Cognitive Psychology **43,** 1–22 (2001)
doi:10.1006/cogp.2001.0748, available online at http://www.idealibrary.com on IDEAL®

Does Language Shape Thought?: Mandarin and English Speakers' Conceptions of Time

Lera Boroditsky

Stanford University

Does the language you speak affect how you think about the world? This question is taken up in three experiments. English and Mandarin talk about time differently—English predominantly talks about time as if it were horizontal, while Mandarin also commonly describes time as vertical. This difference between the two languages is reflected in the way their speakers think about time. In one study, Mandarin speakers tended to think about time vertically even when they were thinking for English (Mandarin speakers were faster to confirm that March comes earlier than April if they had just seen a vertical array of objects than if they had just seen a horizontal array, and the reverse was true for English speakers). Another study showed that the extent to which Mandarin–English bilinguals think about time vertically is related to how old they were when they first began to learn English. In another experiment native English speakers were taught to talk about time using vertical spatial terms in a way similar to Mandarin. On a subsequent test, this group of English speakers showed the same bias to think about time vertically as was observed with Mandarin speakers. It is concluded that (1) language is a powerful tool in shaping thought about abstract domains and (2) one's native language plays an important role in shaping habitual thought (e.g., how one tends to think about time) but does not entirely determine one's thinking in the strong Whorfian sense.

Key Words: Whorf; time; language; metaphor; Mandarin.

Does the language you speak shape the way you understand the world? Linguists, philosophers, anthropologists, and psychologists have long been interested in this question. This interest has been fueled in large part by the observation that different languages talk about the world differently. Does the fact that languages differ mean that people who speak different languages

This research was funded by an NSF Graduate Research Fellowship to the author. Foremost, I thank Jennifer Y. Lee, who has made countless contributions to this research and has been an invaluable source of information about the Mandarin language. I thank Barbara Tversky, Gordon Bower, and Herbert Clark for many insightful discussions of this research and Michael Ramscar for comments on an earlier draft of this article.

Address correspondence and reprint requests to Lera Boroditsky, Department of Psychology, Bldg. 420, Stanford University, Stanford, CA 94305-2130. E-mail to lera@psych.stanford.edu.

0010-0285/01 $35.00

think about the world differently? Does learning new languages change the way one thinks? Do polyglots think differently when speaking different languages? Although such questions have long been issues of interest and controversy, definitive answers are scarce. This article briefly reviews the empirical history of these questions and describes three new experiments that demonstrate the role of language in shaping habitual thought.

The doctrine of Linguistic Determinism—the idea that thought is determined by language—is most commonly associated with the writings of Benjamin Lee Whorf. Whorf, impressed by linguistic diversity, proposed that the categories and distinctions of each language enshrine a way of perceiving, analyzing, and acting in the world. Insofar as languages differ, their speakers too should differ in how they perceive and act in objectively similar situations (Whorf, 1956). This strong Whorfian view—that thought and action are entirely determined by language—has long been abandoned in the field. Particularly effective in undermining the strong view was work on color perception demonstrating that the Dani (a tribe in New Guinea) had little trouble learning the English set of color categories, despite having only two words for colors in their language (Heider, 1972; Rosch, 1975, 1978; but see Lucy & Shweder, 1979; Kay & Kempton, 1984).

Although the strong linguistic determinism view seems untenable, many weaker but still interesting formulations can be entertained. For example, Slobin (1987, 1996) has suggested that language may influence thought during ''thinking for speaking.'' Languages force us to attend to certain aspects of our experience by making them grammatically obligatory. Therefore, speakers of different languages may be biased to attend to and encode different aspects of their experience while speaking. In a similar vein, Hunt and Agnoli (1991) reviewed evidence that language may influence thought by making habitual distinctions more fluent.

Since Rosch's work on color, several lines of research have explored domains that appear more likely to reveal linguistic influences than such low-level domains as color perception. Among the evidence are cross-linguistic differences in the object-substance distinction in Yucatec Mayan and Japanese (e.g., Gentner & Imai, 1997; Lucy, 1992), effects of grammatical gender distinctions in Spanish (Sera, Berge, & del Castillo, 1994), cross-linguistic differences in spatial thinking (e.g., Bowerman, 1996; Levinson, 1996), and evidence suggesting that language influences conceptual development (e.g., Markman & Hutchinson, 1984; Waxman & Kosowski, 1990).

LIMITATIONS OF RECENT EVIDENCE

Although the evidence so far is suggestive, there are serious limitations common to most recent studies of linguistic determinism. First, speakers of different languages are usually tested only in their native language. Any differences in these comparisons can only show the effect of a language on thinking for that particular language. These studies cannot tell us whether experience with a language affects language-independent thought such as thought for other languages or thought in nonlinguistic tasks.

Second, comparing studies conducted in different languages poses a deeper problem: There is simply no way to be certain that the stimuli and instructions are truly the same in both languages. This problem remains even if the verbal instructions are minimal. For example, even if the task is nonlinguistic and the instructions are simply ''which one is the same?'', one cannot be sure that the words used for ''same'' mean the same thing in both languages. If in one language the word for ''same'' is closer in meaning to ''identical,'' while in the other language it is closer to ''relationally similar,'' speakers of different languages may behave differently, but due only to the difference in instructions, not because of any interesting differences in thought. There is no sure way to guard against this possibility when tasks are translated into different languages. Since there is no way to know that participants in different languages are performing the same task, it is difficult to deem the comparisons meaningful.

A third limitation is that even when nonlinguistic tasks (such as sorting into categories or making similarity judgments) are used, the tasks themselves are quite explicit. Sorting and similarity judgment tasks require participants to decide on a strategy for completing the task. How should I divide these things into two categories? What am I supposed to base my similarity judgments on? When figuring out how to perform a task, participants may simply make a conscious decision to follow the distinctions reinforced by their language. For this reason, evidence collected using such explicit measures as sorting preferences or similarity judgments is not convincing as nonlinguistic evidence.

Showing that experience with a language affects thought in some broader sense (other than thinking for that particular language) would require observing a cross-linguistic difference on some implicit measure (e.g., reaction time) in a non-language-specific task. The studies described in this article do just that. They show an effect of first-language thinking on second-language understanding using the implicit measure of reaction time. In particular, the studies investigate whether speakers of English and Mandarin Chinese think differently about the domain of time even when both groups are ''thinking for English.''

TIME

How is the domain of time learned, represented, and reasoned about? Certainly some elements of time are apparent in our experience with the world. From experience, we know that each moment in time only happens once, that we can only be in one place at one time, that we can never go back, and that many aspects of our experience are not permanent (i.e., faculty meetings are not everlasting, but rather begin and end at certain times). In other words, our experience dictates that time is a phenomenon in which we, the observer, experience continuous unidirectional change that may be marked by appearance and disappearance of objects and events. These aspects of conceptual time should be universal across cultures and languages. Indeed, this appears to be the case. In order to capture the sequential order of events, time is generally conceived as a one-dimensional, directional entity. Across languages, the spatial terms imported to talk about time are also one-dimensional, directional terms such as *ahead/behind* or *up/down* rather than multidimensional or symmetric terms such as *narrow/wide* or *left/right* (Clark, 1973; Traugott, 1978). Aspects of time that are extractable from world experience (temporally bounded events, unidirectional change, etc.) appear to be universal across cultures and languages.

However, there are many aspects of our concept of time that are not observable in the world. For example, does time move horizontally or vertically? Does it move forward or back, left or right, up or down? Does it move past us, or do we move through it? All of these aspects are left unspecified in our experience with the world. They are, however, specified in our language—most often through spatial metaphors. Across languages people use spatial metaphors to talk about time. Whether they are looking *forward* to a brighter tomorrow, proposing theories *ahead* of their time, or falling *behind* schedule, they rely on terms from the domain of space to talk about time (Clark, 1973; Lehrer, 1990; Traugott, 1978). Those aspects of time that are not constrained by our physical experience with time are free to vary across languages and our conceptions of them may be shaped by the way we choose to talk about them. This article focuses on one such aspect of time and examines whether different ways of talking about time lead to different ways of thinking about it.

Time in English

In English, we predominantly use front/back terms to talk about time. We can talk about the good times *ahead* of us or the hardships *behind* us. We can move meetings *forward,* push deadlines *back,* and eat dessert *before* we are done with our vegetables. On the whole, the terms used to order events are the same as those used to describe asymmetric horizontal spatial relations (e.g., ‘‘he took three steps *forward*’’ or‘‘the dumpster is *behind* the store’’).

Time in Mandarin Chinese

In Mandarin, front/back spatial metaphors for time are also common (Scott, 1989). Mandarin speakers use the spatial morphemes *qián* (‘‘front’’) and *hòu* (‘‘back’’) to talk about time. Examples in Fig. 1 show parallel uses of *qián* and *hòu* in their spatial and temporal senses. Example sentences and their English glosses were taken from Scott (1989).

What makes Mandarin interesting for present purposes is that Mandarin

(1) SPACE

zài zhuōzi qián-bian zhàn-zhe yī ge xuésheng
there is a student standing in front of the desk

TIME

hǔ nián de qián yī nián shì shénme nián?
what is the year before the year of the tiger?

(2) SPACE

zài zhuōzi hòu-bian zhàn-zhe yī ge lǎoshi
there is a teacher standing behind the desk

TIME

dàxué bìyè yǐ-hòu wǒ yòu jìn le yánjiūyuàn
after graduating from university, I entered graduate school

FIG. 1. Example spatial and temporal uses of *front/back* terms *qián* and *hòu* in Mandarin and their English glosses.

speakers also systematically use vertical metaphors to talk about time (Scott, 1989). The spatial morphemes *shàng* (‘‘up’’) and *xià* (‘‘down’’) are frequently used to talk about the order of events, weeks, months, semesters, and more. Earlier events are said to be *shàng* or ‘‘up,’’ and later events are said to be *xià* or ‘‘down.’’ Examples in Fig. 2 show parallel uses of *shàng* and *xià* to describe spatial and temporal relations (examples taken from Scott, 1989).

Although in English vertical spatial terms can also be used to talk about time (e.g., ‘‘hand *down* knowledge from generation to generation’’ or ‘‘the meeting was coming *up*’’), these uses are not nearly as common or systematic as is the use of *shàng* and *xià* in Mandarin (Chun, 1997a, 1997b; Scott, 1989).

In summary, both Mandarin and English speakers use horizontal terms to talk about time. In addition, Mandarin speakers commonly use the vertical terms *shàng* and *xià*.[1]

[1] The closest English counterparts to the Mandarin uses of *shàng* and *xià* are the terms *next (following)/last (previous)* and *earlier/later*. *Earlier* and *later* are similar to *shàng* and *xià* in that they use an absolute framework to determine the order of events. In Mandarin, *shàng* always refers to events closer to the past, and *xià* always refers to events closer to the future. The same is true in English for *earlier* and *later* terms. This is not true, however, for the other English terms for time. Terms like *before/after, ahead/behind,* and *forward/back* can be used not only to order events relative to the direction of motion of time, but also relative to the observer. When ordering events relative to the direction of motion of time, we can say that

(1) SPACE

mão shàng shù
cats climb trees

TIME

shàng ge yuè
last (or previous) month

(2) SPACE

tā xià le shān méi yǒu
has she descended the mountain or not?

TIME

xià ge yuè
next (or following) month

FIG. 2. Example spatial and temporal uses of *up/down* terms *shàng* and *xià* in Mandarin and their English glosses.

DOES LANGUAGE SHAPE THOUGHT?

So, do the differences between the English and Mandarin ways of *talking* about time lead to differences in how their speakers *think* about time? This question can be expanded into two separate issues: (1) Does using spatial language to talk about time have short-term implications for on-line processing? and (2) Does using spatial language to talk about time have long-term implications?

Does Metaphor Use Have Implications for Online Processing?

Recent evidence suggests that people do not just *talk* about time in spatial terms, but that they also use their spatial knowledge to *think* about time. Boroditsky (2000) showed that people are able to reuse relational information made available by spatial primes to think about time. For example, priming a particular perspective for thinking about space biased how people later

Thursday is *before* Friday. Here, *before* refers to an event that is closer to the past. But, we can also order events relative to the observer, as in "The best is *before* us." Here, *before* refers to an event closer to the future. The same is true for *ahead/behind* and *forward/back*. *Qián* and *hòu,* the horizontal terms used in Mandarin to talk about time, also share this flexibility. Unlike *before/after, ahead/behind,* and *qián/hòu,* terms like *earlier/later* and *shàng/xià* are not used to order events relative to the observer. For example, one cannot say that "the meeting is *earlier* than us" to mean that it is further in the future. *Earlier/later* and *shàng/xià* are absolute terms.

interpreted an ambiguous question about time. Also, spatial relational information was found to be just as useful for thinking about time as temporal information—in answering questions about time, subjects benefited equally from a spatial prime (129-ms benefit) as from a temporal prime (130-ms benefit). It appears that spatiotemporal metaphors do have implications for online conceptual processing.

Does Metaphor Use Have Long-Term Implications for Processing?

How could one's choice of spatiotemporal metaphors affect thinking about time in the long run? Boroditsky (2000) argued that spatial metaphors can provide relational structure to those aspects of time where the structure may not be obvious from world experience (e.g., whether time should be vertical or horizontal). Using spatial metaphors to describe time encourages structural alignment between the two domains and may cause relational structure to be imported from space to time. The mechanism for this type of metaphoric structuring may be the same as that used in analogical inference (Gentner, Bowdle, & Wolff, in press; Gentner & Wolff, 1997). Language-encouraged mappings between space and time may then come to be stored in the domain of time. That is, frequently invoked mappings may become habits of thought. For example, because English speakers often use horizontal metaphors to talk about time, they might grow to think about time horizontally even when not explicitly processing a spatiotemporal metaphor (e.g., when understanding a sentence phrased in purely temporal terms like *earlier* and *later*). For the same reasons, Mandarin speakers might grow to think about time vertically.

Experiment 1 was designed to test whether using spatial metaphors to talk about time can have both immediate and long-term implications for how people think about time. Mandarin and English speakers were asked to answer a spatial priming question followed by a target question about time. The spatial primes were either about horizontal spatial relations between two objects (see Fig. 3a) or about vertical relations (see Fig. 3b). After solving a set of two primes, participants answered a TRUE/FALSE target question about time. Half of the target questions were designed to test the immediate effect of metaphors on processing and so used a horizontal spatiotemporal

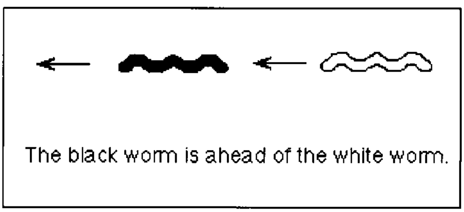

FIG. 3a. Example of a horizontal spatial prime used in Experiments 1 and 3.

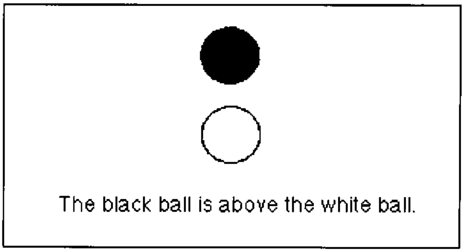

FIG. 3b. Example of a vertical spatial prime used in Experiments 1 and 3.

metaphor (e.g., ''March comes *before* April.''). If horizontal spatiotemporal metaphors are processed by activating horizontal spatial knowledge, then people should be faster to understand such a metaphor if they have just seen a horizontal spatial prime (Fig. 3a) than if they have just seen a vertical prime (Fig. 3b). The other half of the target questions were designed to test the long-term effects of metaphor use on thinking about time and so did not use a metaphor, but instead used the purely temporal terms *earlier* and *later* (e.g., ''March comes *earlier* than April''). If the metaphors frequently used in one's native language do have a long-term effect on how one thinks about time, then even when people are not trying to understand a metaphor (e.g., when deciding whether ''March comes *earlier* than April'') they may still use spatial knowledge to think about time in a way that is consistent with (and encouraged by) the particular metaphors popular in their language.

If one's native language does have a long-term effect on how one thinks about time, then Mandarin speakers should be faster to answer purely temporal target questions (e.g., ''March comes *earlier* than April'') after solving the vertical spatial primes than after the horizontal spatial primes. English speakers, on the other hand, should be faster after horizontal primes because horizontal metaphors are predominantly used in English. Since both English and Mandarin speakers completed the task in English, this is a particularly strong test of the effect of one's native language on thought. If Mandarin speakers do show a vertical bias in thinking about time even when they are ''thinking for English,'' then language must play an important role in shaping speakers' thinking habits.

EXPERIMENT 1

Method

Participants

Twenty-six native English speakers and 20 native Mandarin speakers participated in this study. All participants were graduate or undergraduate students at Stanford University and received either payment or course credit in return for their participation. All of the Mandarin speakers had Mandarin as their first language. It was also their only language until at least the age of 6 years, with a mean age at the onset of English acquisition of 12.8 years.

Design

Participants answered spatial prime questions followed by questions about time. Primes were spatial scenarios accompanied by a sentence description and were either horizontal (see Fig. 3a) or vertical (see Fig. 3b). Targets were statements about time: either *before/after* statements (e.g., ‘‘March comes *before* April’’) or *earlier/later* statements (e.g., ‘‘March comes *earlier* than April’’). Each participant completed a set of 6 practice questions and 64 experimental trials. Each experimental trial consisted of two spatial prime questions (both horizontal or both vertical) followed by one target question about time. The experimental trials were arranged such that the first prime question was FALSE, the second was TRUE, and the target question was TRUE. Participants were not told that the experiment was arranged into such trials, and because randomly arranged filler trials were extensively interspersed throughout the experiment, participants were not able to figure out the trial structure in the course of the experiment. Participants answered each target question twice—once after each type of prime. The order of all trials was randomized for each participant. Overall, the experiment had a fully crossed within-subject 2 (prime-type) × 2 (target-type) design with native language as the only between-subjects factor.

Materials

A set of 128 primes and 32 targets, all TRUE/FALSE questions, was constructed.

Primes. One hundred twenty-eight spatial scenarios were used as primes. Each scenario consisted of a picture and sentence below the picture. Half of these scenarios were about horizontal spatial relations (see Fig. 3a), and the other half were about vertical spatial relations (see Fig. 3b). Half of the horizontal primes used the ‘‘X is *ahead* of Y’’ phrasing and half used the ‘‘X is *behind* Y’’ phrasing. Likewise, half of the vertical primes used the ‘‘X is *above* Y’’ phrasing and have used the ‘‘X is *below* Y’’ phrasing. Primes were equally often TRUE and FALSE. All of these variations were crossed into eight types of primes. In addition, the left/right orientation of the horizontal primes was counterbalanced across variations.

Targets. Sixteen statements about the order of the months of the year were constructed. Half used the spatiotemporal terms *before* and *after* (e.g., ‘‘June comes *before* August’’), and half used the purely temporal terms *earlier* and *later* (e.g., ‘‘August comes *later than* June’’). All four terms were used equally often. All target statements were ‘‘TRUE.’’

Fillers. Sixteen additional statements about months of the year were used as fillers. These statements were similar in all respects to the targets except that all of the filler statements were ‘‘FALSE.’’ Filler statements were constructed by reversing the relation in each of the target statements. Filler time questions (along with filler spatial scenarios drawn randomly from the list of all spatial primes) were inserted randomly in-between experimental trials to ensure that participants did not deduce the trial structure of the experiment. Responses to filler trials were not analyzed.

Procedure

Participants were tested individually. All participants were tested in English with English instructions. Questions were presented on a computer screen one at a time. For each question, participants needed to respond TRUE or FALSE as quickly as possible (and within a 5-s deadline) by pressing one of two keys on a keyboard. Response times were measured and recorded by the computer. Participants received no feedback for the experimental trials.

Results

As predicted, English and Mandarin speakers were affected differently by the spatial primes. Both English and Mandarin speakers answered spatiotemporal *before/after* questions faster after horizontal primes than after vertical primes (see Fig. 4a). This confirms the earlier findings that spatial knowledge can be used in the online processing of spatiotemporal metaphors. However, when it came to purely temporal *earlier/later* questions, English and Mandarin speakers looked very different (see Fig. 4b). As predicted, English speakers answered purely temporal questions faster after horizontal primes than after vertical primes. This pattern was predicted by the preponderance of horizontal spatial metaphors used to describe time in English. The data from Mandarin speakers looked quite different. When answering questions phrased in purely temporal *earlier/later* terms, Mandarin speakers were faster after vertical primes than after horizontal primes. This pattern was predicted by the fact that in Mandarin vertical metaphors are often used to talk about time. Descriptive statistics and analyses are reported below.

Only responses to target time questions were analyzed. Response times exceeding the deadline and those following an incorrect response to a priming question and incorrect responses were omitted from all analyses (7% of all responses were omitted). Error rates did not differ by native language (7.1% for English speakers and 6.9% for Mandarin speakers) or prime type (7.3% after horizontal primes and 6.7% after vertical primes). Both English

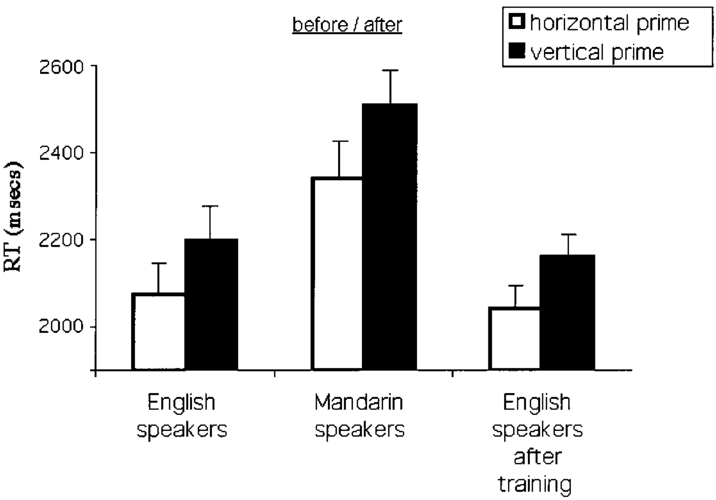

FIG. 4a. Experiments 1 and 3: Response times to spatiotemporal *before/after* questions about time following either a horizontal or a vertical prime are plotted for English speakers, Mandarin speakers, and English speakers who had been trained to talk about time vertically.

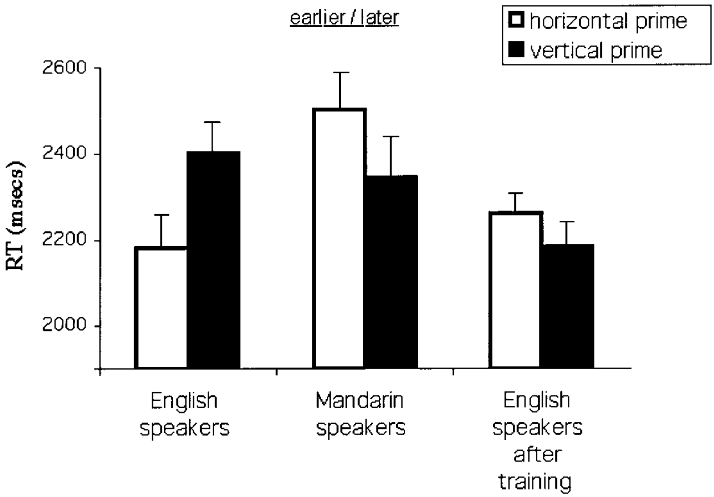

FIG. 4b. Experiments 1 and 3: Response times to purely temporal *earlier/later* questions about time following either a horizontal or a vertical prime are plotted for English speakers, Mandarin speakers, and English speakers who had been trained to talk about time vertically.

and Mandarin speakers made slightly more errors on *earlier/later* targets (8.6%) than on *before/after* targets (5.4%), $\chi^2 = 4.82, p < .05$. Separate 2 (prime type) × 2 (target type) repeated-measures ANOVAs were conducted for data from English and Mandarin speakers.

Native English speakers. As predicted, Native English speakers answered time questions faster after horizontal primes (2128 ms) than after vertical primes (2300 ms), $F(1, 25) = 13.76, p < .01$. Reaction times were also shorter for questions phrased in *before/after* terms (2135 ms) than for those phrased in *earlier/later* terms (2294 ms), $F(1, 25) = 8.23, p < .01$. This difference is most likely due to an uninteresting difference in reading time between the two types of targets; *earlier/later* targets were one to two syllables longer than the *before/after* targets. There was no interaction between prime type and target type, $F(1, 25) = .75, p = .40$. English speakers were faster to solve all questions about time if they followed horizontal primes than if they followed vertical primes.

Native Mandarin speakers. Overall, Mandarin speakers answered time questions just as quickly after horizontal primes (2422 ms) as after vertical primes (2428 ms), $F(1, 19) = .01, p = .92$. However, there was a big difference in how primes affected response times to the two types of targets. Like the English speakers, Mandarin speakers answered the *before/after* target questions faster after horizontal primes (2340 ms) than after vertical primes (2509 ms). When it came to the purely temporal *earlier/later* targets, however, the pattern was exactly reversed. Unlike the English speakers, Mandarin speakers solved purely temporal targets faster after vertical primes (2347 ms) than after horizontal primes (2503 ms). These differences were confirmed as an interaction between prime type and target type, $F(1, 19) = 4.55$, $p < .05$.

Comparing English and Mandarin speakers. Overall, English speakers were not significantly faster to answer target questions than Mandarin speakers (2214 and 2425 ms respectively), $F(1, 44) = 2.01$, $p = .16$. The effect of prime was different for the two language groups; there was an overall effect of prime for English speakers but not for Mandarin speakers, $F(1, 44) = 4.89$, $p < .05$. The critical predicted difference between the two language groups was in the interaction of prime and target. This difference was confirmed as a three-way interaction in a $2_{\text{prime}} \times 2_{\text{target}} \times 2_{\text{language}}$ ANOVA, $F(1, 44) = 5.24$, $p < .05$.

Discussion

In this experiment, native English and native Mandarin speakers were found to think differently about time. This was true even though both groups were tested in English. English speakers were faster to verify that ‘‘March comes *earlier* than April’’ after horizontal primes than after vertical primes. This habit of thinking about time horizontally was predicted by the preponderance of horizontal spatial metaphors used to talk about time in English. The reverse was true for Mandarin speakers. Mandarin speakers were faster to verify that ‘‘March comes *earlier* than April’’ after vertical primes than after horizontal primes. This habit of thinking about time vertically was predicted by the preponderance of vertical time metaphors in the Mandarin. In short, it appears that habits in language encourage habits in thought. Since Mandarin speakers showed vertical bias even when thinking for English, it appears that language-encouraged habits in thought can operate regardless of the language that one is currently thinking for.

These results suggest that experience with a language can shape the way one thinks. Experiment 2 was designed to further test the relationship between language experience and patterns in thinking. How much and in what ways does learning new languages influence one’s way of thinking? Mandarin–English bilinguals were tested in a task similar to that described in Experiment 1. All of the participants were Mandarin–English bilinguals whose first language was Mandarin. In order to be able to assess the effects of second-language learning on thought, this group of participants was chosen to vary much more in how early in life they began to learn English than did the participants in Experiment 1. If learning new languages does change the way one thinks, then participants who learned English early on or had more English experience should show less of a ‘‘Mandarin’’ bias to think about time vertically.

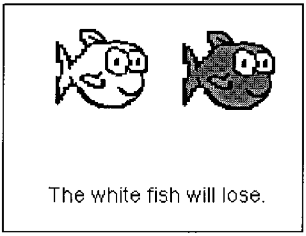

FIG. 5a. Example of a horizontal spatial prime used in Experiment 2.

EXPERIMENT 2

Method

Participants

Twenty-five Mandarin–English bilinguals (with varying degrees of experience with Mandarin and English) participated in this study. All participants were graduate or undergraduate students at Stanford University and received payment in return for their participation. Participants ranged in age from 18 to 28 years (M = 23.4 years, SD = 2.5 years). All participants had acquired Mandarin prior to English. They varied in the age at which they first began to learn English (Age of Acquisition) from 3 to 13 years of age (M = 9.4 years, SD = 3.3 years). All had at least 10 years of Exposure to English (current age minus Age of Acquisition) (M = 14.0 years, SD = 2.3 years).

Design

Just as in Experiment 1, participants answered spatial priming questions followed by target questions about time. Primes were spatial scenarios accompanied by a sentence description and were either horizontal (see Fig. 5a) or vertical (see Fig. 5b). Unlike Experiment 1, all targets were *earlier/later* statements about time (e.g., ‘‘March comes *earlier* than April’’). Because the critical measure was the amount of vertical bias in response to the *earlier/later* targets, the *before/after* targets were not included in this experiment. Each participant completed 80 experimental trials and 240 filler questions. Each experimental trial consisted of two

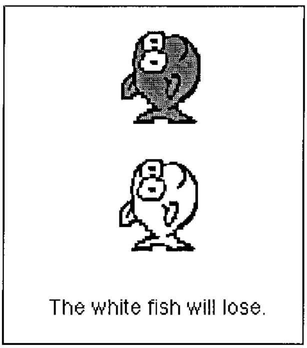

FIG. 5b. Example of a vertical spatial prime used in Experiment 2.

spatial prime questions (both horizontal or both vertical) followed by one target question about time. Participants were not told that the experiment was arranged into such trials, nor did they figure it out in the course of the experiment. They answered each target question twice, once after each type of prime. The order of all trials was randomized for each participant.

Materials

A set of 96 primes and 40 targets, all TRUE/FALSE questions, was constructed.

Primes. Ninety-six spatial scenarios were used as primes. Each scenario consisted of a picture and sentence below the picture. Half of these scenarios were about horizontal spatial relations (see Fig. 5a), and the other half were about vertical spatial relations (see Fig. 5b). The left/right and up/down orientation in horizontal and vertical primes respectively was counterbalanced. Half of the primes used the ''X will win'' phrasing and half used the ''X will lose'' phrasing. Primes were equally often TRUE and FALSE. All of these variations were fully crossed.

Targets. Forty statements about the order of the months of the year were constructed. All of these statements used the purely temporal terms *earlier* and *later* (e.g., ''March comes *earlier* than April''). Both terms were used equally often. All target statements were ''TRUE.''

Fillers. Forty additional statements about months of the year were used as fillers. They were similar to the targets in all respects except they were ''FALSE.'' This was done to insure that participants were alert and did not simply learn to answer ''TRUE'' to all questions about time. Filler time questions (along with filler spatial scenarios drawn randomly from the list of all spatial primes) were inserted randomly between experimental trials to ensure that participants did not deduce the trial structure of the experiment. Responses to filler questions were not analyzed.

Procedure

Participants were tested individually. All were tested in English with English instructions. The procedure was the same as in Experiment 1.

Results and Discussion

The bias to think about time vertically was greater for Mandarin speakers who started learning English later in life. Surprisingly, vertical bias appeared independent of the length of Exposure to English.

Vertical Bias was calculated for each participant by subtracting their mean RT for targets following a vertical prime from that for targets following a horizontal prime (mean Vertical Bias = 54 ms). Each participant also received a score on two predictor variables: Age of Acquisition of English and Years of Exposure to English. As before, response times exceeding the deadline, those following an incorrect response to a priming question, and incorrect responses were omitted from all analyses (10.9% of all responses were omitted). Error rates did not differ by prime type (10.7% after horizontal primes and 11.0% after vertical primes).

As predicted, the Age of Acquisition of English was positively correlated with Vertical Bias with $r = .47$, $p < .01$.[2] Participants who started learning

[2] There was one outlier participant with an unusually high vertical bias of 582 ms. The removal of this outlier only served to increase the correlation between Vertical Bias and Age of Acquisition to $r = .50$, $p < .01$.

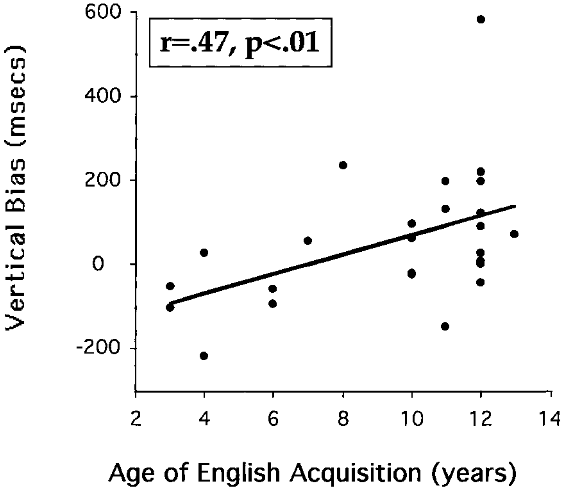

FIG. 6. Experiment 2: Results from 25 Mandarin speakers. Vertical Bias in milliseconds is plotted as a function of Age of Acquisition of English in years. Vertical Bias equals the difference in reaction time between targets following horizontal primes and targets following vertical primes.

English at a younger age showed less Vertical Bias (a less ‘‘Mandarin’’ way of thinking) than participants who started learning English later in life (see Fig. 6). That is, the longer a child was speaking only Mandarin, the greater his or her Vertical Bias score in English. Interestingly, the Length of Exposure to English did not predict Vertical Bias, r $= -.12$, $p = .29$. Acquiring thinking habits promoted by a language (assessed here as Vertical Bias) appears to depend primarily on how early one starts to learn that language and not on the amount of exposure to that language. This dissociation is particularly surprising since (as would be expected for college-age participants) the Age of Acquisition and the Length of Exposure were highly correlated, with r $= -.66$, $p < .01$.

The partial correlations between Vertical Bias and these two factors followed the same pattern as the full correlations. After controlling for the effect of Length of Exposure, there was still a strong correlation between Vertical Bias and Age of Acquisition, with r $= .52$, $p < .01$. There was still no significant correlation between Vertical Bias and Length of Exposure after the effect of Age of Acquisition had been controlled for, r $= .29$, $p = .17$.

These findings bear a conspicuous resemblance to those reported by Johnson and Newport (1989) regarding the acquisition of grammatical proficiency in a second language. In accord with the current results, grammatical proficiency in a second language is better the earlier the immersion in that language, but is nearly independent of the length of exposure to it (Johnson & Newport, 1989). It is striking that the acquisition of semantic biases (here measured as Vertical Bias) is affected by the same variables as the acquisition of basic language skills like grammatical proficiency.

Participants were also asked to provide an intuitive rating of how often they ''think in Mandarin'' compared to English on a scale of 1 to 5 (1 = *I think almost always in English* to 5 = *I think almost always in Mandarin*). This introspective measure was also correlated with Vertical Bias, r = .37, $p < .05$, though not as strongly as Age of Acquisition. Once the effect of Age of Acquisition was controlled for, this introspective measure was no longer correlated with Vertical Bias, r = −.08, $p = .36$. It is reasonable to suppose that the Age of Acquisition is the causal variable driving both Vertical bias and this introspective assessment.

Overall, Mandarin speakers who learned English later in life were more likely to think about time vertically. The propensity to think about time vertically was related to the length of pure Mandarin experience (before any English was learned), but not to the length of English experience.

Although these results strongly suggest an effect of language on habitual thought, there is still one concern. The difference in time metaphors used in English and Mandarin is clearly not the only difference between English speakers and Mandarin speakers. Other cultural factors could conceivably have led to the observed differences. One important factor to consider is that of writing direction.[3] Whereas English is written horizontally from left to right, Mandarin is traditionally written in vertical columns that run from right to left. Beyond writing direction, there may be many other cultural differences between native English and native Mandarin speakers that may have lead to the differences observed in Experiment 1. Experiment 3 was designed to minimize differences in nonlinguistic cultural factors while preserving the interesting difference in language.

In Experiment 3, native English speakers learned to use vertical spatial terms (*above, below, higher than,* and *lower than*) to talk about time. For example, they learned to say that ''cars were invented *above* fax machines'' and that ''Wednesday is *lower than* Tuesday.'' The use of the vertical terms *above/below* and *higher than/lower than* in this training was similar to the use of *shàng* and *xià* in Mandarin. Earlier events were always said to be *above* or *higher than,* and later events were always said to be *below* or *lower than*. This training was designed to alter (temporarily) the English speakers' habit of thinking about time horizontally by making the vertical metaphor highly available in memory. After the training, participants completed ex-

[3] Although this difference is interesting, it cannot explain the results of Experiment 1. The writing direction explanation would predict that—since Mandarin is written vertically—Mandarin speakers should always be faster to answer time questions after vertical than after horizontal primes. This prediction was not borne out by data. Mandarin speakers showed an interaction (faster after vertical primes for earlier/later sentences, but faster after horizontal primes for before/after sentences) and not the main effect predicted by writing direction. Writing direction cannot be responsible for the differences observed in this experiment.

actly the same experiment as in Experiment 1. If it is indeed language (and not other cultural factors) that led to the differences between English and Mandarin speakers in Experiment 1, then the ''Mandarin'' linguistic training given to English speakers in Experiment 3 should make their results look more like those of Mandarin speakers than those of English speakers.

EXPERIMENT 3

Method

Participants

Seventy Stanford University undergraduates, all native English speakers, participated in this study for course credit.

Materials and Design

Participants were told they would learn ''a new way to talk about time.'' They were given a set of five example sentences that ''used this new system'' (e.g., ''Monday is *above* Tuesday'' or ''Monday is *higher than* Tuesday'') and had to figure out on their own how the system worked. The new system used *above/below* and *higher than/lower than*. Events closer to the past were always said to be *above* or *higher than,* and events closer to the future were always said to be *below* or *lower than*. Participants were then tested on a set of 90 questions that used these vertical terms to talk about time (e.g., ''Nixon was president *above* Clinton'' or ''WWII happened *lower than* WWI''). These test questions were presented on a computer screen one at a time, and participants responded TRUE or FALSE to each statement by pressing one of two keys on the keyboard.

Half of the participants learned a system that used the terms *above* and *below,* and half learned a system that used *higher than* and *lower than*. Two different training systems were used in order to equate syntactic similarity between the training phrases and the two types of targets used in this experiment. The *above/below* phrasings were syntactically similar to the *before/after* targets, and the *higher than/lower than* phrasings were similar to the *earlier than/later than* targets. This was done to make sure that any differential transfer from the training phase to the experiment would not be due to simple syntactic priming.

Immediately after the training, participants went on to complete the experiment described in Experiment 1. After the initial training, all materials, instructions, and procedures were identical to those used in Experiment 1.

Results and Discussion

After the short training, results of native English speakers looked more like those of Mandarin speakers than those of untrained English speakers. Results are summarized in Fig. 4.

Unlike untrained English speakers in Experiment 1, trained English speakers were not faster to answer time questions after horizontal primes (2151 ms) than after vertical primes (2170 ms), $F(1, 68) = .53, p = .47$. However, just as was the case with Mandarin speakers, primes affected response times differently for the different targets. For *before/after* targets, response times were shorter after horizontal primes (2040 ms) than after vertical primes (2156 ms). For purely temporal *earlier/later* targets, however, the pattern was exactly reversed; response times were shorter after vertical primes (2185 ms) than after horizontal primes (2262 ms). These differences were confirmed as an interaction between prime type and target type, $F(1, 68) = 10.25$, $p < .01$.

There were no differences between the two training types. This confirms that the effect of training was not simply that of syntactic priming. Also, just as observed for untrained English speakers, response times were shorter for questions phrased in *before/after* terms (2098 ms) than for those phrased in *earlier/later* terms (2223 ms), $F(1, 68) = 11.03$, $p < .01$. As before, this difference is most likely due to an uninteresting difference in reading time between the two types of targets; *earlier/later* targets were one to two syllables longer than the *before/after* targets.

Comparing Trained and Untrained English Speakers

Mean response times did not differ between trained and untrained English speakers, but the effect of prime was different for the two groups; there was an overall effect of prime for untrained English speakers but not for trained English speakers, $F(1, 94) = 4.69$, $p < .05$. The critical predicted difference between the two groups was in the interaction of prime and target. This difference was confirmed as a three-way interaction in a $2_{\text{prime}} \times 2_{\text{target}} \times 2_{\text{training}}$ ANOVA, $F(1, 94) = 5.65$, $p < .05$. These are the very same differences as were observed between English and Mandarin speakers in Experiment 1.

Comparing Trained English Speakers and Mandarin Speakers

Overall, trained English speakers answered targets faster than Mandarin speakers (2161 and 2425 ms respectively), $F(1, 88) = 4.68$, $p < .05$. This was the only difference between the two groups. None of the differences observed between English speakers and Mandarin speakers in Experiment 1 were present after English speakers had been trained to talk about time in a ‘‘Mandarin’’ way.

Overall, English speakers who were trained to talk about time using vertical terms showed a pattern of results very similar to that of Mandarin speakers. These results confirm that, even in the absence of other cultural differences (e.g., writing direction), differences in talking do indeed lead to differences in thinking.

GENERAL DISCUSSION

One’s native language appears to exert a strong influence over how one thinks about abstract domains like time. In Experiment 1, Mandarin speakers relied on a ‘‘Mandarin’’ way of thinking about time even when they were thinking about English sentences. Mandarin speakers were more likely to think about time vertically when deciding whether ‘‘March comes *earlier* than April.’’ This result is predicted by the way Mandarin talks about time;

the fact that vertical terms are commonly used to talk about time predicts that Mandarin speakers would find it more natural to construct a vertical time line when thinking about purely temporal relations. English speakers were more likely to think about time horizontally because horizontal spatial terms predominate in English temporal descriptions.

Experiment 2 showed that the acquisition of semantic biases (such as a habit of thinking about time vertically or horizontally) decreases with the age at which second-language exposure begins. Further, the acquisition of semantic biases is affected by the same variables as the acquisition of basic language skills.

In Experiment 3, native English speakers who had just been briefly trained to talk about time using vertical terms produced results very similar to those of Mandarin speakers. This finding confirms that the effect observed in Experiment 1 was driven by differences in language and not by other cultural differences. Learning a new way to talk about a familiar domain can change the way one thinks about that domain. Taken together these findings make a strong case for language shaping habitual thought.

However, there is an interesting discrepancy between the findings described here on time and those of Rosch and colleagues on color. Why would there be such strong evidence for universality in thought for domains like color perception, but quite the opposite for time? One possibility is that—since color perception predates language both in evolution and in development—children's perceptually based concepts (like colors) may be relatively fixed before they learn language.

A second possibility is that language is most powerful in determining thought for domains that are more abstract, that is, ones that are not so reliant on sensory experience. Gentner and Boroditsky (2001) have argued that the effect of language should be most apparent in the conceptualization of relations (typically encoded by verbs and spatial prepositions) as opposed to objects. Whereas object-concepts are easily individuable from perceptual experience, learning the extent and generality of a relational concept requires considerable experience with language. In one study, adults watched silent films of mothers talking to their children and tried to guess what was being said (Gillette, Gleitman, Gleitman, & Lederer, 2001). Given only the silent film, adult participants were able to correctly guess nouns three times more often than verbs (45 and 15% correct respectively). Further, concrete activity verbs like ‘‘push’’ were much more easily guessed from silent observation than from the syntactic frames in which they were used (50 and 15% respectively), whereas verbs that denote more abstract activities like ‘‘think’’ were much more easily guessed from syntax than from observation (90 and 0% respectively).

In general, the referents of abstract terms are difficult or impossible to pick out just from observing the context in which they are used. Imagine trying to learn to pick out instances of ‘‘idea,’’ ‘‘tomorrow,’’ or ‘‘justice’’ just from immediate interaction with the physical world. One consequence of this is that, in acquiring their first language, children take longer to learn relational terms than object-reference terms (because more language experience is needed to parcel out relational concepts) (Au, Dapretto, & Song, 1994; Gelman & Tardif, 1998; Gentner, 1982; Gentner & Boroditsky, 2001; Macnamara, 1972; Nelson, 1973).[4] Another consequence is that the lexicalization of abstract and relational concepts varies cross-linguistically much more than that of concrete object concepts. It appears that acquiring abstract concepts requires experience with language and that the eventual form of these concepts is largely shaped by the language experience.

But how does language affect thought? Let us again consider the domain of time. How do spatiotemporal metaphors affect thinking about time? Spatial metaphors can provide relational structure to those aspects of time where the structure may not be obvious from world experience (Boroditsky, 2000). In the case of space and time, using spatial metaphors to describe time encourages structural alignment between the two domains and may cause relational structure to be imported from space to time. The mechanism for this type of metaphoric structuring may be the same as that used in analogical inference (Gentner, Bowdle, & Wolff, 2001; Gentner & Wolff, 1997). Language-encouraged mappings between space and time come to be stored in the domain of time. Hence, when spatiotemporal metaphors differ, so may people's ideas of time.

Language can be a powerful tool for shaping abstract thought. When sensory information is scarce or inconclusive (as with the direction of motion of time), languages may play the most important role in shaping how their speakers think.

REFERENCES

Au, T., Dapretto, M., & Song Y. (1994). Input vs constraints: Early word acquisition in Korean and English. *Journal of Memory and Language,* **33,** 567–582.

Boroditsky, L. (1998). Evidence for metaphoric representation: Understanding time. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences*. Sofia, Bulgaria: New Bulgarian Univ. Press.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition,* **75**(1), 1–28.

Boroditsky, L. (in preparation). Which way to the present? Cross-linguistic differences in thinking about time.

Bowerman, M. (1996). The origins of children's spatial semantic categories: cognitive versus linguistic determinants. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 145–176). Cambridge, MA: Cambridge Univ. Press.

[4] See Gopnik and Choi (1995), Choi and Gopnik (1995), and Tardif (1996) for counterevidence to this claim, and Gentner and Boroditsky (2001) for discussion.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. Special Issue: Lexical and conceptual semantics. *Cognition,* **41**(1–3), 83–121.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language,* **22**(3), 497–529.

Chun, L. (1997a). *A cognitive approach to UP metaphors in English and Chinese: What do they reveal about the English mind and the Chinese mind*? In Research degree progress report for Hong Kong Polytechnic University (pp. 125–140).

Chun, L. (1997b). Conceptualizing the world through spatial metaphors: An analysis of UP/DOWN vs. SHANG/XIA metaphors. In *Proceedings of the 19th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Clark, H. (1973). Space, time semantics, and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.

Gelman, S., & Tardif, T. (1998). Acquisition of nouns and verbs in Mandarin and English. In *Proceedings of the Twenty-Ninth Annual Child Language Research Forum* (pp. 27–36).

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Language, thought, and culture (Vol. 2, pp. 301–334)*. Hillsdale, NJ: Erlbaum.

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

Gentner, D. & Boronat, C. (1992, May). *Metaphor as mapping*. Paper presented at the Workshop on Metaphor, Tel Aviv, Israel.

Gentner, D., Bowdle, B., & Wolff, P. (2001). Metaphor is like analogy. In D. Gentner, K. J. Holyoak, & B. Kokinov (Eds.), *The analogical mind: Theory and phenomena*. Cambridge, MA: MIT Press.

Gentner, D., & Imai, M. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* **62**(2), 169–200.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*.

Gopnik, A., & Choi, S. (1995). Names, relational words, and cognitive development in English and Korean Speakers: Nouns are not always learned before verbs. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 63–80). Hillsdale, NJ: Erlbaum.

Heider, E. (1972). Universals in color naming and memory. *Journal of Experimental Psychology,* **93,** 10–20.

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review,* **98,** 377–389.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology,* **21,** 60–99.

Kay, P., & Kempton, W. (1984). What is the Sapir–Whorf hypothesis? *American Anthropologist,* **86,** 65–79.

Lehrer, A. (1990). Polysemy, conventionality, and the structure of the lexicon. *Cognitive Linguistics,* **1,** 207–246.

Levinson, S. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. Bloom & M. Peterson (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.

Lucy, J. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge Univ. Press.

Lucy, J., & Shweder, R. (1979). Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist,* **81,** 581–618.

Macnamara, J. (1972). Cognitive basis of language learning in infants. *Psychological Review,* **79,** 1–13.

Markman, E., & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology,* **16,** 1–27.

Nelson, K. (1973). Some evidence for the cognitive primacy of categorization and its functional basis. *Merrill-Palmer Quarterly,* **19,** 21–39.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General,* **104,** 192–233.

Rosch, E. (1978). Principles of categorization. In R. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.

Scott, A. (1989). The vertical dimension and time in Mandarin. *Australian Journal of Linguistics,* **9,** 295–314.

Sera, M., Berge, C., & del Castillo, J. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development,* **9**(3), 261–292.

Slobin, D. (1987). Thinking for speaking. *Proceedings of the Berkeley Linguistic Society,* **13,** 435–445.

Slobin, D. (1996). From ''thought and language'' to ''thinking for speaking.'' In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, MA: Cambridge Univ. Press.

Tardif, T. (1996). Nouns are not always learned before verbs, but why? Evidence from Mandarin Chinese. *Developmental Psychology,* **32**(3), 492–504.

Traugott, E. (1978). On the expression of spatiotemporal relations in language. In J. H. Greenberg (Ed.), *Universals of human language: Word structure* (Vol. 3, pp. 369–400). Stanford, CA: Stanford Univ. Press.

Waxman, S., & Kosowski, T. (1990). Nouns mark category relations: Toddlers' and Preschoolers' word-learning biases. *Child Development,* **61,** 1461–1473.

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

(Accepted June 10, 1999, published online May 3, 2001)