When *Tuesday* comes before *Threesday*: Cross-linguistic differences in numerical transparency of time words predicts temporal reasoning strategy and performance

Nian Liu (nian.liu@ou.edu)
Department of Modern Languages, Literatures, and Linguistics, University of Oklahoma
780 Van Vleet Oval, Norman, OK 73072, USA

Benjamin Bergen (bkbergen@cogsci.ucsd.edu)
Department of Cognitive Science, UC San Diego
9500 Gilman Drive, La Jolla, CA 92093, USA

Abstract

Time concepts are named differently across the world's languages. In English, the names for days of the week and months of the year are opaque—to people learning and using English, there's no obvious reason why Friday or September have the names they do. But in other languages, like Chinese, time concepts have numerically transparent names—the days of the week and months of the year are named using sequential numbers. We investigated whether having opaque versus mathematically transparent time concepts affects how people reason about time. Results show that Chinese speakers are more likely to spontaneously employ arithmetic when doing temporal calculations, which in turn improves the speed and accuracy of some time calculations. English speakers appear to use other strategies, such as sequential recitation.

Keywords: time concepts; temporal reasoning, mathematical ability; linguistic relativity.

Introduction

The world's languages encode time terms in different ways. In English, names of days of the week (DOW) or months of the year (MOY) are derived from planetary or mythical terms (Zerubavel, 1985), and as a result are largely opaque to contemporary users of the language—why for instance do Wednesday or April have the names they have? By contrast, many other languages exhibit more numerically transparent naming systems. Chinese Mandarin is exemplary of this pattern. It uses numbers in the names of months and days. For DOW, this system begins with Monday as the first day of the week. Thus, *xingqi yi*, "weekday one", is Monday, *xingqi er*, "weekday two", is Tuesday, etc. The only exception is Sunday, the last day of the week, *xingqi ri*, which translates to "weekday of sun". Similarly, months in Chinese are numbered from one to 12, with "month one" denoting January, "month two" February, and so on.

Do differences in how languages name time concepts cause differences in how speakers of those languages reason about time? For instance, does the numerical transparency of DOW and MOY terms affect the kinds of cognitive mechanisms people use to make temporal calculations? There has been little work on this issue. The most relevant line of work has focused on cross-linguistic effects of transparency of number systems themselves, showing that differences in number naming systems can affect cognitive development and non-linguistic performance. Most relevantly, acquisition studies (Miura et al., 1993, 1994; Miller et al., 1995; Paik & Mix, 2003) have found that preschool-aged children whose native languages employ more systematic naming systems for their numbers outperform their counterparts who speak languages that use less transparent number naming systems, on both number matching and number identification tasks. When asked to demonstrate numbers using combinations of individual unit cubes representing the quantity one and long blocks representing ten, Asian children whose languages use numerical names that are congruent with base 10 numeration systems (Fuson, 1990) were much more likely to use the blocks of ten in constructing multi-digit numbers than their counterparts, whose native naming systems were not similarly transparent. This led the authors of that study to argue that "numerical language characteristics may have a significant effect on cognitive representation of numbers" (Miura et al., 1994, p. 410), which in turn may enhance the performance of Asian-language-speaking children on tasks involving the concept of place value.

The types of names given to various symbolic systems, such as numbers, have also been shown to affect the problem solving abilities of competent symbol users. Seron and Fayol (1994) noticed that the number naming system in French-speaking Belgium is simpler than the one used in France (in Belgium, 98 is roughly "ninety-eight" but in France, it's "four-twenty-eighteen"). They reported that second-graders in France made more errors in number production than their Belgian counterparts. The effects of naming system also extend into adulthood and mathematical performance. For instance, adult English speakers have more difficulty reversing two-digit numbers ending in 1 (e.g., saying "14" when shown "41") than Chinese speakers do, presumably a result of English's idiosyncratic rules for naming numbers between 11 and 19 (Miller & Zhu, 1991).

In sum, differences in number naming systems affect the acquisition and use of number concepts. The current study investigated whether the same is true for the naming of time concepts (in this case, DOW and MOY). In particular, we asked whether the mathematically transparent naming of time concepts confers advantages on acquisition of time concepts and reasoning about time. There has been limited work on this question. Kelly et al. (1999) is the only systematic experimental investigation. They asked college students in China and the United States to name the day or month that occurs a specified length of time before or after another given day or month. Chinese college students performed these calculations faster than American college students. Kelly and colleagues argued that the difference resulted from the use of different strategies as a consequence of the naming systems used in Chinese and English. However, this argument about the mechanisms used was based on participants' self-reports. Moreover, the calculation distance in Kelly et al.'s (1999) experiment was held constant—4 for the DOW task and 7 for the MOY task. The use of the constant distance caused some participants to predict the answers; one of the strategies participants reported was memorizing specific pairs of items. The experiment described below, in which distances vary, indirectly assesses the mechanisms adults use to perform temporal reasoning tasks, and whether these mechanisms vary with native language.

Testing mechanisms indirectly

Kelly et al., (1999) reported that adult Chinese speakers outperform their English-speaking counterparts in time calculation tasks. This could be due to the differences in the numerical transparency of time words in the languages they speak. That is, the transparent numerical structure of Chinese time words might facilitate time calculation, by allowing Chinese speakers to employ arithmetic strategies made possible by the use of numerical names. For example, "Four days after Monday is what day?" translates to "*Four* days after Weekday *one* is what day?" To the Chinese speaker, this directly evokes arithmetic, 4+1, and as a result, they might be able to use arithmetic quickly to produce the answer: "Weekday five" (Friday). The same should be true of making calculations about months. By contrast, English speakers do not have the arithmetic laid out for them in tasks like this, so they might rely on alternate strategies, like reciting a sequence of days or months.

But Kelly et al.'s (1999) results could alternatively be the result of other cultural differences, for instance, in the depth, length, and nature of the math education each group receives. It might be that the Chinese population that was sampled simply was better at performing mental calculations than their English-speaking counterparts.

Our experiment teases apart these two possibilities with a nuanced design, based on three predictions. First, if the use of different arithmetic strategies by Chinese and English speakers is responsible for the performance difference, then the Chinese advantage should disappear in cross-week or cross-year calculations. Chinese speakers may encounter difficulties in calculating distances across boundaries since they have to convert the answers into modulo-7 or modulo-12 systems; 3 days after "weekday 5" is not "weekday 8", but rather 8 modulo 7, thus "weekday 1"; 3 months after "month 11" is not "month 14", but 14 modulo 12, thus "month 2". If English speakers use a non-arithmetic strategy by default, then they should exhibit less increase in difficulty when calculating across boundaries.

A second prediction of the hypothesis that Chinese speakers use arithmetic more than English speakers is that the calculation of distances from Sunday (which is called "weekday sun" in Chinese) should be more difficult for Chinese speakers than from any other day in a week, since number is not used in naming this day. This irregularity may cause trouble in applying the arithmetic strategy, thus slowing down Chinese speakers' calculations, compared to calculation involving other days of the week. Again, by contrast, English speakers should show no increased difficulty when making calculations relative to Sunday.

A third prediction is that the Chinese speakers' speed of calendar calculating should not vary much with longer temporal distance if they primarily use an arithmetic method in the calculation, while the English speakers' performance may be negatively affected by increases in temporal distance to be calculated if they are reciting sequences.

Method

Participants

Thirty-two (22 female and 10 male) native Chinese speakers, college students from Beijing United College, ranging in age from 18 to 31 (M = 21, SD = 4.9), and 40 native English speakers (19 female and 21 male), undergraduate students at the University of Hawaiʻi, aged 18 to 29 (M = 21, SD = 2.5), participated in the experiment either for extra credit in an introductory linguistics class or for five dollars or the equivalent.

Materials and design

Each participant completed two temporal reasoning tasks, pertaining to DOW and MOY respectively. For each, we manipulated two factors—Boundary Type (Within/Across boundary) and Direction (Forward/Backward), which produced four question types, as below (showing only DOW). Both factors were manipulated within participants. Language was a between-participants factor. An additional factor, Sunday, applied only to the DOW blocks. Half of the DOW questions involved Sunday, to reveal eventual effects of this non-numerically-named day in Chinese.

(1a) If today is Monday, three days from now is what day? (Within/Forward)
(1b) If today is Thursday, two days ago was what day? (Within/Backward)
(2a) If today is Saturday, three days from now is what day? (Across/Forward)
(2b) If today is Tuesday, five days ago was what day? (Across/Backward)

Calculation distance differed across conditions; distance ranged from 1 to 7 for DOW questions (1–4 for Within and 2–7 for Across) and 2 to 12 for MOY questions (2–10 for Within and 2–12 for Across) in order to match the cyclical nature of the weeks and months. There were 32 questions in the DOW block and 48 in the MOY block.

Procedure

Participants were seated in front of a computer. Each trial began with a fixation cross in the center of the screen. Participants initiated the trial by pressing the SPACE bar when they were ready. They heard a recorded voice read a question to them over headphones. They were instructed to speak the answer as quickly and accurately as possible after hearing the question into a microphone, which was used as a voice key, connected to an E-Prime SR-BOX. A digital recorder was also used to record the answers.

Four practice items preceded the experiment, which was divided into two blocks, one for DOW and one for MOY. Block order was counterbalanced across participants. The questions were randomized within each block. There was a short break between blocks.

Measures

Reaction times in milliseconds were measured from the offset of the question to the onset of the participant's answer. Accuracy (individual proportions of correct answers were arcsine transformed to comply with the normal distribution premise) were calculated as another dependent measure.

Results

All filler syllables, tongue clicks, partial responses, and repeated responses (due to failure to trigger the voice key) were manually excluded. No participants were excluded because of outlying RTs or low accuracy.

Error analysis

Figure 1 shows mean accuracy for each Boundary Type and Direction by language in the MOY and DOW tasks. We performed separate repeated-measures ANOVAs with participants (F_1) and items (F_2) as random factors. In the participants analysis, Language (Chinese/English) was between participants but Direction (Forward/Backward), and Boundary Type (Within/Across) were within subjects. For items, Language was a within-items factor and the other factors (Boundary Type and Direction) were between-items.

As predicted, Chinese speakers made fewer errors than English speakers on the MOY task—a main effect of Language: $F_1(1,68) = 26.846$, $p < 0.001$, $F_2(1, 44) = 54.265$, $p < 0.001$. However, as Figure 1 shows, this advantage was carried by the Within Boundary questions, confirmed by an interaction between Language and Boundary Type, F_1 (1, 68) = 34.842, $p < 0.001$; F_2 (1, 44) = 44.876, $p < 0.001$.

As the second part of Figure 1 shows, Language was a less robust factor in the DOW task. The main effect of Language did not reach significance by participants analysis, but it did by items, $F_2(1, 28) = 5.177$, $p = 0.031$. This indicates that Chinese speakers did not have a global accuracy advantage over English speakers. The lack of Language effect was probably driven by the Chinese speakers' relatively poor performance in the Across-Boundary condition. As with MOY, an interaction effect was found between Language and Boundary Type, $F_1(1,68) = 18.537$, $p < 0.001$; $F_2(1,28) = 24.596$, $p < 0.001$. Once again, Chinese speakers produced more errors in cross-week calculations compared to their English-speaking counterparts.

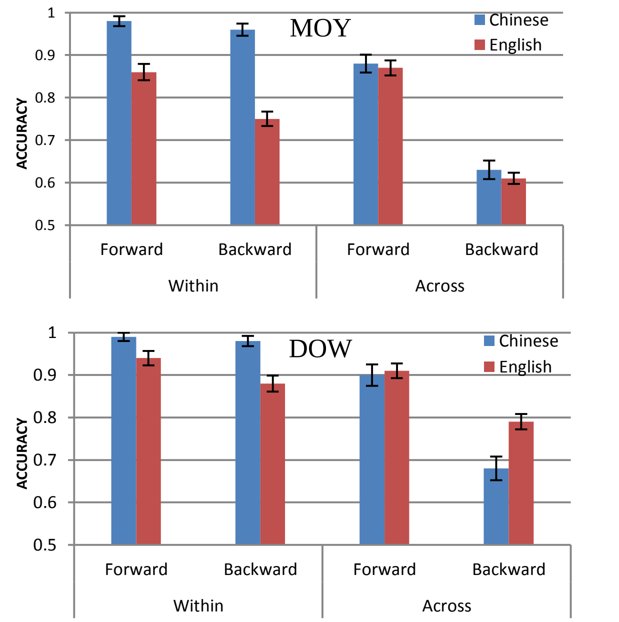

Figure 1: Accuracy rate on the MOY (upper) and DOW (lower) task for Chinese and English speakers.

Questions involving Sunday As discussed in the Method section, half of the DOW questions were Sunday-related questions. These questions could have been the cause of the Chinese speakers' low DOW accuracy if the irregular Sunday term ("weekday sun" instead of "weekday seven") affected the use of an calculation strategy. Statistical analysis confirms this. Chinese speakers were significantly less accurate on Sunday questions than non-Sunday questions, $F_1(1, 29) = 2.38$, $p = 0.024$; $F_2(1, 30) = 2.83$, $p = 0.031$, while English speakers were not sensitive to Sunday and showed no difference in accuracy on Sunday questions (see Figure 2).

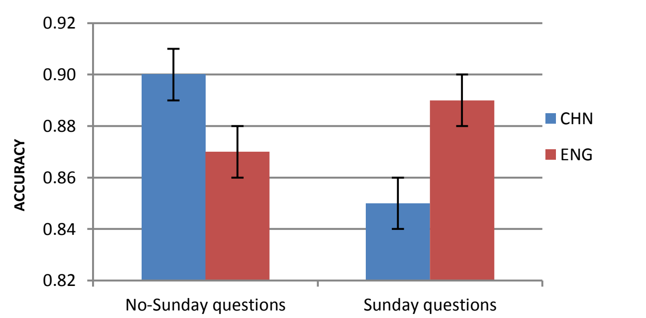

Figure 2: Chinese- and English-speaking participants' accuracy in different Sunday conditions.

Reaction time analysis

All trials with incorrect responses were removed (16.7% of the data). We also removed all responses that were greater than 2.5 standard deviations from the mean of all responses in each of the four conditions (Boundary Type x Direction) for each language group. This excluded another 2.72% of the data. No participants or items were removed for reasons of accuracy or outlying *SD*. The reaction time data approximate normal distribution after this cleaning. Reaction times for correct responses were analyzed using a repeated-measures ANOVA with Boundary Type (Across/Within) and Direction (Forward/Backward) as within-subject factors and Language (Chinese/English) as a between-subject factor. An Item analysis took Language as a within-items factor and Boundary Type and Direction as between-items factors. The results in the MOY and DOW tasks, seen in Figure 3, show mean reaction time for each Boundary Type and Direction by language.

The results from MOY block mirrored the error analysis. A large main effect of Language, $F_1(1, 68) = 36.155$, $p < 0.001$, $F_2(1, 44) = 134.986$, $p < 0.001$, confirmed that Chinese speakers were faster than English speakers. There were also significant interactions between Language and Boundary Type, $F_1(1, 68) = 35.809$, $p < 0.001$, $F_2(1, 44) = 9.613$, $p < 0.001$, and between Language and Direction, $F_1(1, 68) = 13.\ 423$, $p < 0.001$, $F_2(1, 44) = 8.407$, $p = 0.006$. These effects showed that Chinese speakers gained more of a speed advantage from Within month questions than English speakers did, but that English speakers gained more of an advantage from Forward calculations.

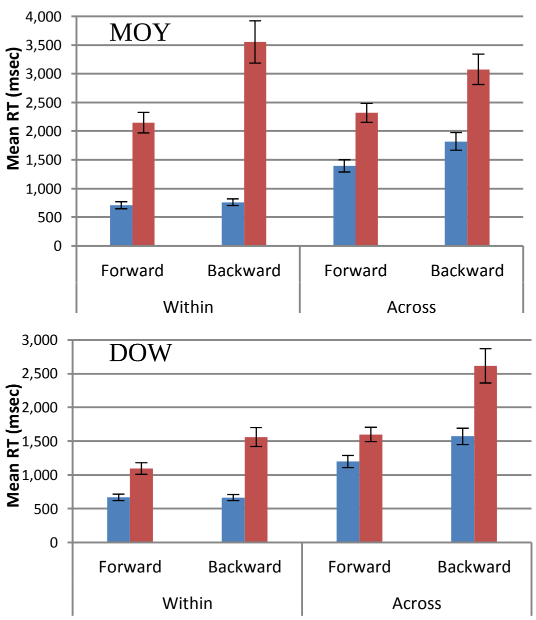

Figure 3: Reaction times on the MOY (upper) and DOW (lower) task for Chinese and English speakers.

Reaction time results from the DOW task are more complicated because of the presence of Sunday. The Chinese speakers were consistently faster than their English-speaking counterparts, as shown by a main effect of language, $F_1(1,68) = 20.617$, $p < 0.001$, $F_2(1,28) = 63.776$, $p < 0.001$. Both language groups answered Within questions faster than the Across ones, $F_1(1,68) = 161.850$, $p < 0.001$, $F_2(1,28) = 27.446$, $p < 0.001$, and the Forward questions faster than the Backward ones, $F_1(1,68) = 42.243$, $p < 0.001$, $F_2(1,28) = 11.496$, $p = 0.002$. There was also an interaction between Language and Direction, $F_1(1,68) = 15.407$, $p < 0.001$, $F_2(1,28) = 9.244$, $p = 0.005$. English speakers were much faster with Forward questions, $F_1(1,39) = 86.977$, $p < 0.001$, $F_2(1,28) = 13.090$, $p = 0.001$. Chinese speakers were also faster in answering Forward questions than Backward questions, but the difference was much smaller, $F_1(1,29) = 13.460$, $p = 0.001$, $F_2(1,28) = 4.933$, $p = 0.035$.

Questions involving Sunday The analysis of the accuracy data above showed that Chinese speakers made more errors on questions involving Sunday. But their reaction times were unaffected, as Figure 4 shows. A two-way ANOVA showed a main effect of Language, $F_1(1,68) = 15.681$, $p < 0.001$, $F_2(1, 30) = 34.942$, $p < 0.001$; Chinese speakers were faster than English speakers. But there was no interaction.

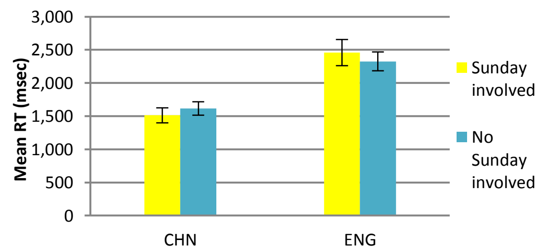

Figure 4: Reaction times in Sunday and Non-Sunday conditions for Chinese and English speakers.

Calculation Distance analysis

Because the mean distance varied across conditions, we also did exploratory analyses by calculation distance, to see if different calculation difficulty revealed strategy differences. Reaction times by distance are presented in Figure 5.

As shown in Figure 5, the two language groups showed different level of sensitivity to the distance calculated. The Chinese speakers' reaction times were far less sensitive to either distance or direction, especially in the within-week condition. This is consistent with the hypothesis that they were using addition and subtraction of small numbers. The across-week condition is more complicated, as an extra step of modulo calculation could have been involved for the Chinese speakers.

The English speakers' reaction times were more strongly affected by the length of temporal distances. As can be seen on the right side of Figure 5, questions that required backwards calculations took much longer than Forward questions, presumably because counting backwards is less familiar. Moreover, reaction times increase steadily as the distances increase. The rise and drop of reaction times with longer distances suggests that the English speakers were applying different approaches when encountering calendar questions with different distances. When the distances grew longer, which makes verbal list counting a less efficient strategy, they may have flexibly and strategically switched to methods such as using numerical equivalents as a shortcut to reciting the list, such as "counting forward by the

12's or 7's complement to solve backward problems" (Kelly et al., 1999). The partial use of an arithmetic strategy was self-reported by some of the English participants in Kelly et al.'s work, and is consistent with the reaction time evidence we find here.

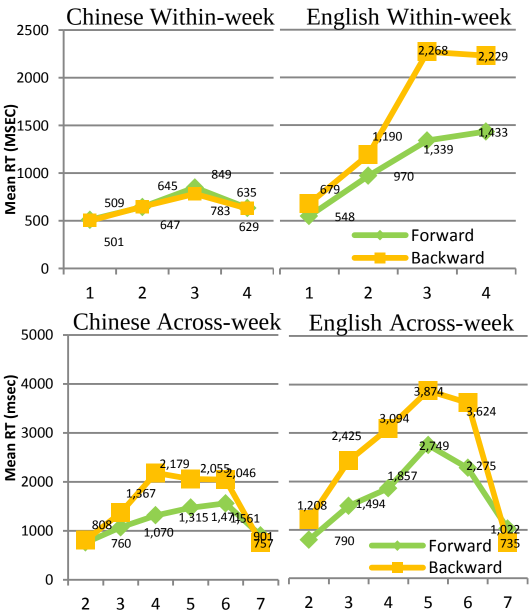

Figure 5: Reaction times of Chinese (left) and English (right) in Within (upper) and Across (lower) Week calculations (The *x*-axis shows calculation distances.)

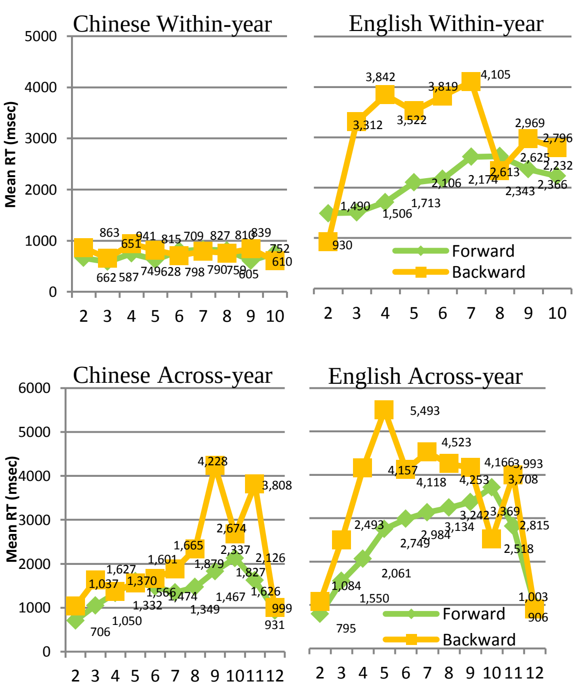

Figure 6: Reaction times by Chinese (left) and English (right) speakers in Within (upper) and Across (lower) Year calculations (The *x*-axis shows calculation distances.)

The year cycle has longer distances and so can provide more information about strategies used in response to possible calculation difficulties. The two language groups' reaction time data were categorized by distance for the Within task and for the Across task, as shown in Figure 6.

Again, the Chinese speakers' reaction times were not affected nearly as much as the English speakers' were by calculation distance or direction, especially in the Within Year tasks, again consistent with the possibility that they are calculating arithmetically. On the other hand, the English speakers present a more complicated picture of responding to different distances. They were generally faster with Forward than Backward questions, and they spent more time when the distances got longer, except for the longest distances (10 through 12), with which they might be using strategies other than reciting due to the difficulty of counting all the way through the month list.

Discussion

Languages differ in how numerically transparent their time words are, and we hypothesized that these differences would affect the temporal cognition of adult speakers of languages with distinct systems. Our results provide several types of experimental evidence that Chinese and English speakers use different strategies in temporal calculation tasks. As a result, Chinese speakers determine temporal distance calculations faster and more accurately. These findings are not consistent with the hypothesis that other factors, such as general mathematics ability, cause differences in overall accuracy or speed.

There are three pieces of relevant evidence. First, Chinese speakers were consistently more accurate in the Within Week and Within Year tasks, but their accuracy dropped significantly and their advantage disappeared with calculations across week or year boundaries. In contrast, the English speakers' answers were not particularly sensitive to boundary crossings. For the reaction time results, although the Chinese speakers were still faster than the English speakers in the Across Boundary calculations, they were slower compared to their own Within Week calculations. So although arithmetic strategies for time may be advantageous for some local calculations, the Chinese speakers' advantage in accuracy and speed diminishes when answering questions that involve temporal boundary crossings.

Second, the results confirmed that the irregularity of Sunday's Chinese name causes trouble in applying the arithmetic strategy, resulting in more mistakes by Chinese speakers in answering questions involving Sunday compared to calculations involving other days of the week. By contrast, English speakers showed no increased difficulty when making calculations relative to Sunday, and their accuracy rates were not affected by questions involving Sunday.

Finally, Chinese speakers' reactions seem not to systematically relate to distance or directions of calculation, whereas the English speakers' reaction times for Forward calculations increase steadily and substantially, implying an increase in time spent reciting day or month lists, and their pattern on Backward questions were more complicated. We hypothesized that versatile/mixed strategies might be used due to the difficulty of reciting lists backwards.

In sum, these results suggest that differences in calendar terms between the two languages lead to dominant arithmetic or list-reciting strategies for speakers of those languages. The verbal list strategy is substantially slower than the number-transferring one, yet the latter results in more possible errors across boundaries because of required modulo 7 and modulo 12 calculations.

Previous cross-linguistic studies have shown that there is at least some causal influence from language to non-verbal cognition and unconscious habitual thought (Kay & Kempton, 1984; Lucy, 1992; Gumperz & Levinson, 1996, etc.). More specifically, the way time is described in a language can affect its speakers' conceptualization of time (Boroditsky, 2001) and can even shape low-level mental processes in psychophysical tasks (Casasanto et al., in revision). The current study adds to this line of study by showing that transparent numerical structure of the calendar might facilitate calendar calculation, causing Chinese-speaking adults to outperform their English-speaking counterparts in time calculation tasks, by exhibiting shorter reaction times and making fewer errors. We conclude that adults' temporal reasoning abilities differ, depending on the transparency of the naming systems that their languages employ for time sequences. In general, such a finding supports the hypothesis that linguistic differences can produce non-linguistic consequences, in this case in affecting people's reasoning about time (Boroditsky, 2000, 2003; Boroditsky & Ramscar, 2002, Matlock et al., 2005; Núñez and Sweetser, 2006; Casasanto & Boroditsky, 2008).

Acknowledgments

This study was funded by National Science Foundation Doctoral Dissertation Grant #1124006.

References

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

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

Boroditsky, L. (2003). Linguistic relativity. *Encyclopedia of Cognitive Science*. MacMillan Press.

Boroditsky, L. & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psych. Science*, 13(2), 185-188.

Casasanto, D., & Boroditsky, L. (2008).Time in the mind: Using space to think about time. *Cognition 106*, 579-593.

Casasanto, D., Fotakopoulu, O., Pita, R., & Boroditsky, L. (In Revision). How deep are effects of language on thought? Time estimation in speakers of English and Greek. *Cognition*.

Friedman, W. J. (1990). *About time: Inventing the fourth dimension*. Cambridge, MA: MIT Press.

Fuson, K.C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction*, 7, 343-403.

Gumperz, J. & Levinson, S. (1996). *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.

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

Kelly, M., Miller, K., Fang, G., & Feng, G. (1999). When days are numbered: calendar structure and the development of calendar processing in English and Chinese. *J. of Experimental Child Psych. 73*, 289-314.

Lucy, J. (1992). Language diversity and thought. *A reformulation of the linguistic relativity hypothesis.* Cambridge University Press.

Matlock, T., Ramscar, M., & Boroditsky, L. (2005). On the experiential link between spatial and temporal language. *Cognitive Science*, 29, 655-664.

Miller, K. F., & Zhu, J. (1991). The trouble with teens: Accessing the structure of number names. *Journal of Memory and Language*, 30, 48-68.

Miller, K. F., Smith, C. M., Zhu, J. J., & Zhang, H. C. (1995). Preschool origins of cross-national differences in mathematical competence—the role of number-naming systems. *Psychological Science*, 6, 56-60.

Miura, I., Okamoto, Y., Kim, C., Steere, M., & Fayol, M. (1993). First graders' cognitive representation of number and understanding of place value: Cross-national comparisons—France, Japan, Korea, Sweden, and the United States. *J. of Educational Psychology*, 85, 24-30.

Miura, I., Okamoto, Y., Kim, C., Chang, C-M., Steere, M., & Fayol, M. (1994). Comparisons of children's cognitive representation of number: China, France, Korea, Sweden, and the United States. *International Journal of Behavioral Development*, 17, 401-411.

Núñez, R., & 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.

Paik, Jae H., & Mix, Kelly S. (2003) U.S. and Korean children's comprehension of fraction names: A reexamination of cross-national Differences. *Child Development*, 74(1), 144-154.

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology*, 12, 281-300.

Zerubeval, E. (1985). *The seven day circle: The history and meaning of the week*. Chicago: Univ. of Chicago Press.