

Simulating Realism in Language Comprehension

Kevin J. Holmes (kevin.holmes@emory.edu)

Phillip Wolff (pwolff@emory.edu)

Department of Psychology, Emory University

36 Eagle Row, Atlanta, GA 30322

Abstract

Do mental simulations of perceptual information look more like high-resolution photographs or sketchy line drawings? In the domain of language comprehension, considerable evidence suggests that words and sentences trigger simulations, but it is unclear to what extent they resemble perceptual reality. We explored the possibility that different types of language may be associated with simulations at different levels of realism. In Experiment 1, participants judged whether an object depicted in a photograph or a line drawing had been mentioned in a preceding sentence. Object recognition was faster for photographs after sentences containing adjectives than sentences containing spatial terms, and this difference was greater than for line drawings. In Experiment 2, recognition performance for color drawings was intermediate to that of photographs and line drawings, pointing to a continuum of realism from schematic to photorealistic. The results suggest that perceptual simulation is not monolithic in nature, and that language, by eliciting simulations capturing different levels of realism, may induce different ways of conceptualizing the world.

Keywords: realism; mental simulation; language comprehension; word meaning.

Introduction

A growing body of evidence suggests that the process of language comprehension spontaneously triggers representations much like those derived through direct perceptual experience. A number of studies have demonstrated that processing words or sentences activates perceptual information such as shape, orientation, and motion (see Zwaan & Madden, 2005, for a review), suggesting that language comprehenders spontaneously construct perceptual simulations to represent the meaning of linguistic expressions (Barsalou, 1999; Zwaan, 2004). Although such simulations are assumed to bear an analog relationship to the objects and events they represent, little is known about the precise nature of their perceptual properties. In particular, it is unclear whether simulations capture the full richness of perceptual reality, as opposed to a coarser or more schematic level of detail. In this research, we investigate the possibility that simulations can vary in realism, ranging from virtually photorealistic to highly schematic, depending on the type of language processed. To the extent that realistic and schematic perceptual forms afford different mental operations, a simulation's level of realism may provide important clues as to how meaning is represented from linguistic input and how such representations are further used in thinking and reasoning.

Although research on simulation in language comprehension has not directly examined realism, a schematic level of representation has typically been assumed. It has been suggested that because simulations reflect partial reactivations of previously experienced perceptual states (Barsalou, 1999) and depend on limited attentional resources (Zwaan, Stanfield, & Yaxley, 2002), they must be relatively schematic in nature. This assumption converges with the cognitive linguistic theories of Talmy (1983) and Langacker (1987), which hold that semantic knowledge draws on image schemas that capture only select components of experience (e.g., paths, spatial relations) and omit irrelevant information. In line with this view, studies examining the psychological reality of image schemas have tended to use highly impoverished materials, consistent with the possibility that the underlying representations lack fine detail. Richardson et al. (2004), for example, used simple geometric shapes and arrows to investigate spatial aspects of verb representations. Similarly, work on perceptual simulation by Zwaan and colleagues (e.g., Stanfield & Zwaan, 2001; Zwaan et al., 2002) has focused largely on properties (e.g., shape, orientation) that can be readily depicted in line drawings (but see Yaxley & Zwaan, 2007, for findings using more realistic materials).

The assumption that limited attention renders simulations schematic suggests that a simulation's level of realism may depend on how deeply the meaning of a linguistic expression is processed. On this view, processing language deeply results in relatively rich simulations. Recent work by Holmes and Wolff (2010) raises an interesting alternative possibility. Rather than being tied to depth of processing, the simulation of realism may depend on the type of language processed, with certain types of language promoting more realistic simulations than others. This possibility was supported by a series of experiments on the perception of implied motion in realistic and schematic scenes. Holmes and Wolff found that when an object's support was suddenly removed (e.g., a pedestal beneath a potted plant disappeared), people appeared to simulate the effect of gravity, showing insensitivity to downward changes in the object's position. Downward motion was much more likely to be simulated, however, when people viewed scenes that resembled line drawings than when they viewed scenes that resembled photographs. Holmes and Wolff also observed that linguistic processing of the scenes influenced subsequent simulation. After writing a verbal description of the scenes, participants also simulated motion in response to the photorealistic scenes. Further, the magnitude of the simulation effect was positively correlated

with the proportion of terms referring to spatial relations (e.g., prepositions) in participants' descriptions, and negatively correlated, albeit not significantly, with the proportion of adjectives. Thus, relational language seemed to encourage participants to conceptualize photorealistic scenes schematically (i.e., as if they were line drawings), while richer perceptual language promoted more realistic construals. As all participants spent the same amount of time writing their descriptions, it is unlikely that the more realistic construals were the result of deeper linguistic processing. These findings are consistent with a proposal by Landau and Jackendoff (1993), who hypothesized that the meanings of prepositions, which encode coarse spatial information (e.g., points, planes, axes), are represented more schematically than the meanings of object nouns and, presumably, terms that encode even finer-grained detail (e.g., color, texture) such as adjectives.

In the present research, we provide a more direct test of the potential influence of language on conceptualization than was provided in Holmes and Wolff (2010). We do so by manipulating the type of language processed and assessing the level of realism that is simulated. We accomplished this by adapting the sentence-picture matching paradigm developed by Zwaan and colleagues (e.g., Zwaan et al., 2002). In this paradigm, participants are presented with a sentence (e.g., *The ranger saw the eagle in the sky*) describing an object in a manner that implies a given perceptual property (e.g., shape), followed by a picture of the object that either matches (e.g., an eagle with outstretched wings) or mismatches (e.g., an eagle with folded wings) the implied property. Participants are asked to make speeded recognition judgments as to whether the object in the picture was mentioned in the preceding sentence. If judgments are faster to matching than to mismatching pictures, it can be concluded that language comprehenders simulate the implied property.

Here we investigate whether language comprehenders simulate realism, and whether realism varies with type of language. In Experiment 1, we presented participants with sentences that described objects using adjectives (e.g., *The watermelon is crunchy and sweet*) or spatial terms (e.g., *The watermelon is in the basket*), and then asked them to make recognition judgments to photographs or line drawings. We predicted that people would be relatively faster to respond to objects shown in photographs after sentences containing adjectives than after sentences containing spatial terms. This was expected because adjectives should trigger more realistic simulations than spatial terms. Further, we expected that there would be a smaller relative difference in performance across the two types of sentences when people responded to objects shown in line drawings. This was expected because spatial terms should trigger less realistic simulations than adjectives. Indeed, we might have gone so far as to predict that responses to objects in line drawings would be faster after sentences containing spatial terms than after sentences containing adjectives, but such a finding would not necessarily be expected, given that there are many perceptual forms that are

even less realistic than line drawings (see General Discussion). As will be shown, responses are faster overall after sentences containing adjectives than sentences containing spatial terms. Thus, a relatively smaller adjective advantage for line drawings than for photographs would be sufficient to indicate that spatial terms trigger less realistic simulations than adjectives.

Experiment 1

The stimuli in Experiment 1 were pairs of sentences and pictures. The sentences described objects and included either adjectives or spatial terms, and the pictures were either photographs or line drawings. Importantly, the perceptual properties described by the two types of sentences could be captured by both types of pictures (e.g., no color terms were used, since line drawings have no color). Consider, for example, the sentence *The watermelon is crunchy and sweet*. A watermelon depicted in a line drawing could be just as crunchy and sweet as one depicted in a photograph; nothing about the terms *crunchy* and *sweet* explicitly points to properties that a schematically rendered watermelon could not conceivably possess. Similarly, the spatial relations encoded by spatial terms in sentences such as *The watermelon is in the basket* were just as likely to be present in photographs as in line drawings. Hence, it is unlikely for any differences observed across sentence types to be due to cueing of specific perceptual properties by individual words. Of interest was relative recognition performance for objects in photographs and line drawings following the two types of sentences. If simulated realism does not depend on type of language, object recognition performance should not depend on whether objects are mentioned in sentences containing adjectives or spatial terms. If, however, different types of language induce different levels of simulated realism, participants should be faster to respond to photographs after sentences containing adjectives than sentences containing spatial terms, and this difference should be greater than for line drawings.

Method

Participants. Sixty-five Emory University undergraduates participated for course credit. The data of three participants were discarded because of long mean response latencies (> 700 ms).

Materials. Eighty picture pairs were used, with each pair consisting of one black-and-white line drawing of an object from the Snodgrass and Vanderwart (1980) set of normed pictures and one full-color photograph of the same object obtained from the Web, selected to closely match its



Figure 1: Example photograph (left) and line drawing (right) from Exp. 1-2 and color drawing (center) from Exp. 2.

counterpart (see Figure 1). The 80 pairs comprised 10 superordinate categories (birds, clothing, fruit, furniture, kitchen utensils, mammals, musical instruments, tools, vegetables, vehicles), with 8 pairs from each category. The two pictures in each pair were the same size, and all fell within a square of approximately 10 cm on the screen. An additional 8 pictures (4 photographs, 4 line drawings) were used on practice trials.

One hundred sixty sentence pairs were generated to accompany the pictures: 80 pairs in which both sentences mentioned the object in the subsequent picture ("yes" response) and 80 pairs in which both sentences mentioned an object other than the one in the subsequent picture ("no" response). Of the two sentences in each pair, one described an object using adjectives and the other described the same object using spatial terms. To ensure that participants processed the entire sentence, the object noun occupied either the subject position (e.g., *The swan was beautiful and serene*) or the object position (e.g., *The man saw the beautiful, serene swan*) in the sentence. An additional 8 pictures (4 photographs, 4 line drawings), paired with an additional 8 sentences (4 with adjectives, 4 with spatial terms), were used on practice trials.

Design. Each participant was presented with trials from one of eight lists, each including either the photograph or line drawing version of all 80 objects. Within each list, the factors of response type (yes or no), sentence type (adjective vs. spatial), picture realism (photograph vs. line drawing), noun position (subject vs. object), and superordinate category were counterbalanced. Each list consisted of 80 sentence-picture pairs (40 "yes" responses, 40 "no" responses), with 20 trials each of the following combinations: adjective/photograph, adjective/line drawing, spatial/photograph, spatial/line drawing. Practice trials were identical across lists.

Procedure. Participants were instructed to read each sentence, and then to decide whether the object in the subsequently presented picture had been mentioned in the sentence. Each trial began with a sentence, center-justified on the screen. Participants pressed the space bar when they

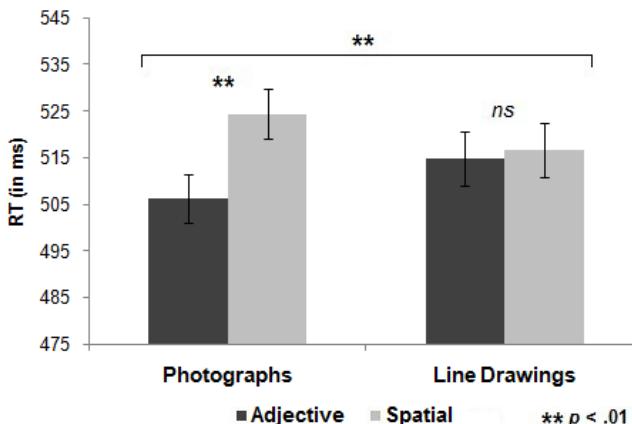


Figure 2: Response times to photographs and line drawings across sentence types in Experiment 1. Error bars are 95% within-subjects confidence intervals.

felt they had understood the sentence. Then a fixation point appeared for 250 ms, followed by a picture. Participants recorded their response by pressing one of two computer keys ("P" key for "yes" responses, "Q" key for "no" responses). The picture remained on the screen until a response was recorded or until 1500 ms elapsed. The next trial began 1000 ms later.

There were 8 practice trials after which participants received feedback on their accuracy and response speed, and 80 test trials in which no feedback was given. Instructions emphasized both speed and accuracy.

Results and Discussion

The findings supported the prediction that sentences with adjectives would produce more realistic simulations than sentences with spatial terms. As shown in Figure 2, responses to photographs were faster after sentences containing adjectives than sentences containing spatial terms, whereas responses to line drawings did not differ across the two types of sentences.

These findings were supported by analyses of object recognition performance across conditions. Preliminary analyses indicated that response type (yes vs. no), noun position (subject vs. object), and list did not interact significantly with these factors (all $p > .05$), so they were not included in subsequent analyses. We conducted 2 (sentence type: adjective vs. spatial) \times 2 (picture realism: photograph vs. line drawing) repeated-measures analyses of variance (ANOVAs) on the accuracy and reaction time (RT) data by participants (F_1) and by items (F_2). In the accuracy analyses, there were no significant main effects of sentence type or realism, and no interaction (all $p > .06$). Mean accuracy was 96.3% overall ($SD = 2.3\%$) and above 95% in all conditions.

In the RT analyses, trials in which participants responded incorrectly (3.7%) or in which RTs were greater than 2.5 SD from individual means (3.3%) were excluded. There was a significant main effect of sentence type, with faster responses to sentences containing adjectives than sentences containing spatial terms, $F_1(1,61) = 6.99, p = .01; F_2(1,79) = 6.63, p = .01$. While there was no main effect of realism ($p > .9$), the interaction between sentence type and realism was significant, $F_1(1,61) = 11.54, p = .001; F_2(1,79) = 6.19, p = .01$. In the case of photographs, sentences containing adjectives resulted in significantly faster response times than sentences containing spatial terms, $t_1(61) = 4.39, p < .0001; t_2(79) = 3.37, p = .001$; response times to line drawings did not differ across the two sentence types ($p > .4$).

These results suggest that participants simulated realism, and that the level of realism depended on the type of sentence processed. A comparison of sentence reading times suggests that more realistic simulation was not due to deeper linguistic processing. Although sentences containing adjectives produced faster responses to photographs than sentences containing spatial terms, these sentences were also read faster (1423 vs. 1466 ms), $t(61) = 2.45, p = .02$. Thus, more realistic simulations occurred for sentences that were processed, if anything, less deeply. We suspect,

however, that depth of processing was comparable across the two sentence types because the sentences containing adjectives were slightly shorter on average (6.5 vs. 7.1 words), and hence could be read faster.

The findings of Experiment 1 suggest that the simulations associated with adjectives are more like photographs than line drawings. The simulations associated with spatial terms, though less realistic than those for adjectives, showed no clear association with one type of picture over the other. However, as there was a larger advantage for adjectives over spatial terms in the case of photographs than in the case of line drawings, the level of simulated realism triggered by spatial terms can be characterized as relatively less realistic than that triggered by adjectives.

Experiment 2

Besides photographs and line drawings, there are countless other perceptual forms that may fall at different points on a continuum from schematic to photorealistic. Images that contain information about an object's surface details (e.g., color, texture), without necessarily being faithful to how the object would be perceived by the visual system, serve as an interesting test case. On the one hand, such images are like photographs in representing fine-grained details of objects that might be omitted from more schematic renderings. On the other hand, the presence of color and texture alone do not ensure that an image will look realistic. Experiment 2 further examined the nature of simulated realism by comparing recognition performance for objects shown in color drawings (including some texture information) to that for objects shown in photographs and in line drawings. If simulated realism is simply a matter of representing surface details, participants should be faster to respond to objects shown in color drawings after sentences with adjectives than sentences with spatial terms, and the magnitude of this effect should be comparable to that for photographs (and larger than that for line drawings). If, however, simulated realism is capable of capturing the rich detail of perceptual reality, the difference in performance for color drawings across the two types of sentences should fall somewhere between that for photographs and line drawings, reflecting a relatively intermediate level of realism.

Method

Participants. Sixty-three Emory University undergraduates participated for course credit. The data of 11 participants were discarded because of high error rates ($> 20\%$; $N = 1$) or long mean response latencies (> 700 ms; $N = 10$).

Materials. The same sentences and pictures as in Experiment 1 were used. An additional 80 drawings of the same objects (Rossion & Pourtois, 2004) were added to the set of pictures. These drawings were augmented versions of Snodgrass and Vanderwart's (1980) original stimuli, with color information and some texture detail added. While the new stimuli were designed to appear more realistic than their line drawing counterparts (Rossion & Pourtois, 2004), they are noticeably less lifelike than photographs (see Figure 1).

Design and Procedure. Each participant was presented with trials from one of 12 lists, each including one of the three versions (photograph, color drawing, or line drawing) of the 80 objects. As in Experiment 1, all within-subjects factors (response type, sentence type, picture realism, noun position, and superordinate category) were fully counterbalanced within each list. All other aspects of the design and procedure were identical to Experiment 1.

Results and Discussion

As shown in Figure 3, the results for photographs and line drawings replicated Experiment 1. Responses to photographs were faster after sentences containing adjectives than sentences containing spatial terms, and the advantage for sentences containing adjectives was somewhat larger for photographs than for line drawings. We also found that color drawings, like photographs, showed faster responses following sentences containing adjectives than sentences containing spatial terms, but this difference across sentence types did not differ significantly from that for photographs or for line drawings. Hence, color drawings might be characterized as representing an intermediate level of realism, between highly realistic photographs and more schematic line drawings.

These findings were supported by 2 (sentence type: adjective vs. spatial) \times 3 (picture realism: photograph vs. color drawing vs. line drawing) repeated-measures ANOVAs on the RT and accuracy data by participants and by items. In the accuracy analyses, there were no significant main effects of sentence type or realism, and no interaction (all $p > .4$). Mean accuracy was 97.4% overall ($SD = 2.3\%$) and above 96% in all conditions. In the RT analyses, trials in which participants responded incorrectly (2.6%) or in which RTs were greater than 2.5 SD from individual means (2.9%) were excluded. While there was a significant main effect of sentence type, with faster responses to sentences containing adjectives than sentences containing spatial terms, $F_1(1,51) = 9.11$, $p = .004$; $F_2(1,79) = 9.35$, $p = .003$, neither the main effect of realism nor the interaction between sentence type and realism was significant ($p > .2$).

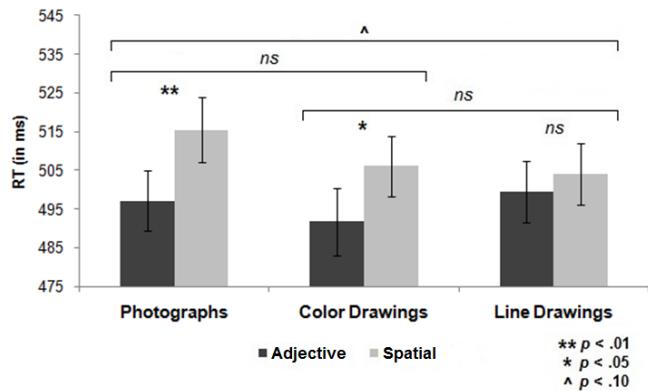


Figure 3: Response times to photographs, color drawings, and line drawings across sentence types in Experiment 2. Error bars are 95% within-subjects confidence intervals.

Figure 3 suggests that the lack of significant interaction may have been due to the intermediate pattern of performance for color drawings relative to photographs and line drawings. Indeed, when trials with color drawings were not included, the interaction became marginally significant in the analysis by participants, $F_1(1,51) = 2.85, p = .098; F_2 < 2$. Sentences containing adjectives resulted in significantly faster response times than sentences containing spatial terms in the case of photographs, $t_1(51) = 2.87, p = .006; t_2(79) = 2.31, p = .02$, but not line drawings ($p > .4$), replicating Experiment 1. Color drawings appeared to pattern after photographs, showing significantly faster responses following sentences containing adjectives than sentences containing spatial terms, $t_1(51) = 2.14, p = .04; t_2(79) = 2.11, p = .04$. However, pairwise comparisons of difference scores across sentence types (mean spatial RT – mean adjective RT) for the three types of pictures revealed that while the advantage for sentences containing adjectives (relative to sentences containing spatial terms) was marginally significant for photographs compared to line drawings, $t(51) = 1.69, p = .098$, color drawings did not differ significantly from either of the other two picture types [color drawings vs. photographs: $t(51) = .53, p > .5$; color drawings vs. line drawings: $t(51) = 1.07, p > .2$]. As in Experiment 1, sentence reading times were faster for sentences containing adjectives than sentences containing spatial terms (1282 vs. 1324 ms), $t(51) = 3.07, p = .003$, inconsistent with a depth of processing explanation.

The results of Experiment 2 provide further evidence that simulations vary in realism according to type of language, with adjectives inducing relatively more realistic simulations than spatial terms. Indeed, the difference between the two sentence types was largest for photographs and smallest for line drawings. Color drawings fell somewhere in the middle; while more strongly associated with adjectives than spatial terms, they nonetheless appeared to reflect an intermediate level of realism. Simulated realism may thus be more continuous than categorical; simulations can be more or less realistic, with the most realistic of simulations resembling the level of detail captured in a photograph, not merely that in a color drawing. Moreover, different types of language may be represented at different points on this continuum of realism.

General Discussion

Previous research on perceptual simulation during language comprehension has considered simulation in essentially monolithic terms. That is, many studies have served as existence proofs that simulation occurs from linguistic input (see Zwaan & Madden, 2005), without examining the conditions under which the perceptual properties of simulations may vary. While effective in challenging the idea that language comprehension relies exclusively on amodal propositions (Barsalou, 1999; Zwaan et al., 2002), this approach may underestimate the range of representational diversity that simulations may be capable of capturing. Our findings suggest that one aspect of such

diversity is the extent to which simulations reflect perceptual reality. Across two experiments, language referring to relatively more fine-grained information about objects resulted in more realistic simulations (i.e., akin to photographs rather than line drawings) than language referring to coarser spatial properties. Moreover, these results could not be explained simply in terms of how deeply participants processed the linguistic stimuli.

As discussed previously, our findings are unlikely to be due to cueing of perceptual properties by individual words, such as *crunchy* and *sweet*. Instead, we suggest that these adjectives induced a more realistic simulation precisely because they are adjectives; processing such terms may lead to the expectation of greater realism more generally, beyond the meanings of the words themselves. This idea is supported by the observation that the results were not contingent on whether the depicted object had in fact been mentioned in the preceding sentence. As this factor did not interact with sentence type or realism, it can be concluded that faster responses to photographs following sentences containing adjectives occurred not only when participants confirmed objects that had been mentioned (e.g., a watermelon after *The watermelon is crunchy and sweet*), but also when they disconfirmed objects that had not been mentioned (e.g., a flute after the same sentence). Thus, realism might be characterized as a mode of simulation induced by language more generally (Wolff & Holmes, 2011), rather than a property restricted to the representation of objects explicitly described.

Although we focused on adjectives and spatial terms, other types of language may be associated with different levels of realism as well. Landau and Jackendoff's (1993) account of the potential neural underpinnings of word meaning contrasts prepositions with object nouns, linking the latter to processing in the ventral stream of the visual system. More recent neural evidence suggests, however, that the dorsal stream also shows sensitivity to properties of objects, namely shape (Chandrasekaran et al., 2006). As processing in the dorsal stream has been characterized as schematic relative to that in the ventral stream (Farivar, 2009), nouns may be more likely to pattern after prepositions than adjectives. Thus, it may be useful to examine the level of realism associated with bare nouns (e.g., *watermelon*), given no other descriptive information, perhaps as a means of specifying the nature of object representations devoid of situational context. Even more schematic than the simulations associated with nouns might be those for verbs, given that verb meanings have often been characterized in terms of semantic components (e.g., CAUSE, CONTACT, PATH) that are highly abstract and relational and for which a common perceptual instantiation is difficult to specify (Talmy, 1988; Wolff & Song, 2003).

Photographs and line drawings clearly differ in realism, but they do not necessarily represent the endpoints of the realism continuum. However, photographs might be relatively close to the realistic endpoint, given that they capture most, if not all, of the perceptual properties

processed by the visual system. Still, images that amplify 3-D depth cues or contain extreme variation in texture and lighting might suggest how far the boundaries of simulated realism can be pushed. Conversely, images that reduce objects to simple geometric forms, such as those used in research on image schemas (e.g., Richardson et al., 2004), are considerably more schematic than line drawings. Note that our results showed an advantage for adjectives over spatial terms for more realistic images (photographs and color drawings), but no corresponding advantage for spatial terms over adjectives for more schematic images (line drawings). If even more schematic images (e.g., geons; Biederman, 1987) were used, such a schematic advantage might be realized, so long as the objects could still be reliably identified. Manipulating perceptual properties in this manner suggests a potentially powerful diagnostic for the content of mental simulations.

In previous work (Holmes & Wolff, 2010), we found that the realism of visual scenes influenced the mental operations used to process them. In particular, more schematic scenes were found to produce greater effects of implied motion, suggesting that stimuli lacking in realism may be especially likely to promote dynamic processing. Together with additional evidence that such processing was correlated with the use of certain types of language, the current findings point to a role for language in modulating the realism perceived in everyday stimuli. If different types of language shunt processing in the direction of greater or lesser realism, the referents of such language may be perceived as more or less realistic than they actually are. Further, non-linguistic mental operations associated with the level of simulated realism (e.g., implied motion, in the case of more schematic simulations), may become more likely to be recruited. In this way, language may induce cognitive processes that continue to be engaged even after language is no longer in use (Wolff & Holmes, 2011). Just as artists capture different levels of realism in their works (e.g., the rich color and heightened intensity of Baroque painting vs. the abstract geometry of Cubism), so too may language give rise to different images in the mind, perhaps offering different affordances for thinking and reasoning. In tapping multiple layers of perceptual experience, language may enable the visual world to be represented in grainy black and white, full technicolor, and everything in between.

Acknowledgments

We thank Kelsey Hodge, David Molho, and Sam Ritter for assistance with stimulus preparation and data collection. This research was supported by a graduate fellowship from the William Orr Dingwall Foundation to Kevin J. Holmes.

References

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

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-147.

Chandrasekaran, C., Canon, V., Dahmen, J. C., Kourtzi, Z., & Welchman, A. E. (2006). Neural correlates of disparity-defined shape discrimination in the human brain. *Journal of Neurophysiology*, 97, 1553-1565.

Farivar, R. (2009). Dorsal-ventral integration in object recognition. *Brain Research Reviews*, 61, 144-153.

Holmes, K. J., & Wolff, P. (2010). Simulation from schematics: Dorsal stream processing and the perception of implied motion. In R. Catrambone & S. Ohlsson (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2704-2709). Austin, TX: Cognitive Science Society.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16, 217-265.

Langacker, R. W. (1987). An introduction to cognitive grammar. *Cognitive Science*, 10, 1-40.

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2004). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, 27, 767-780.

Rosson, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33, 217-236.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174-215.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153-156.

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research, and application*. New York: Plenum Press.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, 12, 49-100.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 253-265.

Wolff, P., & Song, G. (2003). Models of causation and the semantics of causal verbs. *Cognitive Psychology*, 47, 276-332.

Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition*, 105, 229-236.

Zwaan, R. A. (2004). The immersed experient: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation*. New York: Academic Press.

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge, UK: Cambridge University Press.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Do language comprehenders routinely represent the shapes of objects? *Psychological Science*, 13, 168-171.