Spatial Language Influences Memory for Spatial Scenes

Michele I. Feist

University of Louisiana at Lafayette

Dedre Gentner

Northwestern University

Corresponding author:

Michele I. Feist,

University of Louisiana at Lafayette,

Institute of Cognitive Science,

P. O. Drawer 43772,

Lafayette, LA 70504-3772

(telephone: 337-482-1133) (email feist@louisiana.edu)
Does language influence recognition for spatial scenes? In Experiments 1 and 2, participants viewed ambiguous pictures, with or without spatial sentences. In a yes-no recognition task only the spatial sentences group made more false alarms towards the center of the spatial category than in the other direction; three other comparison groups showed no such tendency. This shift towards the core of the semantic category suggests that spatial language interacted with perceptual information during encoding. Experiment 3 varied the materials to test the above Interactive Encoding account against a Separate Encoding account in which separately stored sentences are accessed during picture recognition. The results support the Interactive Encoding account in which spatial language influences the encoding and memory of spatial relations.
Spatial Language Influences Memory for Spatial Scenes

Introduction

In recent times there has been a resurgence of interest in the question of whether and how language influences thought (Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996). Some of this work is in the classic tradition of the strong version of the language-and-thought hypothesis, linguistic determinism: the thesis that the language a person speaks determines the way in which s/he perceives and encodes the world. This version is expressed in Whorf’s (1956) quote of Sapir: “[w]e see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation” (p. 134). For example, in one influential line of research, Levinson and his colleagues (Levinson, 1996, 2003; Levinson, Kita, Haun & Rasch, 2002; Pederson, Danziger, Wilkins, Levinson, Kita, & Senft, 1998) have suggested that the spatial frames of reference employed by a given language influence its speakers’ performance on nonlinguistic tasks. They report that speakers of languages that predominantly use egocentric spatial terms (such as left-right in English) attend greatly to the left-right order of objects with respect to the speaker. In contrast, speakers of languages like Tzeltál, which predominantly use absolute spatial terms (the equivalent of north-south), encode positions with respect to a global framework. Levinson et al.’s conclusion that performance in simple non-linguistic tasks reflects the dominant spatial frame in a participant’s language has generated considerable interest as well as some challenges. Li and Gleitman (2002) argued that an array of strategies can also be induced in monolingual English speakers by changing the task context, to which Levinson et al. (2002) countered that within-language context effects are not inconsistent with substantial cross-linguistic
differences in default patterns of usage. Such within-language effects have been reported by Taylor and Tversky (1996), who found that changes in aspects of the scene to be described resulted in changes in English speakers’ choice of linguistic frame of reference (perspective, in Taylor and Tversky’s terms). Other investigations of the strong Whorfian hypothesis have likewise generated a mix of positive (e.g., Boroditsky, 2001; Lucy, 1992) and negative findings (e.g., Gennari, Sloman, Malt & Fitch, 2002; Malt, Sloman, & Gennari, 2003; Papafragou, Massey, & Gleitman, 2002).

Questions like the above may require considerable research to be settled. However, current discussions have also refined and differentiated the language and thought question (Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003). One moderate version is that habitual ways of talking, and the construals habitually made for linguistic purposes, become especially available in non-linguistic contexts (Hunt & Agnoli, 1991). A still weaker version is Slobin's (1996, 2003) thinking-for-speaking hypothesis, which states that linguistic influences occur when language is used during a task. The idea is that, in speaking, we are induced by the language we use to attend to certain aspects of the world while disregarding or de-emphasizing others.

Our interest here is in this more modest form of the hypothesis, Slobin's thinking-for-speaking hypothesis. This view may seem unremarkable, but the fact that we spend so much of our time using language (both when communicating with others and when thinking within ourselves) suggests that such an influence may in fact be quite pervasive. In addition, this version of the hypothesis has the attractive feature that it permits testing for an effect of language on thought within a language, allowing researchers to factor out effects of culture that can muddy the interpretation of cross-linguistic studies.
A number of studies have reported effects of language on cognitive measures when language was used during the task. For example, Billman and Krych (1998) found that hearing path verbs (e.g., *enter*) or manner verbs (e.g., *walk*) while watching videotaped motion events influenced English-speaking subjects’ subsequent recognition of variants of the events: they were more likely to notice a change in manner if they had heard a manner verb, and likewise for path. Gentner and Loftus (1979) found that performance on a picture recognition task was influenced by having matched the picture to a verb one week prior to the recognition test. Participants who had matched a specific verb (e.g., *hiking*) to a general picture (a woman walking) were highly likely to choose the specific picture (a woman hiking) over the general picture they had actually received. Both these studies suggest that attending to language during the encoding of scenes can influence scene encoding and subsequent recognition. Further, these effects are not limited to motion verbs. Loewenstein and Gentner (2001, in press) found that preschool children performed better in a difficult spatial mapping task when they had previously heard spatial terms describing the arrays. Hermer-Vazquez and her colleagues (Hermer-Vazquez, Moffet, & Munkholm, 2001; Hermer-Vazquez, Spelke, & Katsnelson, 1999) found that performance on a difficult spatial retrieval task was correlated with the ability to use the spatial language relevant to the task (but see Cheng & Newcombe, in press and Learmonth, Nadel, & Newcombe, 2002). Such findings suggest a role for spatial relational language in supporting spatial cognition.

Our research asks whether the presence of spatial language—specifically English spatial prepositions—can influence the encoding and memory of pictorial scenes. In answering this question, we sought evidence bearing on two questions: (1) whether participants’ memories showed language-related alterations (Experiments 1 and 2), and (2)
whether these alterations resulted from an active comparison of language and the picture in pursuit of a common construal of the two, a process we refer to as Interactive Encoding (Experiment 3).

We chose spatial prepositions for several reasons. First, spatial prepositions themselves have been the focus of much recent research (e.g., Bowerman & Pederson, 1992, 1996; Coventry, Carmichael, & Garrod, 1994; Coventry & Garrod, 2003; Levinson, Meira, & the Language and Cognition Group, 2003; Pederson et al, 1998; Regier, 1996; Vandeloise, 1991). Second, spatial prepositions are used with very high frequency even in non-spatial domains because space is a base domain for many metaphors (e.g., Boroditsky, 2001; Gentner & Imai, 1992; Gentner, Imai, & Boroditsky, 2002). Third, there has been comparatively little work on the possible effects of prepositions on the encoding of static spatial relations to date. Yet, spatial prepositions exhibit considerable variability across languages (Bowerman & Pederson, 1992, 1996; Feist, 2000; Levinson et al., 2003). Coupled with evidence of very early learning (Bowerman & Choi, 2003; Choi & Bowerman, 1991) this variability raises the possibility that the semantic encodings in spatial prepositions might be especially likely to influence speakers’ construals of the world.

Design of the Experiments

The basic plan is as follows. For each of the prepositions tested, we created a sentence (e.g. The block is on the building) and a triad of pictures that varied in how well they fit the sentence (see Figure 1 below). The standard pictures were designed to be possible borderline exemplars of the spatial term, but not to be central exemplars. (In Norming Study 2b we used spontaneous verbal descriptions to confirm that the standards did not by themselves give rise to the key prepositions.) For each standard, there were two
variants, one of which (the \textit{plus} variant) was a better exemplar of the spatial term, and one of which (the \textit{minus} variant) was a poorer exemplar. Thus, the standard was designed so that the spatial preposition could apply to it, and the two variants were either \textit{more} typical of the core prepositional category or \textit{less} so. Participants were shown the set of standards and later asked to recognize which picture they saw.

If participants apply the semantic categories of the spatial terms when encoding the standard, they should be more likely to false alarm to the \textit{plus} variant than to the \textit{minus} variant—that is, they should show a shift in recognition towards the category’s center. Across three studies we varied the input conditions and asked under what conditions participants would show this effect. If we see language effects only when people are provided with language at encoding, this will provide support for a thinking-for-language hypothesis—a generalization of Slobin’s thinking-for-speaking hypothesis to encompass comprehending as well as producing language. If, on the other hand, we see spatial category effects even without the presentation of language at encoding, this will support the possibility that language influences cognition in a more far-reaching manner.

The goal of the first two studies is to test whether the presence of spatial language during encoding influences picture recognition. Experiment 3 is aimed at extending the scope of the findings, as well as clarifying the explanation. We test whether picture encoding can be influenced not just according to the specific terms, but according to the dimensions on which the terms operate. In other words, instead of asking whether the use of \textit{on} will shift the encoding of an ambiguous picture toward the center of the category named by \textit{on}, Experiment 3 asks whether the use of any spatial term will highlight the spatial semantic system, shifting the encoding of all ambiguous pictures toward the centers
of the nearest spatial semantic categories. The study also allows us to test whether our findings are explainable by intrusions from a separate encoding of the language used in the task, instead of by an interaction between language and the pictures (as we propose).

Experiment 1

In Experiment 1 we addressed the question of whether the presence of spatial language at encoding would influence recognition memory for simple spatial scenes. Participants viewed pictures depicting static spatial relations - e.g., a marionette standing on a table or a coin in the palm of a hand. After participating in a ten-minute distractor task, participants performed a recognition task that included the original pictures and two variants.

---Insert Figure 1 about here---

In Experiment 1a, half the participants received two sentences along with each picture during the study portion and chose which best described the picture; the other half of the participants received only the pictures. One sentence (the target sentence) was plausible and the other (the distractor sentence) was designed to be obviously wrong; its purpose was simply to cause participants to read the correct sentence and encode the target spatial relational term. To forestall inducing a strategy of focusing directly on the preposition, only the nouns were changed in the distractor sentences. Thus, for the target sentence “The puppet is on the table”, the distractor sentence was “The computer is on the desk.” In constructing the materials, we ensured that the distractor sentence involved the same sense of the preposition as the target sentence (e.g., on to describe support from below and not support via adhesion.). We also ensured that the standards did not by themselves give rise to the key spatial semantic category (see Norming Study 2b).
Participants in Experiment 1b were instructed to attend carefully to the pictures, while participants in Experiment 1c were presented with the pictures along with the sentences from Experiment 1a, with the prepositions removed.

The recognition test included all three pictures—the standard, the plus variant, and the minus variant—presented individually in random order. To the extent that language influences recognition memory, we should see more plus false alarms than minus false alarms.

*Experiment 1a*

In Experiment 1a half the participants (the Spatial Sentences group) received the sentences along with the pictures; the other half (the Control Group) received only the pictures. Both groups were told to remember the pictures for a later test. If strong linguistic determinism is correct—that is, if the semantic categories of one’s language influence cognitive encoding even when language is not being used—we should see more plus false alarms than minus false alarms for both groups. If thinking-for-language is correct, we should see a plus advantage for the spatial sentences group only. Finally, if there is no effect of language on recognition memory, there should be no difference in the rates of plus and minus false alarms for either group.

*Method*

*Design.* Encoding Condition (Spatial Sentences/Control), a between-subjects factor, was crossed with Recognition Item Type (Plus Variant/Standard/Minus Variant), a within-subject factor.
Participants. Thirty-six Northwestern undergraduates received course credit for their participation in this experiment.

Materials. There were 13 triads of pictures and corresponding sets of sentences as well as 18 filler pictures and sentences. Each triad contained one picture whose Figure-Ground relation was ambiguous with respect to the preposition in the sentence (the standard), one that fell closer to the center of the spatial category named by the preposition (the plus variant), and one that fell outside of the spatial category named by the preposition (the minus variant) (See Figure 1).¹ All three pictures involved the same objects; the only source of variation was the spatial relation between the two objects. In preparing the pictures, every attempt was made to guard against a possible recognition bias for the plus variant (see the norming studies for Experiment 2). The standard from each triad was used for the study portion of the experiment; all three pictures in the triad were used for the recognition test. For each standard there was a pair of sentences. One described the Figure-Ground relation using the key preposition, and the other used the same preposition in the same sense but with two clearly incorrect nouns. The target and distractor sentences are presented in Table 1.

---insert Table 1 about here---

Procedure

Part 1: Study. Twenty-five pictures (thirteen standards and twelve fillers) were randomized and presented individually for five seconds each on a computer screen. All participants were told that they would be tested later on their ability to recognize these pictures.
To ensure that the spatial sentences group processed the sentences, we asked them to choose which of two sentences best described the picture. They were provided with answer sheets with two sentences for each picture: the target sentence and a distractor sentence, as described above. Participants in the control condition were given no additional instructions. After completing the study portion, participants did an unrelated filler task for ten minutes.

**Part 2: Recognition.** All participants received the same yes/no recognition test. All three of the pictures in each triad were presented individually in random order on a computer screen along with twelve fillers (six old and six new), for a total of fifty-one pictures. Participants were asked to indicate on a numbered answer sheet whether they had seen each picture during the earlier study portion. Each picture remained on the screen until the participant pressed a key to continue.

**Results**

The two groups did not differ significantly in their overall error rates, nor in the overall rates of the two kinds of possible errors, misses and false alarms (see Figure 2).

The key prediction was that the pattern of false alarms would be influenced by the presence of spatial language at study. Indeed, participants in the spatial sentences condition were significantly more likely to false-alarm to the plus variant than to the minus variant (see Figure 3), $t(17) = 5.32, p < .0001$. Participants in the control condition showed no such difference in their false alarm rate, $t(17) = -.72, p > .10$. Thus, having spatial language present at encoding led to a skewing of recognition errors towards the core of the spatial category, providing evidence for an effect of spatial language on the encoding of the pictures.
Discriminability Analysis. To further test this claim, two d' measures were calculated for each participant: one for the discriminability of the *minus* variant from the standard and one for the discriminability of the *plus* variant from the standard. The larger of the two was then determined and participants were pooled by condition. Table 2 shows the results (along with those for Experiments 1b and 1c). The control condition shows a symmetric discriminability pattern, with equal numbers of participants showing the two possible discriminability biases. In contrast, the spatial sentences condition shows a strongly skewed discriminability pattern: twelve participants showed greater discriminability for the *minus* variant and none showed greater discriminability for the *plus* variant, a significant difference between conditions, $\chi^2 (2, N = 35) = 9.65, p < .01$. As a further check on the effect of condition on discriminability, the set of d' data was subjected to a 2x2 multivariate analysis of variance. In support of our conclusion, we found an effect of condition, $F(1,34) = 10.72, p = .002$.

---Insert Figure 2 and Table 2 about here---

Discussion

When spatial language was present at encoding, memory for the spatial relations in our pictures was systematically shifted in the direction of the spatial preposition. This is evidence for at least the moderate thinking-for-language version of the Whorfian hypothesis. We argue that this shift reflects an interactive encoding process in which the representation of the pictures is influenced by the spatial sentences. In the next study we asked whether people would spontaneously invoke spatial language if faced with a difficult encoding task. Extrapolating from Brown and Lenneberg’s (1954) finding that the codability of colors influenced people’s ability to remember them later, we reasoned that if...
participants believe they are preparing for a difficult memory task, they might mentally invoke linguistic encodings as a memory aid.

**Experiment 1b**

In this study we asked whether there are language effects on recognition memory without the overt use of language at encoding. We tested the possibility that participants instructed to pay careful attention to the pictures at study might be induced to encode the pictures linguistically. If so, they should be more likely to false alarm to the *plus* variant than to the *minus* variant, like the spatial language participants in Experiment 1a.

**Method**

**Participants.** Eighteen Northwestern undergraduates received course credit for their participation in this experiment.

**Materials.** The materials were those used in Experiment 1a.

**Procedure**

The procedure was as in the control condition in Experiment 1a, except that the participants were instructed to pay careful attention to the pictures because the recognition test would be very difficult. After completing the study portion, participants did an unrelated filler task for ten minutes. Then they were given the same recognition test as in Experiment 1a.

**Results and Discussion**

The overall error rate observed in Experiment 1b (*M* = 0.28, *SD* = 0.09) was lower than that observed in Experiment 1a (*M* = 0.34, *SD* = 0.08), suggesting that participants did
pay more careful attention to the pictures during study (see Figure 2). In particular, participants in Experiment 1b had a lower incidence of false alarms than participants in either of the conditions from Experiment 1a (Experiment 1a control: $M=0.41$, $SD = .10$; Experiment 1a spatial sentences: $M= 0.38$, $SD = .08$; Experiment 1b: $M= 0.29$, $SD = .08$, comparison with control: $t(34) = 3.78$, $p < .001$; comparison with spatial sentences: $t(34) = 2.91$, $p < .01$).

However, we found no evidence that participants in Experiment 1b mentally invoked spatial language as a memory aid. They did not show the error pattern characteristic of spatial language use (more plus than minus false alarms). Instead, like the control group in Experiment 1a, the participants in Experiment 1b showed roughly equal numbers of plus and minus false alarms, $t(17) = -1.27$, $p = .22$.

So far we have evidence for the influence of spatial language when it is explicitly presented, although not for the stronger possibility that language can affect encoding and recognition when it is not overtly present. In Experiment 1c, we tested the specificity of the language effect. If the recognition shift is due to spatial language, then we should not see this shift if participants are given verbal descriptions that do not contain spatial language.

**Experiment 1c**

Experiments 1a and 1b provided evidence for a thinking-for-language version of the language and thought hypothesis: that using spatial language during the encoding of scenes fosters spatial representations congruent with the language. In Experiment 1c, we investigate the reverse possibility: that using language that emphasizes only the objects present—that is, eliminating the spatial language—will invite encodings that are primarily centered on the objects. Therefore, participants given object-centered language will
experience no shift towards the core of the spatial category (and thus will not show any differential tendency to false-alarm to the plus variants over the minus variants). Further, if language about the objects induces object-centered encodings, then we may find an overall loss of accuracy in the ability to retain the spatial relations, leading to depressed performance on the recognition task (in which participants must reject alternatives that share objects but differ slightly in their spatial relations). Although this outcome seems most plausible, there is another possibility: it could be that the presence of any linguistic stimuli at encoding will entrain participants into a linguistic mode and prompt them to think of the semantic category that best fits the standard. If the presence of language induces a general mindset to invoke applicable semantic encodings, participants given language at encoding (even non-spatial language) should show a greater rate of plus false alarms than minus false alarms.

We tested these hypotheses by presenting participants with sentences during encoding that simply pointed out the objects without spatial prepositions. Equal plus and minus false alarms in this case would support a semantically specific version of the thinking-for-language hypothesis, while a prevalence of plus false alarms would support a more general effect of language.

Method

Participants. Eighteen Northwestern undergraduates received course credit for their participation in this experiment.

Materials. The pictures were the same as those in Experiment 1a. The sentences on participants’ answer sheets were modified from those used in Experiment 1a by removing the prepositions, resulting in sentences of the following form:
The picture shows a block and a building.
The picture shows a plant and a shelf.

Procedure

The procedure was identical to that in the spatial sentences condition in Experiment 1a: participants saw pictures and circled the appropriate sentence, then did a ten-minute filler task, and finally were given a recognition test.

Results and Discussion

Participants who were given object sentences at encoding made more errors than did those told to pay close attention (Experiment 1b) or those given spatial sentences (Experiment 1a); in particular, they were more likely to have false alarms than participants in the attention condition (see Figure 2). This recognition disadvantage was confirmed by independent samples t-tests (total errors: spatial vs. object sentences: \( t(34) = -2.04, p < .05 \); total errors: attention vs. object sentences: \( t(34) = -3.35, p < .005 \); false alarms: attention vs. object sentences: \( t(34) = -3.82, p < .001 \)).

However, object sentence participants failed to show any recognition shift towards the core of the spatial category designated by the preposition. They demonstrated equal plus and minus false alarms, \( M=0.43 (SD = .18) \) and \( M=0.47 (SD = .15) \) respectively; \( t(17) = -0.94, p = .36 \), much as did the no-language participants in the previous studies. This pattern differed from the pattern shown by spatial sentences participants in Experiment 1a. This provides support for the suggestion that it is specifically the spatial preposition that is responsible for the pattern of responses observed in the language condition in Experiment
1a, a thinking-for-language result. The complete false alarm breakdown for Experiment 1 is presented in Figure 3.

---Insert Figure 3 about here---

*Discriminability Analysis.* In order to compare across the four conditions, two d' measures were calculated for each individual participant in Experiments 1b and 1c (as in Experiment 1a): one for the discriminability of the *minus* variant from the standard, and one for that of the *plus* variant from the standard. The larger of the two was then determined, and participants were pooled by condition (see Table 2). For both the attention group (Experiment 1b) and the object sentences group (Experiment 1c), we found that participants more often failed to discriminate the *minus* variant (which was designed to be perceptually more similar to the standard) than the *plus* variant; that is, the *minus* variant was more confusable with the standard than was the *plus* variant. This contrasts with the pattern found for the spatial sentences group (Experiment 1a), for which the *plus* variant was more confusable with the standard than was the *minus* variant. This difference in discriminability between conditions is significant, $\chi^2(6, N = 68) = 19.31, p < .01$.

----------Table 2 about here----------

As another comparison across the four conditions, we conducted a 4 (condition) x 2(d' type—plus or minus) analysis of variance over the d' data. As predicted, there was an interaction between condition and d' type, $F(3,64) = 5.67, p = .002$. Post hoc t-tests revealed that the standard was more confusable with the *plus* variant than with the *minus* variant for the spatial sentences group, $t(16) = -5.52, p < .0001$ by a Bonferroni correction; the other three conditions showed no difference. The interaction of condition by d' type did not reach significance in the items analysis, $F(3,36) = 2.34, p = .09$. (However, t-tests
showed the same pattern as for the subjects analysis: the standard was more confusable with the plus variant than with the minus variant for the spatial sentences group only, $t(7) = -4.81, p = .002$.

Experiment 2

In Experiment 1, we found that recognition memory for spatial scenes was influenced by spatial language presented at study. These results provide support for the Interactive Encoding claim that participants adjusted their encoding of the picture to better accord with the spatial preposition.

However, in Experiment 1 participants received all three recognition items (standard, plus and minus) for each triad (randomly ordered) in the yes/no recognition task. This leaves open the possibility that responses to a given item may have been influenced by having seen other items from the same triad during the recognition task. In Experiment 2, we changed the design of the recognition test to eliminate this possibility. We replicated the basic plan of Experiment 1a, but in the recognition test we presented each participant with only one recognition item from each triad. In addition, to ensure that the materials met the requirement of the design, we performed three norming studies.

Method

Design. The factors were Encoding Condition (Spatial Sentences/Control) (between-subjects), Recognition Item Type (Plus Variant/Standard/Minus Variant) (within-subjects), and Assignment condition (between-subjects).
Participants. One hundred eighteen Northwestern undergraduates received course credit for their participation. An additional 86 Northwestern undergraduates participated in the norming studies.

Materials. We preformed three norming studies on the materials used in Experiment 1. In norming study 2a, we asked 24 Northwestern University undergraduates to rate the applicability of the spatial sentences to each of the items in order to verify that the spatial sentences were most applicable to the plus variants and least applicable to the minus variants, with intermediate applicability to the standard pictures. We found that, as expected, participants gave the highest ratings to the plus variants ($M = 5.72, SD = .70$), followed by the standards ($M = 4.47, SD = 1.16$), and then by the minus variants ($M = 2.54, SD = 1.63$), $F(2,36) = 22.35, p < 0.0001$. However, examination of the results for individual triads showed a different pattern for two of the triads, one depicting a coin in a hand and one depicting a firefly in a dish. These sentences were adjusted accordingly for Experiment 2. The changed sentences are indicated with asterisks in Table 1.

In norming study 2b, we collected free descriptions of the standards (from 12 Northwestern University undergraduates) and plus variants (from 10 Northwestern University undergraduates) in order to check whether (1) the plus variants would elicit the target spatial prepositions from Experiment 1a; and (2) the standards would be less likely to elicit consistent spatial prepositions, indicating that they were indeed ambiguous with respect to the spatial sentences. As expected, the spontaneous descriptions varied considerably, particularly for the standards. The rate of use of the key spatial prepositions was .25 in descriptions of the standards ($SD = .25$). Also as expected, there was a trend for
the key prepositions to be used more often in descriptions of the plus variants \((M = .32, SD = .32)\), although this difference did not reach significance, \(t(12) = -1.01, p > .10\).

In norming study 2c, we collected ratings of the degree of similarity between the plus variant and the standard and between the minus variant and the standard from 40 Northwestern University undergraduates to verify that both variants were equally similar to the standard. We were especially concerned to rule out an imbalance in similarity in favor of the plus variant, as this could provide an alternate explanation for the plus false alarms. Although the overall ratings were comparable for the two alternatives, participants gave slightly higher similarity ratings to the minus variants \((M = 4.96, SD = 1.40)\) than to the plus variants \((M = 4.73, SD = 1.51)\), \(t(39) = -.47, p < .0001\). Across individual triads, the minus variant was judged more similar to the standard for five triads, the plus variant was judged more similar for three triads, and they were judged equally similar for five triads.

As a result of the norming studies, we made minor modifications to two of the triads of pictures for use in Experiment 2, and a change of preposition (from in to on) in the sentences corresponding to two others (indicated with asterisks in Table 1). One triad (depicting a balloon on a stick) was eliminated so that the number of triads would be divisible by three as required by the design.

**Procedure**

The study procedure was identical to that of Experiment 1a. Also as before, both conditions received the same yes/no recognition test after a ten-minute filled delay. However, each participant saw only one picture from each triad at recognition (instead of all three as in Experiment 1). These were presented in random order along with twelve fillers (six old and six new) on a computer screen. As in Experiment 1, participants were
asked to indicate on a printed answer sheet whether they had seen each picture during the earlier study portion. Each picture remained on the screen until the participant pressed a key.

Results

As in Experiment 1a, we found the predicted effect of spatial language. Participants in the spatial sentences condition were significantly more likely to false-alarm to the plus variant than to the minus variant, while participants in the control condition showed no such asymmetry (Figure 4). The difference between the plus false alarms and the minus false alarms is significant only in the spatial sentences condition, \( t(57) = 6.87, p < .0001 \). In addition, the difference in the rate of false alarms between the two groups only reaches significance for the responses to the plus variant, \( t(116) = -3.93, p = .0001 \).

---Insert Figure 4 about here---

Also as in Experiment 1a, we found no difference in the miss rates between the control and spatial sentences groups \( (t(116) = 1.46, p > .10) \). In contrast to Experiment 1a, the overall false alarm rates for the two groups in Experiment 2 did differ (Figure 5), with false alarms being more prevalent among the participants in the spatial sentences condition, \( t(116) = -1.77, p < .05 \) (one-tailed). As the analysis of the pattern of false alarms (above) shows, this difference is entirely attributable to a difference in responses to the plus variants.

---Insert Figure 5 about here---

**Discriminability Analysis.** As in Experiment 1, two d’ measures were calculated for each individual participant—one for minus and one for plus. The larger of the two was then determined, and participants were pooled by condition (Table 3).
The results of the d' analysis for Experiment 2 replicate those for Experiment 1. In the spatial sentences condition, the dominant pattern is one of greater discriminability between the standard and the \textit{minus} variant than between the standard and the \textit{plus} variant. However, in the control condition, an equal number of participants had a larger \textit{minus} d' as had a larger \textit{plus} d', a significant difference between conditions, $\chi^2(2, N = 118) = 16.67, p < .0001$. As a further check on the effect of condition on discriminability, the set of d' data was subjected to a 2x2 multivariate analysis of variance. In support of our conclusion, we found an effect of condition, $F(2,115) = 6.63, p = .002$.

Experiment 3

Our studies so far have provided evidence for the hypothesis that language interacts with visual input to influence the encoding of pictures. In Experiment 2, as in Experiment 1, the recognition results showed a shift towards the core of the spatial category when spatial language was present at encoding, consistent with the Interactive Encoding account.

However, before drawing strong conclusions, we need to address a possible alternate explanation for these effects, which we will term the \textit{Separate Encoding} account. Because the effect of language in Experiments 1 and 2 was to shift recognition in the direction of the specific terms present at encoding, it could be that the effect was simply due to memory for the language presented, rather than to differential encoding of the pictures. This line of reasoning has been articulated most clearly in research on memory overlay effects, including effects of verbal questioning on eye witness testimony (Loftus & Palmer, 1974; McCloskey & Zaragoza, 1985) and effects of language (Carmichael, Hogan, & Walter, 1932) and of verbal overshadowing (Schooler & Engstler-Schooler, 1990) on the
reproduction of visual materials. In this arena, the issue of whether language alters prior visual memories or merely exists as an alternate, possibly more available encoding has been the source of much debate. In Loftus and Palmer’s (1974) classic study, participants viewed a filmed automobile accident, then answered questions about what they had viewed. The results showed that the phrasing of the questions (e.g., use of the verb smash rather than the verb collide) influenced the likelihood that participants would falsely claim to have seen a given element (e.g., broken glass), suggesting that the subsequent language had altered participants’ memorial representation of the event. However, McCloskey and Zaragoza (1985) argued that the results could have come about without a direct effect of verbal information on pictorial memory. To demonstrate this, they presented participants with a scene which included a target object, A, and then referred to it as object B in postevent questioning. At test, participants were asked to determine whether they had seen object A, which was present in the scene, or object C, which was neither present in the scene nor mentioned in postevent discussion. They found that misleading information did not result in lower accuracy in this task (in contrast to the case when A was contrasted with B at test), from which they concluded that the verbal information had not changed participants’ memory for the original event. As this research shows, finding effects of language on retrieval of visually presented information does not necessarily imply effects of language on the visual representation itself. Such effects could also arise from two separately encoded and stored memory traces, if participants draw on their verbal memory to supplement their visual memory. On this account, the results of Experiments 1 and 2 (i.e., the elevation in plus false alarms) would be explained by people resorting to their stored language trace
when their memory for a picture is weak, rather than by an interactive encoding process by which language influences the encoding of the pictures.

One important difference from the present research is that McCloskey and Zaragoza’s discussion was directed at studies in which the visual materials were encoded first, with verbal descriptions following later. Thus they asked whether later verbal input can alter a prior visual memory, whereas our studies test for interactions of verbal and visual materials presented together. In other words, we are arguing for an interactive encoding process, as opposed to the retroactive alteration process that was the focus of McCloskey and Zaragoza’s critique. But obviously, the Separate Encoding explanation could still apply.

Experiment 3 was designed with two goals in mind: to test whether our language effect could be accounted for by the Separate Encoding possibility, and (if possible) to extend the scope of the language effect. Happily, there is a way to accomplish both these goals at once. Rather than testing for narrow, direct effects of language—i.e., whether participants’ picture encodings shift in the direction of the specific terms used—we will test for a more global effect, using a task inspired by McCloskey and Zaragoza’s (1985) technique. We ask whether participants will show indirect effects of using nonappropriate spatial terms. Specifically, Experiment 3 addresses the question of whether introducing a nonappropriate spatial term will call attention to spatial semantics more generally and lead to a shift towards the more appropriate spatial category.

Experiment 3a

In Experiment 3a, participants saw the standard pictures immediately preceded by sentences. Half the sentences were the key preposition sentences of the earlier studies. The
other half were clearly not appropriate—that is, they expressed a very different spatial
relation from the one that best fit the scene. For example, for the standard in Figure 1, the
key preposition sentence (the plausible description) was *The block is on the building* and
the nonappropriate preposition sentence was *The block is in the building*. Participants were
told to read the sentences and to remember the pictures for a later recognition task.
(Obviously, given this instruction, participants could have ignored the sentences; in
Experiment 3b we verified that they did not.) Then they were given a two-alternative
forced-choice recognition test between the actually presented standard and the *plus* variant
(incorrect). Based on our prior results, for pictures that were coupled with the key
preposition, we expected participants to (incorrectly) choose the *plus* variant. The question
is what will happen with the nonappropriate (nonapp) prepositions.

Suppose that the Interactive Encoding account is correct and further, that hearing
a spatial term calls forth not just the specific category associated with the term but, more
generally, the semantic set or dimension it belongs to. Then hearing a clearly
nonappropriate spatial preposition may prompt participants to consider which spatial
preposition *should* apply. That is, participants may think (for example) “This isn’t *in*, it’s
*on*.” They will substitute the correct (or most applicable) spatial relation—which, by
design, is the key prepositional category. If this reasoning holds, the effect of being given a
nonapp preposition should be (paradoxically) to *increase* false-alarms to the *plus* variants.

In contrast, the Separate Encoding account predicts the pattern found by
McCloskey and Zaragoza: no rise in false alarms for pictures paired with the nonapp
preposition. The nonapp-preposition sentences do not match *either* of the recognition
pictures. Thus, if participants encode the sentences and pictures separately, and during test
resort to the sentences when they forget the pictures (as in the Separate Encoding account), then for the nonapp preposition sentences, their sentence memory will be irrelevant to the choice between the standard and the \textit{plus} variant. They should therefore show no elevation in false alarms.

Thus, both accounts predict that participants will often false-alarm to the \textit{plus} variant when the picture has been coupled with the key preposition; but they make different predictions for pictures coupled with the nonapp prepositions. The Interactive Encoding account predicts a high false alarm rate for pictures presented with \textit{either} the key prepositions or the nonapp prepositions. The Separate Encoding account predicts a high false alarm rate \textit{only} for pictures presented with the key prepositions.

\textit{Method}

\textit{Participants}. Twenty Northwestern undergraduates received course credit for their participation.

\textit{Design and materials}. The factors were Encoding Condition (Key Preposition/Nonapp preposition—within-subjects) and Assignment condition (between-subjects). Each participant saw twelve standards and eight fillers—a total of twenty trials, of which half contained key preposition sentences and half contained nonapp preposition sentences.

The pictures and the key preposition sentences were as in Experiment 2. The nonappropriate preposition sentences were created from the key preposition sentences by replacing the preposition with another that was incompatible with both the standard picture and the \textit{plus} variant. In selecting these, we avoided nonapp prepositions that could name the \textit{minus} variant. For example, for the block and the building depicted in Figure 1, in
which the plus FA is *on* and the minus FA is *off*, we used *in* as the nonapp preposition. The complete set of sentences is presented in Table 4.

---Insert Table 4 about here---

*Procedure*

*Part 1: Study.* The pictures were presented individually for three seconds each on a computer screen in randomized order. Each picture was preceded by a sentence, which was displayed for one second. Participants were told that they would read sentences followed by corresponding pictures, and that they would later be asked to recognize only the pictures. For half of the pictures, the key spatial sentence (from Experiment 2) was presented; for the other half, a nonapp-preposition sentence was displayed. For the filler pictures, we used incorrect sentences that had the right preposition but the wrong noun (i.e., the distractor sentences from Experiment 1a). This was done to keep people from noticing a pattern that the incorrect sentences always erred in their spatial prepositions.

*Part 2: Recognition.* After a ten-minute filled delay, all participants received the same 2-AFC recognition test. Participants were shown each standard paired with its *plus* variant and asked to click on the picture they had seen during the study portion. The pictures remained on the screen until the participant made a choice.

*Results and Discussion*

The key prediction of the Separate Encoding account is that the false alarm rate will be higher for pictures that were coupled with key spatial prepositions than for those coupled with nonapp prepositions. In fact, there was a nonsignificant trend in the reverse direction (key prepositions: \( M = .12, SD = .17 \); nonapp prepositions: \( M = .16, SD = .15 \), \( t(19) \)
=.79, p > .4. Thus the results argue against the possibility that the language effects in Experiments 1 and 2 were simply due to use of independently encoded key preposition sentences and buttress the Interactive Encoding account of the language effect.

In addition, the results extend the Interactive Encoding account beyond purely local effects of words as category attractors. In this study, hearing nonappropriate spatial prepositions apparently led to an indirect language effect wherein participants shifted to the most appropriate spatial category, thus achieving a common construal of sentence and picture.

**Experiment 3b**

Experiment 3b was a manipulation check, conducted to verify that participants did indeed attend to the sentences. To do this, we gave 20 new participants the same materials as in Experiment 3a, but tested memory for sentences instead of memory for the pictures. The study phase was exactly as in Experiment 3a, including the instruction to remember only the pictures for a recognition test. However, after a ten minute filled interval, we administered a surprise sentence recognition test. Because we wanted to test for memory of all the sentences presented, we used a yes/no recognition test rather than a forced-choice task. Each participant was given a packet containing one sentence per study item, and responded yes or no to the question Did this sentence appear in part 1? Half the sentences were old (i.e., seen during the study phase) and half were new. Within each class, half contained the key preposition and half contained the nonapp preposition.

The results showed high sentence recognition accuracy (mean proportion correct = .83, SD = .16), allowing us to rule out the possibility that the sentences were ignored. Memory for the key preposition sentences (M = .88, SD = .16) was nonsignificantly better
than memory for the nonapp preposition sentences ($M = .78, SD = .13$). Thus the failure to find a difference in the false alarm rates for the two encoding conditions in Experiment 3a is not due to participants’ having failed to process the sentences.

Interestingly, we also found evidence that Interactive Encoding works both ways. People were more likely to misrecognize the sentence when a nonapp preposition had been seen at study ($M = .28, SD = .27$) than when a key preposition had been seen ($M = .08, SD = .21$), $t(19) = 2.26, p < .05$, suggesting that the sentences were adjusted to fit the pictures as well as the pictures to fit the sentences.

Discussion

According to Interactive Encoding, people’s encoding of pictures can be influenced by accompanying language. This hypothesis received initial support from our finding that picture recognition was influenced by the presence of spatial language during encoding in Experiments 1 and 2. In Experiment 3a, we found evidence that language effects can extend beyond the application of specific categories. Our results suggest that the language effect may often be to focus attention on a particular dimension, rather than to impose the stated value on that dimension.

This study also tested a possible alternate explanation for the effect of language, that the pictures and sentences were encoded separately and that the effect of language on recognition resulted from participants’ consulting their sentence memory when pictorial memory is weak. Contrary to the prediction of the Separate Encoding account (but consistent with the Interactive Encoding account), there was no difference in false alarm rates as a function of sentence type; indeed, there was a nonsignificant imbalance in the opposite direction. Experiment 3b verified that the sentences were processed and retained.
The study further showed that sentences that were inconsistent with the pictures (those with nonappropriate prepositions) were more likely to be altered than those that were more consistent with the pictures. Both these findings are in line with the Interactive Encoding hypothesis.

We further examined the data across all three experiments for item effects. However, we found no systematic pattern across studies as to which items showed the greatest tendency for picture memory to be affected by verbal labels, nor for this effect to correlate (or anticorrelate) with the degree to which sentence memory was affected by the spatial scenes.

The results help refine and extend the thinking-for-language account. The effect of language in our studies was to focus attention on the spatial relations and to alter both the encoding of the scene and the encoding of the sentence in the direction of the most consistent prepositional category. When a linguistic assertion disagrees with the perceptual facts, the effect can be to move the encodings to a different value within the same semantic set—one where the language and the scene can agree.

General Discussion

Our studies provide evidence that spatial language can influence the way people encode and remember spatial relations in visual scenes. Our standards were designed to be unlikely members of the key spatial category (indeed, the spatial preposition was rarely used in spontaneous descriptions; see norming study 2b). Yet people given spatial prepositions during encoding of the standards showed a shift in picture recognition towards the core of the spatial category denoted by the preposition (Experiments 1a and 2). That is, they were significantly more likely to false-alarm to pictures closer to the center of the
prepositional category than to those less central. Participants who received the pictures without spatial sentences at study showed no such shift, even when asked to pay close attention to the pictures; their false-alarm rates were symmetrical (Experiments 1b and 2). Finally, in Experiments 3a and 3b, we ruled out the possibility that these language effects simply reflect the use of a separately encoded verbal memory, rather than an interactive encoding of sentence and picture. Contrary to the Separate Encoding prediction, the misrecognition effect was not restricted to the category directly named in the sentence. Rather, we found that hearing even a nonappropriate spatial preposition induced a shift towards the key prepositional category. This finding is incompatible with the use of an independently encoded verbal store. However, this pattern is quite compatible with the general Interactive Encoding account.

These studies link two bodies of work that have largely been considered independently. One is the neo-Whorfian investigations on effects of language on perceptual and conceptual representation and reasoning reviewed in the Introduction. The other is research on memory overlay effects. In this arena, the issue of whether language alters prior visual memories or merely exists as an alternate, possibly more available encoding has been the source of much debate. Our studies show that when spatial language is presented simultaneously (or just prior) to visual materials, the interaction between linguistic categories and visual materials can influence the encoding of spatial relations in both the pictures and the sentences.

Not all spatial tasks are susceptible to language effects. For example, Huttenlocher and her colleagues used a location reproduction task (Huttenlocher, Hedges, & Duncan, 1991; Crawford, Regier, & Huttenlocher, 2000), in which participants saw a dot and a
quadrilateral reference object; the dot (but not the quadrilateral) then disappeared, and participants indicated where on the screen the dot had been. Participants’ errors were skewed toward the center of an inferred geometric category, even after rating the applicability of a (different) linguistic category to the array. One difference between the two studies is that in the Huttenlocher et al. studies, memory was probed immediately after the disappearance of the dot, whereas we imposed a ten-minute delay. This delay may have contributed to the language effect. In addition, we speculate that the language effects may be more likely with naturalistic materials like those used in our task than with simple geometric objects such as those used by Huttenlocher and her colleagues.

Our finding of an interaction between verbal and visual encodings can also be related to recent accounts of language comprehension that propose that sentences are understood by constructing perceptual simulations of the events described (Barsalou, 1999; Glenberg, 1997; Zwaan, 2004). Such accounts propose a close interaction between perceptual and semantic processing. For example, Zwaan and Yaxley (2003) found that spatial iconicity affects semantic relatedness judgments: e.g., people were faster to judge BASEMENT and ATTIC as semantically related words when they saw ATTIC above BASEMENT than when they saw BASEMENT above ATTIC. Zwaan, Madden, Yaxley, & Aveyard (2004), have also shown effects of language on perceptual judgments. For example, people were shown pictures of objects and asked to judge whether the second object was the same as the first. Before each picture pair, participants heard sentences such as "You tossed the beach ball over the sand toward the kids" or "The kids tossed the beach ball over the sand toward you". Responses were faster when the meaning of the sentence matched the size difference of the objects. For example, if the sentence described a ball
moving towards the viewer, people were faster to respond when the second picture showed a ball that was larger than the first. The authors suggest that in this case the thoughts triggered by the sentence were compatible with the size discrepancy between the balls, allowing a fast response. They concluded that "...words can, indeed, move mental representations."

According to this view, we activate positional and orientational information when presented with linguistic stimuli (e.g., Zwaan et al., 2004; Zwaan & Yaxley, 2003). In the case of spatial prepositions, this information might best correspond to core members of the spatial prepositional categories, leading to the observed shift in picture representation towards the center of the prepositional category when a spatial preposition has been processed during encoding.

To return to the neo-Whorfian literature, our findings are most compatible with a thinking-for-language hypothesis (see Slobin, 1996) whereby language influences thought when one is either producing or comprehending language. Across our studies, we saw no evidence that people engaged in the covert use of language to encode the pictures, even when they were told to pay close attention. (Of course, our design, with its deliberately ambiguous standards, may have underestimated the effects of language. The low rate of spontaneous use of the key prepositions in the free description task (norming study 2b) suggests that even if participants did encode the pictures linguistically in the attention task, they were unlikely to have used our spatial sentences.\textsuperscript{4}) However, overall, we have evidence for language effects only when language is present.

*Extended effects.* We found that processing a clearly nonappropriate preposition just prior to a scene led to a kind of contrast effect: participants encoded the scene in terms of a
competing (and more applicable) spatial prepositional category. This suggests an indirect effect of language, rather than a simple direct substitution of the named concept into the scene. Here, the use of a linguistic term appears to call forth a system of semantic categories, from which participants choose the most appropriate member. This effect is both more subtle and potentially more pervasive than a simple direct insertion effect.

However, these results invite further questions. Why did we find indirect effects of extraneous language when McCloskey and Zaragoza did not? (Recall that McCloskey and Zaragoza (1985) found that retroactive false information concerning a presented object did not impair participants’ ability to correctly recognize the object they had viewed, as opposed to a new, previously unmentioned, object.) One possibility is that language effects—both direct insertion effects and indirect semantic system effects—are more likely when language is presented before or during the encoding of the perceptual materials (as in our studies) than when the language occurs after the perceptual materials have been encoded (as in McCloskey and Zaragoza’s studies). This would be consistent with the Interactive Encoding account.

Another intriguing possibility is that a further contributor to the difference between the two studies is the kinds of items used. McCloskey and Zaragoza used object substitution: e.g., they showed a hammer and referred verbally to a screwdriver. In contrast, our manipulation involved a change of spatial relation. There is reason to believe that relational terms such as prepositions (and many verbs and adjectives) are more likely to belong to systems of interrelated semantic categories than are concrete nouns (Gentner & Boroditsky, 2001; Huttenlocher & Lui, 1979). It seems plausible that people are more likely to see such interrelated categories as contrastive and to shift from one to another. We
might therefore expect the extended effects explored here to be strongest for relational
terms, and much weaker (if they occur at all) for object terms. A question for future
research is whether and how the technique explored here would apply across different
kinds of semantic systems.

Finally, we note that the meanings of spatial relational terms incorporate nonspatial
factors such as support and containment relations, intended function, etc. (Bowerman &
Pederson, 1992; Coventry & Garrod, 2003; Coventry, Prat-Sala, & Richards, 2001; Feist,
2000, 2004; Feist & Gentner, 2003; Talmy, 1988; Vandeloise, 1991) in addition to
geometry. Given that spatial meaning is in fact quite complex, a question for further
research is how these factors interact with geometric factors in the encoding of spatial
scenes.

Thinking for language. This work fits with recent discussions that have made finer
distinctions within the language and thought arena. Gentner and Goldin-Meadow (2003)
distinguish three versions of the claim that language influences thought: language as lens;
language as tool kit; and language as category shaper. The language as lens view is the
strong “Whorfian hypothesis” of linguistic determinism: that the language we acquire
determines how we perceive and represent the world. On the language as category shaper
view, human categories are relatively universal, but language can influence the boundaries
(Papafragou et al., 2002), implying enduring if limited effects on cognition. We espouse
the third view, language as tool kit: that acquiring a language provides new
representational resources that augment the capacity for encoding and reasoning (Dennett,
1993; Gentner, 2003; Vygotsky, 1962). On this view, language provides tools that facilitate
forming and holding particular construals, but it does not replace all other encoding formats (Gentner & Goldin-Meadow, 2003).

Slobin’s (1996) thinking-for-speaking (or more broadly, thinking-for-language) hypothesis is essentially orthogonal to these distinctions, as it concerns when language effects occur rather than what the effects are. However, it is most compatible with versions 2 and 3: when conversing, people are particularly likely to encode and reason using the semantic tools provided by their language, and to honor the semantic distinctions of the language. Our findings are compatible with such a thinking-for-language account.

Although this finding is less dramatic than those predicted by the strong Whorfian hypothesis, it still leaves room for pervasive semantic effects, given the ubiquity of external (and internal) language. As Pinker (1989: 360) states "Whorf was surely wrong when he said that one's language determines how one conceptualizes reality in general. But he was probably correct in a much weaker sense: one's language does determine how one must conceptualize reality when one has to talk about it."
References


Coventry, K., Prat-Sala, M., & Richards, L. (2001). The interplay between geometry and function in the comprehension of *over, under, above,* and *below.* *Journal of Memory and Language, 44*, 376-398.


This work was supported by NSF-LIS award SBR-9720313 and NSF-ROLE award 21002/REC-0087516 and completed while the first author was a postdoctoral fellow at Northwestern University. We thank Kathleen Braun and Michelle Osmondson for help in conducting and analyzing the study, Jonathan Cohen and Benjamin Scott-Hopkins for help in preparing the manuscript, and Beth Levin, Jason Jameson, Jeff Loewenstein, and Phillip Wolff for helpful discussions of the ideas and methods. We also thank Satoru Suzuki for the d' analysis program, Jeff Rouder for discussions of the statistical analyses, and Melissa Bowerman for allowing us access to her materials. Finally, we thank Karen Emmorey, Nora Newcombe, Rolf Zwaan and an anonymous reviewer for helpful comments on earlier versions of this paper. Correspondence concerning this article should be addressed to Michele I. Feist, University of Louisiana at Lafayette, Institute of Cognitive Science, P. O. Drawer 43772, Lafayette, LA 70504-3772 (email feist@louisiana.edu).
Footnotes

1 Three of the triads (those depicting a ball under a chair, a chair in a corner, and a hose around a tree trunk) were adapted from drawings in Melissa Bowerman’s topological picture series.

2 d' measures within .25 of one another were considered equal.

3 For the triad depicting a firefly and a wagon wheel, the firefly was changed to make its wings visible, making it appear to be flying, the wood grain on the wheel was changed to be more realistic, and the background color was changed from grey to light blue. For the triad depicting a chair and a corner, the lengths of the walls were adjusted to be equal in all three pictures and the distance between the chair and the corner was increased in the standard and minus variant, making the pictures more discriminable.

4 We thank an anonymous reviewer for suggesting this possibility.
Table 1

*Sentences used in Experiments 1 and 2. Sentences marked with * were altered for Experiment 2.*

<table>
<thead>
<tr>
<th>Target sentences</th>
<th>Distractor sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>The firefly is on the wagon wheel.</td>
<td>The watch is on the nightstand.</td>
</tr>
<tr>
<td>The puppet is on the table.</td>
<td>The computer is on the desk.</td>
</tr>
<tr>
<td>The chair is in the corner.</td>
<td>The house is in the middle.</td>
</tr>
<tr>
<td>The dirt is on the dump truck.</td>
<td>The snow is on the mountain.</td>
</tr>
<tr>
<td>The ball is under the chair.</td>
<td>The table is under the lamp.</td>
</tr>
<tr>
<td>The block is on the building.</td>
<td>The plant is on the shelf.</td>
</tr>
<tr>
<td>The balloon is on the stick.</td>
<td>The kite is on the string.</td>
</tr>
<tr>
<td><em>The firefly is in the dish.</em></td>
<td>The house is in the valley.</td>
</tr>
<tr>
<td><em>The coin is in the hand.</em></td>
<td>The spaceship is in the crater.</td>
</tr>
<tr>
<td>The spider is in the bowl of apples.</td>
<td>The shirt is in the basket of laundry.</td>
</tr>
<tr>
<td>The plane is on the ground.</td>
<td>The woman is on the floor.</td>
</tr>
<tr>
<td>The hose is around the tree trunk.</td>
<td>The armband is around the arm.</td>
</tr>
<tr>
<td>The balloon is on the table.</td>
<td>The antenna is on the roof.</td>
</tr>
</tbody>
</table>
Table 2

Participants Pooled According to the Discriminability Analysis, Experiment 1 abc

<table>
<thead>
<tr>
<th>Group</th>
<th>d’ larger for plus</th>
<th>d’ larger for minus</th>
<th>d’ equal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>4</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Spatial Sentences</td>
<td>0</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Attention†</td>
<td>8</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Object Sentences‡</td>
<td>6</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

† One participant was dropped from the analysis due to a hit rate of 1. ‡ One participant was dropped from the analysis due to a hit rate of 1 and one was dropped due to a plus false alarm rate of 0.
Table 3

*Participants Pooled According to the Discriminability Analysis, Experiment 2*

<table>
<thead>
<tr>
<th>Group</th>
<th>d' larger for <em>plus</em></th>
<th>d' larger for <em>minus</em></th>
<th>d' equal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Spatial Sentences</td>
<td>4</td>
<td>38</td>
<td>16</td>
</tr>
</tbody>
</table>
Table 4

*Sentences used in Experiment 3.*

<table>
<thead>
<tr>
<th>Key preposition sentences</th>
<th>Nonapp preposition sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>The firefly is on the wagon wheel.</td>
<td>The firefly is under the wagon wheel.</td>
</tr>
<tr>
<td>The puppet is on the table.</td>
<td>The puppet is under the table.</td>
</tr>
<tr>
<td>The chair is in the corner.</td>
<td>The chair is below the corner.</td>
</tr>
<tr>
<td>The dirt is on the dump truck.</td>
<td>The dirt is under the dump truck.</td>
</tr>
<tr>
<td>The ball is under the chair.</td>
<td>The ball is on the chair.</td>
</tr>
<tr>
<td>The block is on the building.</td>
<td>The block is in the building.</td>
</tr>
<tr>
<td>The firefly is on the dish.</td>
<td>The firefly is under the dish.</td>
</tr>
<tr>
<td>The coin is on the hand.</td>
<td>The coin is under the hand.</td>
</tr>
<tr>
<td>The spider is in the bowl of apples.</td>
<td>The spider is under the bowl of apples.</td>
</tr>
<tr>
<td>The plane is on the ground.</td>
<td>The plane is beneath the ground.</td>
</tr>
<tr>
<td>The hose is around the tree trunk.</td>
<td>The hose is beneath the tree trunk.</td>
</tr>
<tr>
<td>The balloon is on the table.</td>
<td>The balloon is below the table.</td>
</tr>
</tbody>
</table>
Figure Captions

*Figure 1.* Example picture triads together with the sentence used for each triad. In each case, the *plus* variant is closer than the standard to the center of the spatial category conveyed by the sentence and the *minus* variant is farther.

*Figure 2.* Error rates, Experiment 1abc

*Figure 3.* False alarms by condition, Experiment 1abc

*Figure 4.* False alarms by condition, Experiment 2

*Figure 5.* Error rates, Experiment 2
Figure 1

*Plus variant*  *Standard*  *Minus variant*

The block is on the building.

The balloon is on the table.
Figure 2

![Graph showing rate of occurrence for different categories]

- Error rate
- Miss rate
- Overall false alarm rate

Categories:
- Spatial sentences
- Control
- Attention
- Object sentences
Figure 3

![Graph showing rate of false alarms for spatial sentences, control, attention, and object sentences. The x-axis represents picture variants (plus and minus), and the y-axis represents the rate of false alarms. The graph illustrates the differences in false alarm rates across different conditions.]
Figure 4

![Graph showing the rate of false alarms for control and spatial sentences across plus and minus picture variants. The y-axis represents the rate of false alarms ranging from 0 to 0.9, and the x-axis represents the picture variant (plus and minus). The graph includes error bars for each condition.]
Figure 5

![Error rate, miss rate, overall false alarm rate comparison graph]

- Error rate
- Miss rate
- Overall false alarm rate

Comparison between spatial sentences and control conditions.