Title

New space-time metaphors foster new mental representations of time

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors

Hendricks, Rose K
Boroditsky, Lera

Publication Date

2015

Peer reviewed

New space-time metaphors foster new mental representations of time

Rose K. Hendricks (rhendricks@ucsd.edu)
Lera Boroditsky (lera@ucsd.edu)
Department of Cognitive Science, 9500 Gilman Drive,
La Jolla, CA 92093 USA

Abstract

Can learning new linguistic metaphors foster new non-linguistic representations? We describe a set of studies in which we trained English-speaking participants to talk about time using vertical spatial metaphors that are novel to English. One group learned a mapping that placed earlier events above and the other a mapping that placed earlier events below. After mastering the new metaphors, participants were tested in a non-linguistic space-time implicit association task – the Orly task. This task has been used previously to document cross-linguistic differences in representations of time (Boroditsky et. al 2010; Fuhrman et al 2011). Some participants completed temporal judgments in the Orly task without any other secondary task, while others did so under either verbal or visual interference. Finally, we report data from a serendipitous sample of Chinese-English bilinguals on the same task.

Keywords: metaphor; space; time; learning; language

Introduction

Across many languages people use spatial terms and constructions to talk about time. For example in English, one might say "the best is ahead of us" or "let's move the meeting forward," using the horizontal spatial relation terms "ahead" and "forward" to talk about temporal relations. Across languages, which spatial terms or expressions are borrowed to talk about which temporal relations differs. Some languages rely more on vertical terms than English does; others put the past in front and the future behind (reversing the mapping from the English system) (see Boroditsky 2011 for a review). Prior work has indeed found that cross-linguistic differences in spatio-temporal metaphors predict cross-linguistic differences in people's spatial representations of time (e.g., Boroditsky 2001; Boroditsky et. al 2010; Fuhrman et al 2011; Miles et al, 2011; Núñez & Sweetser, 2006).

For example, while both English and Mandarin use horizontal (front/back) terms to talk about relationships between events in time, Mandarin also commonly employs vertical spatial terms (up/down) to talk about time. Earlier or past events are said to be "up," and later or future events are said to be "down." The critical difference is that Mandarin uses vertical metaphors more frequently than English does. While some vertical metaphors for time do occur in English (e.g., "passing knowledge *down* to the next generation") these metaphors are far less common, productive, and systematic than they are in Mandarin.

The differences observed in language also bear out in how people reason about time in non-linguistic tasks. A number of studies have found differences on an implicit space-time association task done with picture sequences (Boroditsky et. al 2010; Fuhrman et al 2011). We will refer to this task as the Orly task, named after its inventor. On each trial, participants see two pictures presented one at a time in the center of the screen. For example, participants might see a photograph of Julia Roberts in her 20s followed by a photo of Roberts as a young girl (or of her in her 40s). Their job is to indicate whether the second picture shows a conceptually earlier or later point in time than the first picture. To respond participants must press different colored buttons (for example, press blue for "earlier" and red for "later"). For some trials the buttons are arranged so that the "earlier" button is above the "later" button. On other trials the mapping is reversed.

People's patterns of responses on this non-linguistic picture sequence task correspond to patterns found in their linguistic and cultural backgrounds. For example, Mandarin makes common use of vertical spatial metaphors that place earlier events above and later events below. Correspondingly, Mandarin speakers are faster to respond when the "earlier" button is above than when it is below. English on the other hand does not have a strong pattern of vertical metaphors, and correspondingly English speakers show a weaker bias on the vertical axis than Mandarin speakers in this same task.

However, in cross-cultural comparisons like this, it is impossible to cleanly attribute differences in non-linguistic behavior to patterns in language. In addition to differences in metaphor use between English and Mandarin, there are many other potential differences in extra-linguistic aspects of cultural experience. Can learning different spatial metaphors for time (in the absence of other extra-linguistic differences) indeed foster new representations of time? And if so, what is the locus of such effects? Do the newly learned mappings reside strictly in the linguistic sphere? Or can learning new ways of talking about time encourage people to create new non-linguistic representations for time (ones that are not disrupted by verbal interference)?

Here we describe a set of studies in which we train English-speaking participants to talk about time using vertical spatial metaphors that are novel to English. One group learned a mapping that placed earlier events above and the other learned a mapping that placed earlier events below. After mastering the new metaphors, participants were tested in the Orly task described above. Some participants completed the task without any other secondary task, while others did so under conditions of either verbal or visual interference.

In a final section we analyze data on the same tasks from a serendipitous sample of Chinese-English bilinguals. These participants self-identified as native English speakers prior to participating, but in a post-experimental questionnaire it was revealed they also had experience speaking Mandarin and/or Cantonese. How does learning a new metaphor in the lab affect participants who have been exposed to the opposite pattern in their natural language experience?

Participants & Analyses

Participant Inclusion

Forty-one undergraduates at UC San Diego participated in Experiment 1 and 80 participated in Experiment 2. All participants received either payment or course credit for their participation.

Although our experiments were advertised as recruiting only native English speakers, the demographics of the UCSD testing population are such that 3 participants in Experiment 1 and 28 participants in Experiment 2 were bilingual in Mandarin or Cantonese, as revealed in a post-experimental language background questionnaire. These participants were excluded from the main analyses and replaced by participants who did not have exposure to Chinese. This was necessary because Chinese speakers would have already had substantial exposure to vertical metaphors for time as part of their natural language experience. For ease of reference we will refer to participants who reported speaking Mandarin or Cantonese as Chinese-English bilinguals, and participants who did not as English speakers. We analyze the data from the sizeable sample of Chinese-English bilinguals in Experiment 2 separately.

In addition to excluding Chinese-English bilinguals, participants were excluded if they failed to follow the instructions, if their performance on the training, the Orly task, or the interference tasks (in Exp 2) was worse than 3 SDs away from the group mean, or if a computer error prevented data from being saved.

All excluded participants were replaced to arrive at a complete counterbalance of 32 participants in Experiment 1 and 48 participants in Experiment 2.

Inclusion of data points

For analyses of response times, we excluded trials for which the participant provided an incorrect picture sequence judgment, trials for which participants made errors on the interference task (In Experiment 2), and trials for which the reaction time was greater than 3 SDs from the participant's cell mean. This yielded 90.4% of original trials for analysis in Experiment 1 and 75.4% in Experiment 2.

Analyses

We fit the data with mixed effects models using Laplace Approximation using the lmer() function within the lme4 analysis package in R (Bates, Maechler, & Dai, 2008; R Development Core Team, 2008). We modeled both participants and items as random effects, training condition (whether participants learned the "earlier is up" or "earlier is down" metaphors) and task condition (whether the 'earlier' key was above or below) as fully crossed fixed effects, and block order as a main effect. In Experiment 2, the model also included interference condition (none, verbal or visual) as a fully crossed fixed effect. Results are summarized in Figure 1.

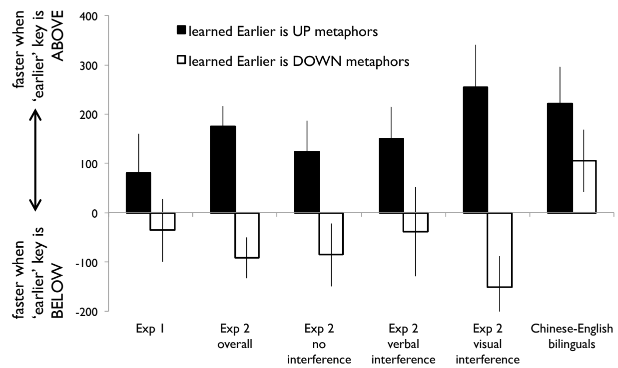

Figure 1. By-subject mean RT difference (in msecs) for picture sequence judgments (RT for 'earlier' key is below minus 'earlier' is above). Error bars show standard error.

Experiment 1

Methods

Procedure

The study proceeded in 4 phases: 1. metaphor training block, 2. Orly task block, 3. second metaphor training block (identical to the first), and 4. a final Orly task block. For a given participant, the two metaphor training blocks were identical, but the Orly task blocks were different. Each participant completed one Orly task block for which the response mapping was congruent with their metaphor training and one for which the response mapping was incongruent. Whether the congruent or the incongruent block occurred first was counterbalanced across participants.

Metaphor training At the start of the experiment, participants were told that they would learn a new way of talking about time. First they read five example sentences showcasing this new system of talking. For half of the participants, earlier events were said to be *above* or *higher* and later events were said to be *below* or *lower* (e.g., Thursday is higher than Friday; When you eat breakfast, dinner is below you). This system was reversed for the other half of participants so that instead earlier events were *below/lower* and later events were *above/higher* (Thursday is lower than Friday; When you eat breakfast, dinner is above you).

After participants said they understood the example sentences, they completed 90 test trials to practice this new system for talking about time. On each trial, they saw a new sentence describing a temporal relation (e.g, Lincoln was president _________ than Carter) and were asked to fill in the correct spatial term (selecting from the two options provided). Participants responded by typing the missing word into the blank. If participants gave the correct answer, the program continued to the next trial. If the answer was incorrect, the computer provided feedback and required the participant to correct their answer before they could continue to the next trial. The order of the testing sentences was generated randomly for each participant. The keyboard for this portion of the experiment was flat on the surface of the table so that participants could type normally.

Orly picture sequence task After they successfully completed the training, participants performed the Orly task: a non-linguistic temporal judgment task. With minor modifications, the task was the same as had been used previously to measure natural implicit space-time associations in English and Mandarin speakers, using the same materials and trial design (Boroditsky et al., 2010; Fuhrman et al., 2011).

In this portion of the experiment participants made their responses on a keyboard mounted perpendicular to the table surface. The keyboard contained three colored keys arranged vertically. The top key was blue, the center was green, and the bottom was red.

To begin each trial, participants were instructed to press and hold the center green key. When they pressed the green key, the first of two images appeared on the screen (e.g., Julia Roberts in her 20s). After 2 seconds, this image was replaced by the second image (e.g., either a younger or older Julia). Participants were instructed to continue holding the green key until they were ready to make a response. For one temporal judgment block participants were instructed to press the blue key if the second image was conceptually earlier, and the red key if it was conceptually later. For the other block, this instruction was reversed, so that participants were asked to press the blue key to indicate "later" and the red key to indicate "earlier." This meant that in one block the key mapping of responses was spatially congruent with the system of metaphors participants had learned, and in one block the key mapping was incongruent.

In the Orly task blocks participants first completed 10 practice trials with feedback, followed by 56 main experimental trials without feedback. Sequences were selected in random order for each participant.

After completing all 56 trials of the Orly task, participants returned to another block of training on the same metaphor, and then did another Orly task block (this time with the key-mapping reversed from their first run). At the end of the experiment, we collected information about participants' language backgrounds.

Results

Overall, participants were faster to respond when the key mapping was congruent with the metaphor they learned (M=1275 ms) than when it was incongruent (M=1333 ms), as confirmed by a significant interaction between task and training conditions ($t(3155.25) = 2.96$, $p = .003$). Participants who had learned metaphors that placed earlier events above responded more quickly in the Orly task when the 'earlier' key was above (M=1312 ms) than when it was below (M= 1393 ms). Participants who had learned metaphors that placed earlier events below, on the other hand, responded more quickly when the 'earlier' key was below (M= 1237 ms) than when it was above (M= 1273 ms).

There was no main effect of training condition ($t(33.78) = 1.51$, $p = .13$) or task condition ($t(3156.23) = 1.15$, $p = .25$). There were no speed-accuracy trade-offs. In addition to being faster on congruent trials, participants were also numerically slightly more accurate on congruent trials (92.5%) than on incongruent trials (91.8%). There were no significant differences in metaphor training or Orly task accuracies by condition (all $p > .28$).

Discussion

In Experiment 1 we saw that learning a new set of metaphors for talking about time affected people's performance in a non-linguistic temporal reasoning task. These results confirm that experience with metaphorical language alone (in the absence of other extra-linguistic cultural differences) can indeed produce the same kinds of differences in thinking as seen between speakers of different languages who have been naturally exposed to different metaphors in their linguistic environment.

These results raise a further question. What is the locus of this language-induced effect? One possibility is that understanding and learning to use new metaphors leads people to create new non-linguistic representations of time. That is, while the creation of these representations is encouraged by experience with language, the new representations themselves are not linguistic in nature. A different possibility is that learning these new metaphors creates new internal *linguistic* routines that people engage when reasoning about time. Experiment 2 was designed to test between these two possibilities.

In Experiment 2 we again taught participants new vertical metaphors for time. Then we tested them on the Orly task with and without verbal interference. We chose a verbal interference procedure known to interfere with internal linguistic routines (Frank et al., 2012). In prior work, Frank et al (2012) demonstrated that applying this verbal interference procedure severely disrupts other internal linguistic routines. For example, otherwise numerate college students were not able to keep a simple exact count of dots while performing this verbal interference task. We replicated the interference procedures from this prior work precisely, using the same code and materials.

If the results of Experiment 1 are due to participants adopting a new internal linguistic routine when reasoning about time, then applying this same verbal interference should wipe out the congruency effects we observed. On the other hand, if learning new metaphors fosters new non-linguistic representations, then verbal interference should not dampen the congruency effects.

Because any secondary task imposes extra general processing costs, we also included a visual interference condition, calibrated for difficulty, as an extra control. Any differences seen under verbal interference that are due to increased general processing demands should also be seen with the visual interference task.

Experiment 2

Methods

Procedure

Participants were randomly assigned to one of three groups: no interference (a replication of Experiment 1), verbal interference, and visual interference. The study proceeded in the following phases: 1. interference calibration block, 2. metaphor training block, 3. Orly task block, 4. second metaphor training block (identical to the first), and 5. a final Orly task block. Only participants in the verbal and visual interference groups completed the interference calibration block. The metaphor training blocks were the same as in Experiment 1. The Orly task blocks were modified to accommodate the three interference conditions. As in Experiment 1, each participant completed one Orly task block for which the response mapping was congruent with their metaphor training and one for which the response mapping was incongruent. Whether the congruent or the incongruent block occurred first was counterbalanced across participants.

Interference task calibration Participants in the interference groups first completed a calibration task to ensure that the interference tasks were properly tuned for individual ability. The materials and calibration procedure were identical to those used in Frank et al (2012). The MATLAB code to run the interference tasks was downloaded from https://github.com/langcog/numint. A full description of the procedure is available in Frank et al (2012). We kept the procedures identical to this prior work because it provides a clear precedent, showing that this same verbal interference task, calibrated in this same way, does strongly interfere with a linguistic cognitive routine like counting. If our congruency effects are also due to participants employing an online linguistic routine, then we should find the same deleterious effects of verbal interference.

Orly picture sequence task For the no interference group, the task was the same as in Experiment 1. For the two interference groups, participants were asked to remember either a string of consonants or a visual pattern while performing temporal judgments in the Orly task. For the verbal interference group, on each trial participants first saw a string of consonants appear sequentially in the same location. They then pressed and held the green button to begin an Orly task trial (meanwhile maintaining the letter sequence they had seen in working memory). After completing the temporal judgment, they were prompted to type in the letter sequence they had been rehearsing. For the visual interference group, on each trial participants first saw a set of blue squares appear sequentially in different locations on a 4x4 grid of white blocks. They then pressed and held the green button to begin the Orly task trial (meanwhile maintaining the visual pattern they had seen in memory). After completing the temporal judgment, they were prompted to click the locations of the blue squares on the grid they had seen at the start of the trial.

Results

Results of Experiment 2 replicated those of Experiment 1. Participants were faster to respond when the spatial arrangement of keys was congruent with the linguistic metaphors they had learned during metaphor training (M=1410 ms) than when it was incongruent (M=1544 ms), as reflected in a significant interaction between task and training conditions ($t(3944.5)= 4.83$, $p < .001$). Participants who had learned metaphors that placed earlier events above were faster to respond in the Orly task when the 'earlier' key was above (M=1386 ms) than when it was below (M= 1562 ms). Participants who had learned metaphors that placed earlier events below, on the other hand, were faster to respond when the 'earlier' key was below (M= 1435 ms) than when it was above (M= 1526 ms).

Importantly, participants responded faster on congruent than on incongruent trials across all three interference conditions. The size of the congruency effect was 105 ms without interference, 94 ms under verbal interference, and 202 ms under visual interference. The congruency effect observed in the verbal interference condition did not differ from that observed without interference (there was no 3-way interaction of training condition by task condition by verbal interference ($t(3948.5) = .91$, $p = .36$)). Analyses of each interference condition separately confirmed that in all three cases, there was the predicted interaction between training and task conditions, such that people were faster when the task and metaphor were congruent than when they were incongruent (all $p < .001$).

In addition to being faster on congruent trials, participants were also numerically slightly more accurate on congruent trials (94.9%) than on incongruent trials (94.7%). There were no significant differences in accuracy across interference type, task condition, congruency, or training condition (all p > .13). In the no interference and verbal interference conditions, accuracy on congruent trials was slightly higher than accuracy on incongruent trials. In the visual interference condition, people were 0.7% more accurate on incongruent than congruent trials, but this difference was not significant (p = .50). We ran an additional analysis including an accuracy covariate in the model described above. The accuracy covariate coded each participant's difference in accuracy on congruent vs incongruent trials. Including this covariate did not change the pattern of results (all results reported as significant remained so, and no newly significant results were revealed).

The analyses revealed main effects of both visual and verbal interference on response times for the Orly task (t(42.8) = 3.12, p = .003 and t(42.3) = 3.50, p = .001). As would be expected with any secondary task, participants were overall slower to make temporal judgments when undergoing interference. There was also a main effect of task condition (t(3944.5) = 2.84, p = .004); people were on average faster to respond when the 'earlier' key was above (M=1456 ms) than when it was below (1498 ms).

Analyses of data from Chinese-English bilinguals

In this section we analyze the data collected from the serendipitous sample of Chinese-English bilinguals (excluded from the main sample analyzed above). These data provide a rare opportunity. In general, when testing bilinguals, one must explicitly advertise for or recruit participants based on language experience, and participants may therefore be more explicitly aware of their own linguistic history and practices when they come to the lab than they might be in their everyday lives. In this case, participants coming into our study did not know that their prior language experience might be of interest to us, and we did not know their language background when we tested them (we conducted the language background questionnaires only after participants completed the task). Further all of the participants self-identified as native English speakers. As a result we have a rare serendipitous sample of bilinguals who were unaware that their language experience was an object of study.

Further, this sample allows us to examine what happens when metaphor patterns in natural language experience and metaphor patterns learned in the lab come into conflict. In this case, participants coming into the lab have had considerable natural language experience with metaphors that place earlier events above. What will happen when some of these participants are taught the opposite pattern in our metaphor training study? Will participants who already have a vertical representation of time be swayed by metaphors learned in the lab?

Participants rated their own language proficiency on a scale of 1 to 5 (with 5=fluent). On average participants rated their English skill at 4.9 (SD=.29) and their skill in Mandarin or Cantonese (whichever was more fluent) as 3.35 (SD=.99). All but two participants reported starting to learn both English and Chinese before age 5 (one began Chinese at 14, and one began English at 10).

Because this was an opportunistic sample, participants were not distributed evenly across conditions by interference type and some cells were missing participants. As a result, we were not able to conduct principled analyses by interference type. We did however have interpretable samples when collapsing across interference conditions. Thus, the mixed effects model included task condition (whether the 'earlier' key was above or below) and training condition (whether participants learned "earlier is up" or "earlier is down" metaphors) as fully crossed fixed effects, block order as a main effect, and participants and items as random effects.

As other participants in Experiments 1 and 2, Chinese-English bilinguals showed a congruency effect as a result of metaphor training. Participants were faster to respond when the key mapping in the Orly task was congruent with the metaphor they learned (M=1521 ms) than when the response mapping was incongruent with the metaphor learned (M=1552 ms). This was confirmed as a significant interaction between training condition and task condition (t(2220.68) = 2.34, p = .02). In addition to being faster on congruent trials, participants were also numerically more accurate on congruent (95.5%) than on incongruent trials (94.8%). There were no significant differences in metaphor training or Orly task accuracies by condition (all p > .23). There was no main effect of training condition (t(19.26) = .03, p=.98).

Interestingly, Chinese-English bilinguals showed a main effect of key-mapping; responses were significantly faster when the 'earlier' response key was above (M=1460 ms) than when it was below (M=1613 ms), as revealed in a main effect of task condition (t(2219.15) = 3.89, p<.001). This finding replicates prior work with Mandarin speakers tested on the same task (Boroditsky et al., 2010; Fuhrman et al., 2011). In Mandarin, naturally occurring vertical time metaphors place earlier events above and later events below.

However, recall that overall in Experiment 2, English speakers also showed a main effect of key-mapping, responding faster when the 'earlier' key was above than when it was below. Numerically, the size of this effect was considerably smaller in English speakers (42 ms) than in Chinese-English bilinguals (153 ms). Is there a significant difference by language background?

Comparing English speakers and Chinese-English bilinguals To analyze data from participants in Experiment 2 and data from Chinese-English bilinguals together, we constructed a mixed effects model that included language background (whether or not a participant spoke any Chinese language), training condition, and task condition as fully crossed fixed effects, block order as a main effect, and subjects and items as random effects.

There was an overall main effect of task condition (t(6217.17) = 3.93, p < .001), reflecting that on average participants responded faster when the ‘earlier’ key was above than when it was below. Importantly, Chinese-English bilinguals showed a significantly stronger “earlier is up” bias than did English speakers, as revealed in an interaction between language background and task condition ($t(6216.18) = 5.75$, $p < .001$). While participants were overall faster when the ‘earlier’ key was up, the Chinese-English bilinguals showed this effect significantly stronger than the English speakers.

There was also an interaction between task and training condition ($t(6216.08) = 8.12$, $p < .001$), which reflected that people were faster when the key-mapping in the Orly task and the trained metaphor were congruent than when they were incongruent. Interestingly, there was also a 3-way interaction between training condition, task condition, and language background ($t(6217.15) = 2.90$, $p = .004$), confirming that Chinese-English bilinguals indeed showed a significantly smaller congruency effect than the English speakers.

The pattern of results from Chinese-English bilinguals reveals an interesting combination pattern. We see influence of both prior language experience and of newly learned metaphors. Chinese-English bilinguals in this sample were faster to reply when the ‘earlier’ key was above than when it was below in all conditions, consistent with the natural pattern in Chinese time metaphors. However, this pattern was also moderated by the metaphors learned in the lab. Those who learned metaphors placing earlier events below showed a smaller “earlier is up” bias than those who learned metaphors placing earlier events above (as in Mandarin).

Further, Chinese-English bilinguals showed a smaller congruency effect than English speakers overall. This pattern suggests that training in the lab has a reduced effect when it must counteract the ballast of many years of natural language experience. For English speakers, the metaphors included in our training were relatively novel, and not in conflict with a strongly established metaphor system in English. For the Chinese-English bilinguals however, one of the metaphor training conditions was in direct conflict with a highly conventional metaphor system that they had already learned through natural language experience. These findings give us a glimpse of how learning new metaphors may interact with existing knowledge when new and old metaphors conflict.

Discussion

Participants learned new vertical metaphors for talking about time, and then completed a nonlinguistic implicit temporal judgment task. Although all stimuli and responses in this task were non-linguistic, they responded faster when the spatial arrangement of response keys was congruent with their newly learned metaphors than when they were incongruent. These results confirm that experience with metaphorical language alone (in the absence of other extra-linguistic cultural differences) can produce the same differences in temporal reasoning as seen between speakers of different languages.

Further this congruency effect was not disrupted by verbal interference. The locus of this language-induced effect does not appear to be in the creation of new internal *linguistic* routines (as for example is the case for counting). Instead, it appears that learning new metaphors for time fosters new non-linguistic representations. Results from the Orly task suggest that these representations contain some spatial properties, such as axis (vertical) and orientation (earlier is above or below). Future work may examine whether these representations also include metric properties like distance.

These results suggest that experience with language can affect thinking through qualitatively different routes. One type of influence is through the creation of internal linguistic routines that can participate online in the course of thinking (as for example is the case with counting). In other cases, experience with language can encourage people to construct and store new non-linguistic representations. Future work combining cross-linguistic approaches with controlled language learning studies will shed further light on the mechanisms through which language experience can foster new ways of thinking.

References

Bates, D., M. Maechler and B. Dai. 2008. *lme4: Linear mixed-effects models using S4 classes*. [R version 3.1.0 (2014-04-10).]

Boroditsky, L. (2011). How Languages Construct Time. In *Space, Time and Number in the Brain* (pp. 333–341). Elsevier.

Boroditsky, L., Fuhrman, O., & McCormick, K. (2010). Do English and Mandarin speakers think about time differently? *Cognition*, *118*(1), 123–129. doi:10.1016/j.cognition.2010.09.010

Frank, M. C., Fedorenko, E., Lai, P., Saxe, R., & Gibson, E. (2012). Verbal interference suppresses exact numerical representation. *Cognitive Psychology*, *64*(1-2), 74–92. doi:10.1016/j.cogpsych.2011.10.004

Fuhrman, O., McCormick, K., Chen, E., Jiang, H., Shu, D., Mao, S., & Boroditsky, L. (2011). How Linguistic and Cultural Forces Shape Conceptions of Time: English and Mandarin Time in 3D. *Cognitive Science*, *35*(7). doi:10.1111/j.1551-6709.2011.01193.x

Miles, L. K., Tan, L., Noble, G. D., Lumsden, J., & Macrae, C. N. (2011). Can a mind have two time lines? Exploring space–time mapping in Mandarin and English speakers. *Psychonomic Bulletin & Review*, *18*(3), 598–604. doi:10.3758/s13423-011-0068-y

Núñez, R. E., & Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science*, *30*(3), 401–450.