30 context words, each paired with one upper and one lower cue word, thus yielding 60 spatial cues. Thirty yoked pairs of non-spatial filler cues (e.g., “chocolate powder”, “chocolate shavings”) were also included. Participants initiated each trial by pressing the spacebar, 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 the 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, which 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 high cues followed by top targets and low cues followed by bottom targets, whereas the "Atypical" condition included high cues followed by bottom targets and low cues followed by top targets. Data from each of the experiments reported herein were analyzed as follows. Response times on incorrect trials were removed from all analyses. Response times greater than 1000 ms were also removed, resulting in the exclusion of 1% to 4% of trials (across experiments). Data were analyzed using ANOVA across participants (F_1) and items (F_2). Figure 1 illustrates the mean response times and error rates for all experiments.

Targets were identified more slowly (Figure 1A) and less accurately (Figure 1B) when appearing in the typical location of the preceding object word. In response times, this effect was significant [$F_1(1, 17) = 40.19$, $p_{rep} \approx 1.00$; $F_2(1, 58) = 41.54$, $p_{rep} \approx 1.00$] and robust (37 ms, $\eta^2 = .70$). In error rates, this effect was also significant [$F_1(1, 17) = 9.33$, $p_{rep} = .97$; $F_2(1, 58) = 18.96$, $p_{rep} \approx 1.00$] 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 interference. To test this prediction, Experiment 2 replicated the procedure of Experiment 1, but included a condition in which both target locations were visually masked prior to target presentation. Another possibility is that interference could reflect *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. This is well within the timeframe of IOR, which has a typical onset around 225 ms post-stimulus (see Klein, 2000). To test this explanation, the cue-target asynchrony was reduced to 150 ms in Experiment 2. If interference occurs at this brief delay, it cannot be attributed to IOR.

Method

Fifty-nine undergraduates were randomly assigned to an Unmasked or a Masked condition. The stimuli and procedure of the Unmasked condition were identical to 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 (approx. 3° × 5.25°)—appeared at both target locations simultaneously on each trial.

Results and Discussion

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) × 2 (Mask: unmasked, masked) mixed ANOVA.

As Figures 1C and 1D illustrate, the unmasked condition replicated the perceptual interference effect in both response times and error rates. In the masked condition, however, this effect was attenuated (Figures 1E and 1F). For response times, the main effect of Target Location was significant [$F_1(1, 51) = 18.28$, $p_{rep} \approx 1.00$; $F_2(1, 58) = 155.18$, $p_{rep} \approx 1.00$]; target identification was again slower in the typical location of the preceding object word. The main effect of Mask was also significant [$F_1(1, 51) = 16.54$, $p_{rep} = .99$; $F_2(1, 58) = 509.41$, $p_{rep} \approx 1.00$]; visual masking slowed target identification. Most importantly, however, Target Location and Mask interacted, with a larger interference effect in the unmasked condition than in the masked condition [$F_1(1, 51) = 4.31$, $p_{rep} = .89$; $F_2(1, 58) = 32.70$, $p_{rep} \approx 1.00$]. For error rates, only the main effect of Target Location was significant [$F_1(1, 51) = 8.01$, $p_{rep} = .96$; $F_2(1, 58) = 13.21$, $p_{rep} =.99$]. The unmasked and masked conditions are analyzed separately below.

Unmasked. In response times, perceptual interference was significant [$F_1(1, 25) = 19.65$, $p_{rep} = .99$; $F_2(1, 58) = 176.11$, $p_{rep} \approx 1.00$] and robust (74 ms, $\eta^2 = .44$). In error rates, interference was also significant [$F_1(1, 25) = 7.24$, $p_{rep} = .94$; $F_2(1, 58) = 11.49$, $p_{rep} = .98$] and robust (1.60%, $\eta^2 = .23$).

Masked. Response times yielded only mixed evidence of perceptual interference [$F_1(1, 26) = 2.48$, $p_{rep} = .79$ and $F_2(1, 58) = 27.90$, $p_{rep} \approx 1.00$]. In error rates the interference was nonsignificant.

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 both conditions, the significant interaction in response times indicates that these patterns differed quantitatively. Furthermore, the interference effect was obtained for errors 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, object words appear to elicit a perceptual simulation in the object's typical location. This perceptual simulation interferes with perceiving a target in that location. A visual mask in that target 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 is unrelated to the context words, in Experiment 3 only the location cues were presented.

Method

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

Results and Discussion

Three outlying participants (2.5 standard deviations beyond the mean) were excluded. In response times (Figure 1G), the interference effect was significant [$F_1(1, 26) = 22.59$, $p_{rep} \approx 1.00$; $F_2(1, 58) = 25.15$, $p_{rep} \approx 1.00$] and robust (32 ms, $\eta^2 = .47$). In error rates (Figure 1H), interference was also significant [$F_1(1, 26) = 12.77$, $p_{rep} = .98$; $F_2(1, 58) = 13.04$, $p_{rep} = .99$] and robust (3.21%, $\eta^2 = .33$). Thus, individual object words orient attention to and evoke a perceptual simulation in their 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 low objects (e.g., “boot”, “puddle”) hindered identification of targets at the bottom. This perceptual interference is attributable to attentional orienting and perceptual simulation. First, object words orient attention toward the object’s typical location. To illustrate, because birds typically are seen in high places, they become associated with the upper visual field (cf. Richardson & Spivey, 2000). Consequently, the word “bird” elicits an upward shift of attention. This attentional orienting occurred in the present experiments even though the object words did not predict target location. These results therefore suggest that orienting to object words is automatic, or reflexive.

Second, object words activate perceptual simulations. Much evidence indicates that words automatically activate the neural systems associated with the perceptual features of their referents (Barsalou, 2008; Martin, 2007). When a simulated object (e.g., a bird) and a perceptual target (e.g., "X") share few features, perception of the target requires inhibition of the neural circuits activated during object simulation, thereby slowing target identification at the attended location. At the unattended location, the perceptual mechanisms are not engaged in object simulation, and hence target identification proceeds without interference. When a visual mask appears at both target locations, the features of the mask must be inhibited, thus slowing target identification at both locations.

Reflexive orienting and perceptual simulation together explain why some linguistic cues hinder perception whereas others facilitate it. When the linguistic cue is infelicitous, perception is delayed. In the present experiments, for example, the cue words were unrelated to the target letters. 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 infelicitous visual target (i.e., a square or circle) on the vertical or the horizontal axis. Targets were identified more slowly on the axis associated with the preceding sentence. Kaschak and colleagues (2005) reported an analogous pattern of interference between unrelated visual depictions and sentential descriptions of motion. The present results suggest that participants in all these studies attended to the visual location, axis or direction of motion implied by the sentence, and simulated the described activity. Interference occurred because the simulation and the visual stimulus activated different perceptual representations.

In contrast, when a linguistic cue is felicitous (i.e., it accurately describes the visual 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*”), followed by an image of the target object in one of these orientations (i.e., a nail oriented horizontally or vertically). Judgments of the target object were faster when it matched the orientation implied by the sentence. Zwaan, Madden, Yaxley, and Aveyard (2004) similarly found that a ball in motion was processed faster when preceded by a sentence describing a ball moving in the same direction. In both cases, facilitation occurred because the simulation and the visual stimulus activated similar perceptual representations. Thus, when attention is oriented to the target location and the perceptual simulation resembles the visual target, perception is facilitated. When either attentional orienting or perceptual simulation is inconsistent with the visual target, interference occurs instead.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577-660.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59.*

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., Therriault, 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. Houwhuis (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.

Vandenberghe, 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.

Figure 1. Response times and error rates (*M* ± *SE*) as a function of target location (typical, atypical) in Experiment 1 (Panels A and B), Experiment 2 (Unmasked: Panels C and D; Masked: Panels E and F), and Experiment 3 (Panels G and H). * $p < .05$, ** $p < .01$, *** $p < .001$