Feist, M. I., & Gentner, D. (2001). An influence of spatial language on recognition memory for spatial scenes. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 279-284). Mahwah, NJ: Lawrence Erlbaum Associates.

An Influence of Spatial Language on Recognition Memory for Spatial Scenes

Michele I. Feist (m-feist@northwestern.edu)
Department of Psychology, Northwestern University
2029 Sheridan Road, Evanston, IL 60208 USA

Dedre Gentner (gentner@northwestern.edu)
Department of Psychology, Northwestern University
2029 Sheridan Road, Evanston, IL 60208 USA

Abstract

Whether and how much the routine use of language influences thought is a perennially fascinating question in cognitive science. The current paper addresses this issue by examining whether the presence of spatial language influences the encoding and memory of simple pictures.

Introduction

In the last few years there has been a resurgence of interest in the question of whether and how much language influences thought. As Billman and Krych (1998) point out, this is a question that can be asked either at the level of the language system, or at the level of the linguistic form.

At the level of the language system, one can ask whether cognitive differences can be explained via cross-linguistic differences. The strong version of this hypothesis is well expressed in Whorf's (1956, p. 134) 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." Other scholars suggest a weaker version of the hypothesis, namely that language, while not determining thought, nonetheless influences how one thinks. Slobin's (1996) *thinking-for-speaking* hypothesis states that linguistic influences exist only when one performs a linguistically-mediated task (cf., Slobin, 1996).

Evaluation of the hypothesis at the level of the language system involves an examination of performance on non-linguistic tasks by speakers of different languages in order to determine whether there are language-related differences. Such examinations have yielded mixed results. Pederson and his colleagues (1998) and Levinson (1996) found that speakers of different languages performed differently on nonlinguistic tests of visual memory, including reconstruction of an array of objects, a clearly Whorfian result. Malt, Sloman, and Gennari (in press), on the other hand, found that Spanish speakers' judgments of similarity of videotaped motion events conformed to normal verb use in Spanish, but only when participants were instructed to use linguistic descriptions during the encoding phase of the experiment. This is consistent with a thinking-for-speaking (Slobin 1996) version of the Sapir-Whorf hypothesis. Furthermore, the language effect did not appear for the English-speaking participants, nor did Malt and her colleagues find a language effect on similarity judgments for artifacts, nor on recognition memory.

The other level at which language could influence thought is that of linguistic forms within a language. Evaluation of the hypothesis at this level involves comparing performance on non-linguistic tasks by speakers of the same language in conditions that invite different forms within the language. For example, Bower, Karlin, and Dueck (1975) found that participants rated new pictures as more similar to the one they had seen during encoding if they conformed to the linguistic description presented at encoding. Gentner and Loftus (1979) found an influence of the language presented at encoding on participants' recognition memory for pictures of events. Billman and Krych (1998) found effects of verbs present at encoding on recognition of videotaped motion events (but see Malt et al., in press).

Our research asks whether spatial prepositions can influence the way people encode and remember spatial relations. We chose spatial prepositions for several reasons. First, while many studies of the Whorfian question have focused on possible effects of verbs of motion on the encoding of events, there has been comparatively little work on the possible effects of prepositions on the encoding of static spatial relations. Spatial prepositions exhibit striking cross-linguistic variability, as demonstrated by Bowerman and Pederson's (in preparation) comparative study of the semantics of '*on*-terms' – terms related to contact and support. As Gentner (1981; Gentner & Boroditsky, 2001) points out, relational terms such as verbs and prepositions are a promising arena in which to seek Whorfian evidence. Relational terms are more variable cross-linguistically than nominal terms of comparable concreteness. This semantic variability suggests that there is a wide variety of plausible encodings consistent with the perceptual input. Thus, this arena may provide fruitful ground for the investigation of Whorfian effects.

In this research, we showed people spatial scenes under different linguistic encoding conditions, and later tested their recognition memory. Our goal was to determine (1) whether

spatial language influences spatial encoding and memory and (2) whether such influence occurs when there is no overt use of language, or is restricted to the case when spatial language is explicitly present. If we see language effects only when people are encouraged to utilize language at encoding, this will provide support for a thinking-for-speaking or, in our case, *thinking-for-comprehending* hypothesis. If, on the other hand, we see language effects under other conditions, this would leave open the possibility of language influencing cognition in a more comprehensive manner.

The logic of our studies is as follows. For each of the prepositions, we created a sentence and a triad of pictures that ranged in how well they fit the sentence (see Figure 1). The standard picture (the *initial* picture) was acceptably described. For each standard, there were two variants: the *plus* variant, which was a better exemplar of the spatial term, and the *minus* variant, which was a poorer exemplar (see Figure 1 below). Thus, the initial picture was somewhat ambiguous, but was designed so that the spatial term could apply to it, and the two variants were either more typical of the core prepositional category or less so. All of the 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 Experiment 2).

Experiment 1a

Participants viewed pictures depicting static spatial relations - e.g., a marionette standing on a table or a coin in a hand. Half the participants read a descriptive sentence at the time that the pictures were encoded. After participating in unrelated experiments for about fifteen minutes, participants performed a recognition task that included the original pictures and two variants.

The recognition test included all three pictures - the initial picture, the plus variant, and the minus variant. If the presentation of language at encoding influences recognition memory, there should be different patterns of false alarms for the two groups. The group provided with sentences at encoding should be more likely than the control group to falsely claim that they had previously seen the plus variants of the pictures.

Method

Design. Encoding Condition (Spatial Sentences/Control), a between-subjects variable, was crossed with Recognition Item Type (Plus Variant/Initial Picture/Minus Variant), a within-subject factor.

Subjects. Thirty-six Northwestern undergraduates received course credit for their participation in this experiment. All reported being fluent speakers of English.

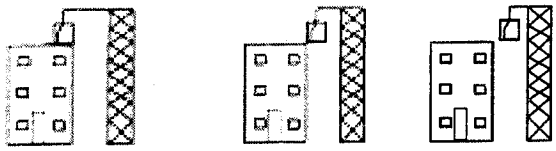
Stimuli. Thirteen triads of pictures and corresponding sets of sentences were created for this experiment. As discussed above, the pictures were created such that one might be well described by a target sentence, one passably described, and one poorly described. Each triad of pictures was associated with a pair of sentences: the target sentence that described the picture as outlined above, and a distracter sentence in which only the nouns were changed. The distracter sentence was meant to be obviously wrong; its purpose was simply to force participants to read the correct sentence and encode the target spatial relational term. For example, for the picture in Figure 1, participants chose between *The block is on the building* and *The plant is on the shelf.*

The initial picture from each triad was used for the study portion of the experiment; all three pictures in the triad were used for the recognition task.

Procedure

Part 1: Study. Twenty-five pictures (thirteen targets and twelve distracters) were randomized and presented individually for five seconds each on a computer screen. All participants were told that this was part one of a two-part experiment.

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 distracter sentence. Participants in the control condition were given no additional instructions.



Plus variant *Initial picture* *Minus variant*

Figure 1: Triad of pictures corresponding to the sentence "The block is on the building."

Part 2: Recognition. All participants received the same yes/no recognition task. All three of the pictures in each triad were presented individually in random order along with twelve distracters (six old and six new). Participants were asked to indicate on a numbered answer sheet whether or not they had seen each picture during the earlier study portion. Each picture remained on the screen until the participant pressed the "c" key, indicating that they were ready to continue.

Results

As predicted, we found that participants' recognition memory was influenced by whether a linguistic description was presented during study. Participants in the spatial sentences condition were significantly more likely to false-alarm to the *plus* variant than to the *minus* variant. (Figure 2). The difference between the false alarms in response to the *plus* variant and the false alarms in response to the *minus* variant differs significantly in the spatial sentences condition, as confirmed by a paired samples t-test ($t(17) = 5.32$, $p < .0001$). Participants in the control condition showed no such difference in their false alarm rate. Thus,

having spatial language present at encoding led to a skewing of recogntion errors towards the core of the spatial category.

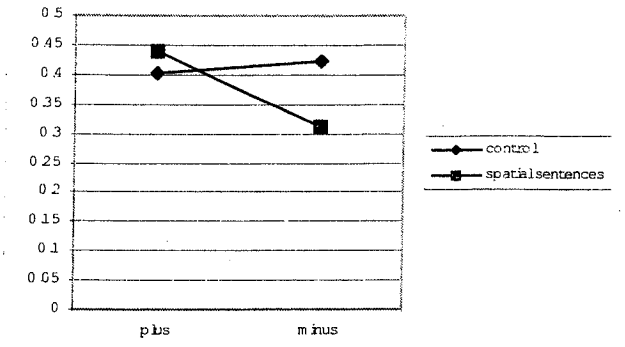


Figure 2: False alarms by condition, Experiment 1a

d' analysis To further test the claim that the presentation of sentences during study influences recognition memory for pictures, two d' measures were calculated for each individual subject. One d' indicates the discriminability of the *minus* variant and the initial picture; the other, the discriminability of the *plus* variant and the initial picture. The larger of the two was then determined, and the participants were pooled by condition, as shown in Table 1.

Table 1: Participants pooled according to the d' analysis, Experiment 1a

	Plus larger	*Minus* larger	Equal[1]
Control	4	4	10
Spatial Sentences	0	12	5

In the spatial sentences condition, but not in the control condition, the discriminability of the *minus* variant is greater than that of the *plus* variant (X^2=9.65, p<.01).

Discussion

We found that when spatial language was present at encoding, memory for the spatial relations in the pictures was systematically shifted in the direction of the spatial preposition. This is evidence for at least the moderate thinking-for- speaking version of the Whorfian hypothesis. In the next study we sought evidence for the strong version of the hypothesis. We hypothesized that if people had to attend closely to the pictures, this might evoke spontaneous linguistic descriptions as a memory aid. We thus examine the effect of more careful attention on recognition memory in Experiment 1b.

[1] d' measures within .25 of one another were considered equal for the analyses discussed in this paper.

Experiment 1b

In this study we asked whether participants instructed to pay careful attention to the pictures at study might be induced to encode the pictures linguistically and, as a result, to display an error pattern similar to that seen in the spatial sentences condition of Experiment 1a.

Method

Subjects Eighteen Northwestern undergraduates received course credit for their participation in this experiment. All reported being fluent speakers of English.

Stimuli The stimuli used were the same as those in Experiment 1a.

Procedure

Part 1: Study The procedure was identical to 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.

Part 2: Recognition The recognition task was the same as that used in Experiment 1a.

Results and Discussion

The error rate observed in Experiment 1b is lower than that observed in Experiment 1a, indicating that participants did pay more careful attention to the pictures during study. However, the pattern of false alarms is the same as that observed for the control subjects from Experiment 1a. Figure 3 shows the results of Experiment 1b along with those of Experiment 1a. These results suggest that more careful attention did not necessarily evoke linguistic encoding.

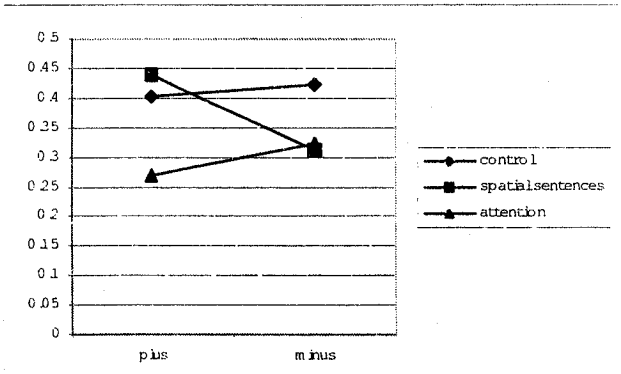


Figure 3: False alarms by condition, Experiments 1a and 1b

So far we have evidence for the influence of spatial language when it is explicitly presented, although not for the stronger possibility that language will affect cognition even when it is not overtly present. In Experiment 1c, we tested the specificity of the language effect. If, as we have assumed, 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

In order to more carefully inspect the source of the language effect from Experiment 1a, we presented participants with sentences without spatial prepositions at encoding. The sentences used named only the objects in the picture. We predict that these sentences, which are missing the hypothesized source of the language effect, will not replicate the effect found in Experiment 1a.

Method

Subjects Nineteen Northwestern undergraduates received course credit for their participation in this experiment. All reported being fluent speakers of English.

Stimuli 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

Part 1: Study The procedure was identical to that in the spatial sentences condition in Experiment 1a. Participants chose which sentence best matched the picture.

Part 2: Recognition The recognition task was the same as that used in Experiment 1a.

Results and Discussion

As predicted, participants failed to show any shift towards the core spatial category designated by the preposition. The participants in Experiment 1c demonstrated the same pattern of equal *plus* and *minus* false alarms as the no-language subjects in the previous studies (the subjects in Experiment 1b and the control subjects in Experiment 1a). This pattern differed significantly from the pattern by spatial sentence subjects in Experiment 1a. Specifically, the two groups differed in their rate of false alarms in response to the minus variant (independent samples t-test: $t(34) = 3.91$, $p<.005$). This provides support for the suggestion that it is specifically the preposition that is responsible for the change in the pattern of responses observed in the spatial sentences condition in Experiment 1a. The complete set of results for Experiment 1 is presented in Figure 4.

d' analysis As in Experiment 1a, two d' measures were calculated for each individual participant in Experiment 1: one indicated the discriminability of the *minus* variant and the initial picture, and one indicated the discriminability of the *plus* variant and the initial picture. The larger of the two was then determined, and the participants were pooled by condition (Table 2).

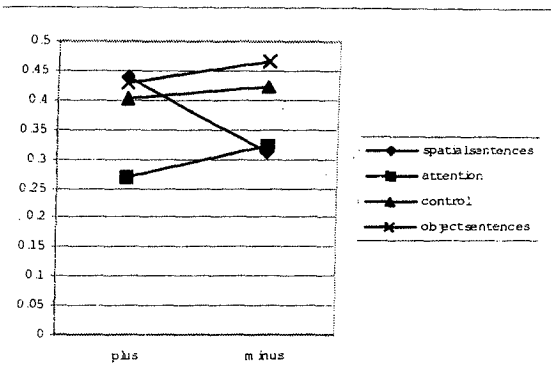


Figure 4: False alarms by condition, Experiment 1

Table 2: Participants pooled according to the d' analysis, Experiment 1

	Plus larger	*Minus* larger	Equal
Control	4	4	10
Spatial sentences	0	12	5
Attention	8	4	5
Object sentences	6	3	7

In the spatial sentences condition only, the discriminability of the *minus* variant is greater than that of the *plus* variant ($X^2=19.31$, $p<.01$). Or to put it more directly, only in the spatial sentences condition is the *plus* version more confusable with the initial picture than the *minus* version.

Experiment 2

This study was done to verify that the spatial sentences applied to the three variants of each picture as expected. We asked participants to rate the applicability of the sentences from the study portion of Experiment 1a to each of the pictures.

Method

Subjects Twenty-four Northwestern undergraduates received course credit for their participation in this experiment. All reported being fluent speakers of English.

Stimuli The pictures used were the same as those in Experiment 1. The sentences used were the correct spatial sentences from Experiment 1a.

Procedure

All three of the pictures in each triad were presented individually in random order along with the twelve distracters from the recognition task from Experiment 1. Participants were asked to rate the applicability of the sentences to the pictures on a scale from one to seven, with seven being the highest rating. Each picture remained on the screen until the participant pressed the "c" key, indicating that they were ready to continue.

Results and Discussion

As expected, participants gave the highest ratings to the *plus* variants (mean rating 5.72), in-between ratings to the initial pictures (mean rating 4.47), and the lowest ratings to the *minus* variants (2.54). This distribution of the ratings suggests that the assignment of pictures to the various categories with respect to the sentences used in the spatial sentences condition of Experiment 1a was indeed appropriate. Examination of the results for individual triads showed that for two of the triads, one depicting a coin in a hand and one depicting a firefly in a dish, the sentences did not fit exactly as predicted. These sentences were adjusted accordingly for Experiment 3.

Experiment 3

This study was a replication of the spatial language condition, with a methodological improvement. In Experiment 1a, participants saw all three versions of each of the pictures (one at a time) during the yes/no recognition task. This leaves open the possibility of carryover effects from one variant to another. In Experiment 3, the study task was that of Experiment 1a, but the recognition task was designed so that each participant was tested on only one version of each picture.

Method

Design. Encoding Condition (Spatial Sentences/Control), a between-subjects variable, was crossed with Recognition Item Type (Plus Variant/Initial Picture/Minus Variant) (within-subjects) and with Assignment condition. This was a between-subjects variable determining which variant in each set was received by a given participant in the recognition test.

Subjects. One hundred eighteen Northwestern undergraduates received course credit for their participation in this experiment. All reported being fluent speakers of English.

Stimuli. The stimuli used were the same as those in Experiment 1, with minor modifications to two of the triads of pictures, and with a change of preposition (from *in* to *on*) in the sentences corresponding to two others. One of the triads used in Experiment 1, depicting a balloon on a stick, was not used for Experiment 3.

Procedure

Part 1: Study The procedure was identical to the study portion of Experiment 1a.

Part 2: Recognition Both conditions received the same yes/no recognition task. One picture from each triad was presented in random order along with twelve distracters (six old and six new). As in Experiment 1, participants were asked to indicate whether or not they had seen each picture during the earlier study portion, and each picture remained on the screen until the participant pressed the "c" key indicating readiness to continue.

Results

As in Experiment 1a, we found that participants' recognition memory was influenced by the presence or absence of spatial language during study. The pattern of false alarms for the spatial sentences condition differs from that in the control condition (Figure 5). As in Experiment 1a, participants in the spatial sentences condition were significantly more likely to false-alarm to the *plus* variant than to the *minus* variant. Participants in the control condition showed no such difference in their false alarm rate. The difference between the false alarms in response to the *plus* variant and the false alarms in response to the *minus* variant differs significantly only in the spatial sentences condition, as confirmed by a paired samples t-test ($t(57) = 2.23$, $p = .047$). In addition, the difference in the rate of false alarms between the two groups only reaches significance for the responses to the *plus* variant, as confirmed by an independent samples t-test ($t(116) = 2.20$, $p = .039$).

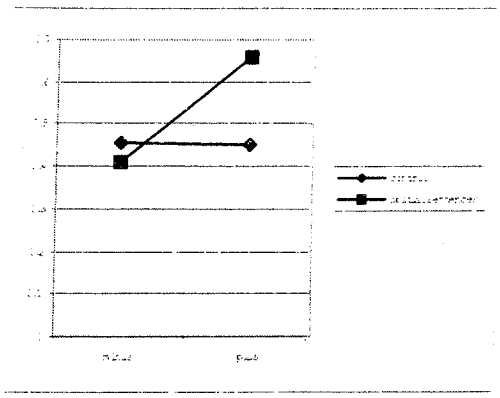


Figure 5: False alarms by condition, Experiment 3

d' analysis As in Experiment 1a, two d' measures were calculated for each individual subject. One d' indicates the discriminability of the *minus* variant and the initial picture; the other, the discriminability of the *plus* variant and the initial picture. The larger of the two was then determined, and the participants were pooled by condition (Table 3).

Table 3: Participants pooled according to the d' analysis, Experiment 3

	Plus larger	*Minus* larger	Equal
Spatial sentences	4	38	16
Control	20	20	20

The results of the d' analysis for Experiment 3 replicate those for Experiment 1: in the spatial sentences condition alone, the discriminability of the *minus* variant is greater than that of the *plus* variant ($X^2 = 16.67$, $p < .0001$).

General Discussion

In these experiments, we examined the question of whether spatial language influences the encoding and memory of spatial relations presented visually. The answer is a qualified yes. Our evidence shows that the use of spatial language during the encoding of a picture can affect recognition memory for the spatial relations in the picture. People given spatial prepositions during encoding showed a shift in recognition towards the core spatial category denoted by the preposition (Experiments 1a and 3). This effect was specific to spatial relational language (Experiment 1c); no such shift was observed for sentences that simply described the objects in the pictures.

However, our evidence that language influenced encoding was limited to the case when overt spatial language was present. We did not find a shift towards the core spatial semantic category when participants were simply instructed to pay close attention to the pictures (Experiment 1b). Thus, our evidence supports the view that language can affect encoding when it is present, but not the strong Whorfian view that non-linguistic perception is shaped by the language one speaks.

There has been much controversy in recent years over whether language exerts an effect on non-linguistic cognition. Our results suggest that language forms do exert an effect on one type of non-linguistic cognition: recognition memory for simple pictures. This suggestion must be qualified, however, as we do not show an effect of language forms in the absence of linguistic descriptions at encoding, which would suggest a stronger influence of language on everyday non-linguistic cognition. Of course, it remains an open question whether in some situations, speakers might prefer encodings that are compatible with their language, resulting in cross-linguistic differences that are habitual though not inescapable.

Our results are compatible with Slobin's (1996) thinking-for-speaking hypothesis and with the results of Malt et al. (in press). They suggest that language can have profound non-linguistic effects when it is used, but that its use is not inevitable. This is consistent with Gentner and Loewenstein's (in press) suggestion that language provides tools that potentiate forming and holding ideas -- the *tools–for-thought* hypothesis. On this view, language potentiates kinds of encodings rather than forcing them.

Acknowledgments

Please address all correspondence and reprint requests to Dedre Gentner, Northwestern University, Department of Psychology, 2029 Sheridan Road, Evanston, IL 60208. This work was supported by NSF-LIS grant SBR-9720313 to the second author.

References

Billman, D., & Krych, M. (1998). Path and manner verbs in action: Effects of "skipping" and "exiting" on event memory. *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Bower, G. H., Karlin, M. B., and Dueck, A. (1975). Comprehension and memory for pictures. *Memory and Cognition, 3 (2), 216-220.*

Bowerman, M., and Pederson, E. (in preparation). Cross-linguistic perspectives on topological spatial relationships.

Gentner, D. (in press). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought.* Cambridge, MA: MIT Press.

Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory,4 (2),* 161-178.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual deveolopment.* Cambridge, England: Cambridge University Press.

Gentner, D., & Loewenstein, J. (in press). Relational language and relational thought. In J. Byrnes & E. Amsel (Eds.), *Language, literacy, and cognitive development.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Gentner, D., & Loftus, E. (1979). Integration of verbal and visual information as evidenced by distortions in picture memory. *American Journal of Psychology, 92 (2),* 363-375.

Levinson, S.C. (1996). "Relativity in spatial conception and description." In Gumperz, J. and Levinson, S. (Eds.), *Rethinking Linguistic Relativity.* Cambridge, England: Cambridge University Press.

Malt, B. C., Sloman, S. A., & Gennari, S. (in press). Speaking vs. thinking about objects and actions. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought.* Cambridge, MA: MIT Press.

Pederson, E., Danziger, E., Wilkins, D., Levinson, S. C., Kita, S., & Senft, G. (1998). Semantic typology and spatial conceptualization. *Language, 74 (3),* 557-589.

Slobin, D. (1996). From "thought and language" to "thinking for speaking." In J. J. Gumperz and S. C. Levinson (Eds.), *Rethinking linguistic relativity.* Cambridge: Cambridge University Press.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf.* J. B. Carroll (Ed). Cambridge, MA: MIT Press.