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

Daniel Casasanto[†]   **Lera Boroditsky**   **Webb Phillips**   **Jesse Greene**

Shima Goswami   **Simon Bocanegra-Thiel**   **Ilia Santiago-Diaz**

MIT Department of Brain & Cognitive Sciences, 77 Massachusetts Avenue NE20-457
Cambridge, MA 02139 USA

Olga Fotokopoulu   **Ria Pita**
Aristotle University of Thessaloniki, Greece

David Gil
Max Planck Center for Evolutionary Anthropology
Jakarta Field Station, Indonesia

Abstract

Do the languages that we speak affect how we experience the world? This question was taken up in a linguistic survey and two non-linguistic psychophysical experiments conducted in native speakers of English, Indonesian, Greek, and Spanish. All four of these languages use spatial metaphors to talk about time, but the particular metaphoric mappings between time and space vary across languages. A linguistic corpus study revealed that English and Indonesian tend to map duration onto linear distance (e.g., a *long* time), whereas Greek and Spanish preferentially map duration onto quantity (e.g., *much* time). Two psychophysical time estimation experiments were conducted to determine whether this cross-linguistic difference has implications for speakers' temporal thinking. Performance on the psychophysical tasks reflected the relative frequencies of the 'time as distance' and 'time as quantity' metaphors in English, Indonesian, Greek, and Spanish. This was true despite the fact that the tasks used entirely non-linguistic stimuli and responses. Results suggest that: (1.) The spatial metaphors in our native language may profoundly influence the way we mentally represent time. (2.) Language can shape even primitive, low-level mental processes such as estimating brief durations – an ability we share with babies and non-human animals.

Introduction

"Are our own concepts of 'time,' 'space,' and 'matter' given in substantially the same form by experience to all men, or are they in part conditioned by the structure of particular languages?" (Whorf, 1939/2000, pg. 138.) This question, posed by Benjamin Whorf over half a century ago, is currently the subject of renewed interest and debate. Does language shape thought? The answer *yes* would call for a reexamination of some foundational theories that have guided Cognitive Science for decades, which assume both the universality and the primacy of non-linguistic concepts (Chomsky, 1975; Fodor, 1975). Yet despite unreserved belief among the general public that people who talk differently also think differently (ask anyone about the Eskimos' words for snow), it has remained widely agreed among linguists and psychologists that they do not.

Skepticism about some Whorfian claims has been well founded. Two crucial kinds of evidence have been missing from many previous inquiries into relations between language and thought: objectively evaluable linguistic data, and language-independent psychological data. A notorious fallacy, attributable in part to Whorf, illustrates the need for methodological rigor. Whorf (1939) argued that Eskimos must conceive of snow differently than English speakers because the Eskimo lexicon contains multiple words that distinguish different types of snow, whereas English has only one word to describe all types. The exact number of snow words the Eskimos were purported to have is not clear. (This number has now been inflated by the popular press to as many as two-hundred.) According to a Western Greenlandic Eskimo dictionary published in Whorf's time, however, Eskimos may have had as few as two distinct words for snow (Pullum, 1991).

Setting aside Whorf's imprecision and the media's exaggeration, there remain two problems with Whorf's argument, which are evident in much subsequent 'Language and Thought' research, as well. First, although Whorf asserted an objective difference between Eskimo and English snow vocabularies, his comparative linguistic data were subjective and unfalsifiable: it is a matter of opinion whether any cross-linguistic difference in the number of snow words existed. As Geoffrey Pullum (1991) points out, English could also be argued to have multiple terms for snow in its various manifestations: *slush, sleet, powder, granular, blizzard, drift, etc.* The problem of unfalsifiability would be addressed if cross-linguistic differences could be demonstrated empirically, and ideally, if the magnitude of the differences could be quantified.

A second problem with Whorf's argument (and others like it in the contemporary Cognitive Linguistics literature) is that it uses purely linguistic data to motivate inferences about non-linguistic thinking. Steven Pinker illustrates the resulting circularity of Whorf's claim in this parody of his reasoning: "[Eskimos] speak differently so they must think differently. How do we know that they think differently? Just listen to the way they speak!" (Pinker, 1994, pg. 61). This circularity would be escaped if non-linguistic evidence could be produced to show that two groups of speakers who talk differently also think differently in corresponding ways.

[†]Corresponding author: Daniel Casasanto (djc@mit.edu)

But what counts as 'non-linguistic' evidence? Recent studies have tested predictions derived from cross-linguistic differences using behavioral measures such as accuracy and reaction time. Oh (2003) investigated whether Korean and English speakers would remember motion events differently, consistent with the way motion is habitually encoded in their native languages (i.e., in terms of 'path' or 'manner' of motion). Participants described videos of motion events, and then took a surprise memory test probing small details of the videos. Oh found that, as expected, English speakers used more manner-of-motion verbs in their video descriptions than Korean speakers. English speakers also performed better than Koreans on the manner-relevant portion of the memory test. Oh refers to the memory task as 'non-linguistic,' yet the questions were posed using motion language, and participants may have recalled their own verbal descriptions of the videos while responding.

Boroditsky (2001) investigated whether speakers of English and Mandarin think differently about time. English typically describes time as horizontal, while Mandarin commonly uses vertical spatiotemporal metaphors. Boroditsky found that English speakers were faster to judge sentences about temporal succession (e.g., *March comes earlier than April*) when primed with a horizontal spatial event, but Mandarin speakers were faster to judge the same sentences when primed with a vertical spatial stimulus. This was true despite the fact that all of the sentences were presented in English.

These studies by Oh and Boroditsky support a version of the Whorfian hypothesis which Slobin (1986) has termed *thinking for speaking*: language can affect thought when we are thinking with the intent to use language, plausibly by directing attention to elements of our experience that are ordinarily encoded in the language we speak. For instance, because English tends to encode information about manner of motion more often than Korean does, Oh's English subjects may have automatically attended to the manner information in the videos more than her Korean subjects did. Some researchers have characterized the effects of thinking for speaking as uninterestingly weak (Pinker, 1994; Papafragou, Massey, & Gleitman, 2002). Results such as Oh's and Boroditsky's suggest otherwise: at minimum, thinking for speaking appears to influence ubiquitous cognitive processes such as attention and memory, and is capable of changing the nature of our abstract mental representations. Furthermore, habits formed while thinking for speaking are likely to be practiced even when people are not explicitly encoding information for language. We never know when we might want to talk about an event at some later point, so it is in our best interest to encode language-relevant details as a matter of policy.

Can the influence of language on thought go beyond thinking for speaking? Much of our mental life is unutterable: what words can capture the sound of a clarinet, or explain the color *red* to a blind person (Locke, 1689/1995; Wittgenstein, 1953)? Can peculiarities of our native language shape even the deep, primitive kinds of representations that we share with pre-linguistic infants and non-human animals? Previous research suggests that language can affect our high-level linguistic and symbolic representations in the abstract domain of time (Boroditsky, 2001). The goal of the present study was to find out whether language can also shape our low-level, non-linguistic, non-symbolic temporal representations. A linguistic study was conducted to investigate a previously unexplored pattern in spatiotemporal metaphors, and to quantify cross-linguistic differences in the way these metaphors are used by speakers of English, Indonesian, Greek, and Spanish (Experiment 1). To determine whether these cross-linguistic differences have consequences for speakers' non-linguistic time representations, the results of the linguistic study were used to predict performance on a pair of psychophysical time estimation tasks, with entirely non-linguistic stimuli and responses (Experiments 2 and 3).

Experiment 1:
Time in a Bottle or Time on the Line?

Linguists have noted that spatial metaphors are often used to talk about non-spatial phenomena -- in particular abstract phenomena such as social rank (e.g., a *high* position), mathematics (e.g., a *low* number), and time (e.g., a *long* vacation) (Clark, 1973; Gibbs, 1994; Jackendoff, 1983; Lakoff & Johnson, 1980). Recently, psychologists have begun to explore the proposal that these metaphors in language provide a window on our underlying mental representations in abstract domains, using the domain of time as a testbed (Boroditsky, 2000, 2001; Boroditsky & Ramscar, 2002; Casasanto & Boroditsky, 2003; Gentner, 2001). In general, this work has focused on how time can be expressed (and by hypothesis conceptualized) in terms of linear space. Linear spatiotemporal metaphors are pervasive in English, and are used to talk about various aspects of time, including succession (e.g., Monday comes *before* Tuesday), motion through time (e.g., Let's move the meeting *forward*), and duration (e.g., a *short* intermission). But is time necessarily conceptualized in terms of uni-dimensional space? English speakers also talk about *oceans of time, saving time in a bottle*, and compare epochs to *sand through the hourglass*, apparently mapping time onto volume.

Experiment 1 compared the use of 'time as distance' and 'time as quantity' metaphors across languages. Every language examined so far uses both distance and quantity metaphors, but their relative prevalence and productivity appear to vary markedly. In English, it is natural to talk about *a long time*, borrowing the structure and vocabulary of a spatial expression like *a long rope*. Yet in Spanish, the direct translation of 'long time', *largo tiempo*, sounds awkward to speakers of most dialects. *Mucho tiempo*, which means 'much time', is preferred.

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' in physical quantity (e.g., much water). Compare how English and Greek typically modify the duration of the following the events (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 *big night* is likely to mean 'an exciting night', and *big relationship* 'an important relationship'. For Greek speakers, however, these phrases communicate duration, expressing time not in terms of uni-dimensional space, but rather in terms of physical quantity (i.e., three-dimensional space).

For Experiment 1, the most natural phrases expressing the ideas 'a long time' and 'much time' were elicited from native speakers of English, Indonesian, Greek, and Spanish (see table 1). The frequencies of these expressions were compared in a very large multilingual text corpus: www.google.com. Each expression in table 1 was entered as a search term. 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.

Table 1: Distance and quantity metaphors for duration.

	Distance	Quantity
English	long time	much time
Indonesian	waktu panjang	waktu banyak
Greek	makry kroniko diatstima	poli ora
Spanish	largo tiempo	mucho tiempo

Results The number of google 'hits' for each expression was tabulated, and the percentage of distance hits and quantity hits was calculated for each pair of expressions, as a measure of their relative frequency (see figure 1). Results showed that in English and Indonesian, distance metaphors were more frequent than quantity metaphors. The opposite pattern was found in Greek and Spanish. A Chi-Square test showed that the relationship between distance and quantity metaphors varied significantly across languages ($X^2=8.5 \times 10^5$, df=3, p<0.0001). These findings corroborate native speakers' intuitions for each language. Additional quantitative studies are in progress to validate these results.
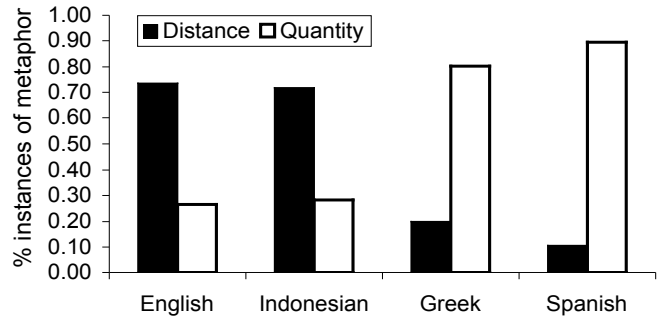

Figure 1: Corpus search results. Black bars indicate the percentage of distance metaphors and white bars the percentage of quantity metaphors found for each language.

Experiments 2 & 3:

Do People Who Talk Differently Think Differently?

How might this difference in the way English, Indonesian, Greek, and Spanish speakers talk about time affect the way they think about it? Linguists and psychologists have argued that our conception of time is intimately dependent on our knowledge of space, noting that in many languages, people can hardly avoid using spatial words when they talk about time (Clark, 1973; Gibbs, 1994; Jackendoff, 1983; Lakoff & Johnson, 1980). Behavioral studies show that changing someone's immediate spatial environment or frame of reference can dramatically change the way they interpret temporal language (Boroditsky, 2000; Boroditsky & Ramscar, 2002). But does space influence our temporal thinking even when we are not *thinking for speaking*?

A recent study by Casasanto & Boroditsky (2003) shows that space influences even our low-level, non-linguistic, non-symbolic representations of time. English speakers watched lines 'growing' across a computer screen, one pixel at a time, and estimated either how far they grew or how much time they remained on the screen. Estimates were made by clicking the mouse to indicate the beginning and end of each spatial or temporal interval. Line distances and durations were varied orthogonally, so there was no correlation between the spatial and temporal components of the stimuli. As such, one stimulus dimension served as a distr actor for the other: an irrelevant piece of information that could potentially interfere with task performance. Patterns of cross-dimensional interference were analyzed to reveal relationships between subjects' representations of space and time. Results showed that subjects were unable to ignore irrelevant spatial information when estimating time (even when they were encouraged to do so). Line stimuli of the same average duration were judged to take a longer time when they grew a longer distance, and a shorter time when they grew a shorter distance. In contrast, line duration did not affect subjects' distance estimates. This asymmetric relation between space and time was predicted based on patterns in language: we talk about time in terms of space more than we talk about space in terms of time (Lakoff & Johnson, 1980).

These findings suggest that the metaphoric relationship between time and space is not just linguistic, it is also conceptual. Not only do people talk about time in terms of space, they also think about time using spatial representations. However, the experiments reported in Casasanto & Boroditsky (2003) leave the Whorfian question unanswered: do metaphors in language merely reflect underlying conceptual structures, or might the metaphors we use also play some role in constructing concepts, or establishing their interrelations?

In the present study speakers of four different languages performed the same pair of non-linguistic psychophysical tasks, which required them to estimate time while overcoming spatial interference. It was reasoned that if people's concepts of time and space are substantially the same universally, irrespective of the languages they speak, then performance on these tasks should not differ between language groups. If, on the other hand, the spatiotemporal metaphors people use affect how they represent time and space non-linguistically, then performance should vary in ways predicted by participants' language-particular metaphors.

For Experiment 2, subjects performed a 'growing line' task similar to the task described above. It was reasoned that the English participants in our previous study may have suffered interference of distance on duration estimation, in part, because these notions are conflated in the English language. It is hard to imagine expressing the idea 'a long time' in English without using an adjective that can also indicate spatial extent. Piaget (1927) made a similar suggestion when he observed that young French speaking children often mistook distance for duration, noting that both of these concepts are commonly described in French using the adjective *longue*. We predicted that speakers of 'Distance Languages' (i.e., English and Indonesian) would show a considerable effect of distance on time estimation when performing the growing line task.

If conflations in language contribute to confusions in thought, can distinctions in language help speakers distinguish closely related concepts? We reasoned that speakers of languages that do not ordinarily express duration in terms of distance might have an easier time distinguishing the spatial and temporal information conveyed in our growing line stimuli. We predicted that speakers of 'Quantity Languages' (i.e., Greek and Spanish) would show only a mild effect of distance on time estimation when performing the growing line task.

For Experiment 3, a task complementary to the growing line task was developed. Subjects watched a schematically drawn container of water filling up, one row of pixels at a time, and estimated either how full it became or how much time it remained on the computer screen, using mouse clicks. We predicted the converse pattern of behavioral results for the filling container task as for the growing line task: speakers of Quantity Languages would show a considerable influence of 'fullness' on time estimation,

whereas speakers of Distance Languages would show a milder effect.

Methods for Experiment 2: Growing Lines

Subjects A total of 65 subjects participated in exchange for payment. Native English and Spanish speaking participants were recruited from the Greater Boston community, and were tested on MIT campus. Native Indonesian speakers were recruited from the Jakarta community, and were tested at the Cognation Outpost in the Jakarta Field Station of the Max Planck Center for Evolutionary Anthropology. Native Greek speakers were recruited from the Thessaloniki community, and tested at the Aristotle University of Thessaloniki.

Materials Lines of varying lengths were presented on a computer monitor (resolution=1024x768 pixels, dpi=72), for varying durations. Durations ranged from 1000 milliseconds to 5000 milliseconds in 500 millisecond increments. Displacements ranged from 100 to 500 pixels in 50 pixel increments. Nine durations were fully crossed with nine displacements to produce 81 distinct line types. Lines started as a single point and 'grew' horizontally across the screen one pixel at a time, from left to right along the vertical midline. Each line remained on the screen until its maximum displacement was reached.

Written instructions were given prior to the start of the task, in the native language of the participant. Care was taken to avoid using distance metaphors for time. The task itself was entirely non-linguistic, consisting of lines (stimuli) and mouse clicks (responses).

Procedure Participants viewed 162 growing lines, one line at a time. Immediately before each trial, a prompt appeared indicating that the subject should attend either to the line's duration or to its spatial displacement. Space trials and time trials were randomly intermixed.

To estimate displacement, subjects clicked the mouse once on the center of an 'X' icon, moved the mouse to the right in a straight line, and clicked the mouse a second time to indicate they had moved a distance equal to the maximum displacement of the stimulus. To estimate duration, subjects clicked the mouse once on the center of an 'hourglass' icon, waited the appropriate amount of time, and clicked again in the same spot, to indicate the time it took for the stimulus to reach its maximum displacement.

All responses were self-paced. Importantly, for a given trial, subjects reproduced either the displacement or the duration of the stimulus, never both. Response data were collected for both the trial-relevant and the trial-irrelevant stimulus dimension, to ensure that subjects were following instructions.

Methods for Experiment 3: Filling Containers

Subjects A total of 74 subjects participated in exchange for payment. Subjects were recruited at the same time as those who participated in Experiment 2, from the same populations.

Materials and procedure The filling container task was closely analogous to the growing line task (Experiment 2). Participants viewed 162 containers, and were asked to imagine that each was a tank filling with water. Containers were simple line drawings, 600 pixels high and 500 pixels wide. Empty containers filled gradually, one row of pixels at a time, for varying durations and 'volumes,' and they disappeared when they reached their maximum fullness. Nine durations were fully crossed with nine volumes to produce 81 distinct trial types. For each trial, participants estimated either the amount of water in the container (by clicking the mouse once at the bottom of the container and a second time at the appropriate 'water level'), or they estimated the amount of time that the container took to fill (by clicking the hourglass icon, waiting the appropriate time, and clicking it again, as in Experiment 2). Durations ranged from 1000 milliseconds to 5000 milliseconds in 500 millisecond increments. Water levels ranged from 100 to 500 pixels, in 50 pixel increments.

As before, written instructions were given prior to the start of the task, in the native language of the participant. Care was taken to avoid using quantity metaphors for time. The task itself was entirely non-linguistic, consisting of containers (stimuli) and mouse clicks (responses).

Results for Experiments 2 and 3

Both time estimates and space estimates were collected for each subject, but since our present hypothesis concerns effects of language and space on time estimation, only data for the time estimation trials are reported here. (See Casasanto & Boroditsky, 2003 for a discussion of subjects' space estimation in a related task.)

Time estimation, within-domain effects Overall, for both Experiments 2 and 3 subjects' time estimates were highly accurate across all language groups, as indicated by the strong correlations between the actual stimulus duration and subjects' estimated stimulus duration. These correlations did not differ significantly between groups or between tasks (see table 2).

Table 2: Time estimation results for Experiments 2 and 3. Slope and r-square values for the correlation between actual stimulus duration and subjects' grand averaged estimates of stimulus duration. (Perfect performance would be indicated by a slope of 1.00 and r^2 of 1.00.)

	Growing Lines		Filling Containers	
	Slope	r2	Slope	r2
English	0.75	0.999	0.76	0.999
Indonesian	0.71	0.996	0.68	0.997
Greek	0.74	0.995	0.77	0.996
Spanish	0.72	0.996	0.71	0.995

Time estimation, cross-domain effects The within-domain results reported in table 2 are important to establish that subjects were able to estimate time well, and importantly, that subjects estimated time about equally well in all groups, and on both tasks. Of principal interest, however, are the cross-domain effects (i.e., effects of actual distance and actual quantity on estimated time), which are summarized in figure 2.

To investigate the effect of spatial interference on time estimation, grand averaged time estimates in milliseconds were plotted as a function of actual stimulus displacement in pixels (i.e., line length or water level). A line of best fit was computed, and the slope was used as an index of effect strength. In our previous experience with similar tasks, we found the strongest linear effects on the dependent variable (i.e., estimated time) near the middle of the range of the independent variable (i.e., actual stimulus displacement), possibly due to 'endpoint effects' commonly observed in magnitude estimation tasks. For the analyses reported here, the outer points were removed, and the middle five points of the correlations were analyzed.

Cross-domain effects varied markedly across language groups. For the growing line task, English and Indonesian speakers showed a strong effect of distance on time estimation (English: Slope=1.49, r^2=0.98; t=8.5; df=3; p<0.001; Indonesian: Slope=1.40, r^2=0.80; t=3.4; df=3; p<0.01). By contrast, Greek and Spanish speakers showed weak, non-significant effects of distance on time estimation (Greek: Slope=0.47, r^2=0.33; t=1.2; df=3; *ns*; Spanish: Slope=0.20, r^2=0.13; t=0.7; df=3; *ns*).

For the filling container task, the opposite pattern of results was found. English and Indonesian speakers showed a weak, non-significant effect of volume on time estimation (English: Slope=0.18, r^2=0.12; t=0.6; df=3; *ns*; Indonesian: Slope=0.13, r^2=0.51; t=1.7; df=3; *ns*), whereas Greek and Spanish speakers showed strong effects of volume on time estimation (Greek: Slope=1.24, r^2=0.95; t=6.9; df=3; p<0.001; Spanish: Slope=1.16, r^2=0.97; t=8.5; df=3; p<0.001).
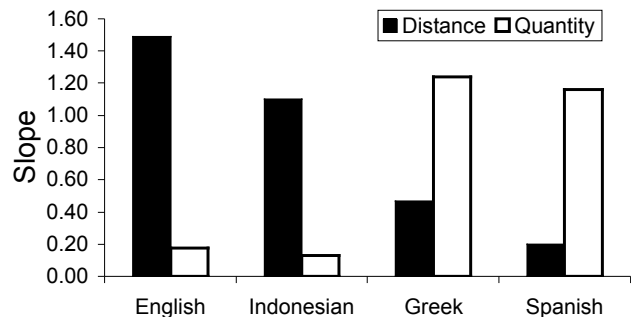

Figure 2: Effects of distance and quantity interference on time estimation.

A 4 x 2 factorial ANOVA with Language and Task as between-subject factors revealed a highly significant Language by Task interaction (F (3,139) = 5.25, p<0.002), with no main effects, signaling a true crossover interaction.

Linear regression analysis revealed a highly significant positive relation between the frequency of Distance and Quantity metaphors in each language (as measured in Experiment 1) and the amount of Distance and Quantity interference on time estimation (as measured in Experiments 2 and 3) (Slope=1.62, r^2=0.84; t=5.6; df=6; p<0.001).

General Discussion and Conclusions

Do people who talk differently also think differently? Performance on a pair of psychophysical time estimation tasks differed dramatically for speakers of different languages, in ways predicted by their language-particular spatiotemporal metaphors. The effects of distance interference and quantity interference on time estimation in speakers of English, Indonesian, Greek, and Spanish corresponded strikingly to the relative prevalence of distance metaphors and quantity metaphors found in these languages (compare figures 1 and 2). This was true despite the fact that the behavioral tasks comprised entirely non-linguistic stimuli and responses.

Returning to the question of Whorf's posed in the introduction, it is possible that our concepts of time and space are "given in substantially the same form by experience" to all of us, *and also* that they are "in part conditioned by the structure of particular languages." Perhaps people learn associations between time and space via physical experience (e.g., by observing moving objects and changing quantities). Since presumably the laws of physics are the same in all language communities, pre-linguistic children's conceptual mappings between time, distance, and quantity could be the same universally. When children acquire language, these mappings could be adjusted, plausibly by a process analogous to Hebbian learning: each time we use a linguistic metaphor, we may invoke the corresponding conceptual mapping. Speakers of Distance Languages then would invoke the time-distance mapping frequently, eventually strengthening it at the expense of the time-quantity mapping (and vice-versa for speakers of Quantity Languages). Alternatively, experience may not teach us to map time onto space. It could be language that causes us to notice structural parallels between these domains, in the first place. On this possibility, language would be responsible for establishing the time-distance and time-quantity conceptual mappings evident in our adult subjects, not just for modifying these mappings. Studies are in progress on young learners of Distance and Quantity languages to explore these possibilities.

The findings we present here are difficult to reconcile with a 'universalist' view of language-thought relations according to which language calls upon pre-formed, antecedently available non-linguistic concepts, which are presumed to be "universal" (Pinker, 1994, pg. 82) and "immutable" (Papafragou, Massey, & Gleitman, 2002, pg. 216). Rather, these results support what we might call a *deep* version of the linguistic relativity hypothesis (to distinguish it from the so-called *weak* version which posits that language affects 'thinking for speaking,' and from *strong* linguistic determinism). The particular languages that we speak can influence not only the representations we build for the purpose of speaking, but also the non-linguistic representations we build for remembering, acting on, and perhaps even perceiving the world around us.

References

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. *Cognit Psychol, 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., & Boroditsky, L. (2003). *Do we think about time in terms of space?* Paper presented at the 25th Annual conference of the Cognitive Science Society, Boston, MA.

Chomsky, N. (1975). *Reflections on Language*. New York: Pantheon.

Clark, H. H. (1973). Space, Time, Semantics and the Child. In T. E. Moore (Ed.), *Cognitive Development and the Acquisition of Language*. New York: Academic Press.

Fodor, J. (1975). *The Language of Thought*. Cambridge: Harvard University Press.

Gentner, D. (2001). Spatial Metaphors in Temporal Reasoning. In M. Gattis (Ed.), *Spatial Schemas and Abstract Thought*. Cambridge: MIT Press.

Gibbs, R. W., jr. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge: Cambridge University Press.

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

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Locke, J. (1689/1995). *An Essay Concerning Human Understanding*. Amherst: Promethius Books.

Oh, K. (2003). *Habitual patterns of language use and thinking for speaking: a Whorfian effect on motion events*. University of California, Berkeley, Berkeley.

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

Piaget, J. (1927/1969). *The Child's Conception of Time*. New York: Ballantine Books.

Pinker, S. (1994). *The Language Instinct*. New York, US: William Morrow and Company.

Pullum, G. K. (1991). *The Great Eskmo Vocabulary Hoax and other irreverent essays on the study of language*. Chicago: University of Chicago Press.

Slobin, D. (1986). From "THOUGHT AND LANGUAGE" to "THINKING FOR SPEAKING" in *Rethinking Linguistic Relativity*, Grumperz & Levinson, eds.

Whorf, B. (2000). *Language, Thought and Reality*. Cambridge, MA: MIT Press.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell