Running head: Space, time, and language

How deep are effects of language on thought? Time estimation in speakers of English and Greek

Daniel Casasanto[1], Olga Fotokopolou[2], Ria Pita[2], and Lera Boroditsky[1]

Department of Psychology, Stanford University[1]
Department of Psychology, Aristotle University of Thessaloniki[2]

Word counts:

Abstract: 228
Main Text & Notes: 7,265

Corresponding Author:
Daniel Casasanto (casasan@stanford.edu)
Department of Psychology
Jordan Hall, Bldg. 420
Stanford University
Stanford, CA 94305
Phone: 650.353.1700

Abstract

Languages differ dramatically in the way they describe the world. Does talking differently lead speakers of different languages to *think* differently, even when they're not using language? Across languages, speakers use spatial metaphors to talk about time, but these metaphors vary: English speakers talk about duration more often in terms of linear distance (e.g., *a long time*), whereas Greek speakers talk about duration more often in terms of amount (e.g., *poli ora*, tr. 'much time'). To determine whether this difference in language leads to a difference in thinking, we compared Greek and English speakers' ability to estimate duration in the presence of distracting information about distance or amount, using simple duration reproduction tasks with non-linguistic stimuli and responses. Participants' non-linguistic duration estimates varied as predicted by the space-time metaphors in their native languages: English speakers' duration estimates were more influenced by irrelevant distance information, and Greek speakers' by irrelevant amount information. Next, we trained English speakers to use Greek-like metaphors for duration (e.g., a week is *more* than a day), which resulted in Greek-like performance on a non-linguistic duration estimation task. These findings demonstrate that (a) people who talk about time differently also think about it differently, and that (b) language not only reflects the structure of our non-linguistic mental representations, it can also shape those representations in fundamental ways that can be observed even low-level perceptuo-motor tasks.

Keywords: Metaphor; Time; Space; Whorfian Hypothesis; Embodied Cognition

Introduction

Languages differ dramatically in the way they describe the world (e.g., Slobin, 1987). Does talking differently lead speakers of different languages to *think* differently, even when they're not using language? The idea that language shapes the way we think, often associated with Benjamin Whorf (1956), was decried for decades on theoretical grounds (Chomsky, 1973; Fodor, 1985; Pinker, 1994), and was considered to be disconfirmed by empirical tests (Au, 1983; Papafragou, Massey, & Gleitman, 2002; Rosch-Heider, 1972). Yet, recent experimental evidence has reopened debate about the extent to which language influences non-linguistic cognition in domains such as *color* (Gilbert, Regier, Kay, & Ivry, 2006; Kay & Kempton, 1984; Robertson, Davies & Davidoff, 2000; Witthoft, et al., 2003), *number* (Casasanto, 2005a; Gordon, 2004; Gelman & Gallistel, 2004; Miller, Major, Shu, & Zhang, 2000; Pica, Lemer, Izard, & Dehaene, 2004; Spelke & Tsivkin, 2001), *space* (Levinson, 1996; Li & Gleitman, 2002), and *time* (Boroditsky, 2001; Chen, in press; January & Kako, in press; Núñèz & Sweetser, 2006). One obstacle to resolving this controversy, particularly in the more abstract conceptual domains of *number* and *time*, has been devising non-linguistic tests to evaluate how speakers of different languages perceive or remember their experiences. For the studies reported here, we developed non-linguistic perceptuo-motor tasks to investigate whether metaphors in language influence our non-linguistic mental representations of time.

Across languages, people use the same words to talk about time that they use to talk about space (Alverson, 1994; H. Clark, 1973; Gruber, 1965; Haspelmath, 1997; Jackendoff, 1983; Lakoff & Johnson, 1980; Traugott, 1978). For example, we might talk about 'a *long* vacation' or 'a *long* rope', and 'moving a meeting *forward*' or 'moving a truck *forward*'. Evidence from behavioral experiments suggests that people not only talk about time using spatial language, they also think about time using mental representations of space (Boroditsky, 2000, 2001; Boroditsky & Ramscar, 2002; Casasanto, 2005b; Casasanto & Boroditsky, 2003; in review; Casasanto & Lozano, 2006; in press; Cohen, 1967; Gentner, 2001; Núñèz & Sweetser, 2006; Piaget, 1927/1969; Torralbo, Santiago, & Lupiáñez, 2006; Tversky, Kugelmass, & Winter, 1991). Although using spatial metaphors for time appears to be universal (Alverson, 1994), the particular mappings from space to time vary across languages. For instance, depending on the language we're speaking we might talk about the future as if it lies *ahead of* us (in English), *behind* us (in Aymara), or *below* us (in Mandarin Chinese). Behavioral studies suggest that speakers of languages that use different spatio-temporal metaphors may indeed think about time differently (Boroditsky, 2001; Núñèz & Sweetser, 2006; Tversky, Kugelmass, & Winter, 1991).

There is, however, an important limitation shared by the cross-linguistic experiments reported to date: subjects were tested on tasks that required them to produce or understand language (e.g., Núñèz & Sweetser's participants were producing co-speech gestures; Boroditsky's were judging sentences containing spatial or temporal language). Perhaps these studies showed relations between spatial and temporal thinking that were consistent with linguistic metaphors only because participants were required to process space or time *in language*. Would the same relations between mental representations of space and time be found if participants were tested on non-linguistic tasks?

The experiments we report here were designed to test whether speakers of languages that use different spatio-temporal metaphors think about time differently even when they're not using language. First, we analyzed previously unexplored cross-linguistic differences in metaphors for duration in English and Greek. Next, we tested whether these linguistic differences correlate with differences in English and Greek speakers' performance on low-level, non-linguistic time estimation tasks. Finally, we conducted a training study to evaluate a causal role for language in shaping time representations.

Experiment 1: Distance and Amount Metaphors for Time

English speakers often rely on linear *distance* metaphors to talk about duration (e.g., a *long* party, a *short* concert, an *extended* period). In addition, duration can also be described in terms of an *amount* of time (e.g., *large amounts* of time, a *lot* of time, a *big chunk* of time, a *little bit* of time).

Languages differ in the extent to which they describe duration in terms of distance as opposed to amount. In English, it is natural to talk about *a long time* or *a long meeting*, borrowing the structure and vocabulary of a linear spatial expression like *a long rope*. In Greek, the words *makris* and *kontos* are the literal equivalents of the English spatial terms *long* and *short*. They can be used in spatial contexts much the way long and short are used in English (e.g., *ena makry skoini* means 'a long rope'). In temporal contexts, however, *makris* and *kontos* are dispreferred in instances where long and short would be used naturally in English. It would be unnatural to translate *a long meeting* literally as *mia makria synantisi*. Rather than using distance terms, Greek speakers typically indicate that an event lasted a long time using *megalos*, which in spatial contexts means physically 'large' (e.g., a big building), or using *poli*, which in spatial contexts means 'much' (e.g., much water). Compare how English and Greek typically modify the duration of the following events (Table 1.)

Table 1. English (e) and Greek (g) expressions for event durations (literal translations in parentheses).

1e. long night
1g. *megali nychta* (big night)

2e. long relationship
2g. *megali schesi* (big relationship)

3e. long party
3g. *parti pou kratise poli* (party that lasts much)

4e. long meeting
4g. *synantisi pou diekese poli* (meeting that lasts much)

In examples 1g. and 2g., the literal translations might surprise an English speaker, for whom a *big night* is likely to mean 'an exciting night', and a *big relationship* 'an important relationship'. For Greek speakers, however, these phrases naturally communicate duration, expressing time not in terms of linear distance, but rather in terms of amount.

As an index of the relative frequency of distance and amount metaphors for duration across languages, the most natural phrases expressing the ideas 'long time' and 'much time' were elicited from native speakers of English (i.e., *long time*; *much time*) and Greek (i.e., *makry kroniko diastima*; *poli ora*). The frequencies of these expressions were compared in a very large multilingual text corpus: www.google.com. Google's language tools were used to find exact matches for each expression, and to restrict the search to web pages written only in the appropriate languages. The number of google 'hits' for each expression was tabulated, and the proportion of distance hits and amount hits was calculated for each pair of expressions. Results showed that in English, distance metaphors were dramatically more frequent than amount metaphors, whereas the opposite pattern was found in Greek (fig. 1a).

Although both languages use *both* distance and amount metaphors for duration, the distribution of 'long time' and 'much time' expressions in this corpus varied significantly between languages (X^2=76029.70, df=1, p<.00001). This simple corpus search by no means captures all of the complexities of how time is metaphorized in terms of distance and amount within or between languages, but these data corroborate native speakers' intuitions for each language, and provide one quantitative linguistic measure on which to base predictions about behavior in non-linguistic tasks.

Whereas the corpus search measured how often English and Greek speakers use distance and amount metaphors to talk about *time*, per se (e.g., *long time*, *much time*, etc.), a questionnaire study was conducted to investigate cross-linguistic differences in the use of distance and amount metaphors to describe the durations of *events* (e.g., *a long night*, *a 'big' night*, etc.).

Methods

Participants A total of 30 participants (21 English, 9 Greek) completed the questionnaire in exchange for payment. Native English speakers were recruited and tested at MIT. Native Greek speakers were recruited and tested at the Aristotle University of Thessaloniki. Participants were considered native speakers of a language only if it was the only language they learned before age 5 y.o. and it was their strongest language at the time of testing, according to a language background questionnaire.

Materials For the English version of the questionnaire, thirty-five words were selected from the combined Kucera & Francis and Wall Street Journal corpora: the twenty-five highest frequency common nouns that could refer to an event (e.g., day, night, party, war, process, etc.), and the 10 highest frequency adjectives or quantifiers that could potentially describe the duration of an event in at least one of the languages investigated (e.g., long, short, big, little, much, etc.) The nouns and adjectives/quantifiers were fully crossed to produce 250 two-word noun phrases (e.g., *short party, big night, much war*). For the Greek version of the questionnaire, English stimuli and instructions were translated by a native Greek speaker. The 10 English adjectives/quantifiers translated to 8 distinct Greek adjectives/quantifiers, which combined with the 25 nouns to produce 200 two-word noun phrases. (For a complete list of materials, see Appendix A.)

Procedure Participants were tested in their native languages. Noun phrases were presented in the middle of a computer screen, one at time, in randomized order. Participants pressed a button to respond to two questions about each of the noun phrases. First, they indicated which property of the event the adjective or quantifier was describing: (a) duration, (b) physical size, (c) importance, (d) how good or bad, (e) none of the above. Second, participants indicated whether they would be likely to use the adjective/quantifier to describe the *duration* of the event (yes or no).

The first question was included to orient participants to the dimension of the event described by the adjective or quantifier, in order to increase the validity of responses on the second question, which was of principal interest. The phrase ‘long meeting’ probably describes the event’s duration, but ‘big meeting’ probably describes its size or importance. Orienting participants to the relevant dimension was intended to reduce spurious ‘yes’ responses on the second question for phrases like ‘big meeting’ which are highly familiar, but do not ordinarily denote duration in English.

Results and Discussion

The proportions of ‘yes’ responses on the second question were computed for noun phrases using distance and amount metaphors, and were compared across groups. English speakers were more likely than Greek speakers to judge distance metaphors as describing the durations of events, whereas Greek speakers were more likely to judge amount metaphors as describing duration (difference of proportions = .37, $t(28) = 11.54$, $p<.00001$; fig. 1b).

These results corroborate those of the Google search. Together, these findings demonstrate a previously unexplored cross-linguistic difference in spatial metaphors for time: English prefers *distance* metaphors whereas Greek prefers *amount* metaphors to describe both the duration of time, per se, and the durations of events.

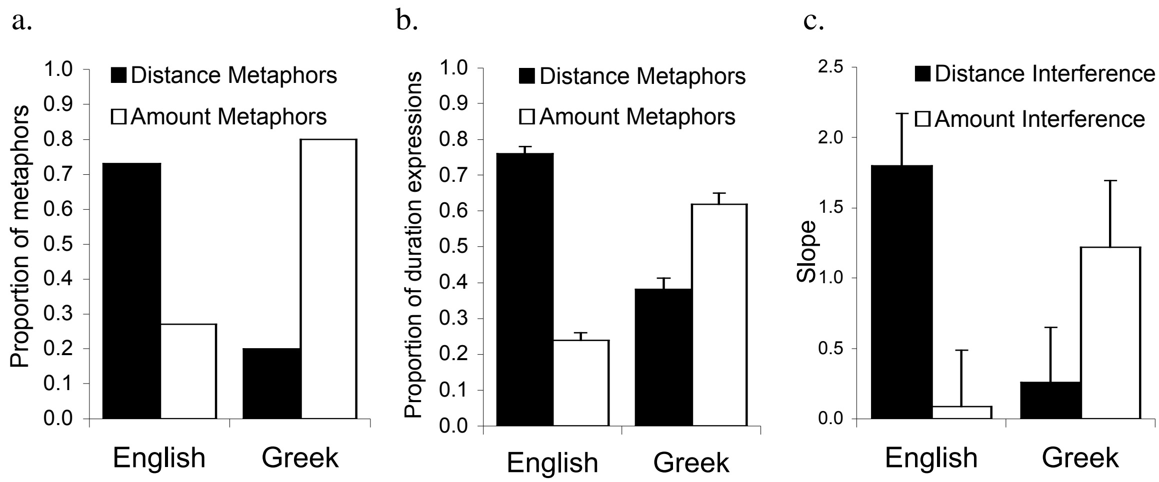

Figure 1. 1a. (left): Results of Experiment 1 corpus search. Black bars indicate the proportion of Google 'hits' for expressions meaning *long time*, and white bars for expressions meaning *much time* in English and Greek. 1b. (middle): Results of Experiment 1 questionnaire. Black bars indicate the proportion of duration expressions using distance metaphors and white bars the proportion of duration expressions using amount metaphors, as judged by English and Greek participants. 1c. (right): Results of Experiment 2. Black bars indicate the slope of the effect of distance on time estimation (Growing Lines) and white bars the slope of the effect of amount on time estimation (Filling Containers) in speakers of English and Greek. (Error bars indicate s.e.m.)

Experiment 2: Do people who talk differently also think differently?

Do these differences in metaphor frequency between languages lead English and Greek speakers to think about time differently, even when they're not using language? In Experiment 2 we asked English and Greek speakers to estimate the duration of events that contained either distracting information about linear distance (distance interference) or distracting information about amount (amount interference). In the distance interference condition, subjects viewed a line "growing" across the screen, and were asked to reproduce the duration of the growing event. The distance that the line grew was varied orthogonally to the duration of the event, and as such served as distracter: a piece of information that was unrelated to the task, but could potentially interfere with task performance. In the amount interference condition, subjects viewed a schematic drawing of a container filling with liquid and were asked to reproduce the duration of the filling event. Analogously to the distance interference condition, the amount of fill varied orthogonally with the event duration, and as such served as a distracter for the subject's task of estimating duration.

In previous studies using a similar distance interference task with English speakers (Casasanto & Boroditsky, 2003; in review), we found that English-speaking participants were unable to ignore the distance that a line grew when estimating its duration, even though distance was irrelevant to the time estimation task. Is this conflation of distance and duration universal to humans, or does it depend in part on the conflation of distance and duration in language? If patterns in language are partly responsible for the distance-duration confusion in English speakers, then irrelevant distance and amount information should interfere with English and Greek participants' duration estimates differently. English speakers should show more interference from distance than amount on their time estimates. Greek speakers should show the opposite pattern, being more distracted by amount than by distance interference.

Methods

Participants

A total of 40 participants performed the distance interference task (20 English, 20 Greek) and 40 performed the amount interference task (20 English, 20 Greek), in exchange for payment. Of these, 9 participants were removed from analyses of the distance interference task (3 English, 6 Greek) and 10 were removed from analyses of the amount interference task (4 English, 6 Greek) for performing the experiments incorrectly (e.g., estimating distance when instructed to estimate time), or for excessively poor performance: Ss were excluded if the slope of the correlation between actual time and estimated time was less than .50 (e.g., if they reported that the 4-second stimuli lasted less than 2-seconds on average), as this was thought to reflect impatience with the repetitive task rather than genuine inaccuracy. Native English speakers were recruited and tested at MIT. Native Greek speakers were recruited and tested at the Aristotle University of Thessaloniki.

Materials

For the distance interference task, subjects were shown lines growing across a computer screen. The growing line events varied in distance and duration. Durations ranged from 1000 milliseconds to 5000 milliseconds in 500 millisecond increments. Distances ranged from 100 to 500 pixels in 50 pixel increments. Nine durations were fully crossed with nine distances to produce 81 distinct line types. Lines ‘grew’ horizontally across the screen one pixel at a time, from left to right, along the vertical midline. Lines were situated in a square box (700 x 700 pixels), to minimize any influence that the asymmetric rectangular shape of the computer monitor might have on perception of horizontally growing lines vs. vertically filling containers. The starting point of each line was jittered with respect to the average starting point (+/- up to 50 pixels), so that the box would not provide a reliable spatial frame of reference. Each line remained on the screen until it reached its designated distance, and then it disappeared.

The amount interference stimuli were constructed analogously. Participants watched schematically drawn containers filling on the computer screen (600 pixels high x 500 pixels wide), and were asked to imagine these were tanks gradually filling with water. Nine durations were fully crossed with nine fill levels to produce 81 distinct trial types. Durations ranged from 1000 milliseconds to 5000 milliseconds in 500 millisecond increments. ‘Fill levels’ ranged from 100 to 500 pixels, in 50 pixel increments. Empty containers filled gradually, one row of pixels at a time, for varying durations and fill levels, and disappeared when they reached their designated fullness.

Procedure

Each participant performed either the distance interference or amount interference task. Instructions were presented on the screen before each task in the participant’s native language. No distance or amount metaphors for time were used in the instructions. The tasks, themselves, were entirely non-linguistic, consisting of growing lines or filling containers (stimuli) and mouse clicks (responses).

For the distance interference task, participants viewed 162 line stimuli, one at a time in random order, from a viewing distance of approximately 50 cm. For each trial, participants reproduced either the displacement or the duration of the stimulus, never both. Before each stimulus, an icon appeared for 2 seconds in the center of the screen alerting that the participant that s/he would need to reproduce either the distance that the line traveled (if an 'X' icon appeared) or the duration for which it remained on the screen (if an 'hourglass' icon appeared). Immediately after each stimulus was shown, the same icon appeared as a response prompt. To estimate distance, subjects clicked the mouse once on the center of the X icon, moved the mouse to the right in a straight line, and clicked the mouse a second time to indicate that they had moved a distance equal to that of the stimulus. To estimate duration, subjects clicked the mouse once on the center of the hourglass icon, waited the appropriate amount of time, and clicked again in the same spot.

The amount interference task was conducted analogously. Before each stimulus, an icon appeared for 2 seconds in the center of the screen alerting the participant that s/he would need to reproduce either the fullness of the container (if an 'X' icon appeared) or the duration for which it remained on the screen (if an 'hourglass' icon appeared). Immediately after each stimulus was shown, the same icon appeared as a response prompt. To estimate fullness, subjects clicked the mouse once on the center of the X icon, moved the mouse up the side of the container in a straight line, and clicked the mouse a second time to indicate that they had moved a distance equal to the 'water level' that had been reached. To estimate duration, subjects clicked the mouse once on the center of the hourglass icon, waited the appropriate amount of time, and clicked again in the same spot, exactly as in the growing line task.

For both tasks, all responses were self-paced. Response data were collected for both the trial-relevant and the trial-irrelevant stimulus dimensions, to ensure that subjects were following instructions. Tasks lasted about 30 minutes.

Results and Discussion

To avoid experimental artifacts, the two highest and lowest durations and distances/amounts were excluded from the analysis, leaving the middle five durations and distances/amounts to be analyzed. Trimming the endpoints of perceptual continua in this way is common practice in magnitude estimation tasks, as endpoints are susceptible to strategic responding on the part of the subject. In this case, the longest and shortest stimuli (in space and time) were most amenable to verbal labeling, whereas the middle stimuli were least likely to be covertly labeled as *long* or *short*, etc.

To evaluate the effects of irrelevant distance and amount information on time estimation, each participant's time estimates in milliseconds were plotted as a function of the actual distances that lines traveled in pixels (for the distance interference task), or the actual change in 'water level' in pixels (for the amount interference task). A line of best fit was computed, and the slope was used as a measure of cross-dimensional interference. The standard deviation of the slopes was computed, and participants whose slope was more than 2 standard deviations from the mean for their task and language group were excluded from further analysis (this resulted in the removal of only one subject, overall). Next, for group comparisons, grand averaged time estimates in milliseconds were plotted as a function of distance or fill level in pixels, and slopes were compared between tasks and between language groups.[1]

As predicted by the patterns in language, English and Greek speakers showed different patterns of interference in the duration estimation: English speakers were more affected by irrelevant distance information than by amount, while Greek speakers were more affected by irrelevant amount information than by distance (Figure 1c).

For the distance-interference task, English speakers showed a strong, significant effect of distance on time estimation, whereas Greek speakers showed a weak, non-significant effect (English: N=17, Slope=1.80, r^2=0.98, t=12.16, df=3, p<0.0001; Greek: N=13, Slope=0.26, r^2=0.08, t=0.54, df=3, *ns*; fig. 2a-b). The opposite pattern of results was found for the amount-interference task. English speakers showed a weak, non-significant effect of amount on time estimation, whereas Greek speakers showed a strong, significant effect (English: N=16, Slope=0.09, r^2=0.005, t=0.12, df=3, *ns*; Greek: N=14, Slope=1.22, r^2=0.83, t=3.80, df=3, p<0.02; fig. 2c-d). A 2 x 2 ANOVA with Language (English, Greek) and Task (distance interference, amount interference) as between-subject factors revealed a highly significant Language by Task interaction, with no main effects (F(1,56)=10.41, p<.002). Post-hoc independent-samples t-tests showed that within language groups, the effect of distance on time estimation was significantly greater than the effect of amount on time estimation for English speakers (difference of slopes=1.72, t=3.15, p<.002), whereas the opposite trend was found for Greek speakers (difference of slopes=0.96, t=1.52, p<.07). Furthermore, between language groups, the effect of distance on time estimation was significantly greater in English speakers than in Greek speakers (difference of slopes=1.54, t=2.78, p<.005), whereas the effect of amount on time estimation was significantly greater in Greek speakers than in English speakers (difference of slopes=1.13, t=1.18, p<.04).

The differences we observed in the effects of distance and amount on duration estimation cannot be attributed to overall differences in duration estimation, per se. The effect of actual duration on estimated duration was compared across both tasks and both groups: all correlation coefficients were above .98, and no significant difference between slopes was found (F(3,56)=0.54, p<.66; fig. 3).[2]

In summary, the pattern of cross-dimensional interference observed in English and Greek speakers on this pair of non-linguistic time estimation tasks was predicted by the pattern of spatio-temporal metaphors found in English and Greek.[3]

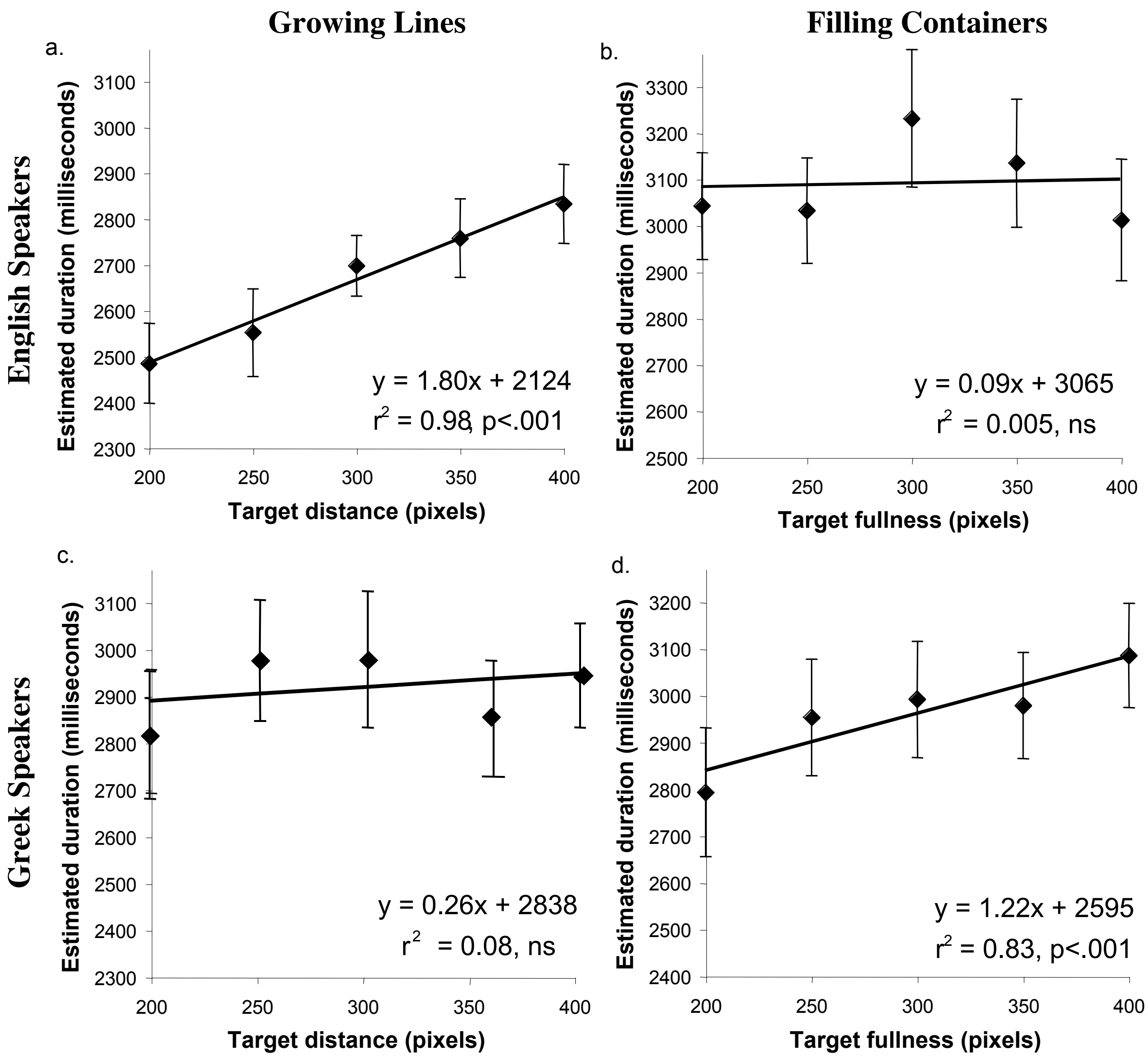

Figure 2. Results of Experiment 2: grand averaged duration estimates as a function of actual line distance or actual tank fullness in English speakers (top) and Greek speakers (bottom). 2a. Effect of distance on time estimation in English speakers. 2b. Effect of tank fullness on time estimation in English speakers. 2c. Effect of distance on time estimation in Greek speakers. 2d. Effect of tank fullness on time estimation in Greek speakers. (Error bars indicate s.e.m.)

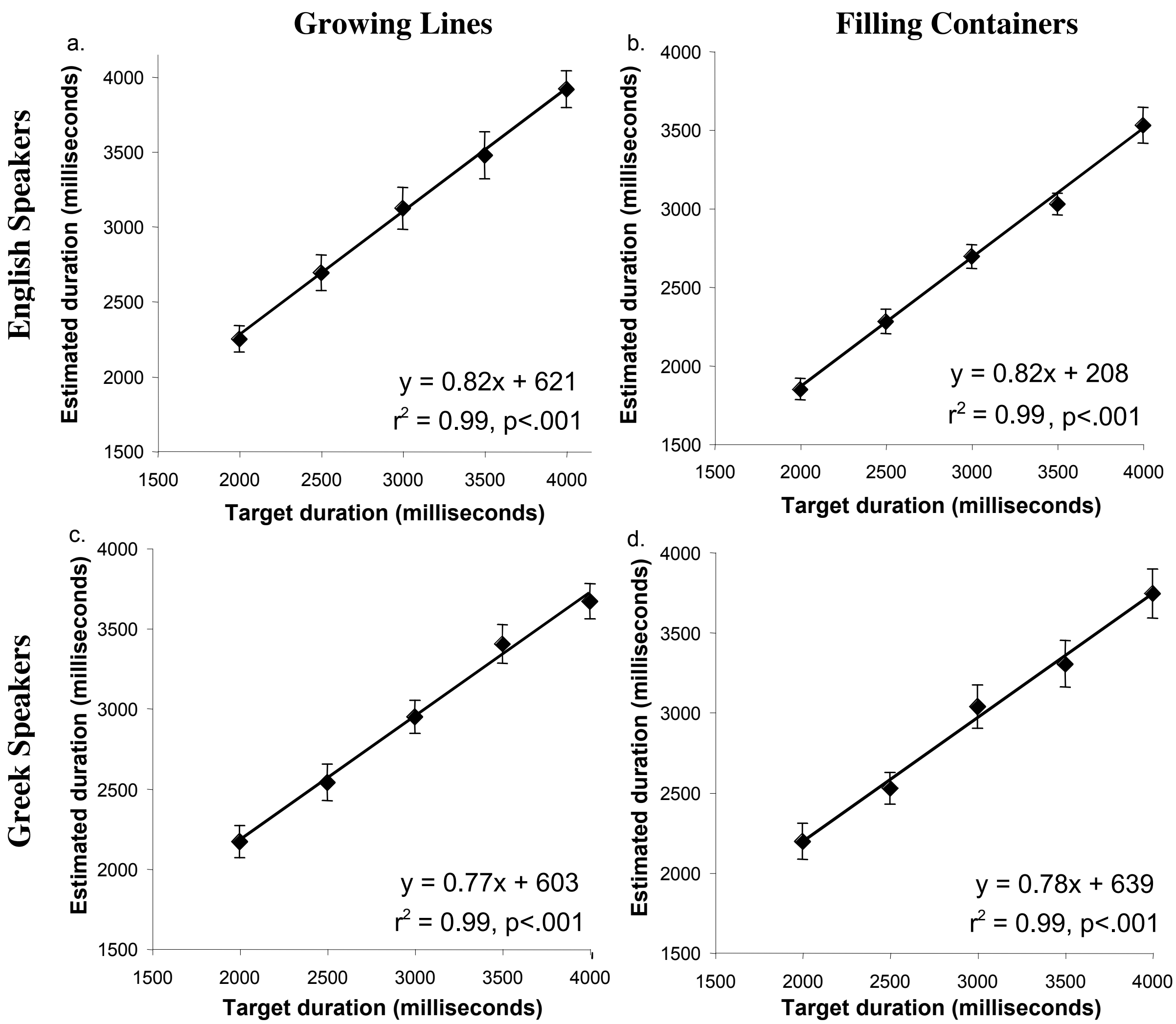

Figure 3. Results of Experiment 2, continued: grand averaged duration estimates as a function of actual target duration in English speakers (top) and Greek speakers (bottom). 3a-b. Effect of actual duration on estimated duration for the distance interference task (left) and amount interference task (right) in English speakers. 3c-d. Effect of actual duration on estimated duration for the distance interference task (left) and amount interference task (right) in Greek speakers . (Error bars indicate s.e.m.)

Experiment 3: Does language shape the way we think?

A perennial complication for studies that aim to show effects of language on thought is that researchers rely on correlation to argue for causation (Casasanto, 2004; Pinker, 1994; Pullum, 1991). Our data so far show that patterns in non-linguistic duration estimation align closely with patterns in language. However, the direction of causation is hard to infer from these data. Did differences in language between English and Greek speakers give rise to the cross-linguistic differences in duration estimation in this case? Or are there some other extra-linguistic cultural factors that cause the differences in time-estimation patterns (and perhaps also give rise to the differences between the languages)? Using cross-linguistic comparisons to test for causal influences of language on thought is problematic because speakers of different languages differ not only in the language they speak, but also in a myriad of other extra-linguistic cultural factors. Because cross-linguistic comparisons afford only quasi-experimental designs (subjects are not randomly assigned to groups), it is not possible to infer causality from such comparisons.

While we cannot randomly assign subjects to be Greek or English speakers, we can do the logical equivalent by giving people different language-training experience in the laboratory. In Experiment 3, we trained English speakers to talk about duration using either distance metaphors (already the dominant metaphor in English) or using amount metaphors (similar to the dominant pattern in Greek). In effect, we randomly assigned one group of English speakers to be English speakers, and another group of English speakers to become Greek.

Native English speakers underwent either ‘distance training’ or ‘amount training.’ Participants completed fill-in-the-blank sentences using the words *longer* or *shorter* for distance training, and *more* or *less* for the amount training task. Half of the sentences compared the length or size of physical objects (e.g., An alley is *longer / shorter* than a clothesline; A teaspoon is *more / less* than an ocean), the other half compared the duration of events (e.g., A sneeze is *longer / shorter* than a vacation; A sneeze is *more / less* than a vacation).

If using a linguistic metaphor activates a corresponding non-linguistic mapping, then training English speakers to talk about duration using amount terms should (transiently) reinforce the mapping between amount and time (which is ordinarily weaker for English speakers than for speakers of a language like Greek). Training English speakers to talk about duration using distance terms should reinforce the already preferred mapping between distance and time (this task served as a control condition). After the linguistic training, all participants performed the non-linguistic amount interference task from Experiment 2. If using one set of linguistic space-time metaphors or another can influence how people mentally represent time, then using amount metaphors to talk about event durations during training should cause English speakers to be more distracted by amount information during the non-linguistic duration estimation task, and to show a pattern of cross-dimensional amount interference similar to that of native Greek speakers.

Methods

Participants

A total of 36 Stanford University students performed the distance training task and 36 performed the amount training task, in exchange for payment. Of these, 12 participants (6 distance trainees, 6 amount trainees) were removed from the analyses reported here for performing the experiment incorrectly, or for excessively poor performance according to the criteria described in Experiment 2. All participants were native monolingual speakers of English, according to a language background questionnaire.

Materials and Procedure

Participants were randomly assigned to perform either the distance training or amount training task. They completed 192 fill-in-the-blank sentences using the words *longer* or *shorter* for distance training, and *more* or *less* for the amount training task. Half of the sentences compared the length or size of physical objects, and the other half compared the duration of events (for a complete list of stimuli, see Appendix B). Sentence types were randomly intermixed. Trials were self-paced, and the training tasks lasted approximately 20 minutes. After training, all participants performed the amount-interference task used in Experiment 2.

Results and Discussion

Results of the training phase showed that participants filled in the blanks with high accuracy for both the distance training task (N=30, %Correct=98%, SD=0.02%) and the amount training task (N=30, %Correct=98%, SD=0.07%). An independent-samples t-test showed that accuracy did not differ between tasks (difference of means=0.05%, t(58)=1.22, *ns*).

Data for the post-training amount-interference task were analyzed as described in Experiment 2. An omnibus one-way ANOVA compared the slopes of the effect of amount on time estimation in distance-trained English speakers and amount-trained English speakers from Experiment 3, as well as untrained English and Greek speakers from Experiment 2 (F(3,86)=3.24, p<.03; fig. 4). Planned independent-samples t-tests showed that, most importantly, the slope of the effect of amount on time estimation was significantly greater after amount training than after distance training (difference of slopes=0.93, $t(58)=2.40$, $p<.01$). Furthermore, in distance-trained English speakers the effect of amount on time estimation was statistically indistinguishable from the effect in untrained English speakers (difference of slopes=0.18, $t(44)=0.33$, *ns*), but was significantly weaker than the effect in untrained Greek speakers (difference of slopes=1.32, $t(42)=2.22$, $p<.02$). By contrast, in amount-trained English speakers the effect of amount on time estimation was significantly stronger than in untrained English speakers (difference of slopes=0.75, $t(44)=1.92$, $p<.04$), but was statistically indistinguishable from the effect in untrained Greek speakers (difference of slopes=0.38, $t(42)=0.91$, *ns*).

Linguistic experience with amount metaphors (but not distance metaphors) caused native English speakers to perform the amount-interference task more like native speakers of Greek than like native speakers of English, demonstrating that language can influence even our low-level, non-linguistic time representations. This training task illustrates a process by which our everyday linguistic experience may shape our metaphorical concepts. If giving English speakers a concentrated ‘dose’ of amount metaphors in the lab can transiently strengthen their conceptual mapping between amount and time, perhaps ordinary doses of Greek have a similar effect on Greek speakers’ conceptual mappings, over a longer timecourse. On this view, to use our native language is to participate in a natural training experiment.

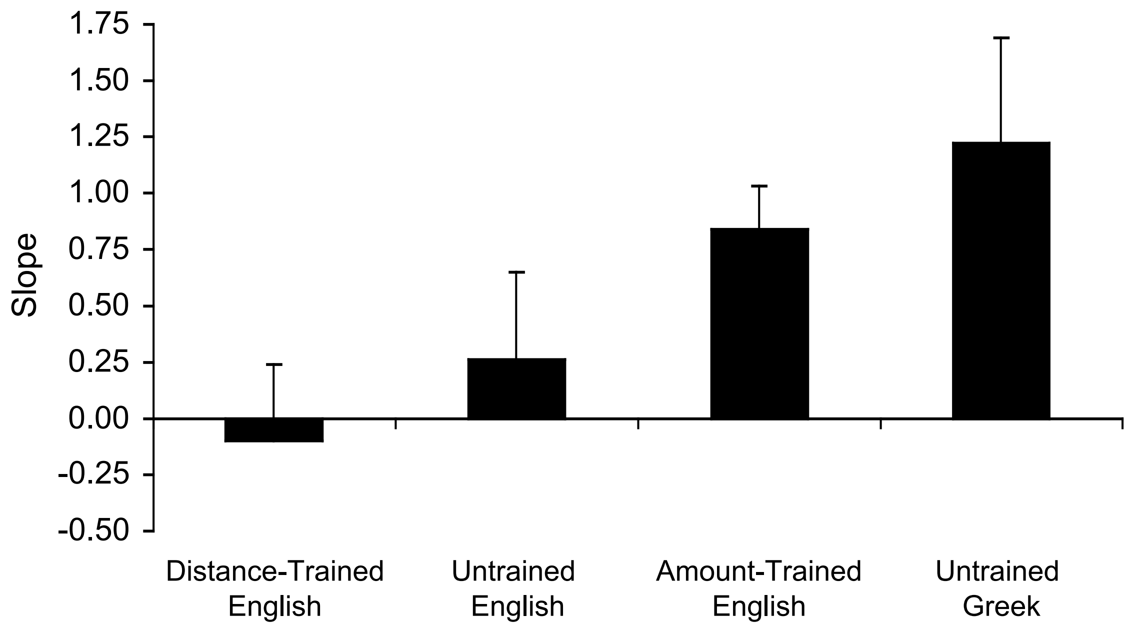

Figure 4 Comparison of the effects of amount interference on non-linguistic time estimation in untrained English and Greek speakers (Experiment 2) and in English speakers trained to use either distance or amount metaphors for duration (Experiment 3). (Error bars indicate s.e.m.)

General Discussion

Previous experiments with linguistic stimuli and responses suggest that people who use different space-time metaphors in their native languages think about time differently (Boroditsky, 2001; Núñèz & Sweetser, 2006; Tversky, Kugelmass, & Winter, 1991). The studies we report here were designed to test whether language shapes the way people mentally represent time even when they're not using language. Specifically, we sought to discriminate two theoretical positions, either of which was tenable given the previous evidence, one positing *shallow* and the other *deep* effects of language on mental representations of time.

One view is that cross-linguistic differences in thinking emerge only when people are actively engaged in the process of producing or understanding language. When people are planning their own speech or processing spoken or written language, their thoughts must be structured, in part, according to their language and its peculiarities (Slobin, 1987). Consequently, speakers of different languages may think differently during language use. But how about when they're not using language? When people are thinking for the purpose of categorizing, remembering, perceiving, or acting, they may activate mental representations that differ from those activated during language use, and that may include information not ordinarily encoded in their language. It is plausible that these non-linguistic representations are closer to being universal than the representations people construct for the purpose of speaking. As such, results might differ dramatically between tests of language-thought relations that require people to use language and those that do not. For tasks that don't require producing or understanding language, people who speak different languages may respond in similar, or even identical, ways. On this *shallow view* of language-thought relations, cross-linguistic typological differences may produce behavioral differences, but only on tasks that require language.

Alternatively, the habit of perceiving, attending, categorizing, remembering, or acting in language-particular ways during language use may lead to *deep* influences of language on non-linguistic cognition. Some cross-linguistic typological differences may produce behavioral differences observable not only during tasks requiring language, but also when subjects perform non-linguistic tests of low-level perceptuo-motor abilities. On this *deep view* of language-thought relations, where languages differ, speakers' corresponding conceptual and perceptual representations may also differ.

The present studies support the *deep view* of language-thought relations. Experiment 1 quantified cross-linguistic differences in English and Greek speakers' reliance on time-as-distance metaphors and time-as-amount metaphors. Whereas English speakers tend to use distance metaphors for duration more than amount metaphors, Greek speakers show the opposite pattern. Experiment 2 showed that these patterns in language predicted differences between English and Greek speakers' performance on simple duration reproduction tasks: English speakers were more strongly affected by irrelevant distance information in the stimuli, whereas Greek speakers were more strongly affected by irrelevant amount information. This was true despite the fact that these perceptuo-motor tasks used entirely non-linguistic stimuli and responses. Experiment 3 demonstrated that experience using one linguistic metaphor or another can influence performance on a simple, non-linguistic duration estimation task. This training study was important for establishing that language and behavior are not merely correlated; rather, linguistic experience can play a causal role in shaping non-linguistic mental representations of time.

Conclusions

People who talk differently about time also think about it differently, in ways that correspond to their language-particular metaphors. Language not only reflects the structure of our underlying mental representations of time, it can also shape those representations. Beyond influencing how people think when they are required to speak or understand language, language can also shape the non-linguistic representations we build for the purpose of remembering, acting on, and perhaps even perceiving the world around us. It may be universal that people conceptualize time according to spatial metaphors, but since these metaphors vary across languages, members of different language communities develop distinctive conceptual repertoires.

References

Alverson, H. (1994). *Semantics and experience: Universal metaphors of time in English, Mandarin, Hindi, and Sesotho*. Baltimore: Johns Hopkins University Press.

Au, T. (1983). Chinese and English counterfactuals: The Sapir-Whort hypothesis revisited. *Cognition, 15*, 155-187.

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

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*(1), 1-22.

Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science, 13*(2), 185-189.

Casasanto, D. (2005a). Crying "Whorf!" *Science, 307*, 1721-1722.

Casasanto, D. (2005b). *Spatial foundations of abstract thought*. Doctoral Dissertation, MIT, Cambridge.

Casasanto, D., & Boroditsky, L. (2003). *Do we think about time in terms of space?* Proceedings of the 25th Annual Conference of the Cognitive Science Society, Boston.

Casasanto, D., & Boroditsky, L. (In review). Time in the mind: Using space to think about time.

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., et al. (2004). *How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish*. Proceedings of the 26th Annual Conference of the Cognitive Science Society. Chicago.

Casasanto, D., & Lozano, S. (2006). *Metaphor in the mind and hands*. Proceedings of the 28th Annual Meeting of the Cognitive Science Society, Vancouver, BC.

Casasanto, D., & Lozano, S. (In press). The meaning of metaphorical gestures. In A. Cienki & C. Muller (Eds.), *Metaphor and Gesture*. Amsterdam: John Benjamins.

Chen, J. (In press). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition, doi:10.1016/j.cognition.2006.09.012*.

Chomsky, N. (1973). Introduction to Language and Cognition. In A. Schaff, *Language and Cognition*. New York, NY: McGraw Hill.

Clark, H. H. (1973). Space, time, semantics and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27-63). New York: Academic Press.

Fodor, J. (1985). Precis of the Modularity of Mind. *Behavioral and Brain Scences, 8*, 1-5.

Gentner, D. (2001). Spatial metaphors in temporal reasoning. In M. Gattis (Ed.), *Spatial Schemas and Abstract Thought* (pp. 203-222). Cambridge: MIT Press.

Gilbert, A., Regier, T., Kay, P., & Ivry, R. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, 103*(2), 489-494.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science, 306*, 496.

Gruber, J. (1965). *Studies in lexical relations*. MIT, Cambridge.

Haspelmath, M. (1997). *From space to time: Temporal adverbials in the world's languages*. Munich & Newcastle: Lincom Europa.

Jackendoff, R. (1983). *Semantics and Cognition*. Cambridge: MIT PRess.

January, D., & Kako, E. (2006). Re-evaluating evidence for linguistic relativity: Reply to Boroditsky (2001). *Cognition, doi:10.1016/j.cognition.2006.07.008*.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist, 86*(1), 65-79.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Levinson, S. C. (1996). Frames of reference and Molyneux's question: Cross-linguistic evidence. In P. Bloom, M. Peterson, L. Nadel & M. Garrett (Eds.), *Language and space*. *Language, speech, and communication* (pp. 385–436). Cambridge: MIT Press.

Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition, 83*, 265–294.

Miller, K., Major, S., Shu, H., & Zhang, H. (2000). Ordinal knowledge: Number names and number concepts in Chinese and English. *Canadian Journal of Experimental Psychology, 54*(2), 129-139.

Núñez, R., & Sweetser, E. (2006). Looking ahead to the past: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science, 30*(3), 401-450.

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n 'roll: The representation of motion in langauge and cognition. *Cognition, 84*, 189-219.

Piaget, J. (1927/1969). *The child's conception of time*. New York: Ballantine Books.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*, 499.

Pinker, S. (1994). Mentalese. In *The Language Instinct* (pp. 55-82). New York: Harper.

Pullum, G. (1991). The great Eskimo vocabulary hoax. Chicago: University of Chicago Press.

Robertson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129*, 369-398.

Rosch-Heider, E. (1972). Universals in color naming and memory. *Journal of Experimental Psychology, 93*(1), 10-20.

Slobin, D. (1987). *Thinking for Speaking*. Paper presented at the Berkeley Linguistics Society, Berkeley, CA.

Spelke, E., & Tsivkin, S. (2001). Language and number: A bilingual training study. *Cognition, 78*, 45-88.

Torralbo, A., Santiago, J., & Lupiáñez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science,(30)*4, 745-758.

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

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 23*, 515-557.

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

Witthoft, N., Winawer, J., Wu, L., Frank, M., Wade, A., & Boroditsky, L. (2003). Effects of language on color discriminability. *Proceedings of the 25th Annual Conference of the Cognitive Science Society,* Boston.

Acknowledgments

The authors thank the citizens of Cognation for discussion, Webb Phillips for programming, and Jesse Greene for help with data collection. This research was supported by NSF CAREER grant #0608514 to L. Boroditsky and an NSF Graduate Research Fellowship to D. Casasanto. These studies constituted part of D. Casasanto's doctoral dissertation at MIT.

Notes

1. Both time estimates and space estimates (i.e., distance or amount) were collected from each subject. However, no significant effects of actual time on estimated distance or amount were found, consistent with expectations based on our previous studies (Casasanto & Boroditsky, 2003; in review). Since our present hypothesis concerns effects of language on time estimation, only data for the time estimation trials are discussed here.

2. One difference between the growing line and filling container tasks was that the lines grew horizontally, but the containers filled vertically. To rule out the spatial orientation of the stimuli and responses as a potential source of the observed cross-linguistic differences in performance, and 'upward growing lines' task was administered to speakers of English and Greek. No significant difference was found in the effect of vertical displacement on time estimation across languages (see Casasanto, 2005b). Results suggested that the orientation of stimuli cannot account for the between-group differences observed in Experiments 2.

3. Preliminary data from experiments corroborating these results in Indonesian and Spanish speakers were reported previously (Casasanto, 2005b; Casasanto, et al. 2004). Studies in progress are further investigating the distribution of distance and amount metaphors in these languages.

Appendix

Appendix A. Stimuli used in Experiment 1 questionnaire study.

Adjectives and Quantifiers in English

Big
Full
Great
Huge
Large
Little
Long
Much
Short
Small

Adjectives and Quantifiers in Greek

GEMATH/ GEMATO/ GEMATOS
KONTH/ KONTO/ KONTOS
LIGH/ LIGO/ LIGOS
MAKRIA/ MAKRU/ MAKRUS
MEGALH/ MEGALO/ MEGALOS
MIKRH/ MIKRO/ MIKROS
POLLH/ POLU/ POLUS
TERASTIA/ TERASTIO/ TERASTIOS

Nouns in English

Action
Case
Contract
Day
Effort
History
Increase
Job
Life
Meeting
Month
Morning
Night
Party
Period
Problem
Process
Production
Program
Question
Report
Time
War
Week
Year

Nouns in Greek

ANAFORA
AUXHSH
CRONOS
DIADIKASIA
DOULEIA
DRASH
EBDOMADA
ERWTHMA
ETOS
GEGONOS
HMERA
ISTORIA
MHNAS
NUCTA
PARAGWGH
PARTU
PERIODOS
POLEMOS
PROBLHMA
PROGRAMMA
PROSPAQEIA
PRWI
SUMBASH
SUNANTHSH
ZWH

Appendix B. Stimuli used in Experiment 3 training study: blanks filled in with *longer*/*shorter* for distance training and *more*/*less* for amount training.

Temporal Sentences

A blink is ____ than a concert.
A blink is ____ than a lunch.
A blink is ____ than a movie.
A blink is ____ than a pregnancy.
A blink is ____ than a semester.
A blink is ____ than a summer.
A blink is ____ than a wedding.
A blink is ____ than a winter.
A breath is ____ than a concert.
A breath is ____ than a lunch.
A breath is ____ than a movie.
A breath is ____ than a pregnancy.
A breath is ____ than a semester.
A breath is ____ than a summer.
A breath is ____ than a wedding.
A breath is ____ than a winter.
A concert is ____ than a blink.
A concert is ____ than a breath.
A concert is ____ than a hiccup.
A concert is ____ than a pregnancy.
A concert is ____ than a semester.
A concert is ____ than a sneeze.
A concert is ____ than a summer.
A concert is ____ than a winter.
A hiccup is ____ than a concert.
A hiccup is ____ than a lunch.
A hiccup is ____ than a movie.
A hiccup is ____ than a pregnancy.
A hiccup is ____ than a semester.
A hiccup is ____ than a summer.
A hiccup is ____ than a wedding.
A hiccup is ____ than a winter.
A lunch is ____ than a blink.
A lunch is ____ than a breath.
A lunch is ____ than a hiccup.
A lunch is ____ than a pregnancy.
A lunch is ____ than a semester.
A lunch is ____ than a sneeze.
A lunch is ____ than a summer.
A lunch is ____ than a winter.
A movie is ____ than a blink.
A movie is ____ than a breath.
A movie is ____ than a hiccup.
A movie is ____ than a pregnancy.
A movie is ____ than a semester.
A movie is ____ than a sneeze.
A movie is ____ than a summer.
A movie is ____ than a winter.
A pregnancy is ____ than a blink.
A pregnancy is ____ than a breath.
A pregnancy is ____ than a concert.
A pregnancy is ____ than a hiccup.
A pregnancy is ____ than a lunch.
A pregnancy is ____ than a movie.
A pregnancy is ____ than a sneeze.
A pregnancy is ____ than a wedding.
A semester is ____ than a blink.
A semester is ____ than a breath.
A semester is ____ than a concert.
A semester is ____ than a hiccup.
A semester is ____ than a lunch.
A semester is ____ than a movie.
A semester is ____ than a sneeze.
A semester is ____ than a wedding.
A sneeze is ____ than a concert.
A sneeze is ____ than a lunch.
A sneeze is ____ than a movie.
A sneeze is ____ than a pregnancy.
A sneeze is ____ than a semester.
A sneeze is ____ than a summer.
A sneeze is ____ than a wedding.
A sneeze is ____ than a winter.
A summer is ____ than a blink.
A summer is ____ than a breath.
A summer is ____ than a concert.
A summer is ____ than a hiccup.
A summer is ____ than a lunch.
A summer is ____ than a movie.
A summer is ____ than a sneeze.
A summer is ____ than a wedding.

A wedding is ____ than a blink.
A wedding is ____ than a breath.
A wedding is ____ than a hiccup.
A wedding is ____ than a pregnancy.
A wedding is ____ than a semester.
A wedding is ____ than a sneeze.
A wedding is ____ than a summer.
A wedding is ____ than a winter.

A winter is ____ than a blink.
A winter is ____ than a breath.
A winter is ____ than a concert.
A winter is ____ than a hiccup.
A winter is ____ than a lunch.
A winter is ____ than a movie.
A winter is ____ than a sneeze.
A winter is ____ than a wedding.

<u>Spatial Sentences</u>

A bathtub is ____ than a keg.
A bathtub is ____ than a kettle.
A bathtub is ____ than a mug.
A bathtub is ____ than a pitcher.
A bathtub is ____ than a shot glass.
A bathtub is ____ than a teacup.
A bathtub is ____ than a teapot.
A bathtub is ____ than a thimble.
A keg is ____ than a bathtub.
A keg is ____ than a lake.
A keg is ____ than a mug.
A keg is ____ than a shot glass.
A keg is ____ than a swimming pool.
A keg is ____ than a teacup.
A keg is ____ than a thimble.
A keg is ____ than an ocean.
A kettle is ____ than a bathtub.
A kettle is ____ than a lake.
A kettle is ____ than a mug.
A kettle is ____ than a shot glass.
A kettle is ____ than a swimming pool.
A kettle is ____ than a teacup.
A kettle is ____ than a thimble.
A kettle is ____ than an ocean.
A lake is ____ than a keg.
A lake is ____ than a kettle.
A lake is ____ than a mug.
A lake is ____ than a pitcher.
A lake is ____ than a shot glass.
A lake is ____ than a teacup.
A lake is ____ than a teapot.
A lake is ____ than a thimble.
A mug is ____ than a bathtub.
A mug is ____ than a keg.
A mug is ____ than a kettle.
A mug is ____ than a lake.

A mug is ____ than a pitcher.
A mug is ____ than a swimming pool.
A mug is ____ than a teapot.
A mug is ____ than an ocean.
A pitcher is ____ than a bathtub.
A pitcher is ____ than a lake.
A pitcher is ____ than a mug.
A pitcher is ____ than a shot glass.
A pitcher is ____ than a swimming pool
A pitcher is ____ than a teacup.
A pitcher is ____ than a thimble.
A pitcher is ____ than an ocean.
A shot glass is ____ than a bathtub.
A shot glass is ____ than a keg.
A shot glass is ____ than a kettle.
A shot glass is ____ than a lake.
A shot glass is ____ than a pitcher.
A shot glass is ____ than a swimming pool.
A shot glass is ____ than a teapot.
A shot glass is ____ than an ocean.
A swimming pool is ____ than a keg.
A swimming pool is ____ than a kettle.
A swimming pool is ____ than a mug.
A swimming pool is ____ than a pitcher.
A swimming pool is ____ than a shot glass.
A swimming pool is ____ than a teacup.
A swimming pool is ____ than a teapot.
A swimming pool is ____ than a thimble.
A teacup is ____ than a bathtub.
A teacup is ____ than a keg.
A teacup is ____ than a kettle.
A teacup is ____ than a lake.
A teacup is ____ than a pitcher.
A teacup is ____ than a swimming pool.
A teacup is ____ than a teapot.
A teacup is ____ than an ocean.

A teapot is ____ than a bathtub.
A teapot is ____ than a lake.
A teapot is ____ than a mug.
A teapot is ____ than a shot glass.
A teapot is ____ than a swimming pool.
A teapot is ____ than a teacup.
A teapot is ____ than a thimble.
A teapot is ____ than an ocean.
A thimble is ____ than a bathtub.
A thimble is ____ than a keg.
A thimble is ____ than a kettle.
A thimble is ____ than a lake.
A thimble is ____ than a pitcher.
A thimble is ____ than a swimming pool.
A thimble is ____ than a teapot.
A thimble is ____ than an ocean.
An ocean is ____ than a keg.
An ocean is ____ than a kettle.
An ocean is ____ than a mug.
An ocean is ____ than a pitcher.
An ocean is ____ than a shot glass.
An ocean is ____ than a teacup.
An ocean is ____ than a teapot.
An ocean is ____ than a thimble.